taxonomic units and estimating species richness, *Appl. Environ. Microbiol.*, **71**, 1501–6.
62. del Campo, R., Bravo, D., Cantón, R., et al. 2005, Scarce evidence of yogurt lactic acid bacteria in human feces after daily yogurt consumption by healthy volunteers, *Appl. Environ. Microbiol.*, **71**, 547–9.
63. Oozeer, R., Leplingard, A., Mater, D.D., et al. 2006, Survival of *Lactobacillus casei* in the human digestive tract after consumption of fermented milk, *Appl. Environ. Microbiol.*, **72**, 5615–7.
64. Marco, M.L., de Vries, M.C., Wels, M., et al. 2010, Convergence in probiotic Lactobacillus gut-adaptive responses in humans and mice, *ISME J.*, **4**, 1481–4.
65. Manichanh, C., Rigottier-Gois, L., Bonnaud, E., et al. 2006, Reduced diversity of fecal microbiota in Crohn's disease revealed by a metagenomic approach, *Gut*, **55**, 205–11.
66. Turnbaugh, P.J., Hamady, M.H., Yatsuneko, T., et al. 2009, A core gut microbiome in obese and lean twins, *Nature*, **475**, 480–4.
67. Kurokawa, K., Itoh, T., Kuwahara, T., et al. 2007, Comparative metagenomics revealed commonly enriched gene sets in human gut microbiomes, *DNA Res.*, **14**, 169–81.
68. Jernberg, C., Löfmark, S., Edlund, C. and Jansson, J.K. 2007, Long-term ecological impacts of antibiotic administration on the human intestinal microbiota, *ISME J.*, **1**, 56–66.
69. Wu, G.D., Chen, J., Hoffmann, C., et al. 2011, Linking long-term dietary patterns with gut microbial enterotypes, *Science*, **334**, 105–8.
70. Sokol, H., Seksik, P., Furet, J.P., et al. 2009, Low counts of *Faecalibacterium prausnitzii* in colitis microbiota, *Inflamm. Bowel Dis.*, **15**, 1183–9.

DNA Research **21**, 15–25, (2014) doi:10.1093/dnares/dst037
Advance Access publication on 7 September 2013

Dysbiosis of Salivary Microbiota in Inflammatory Bowel Disease and Its Association With Oral Immunological Biomarkers

Heba S. Said[1], Wataru Suda[1], Shigeki Nakagome[2], Hiroshi Chinen[3], Kenshiro Oshima[1], Sangwan Kim[1], Ryosuke Kimura[4], Atsushi Iraha[3], Hajime Ishida[4], Jiro Fujita[5], Shuhei Mano[2], Hidetoshi Morita[6], Taeko Dohi[7], Hiroki Oota[8], and Masahira Hattori[1,*]

Department of Computational Biology, Graduate School of Frontier Sciences, The University of Tokyo, Kashiwanoha 5-1-5, Kashiwa, Chiba 277-8561, Japan[1]; Risk Analysis Research Center, The Institute of Statistical Mathematics, 10-3, Midori-cho, Tachikawa, Tokyo 190-8562, Japan[2]; University Hospital, Faculty of Medicine, University of the Ryukyus, Uehara 207, Nishihara, Okinawa 903-0215, Japan[3]; Department of Human Biology and Anatomy, Graduate School of Medicine, University of the Ryukyus, Uehara 207, Nishihara, Okinawa 903-0215, Japan[4]; Department of Infectious, Respiratory, and Digestive Medicine, Control and Prevention of Infectious Diseases, Graduate School of Medicine, University of the Ryukyus, Uehara 207, Nishihara, Okinawa 903-0215, Japan[5]; School of Veterinary Medicine, Azabu University, Fuchinobe 1-17-71, Chuo-ku, Sagamihara, Kanagawa 252-5201, Japan[6]; Department of Gastroenterology, Research Center for Hepatitis and Immunology, Research Institute, National Center for Global Health and Medicine, Kohnodai 1-7-1, Ichikawa, Chiba 272-8516, Japan[7] and Laboratory of Genome Anthropology, Department of Anatomy, Kitasato University School of Medicine, Kitasato 1-15-1, Minami-ku, Sagamihara, Kanagawa 252-0674, Japan[8]

*To whom correspondence should be addressed. Tel. +81 4-7136-4070. Fax. +81 4-7136-4080.
Email: hattori@k.u-tokyo.ac.jp

Edited by Dr Katsumi Isono
(Received 14 July 2013; accepted 12 August 2013)

Abstract

Analysis of microbiota in various biological and environmental samples under a variety of conditions has recently become more practical due to remarkable advances in next-generation sequencing. Changes leading to specific biological states including some of the more complex diseases can now be characterized with relative ease. It is known that gut microbiota is involved in the pathogenesis of inflammatory bowel disease (IBD), mainly Crohn's disease and ulcerative colitis, exhibiting symptoms in the gastrointestinal tract. Recent studies also showed increased frequency of oral manifestations among IBD patients, indicating aberrations in the oral microbiota. Based on these observations, we analyzed the composition of salivary microbiota of 35 IBD patients by 454 pyrosequencing of the bacterial 16S rRNA gene and compared it with that of 24 healthy controls (HCs). The results showed that Bacteroidetes was significantly increased with a concurrent decrease in Proteobacteria in the salivary microbiota of IBD patients. The dominant genera, *Streptococcus*, *Prevotella*, *Neisseria*, *Haemophilus*, *Veillonella*, and *Gemella*, were found to largely contribute to dysbiosis (dysbacteriosis) observed in the salivary microbiota of IBD patients. Analysis of immunological biomarkers in the saliva of IBD patients showed elevated levels of many inflammatory cytokines and immunoglobulin A, and a lower lysozyme level. A strong correlation was shown between lysozyme and IL-1β levels and the relative abundance of *Streptococcus*, *Prevotella*, *Haemophilus* and *Veillonella*. Our data demonstrate that dysbiosis of salivary microbiota is associated with inflammatory responses in IBD patients, suggesting that it is possibly linked to dysbiosis of their gut microbiota.

Key words: Crohn's disease; ulcerative colitis; salivary microbiota; 16S rRNA; pyrosequencing

1. Introduction

Current advances of next-generation sequencing technologies (NGS) have enabled us to acquire massive DNA sequence data from any types of samples.[1] In particular, complex bacterial communities composed of numerous species in various environments including human body has become the practically feasible targets, and the analysis has been shifting to the DNA-based approach in conjugation with bioinformatics for enumerated data of metagenome and 16S rRNA gene (16S) produced by NGS.[2–5] Among these approaches, pyrosequencing-based 16S gene analysis is rapid and cost effective to comprehensively evaluate the overall structure of bacterial communities and to identify species present in them, irrespective of the yet-uncultured species.[6] This method includes targeted PCR amplification of 16S rRNA gene variable regions with appropriate primers, followed by sequencing of the 16S amplicons using 454 pyrosequencer.[7–10] We recently developed the improved analytical pipeline for pyrosequencing data of 16S rRNA gene V1–V2 variable region for human gut microbiota, by reassessing a PCR primer sequence, clustering conditions of error-prone 16S reads, and the quality check process to effectively remove low-quality data, and thereby the pipeline provided the high quantitative accuracy to estimation of the bacterial composition and abundance in the community.[10]

In this study, we applied our improved pipeline to the analysis of the human oral microbiota. The oral cavity is a large reservoir of bacteria of >700 species or phylotypes, and is profoundly relevant to host health and disease.[11–14] Current studies reported that various oral symptoms such as aphthous stomatitis, oral ulcer, dry mouth, and pyostomatitis vegetans are frequently observed in inflammatory bowel disease (IBD) patients.[15–20] IBD, including Crohn's disease (CD) and ulcerative colitis (UC), is a chronic, idiopathic, relapsing inflammatory disorder of the gastrointestinal tract.[21,22] The most widely accepted mechanism of IBD pathogenesis includes inflammation due to altered host immune response in association with continuous stimulation from the resident gut microbiota.[23–28] Many studies also revealed that the gut microbiota of IBD patients significantly differed from that of healthy controls (HCs), and is termed dysbiosis.[29–34]

Similarly, oral manifestations observed in IBD patients suggest the association of oral microbiota with such manifestations, yet-limited information exists about the oral microbiota of IBD patients. We characterized the salivary microbiota of IBD patients and HCs by barcoded pyrosequencing analysis of the bacterial 16S rRNA gene. We observed that the salivary microbiota in IBD patients significantly differed from that of HCs, and found particular bacterial species associated with dysbiosis. We also showed that the observed dysbiosis is strongly associated with elevated inflammatory response of several cytokines with depleted lysozyme in the saliva of IBD patients, some of which showed a strong correlation with the relative abundance of certain bacterial species. Thus, the present study demonstrates an association between dysbiosis of the salivary microbiota and change in the host's physiological state in IBD.

2. Material and methods

2.1. *Patients and control groups*

All participants of the CD, UC, and HC groups were informed of the purpose of this study, and written consent was obtained. This project was approved by the ethical committee of University of the Ryukyus. Metadata collected at the time of sampling included various demographics and a medication history for each patient (Supplementary Tables S1 and S2).

2.2. *Sample collection and DNA extraction*

Unstimulated saliva collected from subjects was immediately frozen by liquid nitrogen and stored in −80°C until use. Salivary genomic DNA was prepared according to the literature with minor modifications.[35]

Bacterial cells were harvested from 1 ml of saliva by centrifugation at 3300*g* for 10 min at 4°C. Bacterial pellets were suspended in 10 mM Tris–HCl/10 mM EDTA buffer and incubated with 15 mg/ml lysozyme (Sigma-Aldrich Co. LLC) for 1 h at 37°C. Purified achromopeptidase (Wako Pure Chemical Industries, Ltd.) was added to a final concentration of 2000 units/ml and samples were further incubated for 30 min. Ten percentage of (wt/vol) sodium dodecyl sulphate (SDS) and proteinase K (Merck Japan) were added to the suspension to final concentrations of 1% and 1 mg/ml, respectively, and samples were further incubated at 55°C for 1 h. The lysate was treated with phenol/chloroform/isoamyl alcohol (Life Technologies Japan, Ltd.) and centrifuged at 3300*g* for 10 min. DNA was precipitated by adding 1/10 volume of 3 M sodium acetate (pH 4.5) and 2 volumes of ethanol (Wako Pure Chemical Industries, Ltd.) to the supernatant. DNA was pelleted by centrifugation at 3300*g* for 15 min at 4°C. DNA pellets were rinsed with 75% ethanol, dried and dissolved in 10 mM Tris–HCl/1 mM EDTA (TE) buffer. DNA was further treated with 1 mg/ml RNase A (Wako Pure Chemical Industries, Ltd.) at 37°C for 30 min, and precipitated by adding equal volumes of 20% PEG solution (PEG6000-2.5M NaCl). DNA was pelleted by centrifugation at 8060*g* at 4°C, rinsed twice with 75% ethanol, dried, and dissolved in TE buffer.

2.3. *Bacterial 16S rRNA gene-based analysis*

2.3.1. PCR amplification of the 16S rRNA gene V1–V2 region and barcoded 454 pyrosequencing The hypervariable V1–V2 region of the 16S rRNA gene was amplified by PCR with barcoded 27Fmod and 338R primers.[10] PCR was performed in 50 μl of 1× Ex Taq PCR buffer composed of 10 mM Tris–HCl (pH 8.3), 50 mM KCl, and 1.5 mM $MgCl_2$ in the presence of 250 μM dNTP, 1 U Ex Taq polymerase (Takara Bio, Inc.), forward and reverse primers (0.2 μM) and ~20 ng template DNA. Thermal cycling consisted of initial denaturation at 96°C for 2 min, followed by 25 cycles of denaturation at 96°C for 30 s, annealing at 55°C for 45 s and extension at 72°C for 1 min, and final extension at 72°C on a 9700 PCR system (Life Technologies Japan, Ltd.). Negative controls were treated similarly, except that no template DNA was added to the PCR reactions. PCR products of ~370 bp were visualized by electrophoresis on 2% agarose gels, while negative controls failed to produce visible PCR products and were excluded from further analysis. PCR amplicons were purified by AMPure XP magnetic purification beads (Beckman Coulter, Inc.), and quantified using the Quant-iT PicoGreen dsDNA Assay Kit (Life Technologies Japan, Ltd.). Equal amounts of each PCR amplicon were mixed and then sequenced using either 454 GS FLX Titanium or 454 GS JUNIOR (Roche Applied Science).

2.3.2 Analysis pipeline for 16S data We developed and used an analysis pipeline for pyrosequencing data of the 16S rRNA gene V1–V2 region generated from oral microbiota. Based on sample specific barcodes, reads were assigned to each sample followed by the removal of reads lacking both forward and reverse primer sequences. Data were further denoised by removal of reads with average quality values <25 and possible chimeric sequences. For chimera checking and taxonomy assignment of the 16S rRNA data, we constructed our own databases from three publically available databases: Ribosomal Database Project (RDP) v. 10.27, CORE (http://microbiome.osu.edu/), and a reference genome sequence database obtained from the NCBI FTP site (ftp://ftp.ncbi.nih.gov/genbank/, December 2011). Reads having BLAST match lengths <90% with the representative sequence in the three databases were considered as chimeras and removed. Finally, filter-passed reads were used for further analysis after trimming off both primer sequences.

All of the 16S rRNA sequence data used in this study were deposited in DDBJ/GenBank/EMBL under accession numbers: DRA000984–DRA000986.

2.3.3. Operational taxonomic unit clustering and UniFrac analysis From the filter-passed reads, 3000 high-quality reads/sample were randomly chosen. The total reads (59 × 3000 reads) were then sorted according to average quality value and grouped into operational taxonomic units (OTUs) using UCLUST (http://www.drive5.com/) with a sequence identity threshold of 96%. Taxonomic assignments were made according to the best BLAST-hit phylotype. Weighted and unweighted UniFrac metrics[36] were used to assess the diversity of the salivary microbiota between the CD, UC, and HC groups. UniFrac distances were based on the fraction of branch length shared between two communities within a phylogenetic tree constructed from the 16S rRNA gene sequences from all communities being compared.

2.4. *Immunoassays*

The centrifugal supernatant of unstimulated saliva was analyzed by the Luminex fluorescence technique, using the Bio-Plex Pro Human cytokine 27-Plex Assay (Bio-Rad Laboratories, Inc.) according to the manufacturer's instructions. LL-37 (cathelicidin, hCAP-18) levels were measured by ELISA using the Human LL-37 ELISA Kit (Hycult Biotech, Uden, The Netherlands). IgA levels were measured using the EIA-sIgA Test (MBL, Nagoya, Japan). Salivary lysozyme levels were measured using turbidimetric technique (SRL Inc., Japan). Total protein concentrations were measured by the Bradford protein assay using bovine serum albumin as the standard. In this study, saliva samples of only 15 HC, 14 CD, and 10 UC subjects were used for the assay of biomarkers, because the saliva from the other subjects was insufficient for measurement of all the indicated biomarkers.

2.5. *Statistical analysis*

All statistical analyses were conducted with R version 2.15.2. Microbial richness, evenness, and diversity were assessed using the R Vegan package. Depending on the normality of the data, the Student's *t*-test or Mann-Whitney's U-test was used to perform statistical analysis. *P*-values were corrected for multiple testing using the Benjamini–Hochberg method. Correlations between relative abundance of genera and immunological markers in saliva were calculated by Pearson correlation coefficients.

3. Results

3.1. *Collection of 16S data*

We surveyed the salivary microbiota of 21 CD patients, 14 UC patients, and 24 HCs, all of whom (including their relatives) are residents, lasting at least three generations, of the Okinawa area in Japan. The general and clinical parameters of the study populations are given in Supplementary Table S1, and individual details are shown in Supplementary Table S2.

Sample-assigned pyrosequencing reads having both forward and reverse primer sequences accounted for ~60% of the total number of reads. The 16S reads having average quality values < 25 and possibly chimeric sequences represented 0.75 and 0.46% of the selected dataset, respectively. Finally, 506 133 high-quality 16S reads were obtained from 59 salivary samples. Sorting of the 16S reads by average quality value prior to clustering enabled selection of the representative sequence with the highest quality value among the 16S reads grouped in each OTU. On the other hand, the primer check step for removing reads lacking both primer sequences[10] had the possibility to incorrectly remove reads containing V1–V2 regions longer than the maximum length of 431 bp in the filter-passed reads. This is because there are a few species with a V1–V2 region >431 bp (e.g. *Campylobacter rectus* has a length of 493 bp). Our primer check step did not significantly affect the present results because only one of the 177 000 raw reads examined hit to Campylobacter. However, to avoid the incorrect filtration of reads, we modified the primer check step so as not to remove reads having a length of >400 bp, even though they may not have both primer sequences.

3.2. *Overall composition of the salivary bacterial communities*

We evaluated the ecological features of the salivary bacterial communities of the CD, UC, and HC groups by a variety of indices at the OTU level.[37,38] The results are summarized in Table 1. Species richness is the observed number of bacterial species assigned by OTUs detected in the samples. Richness estimates were obtained from the observed number of species by the extrapolation method using estimators such as the Chao1 and ACE indices. Evenness is the degree of homogeneity of abundance of the species detected in the samples. Diversity estimates were obtained from species richness and evenness by using several different indices, which exhibit different sensitivities to given factors, to confirm our results. The results suggested that there were no significant differences in the overall configuration of the salivary microbiota among the three groups (Table 1).

Table 1. OTU-based microbial richness and diversity across the HC, CD and UC groups

	HC	CD	UC
Diversity estimates			
Shannon Index	3.4 ± 0.1	3.4 ± 0.1	3.4 ± 0.1
Simpson Index	0.93 ± 0.01	0.93 ± 0.01	0.94 ± 0.01
Invsimpson Index	16.7 ± 1.1	16.7 ± 1.1	17.1 ± 1.4
Fisher alpha Index	26.8 ± 1.4	26.3 ± 1.4	24.8 ± 1.8
Evenness estimate			
Pielou's Index	0.7 ± 0.01	0.7 ± 0.01	0.71 ± 0.01
Richness estimates			
Number of OTUs	126 ± 5	124 ± 5	118 ± 7
chao1 Index	183 ± 8	183 ± 9	164 ± 13
ACE Index	182 ± 8	177 ± 8	165 ± 11

We then compared the overall bacterial community composition using the UniFrac distance metric, a phylogenetic tree-based metric ranging from 0 (distance between identical communities) to 1 (distance between totally different communities). A principal coordinate analysis (PCoA) plot based on the weighted UniFrac metric revealed clear clustering of most IBD samples apart from the HC samples, indicating the difference in microbial communities between the two groups (Fig. 1A). A bar chart more clearly shows the significant difference in microbiota composition between the IBD and HC groups (Fig. 1B). Comparison of the salivary microbiota of HCs with that of the CD and UC groups indicated that the microbiota of HCs significantly differs from both of them, and no significant difference was found between the UC and CD groups (Fig. 1C). Similar results were obtained using the unweighted UniFrac metric with lower statistical significance than that of the weighted UniFrac metric (Supplementary Fig. S1). These data suggest that species abundance, rather than species diversity, largely contributes to the observed differences in salivary microbiota between the HC and IBD groups.

Although the average age was considerably different between HCs and the IBD patients, weighted UniFrac distance analysis of 10 selected healthy subjects (average age 25.0 yr), 10 IBD patients (average age 28.7 yr, which matched the selected HC group), and the remaining 25 IBD patients (average age 54.6 yr) showed results similar to that of the total samples (Supplementary Fig. S2). Moreover, there was no significant difference between the two IBD subgroups. These data suggest that age might not affect the observed dysbiosis of the salivary microbiota of the IBD patients.

3.3. *Differences in salivary microbiota composition between the HC, CD, and UC groups*

The final dataset of the examined CD, UC, and HC groups ($n = 59$) consisted of 177 000 reads and included representatives of 12 bacterial phyla (Fig. 2; Supplementary Fig. S3 and Table S3). The majority of the 16S reads were classified into only five phyla: Firmicutes (46.5%), Bacteroidetes (22.3%), Actinobacteria (13.7%), Proteobacteria (12.5%), and Fusobacteria (4.2%). TM7, SR1, Spirochaetes, Synergistetes, Tenericutes, and Cyanobacteria were also detected and collectively represented <1% of the total reads analyzed. Analysis at the phylum level showed that the relative abundance of Bacteroidetes was significantly higher in both the CD

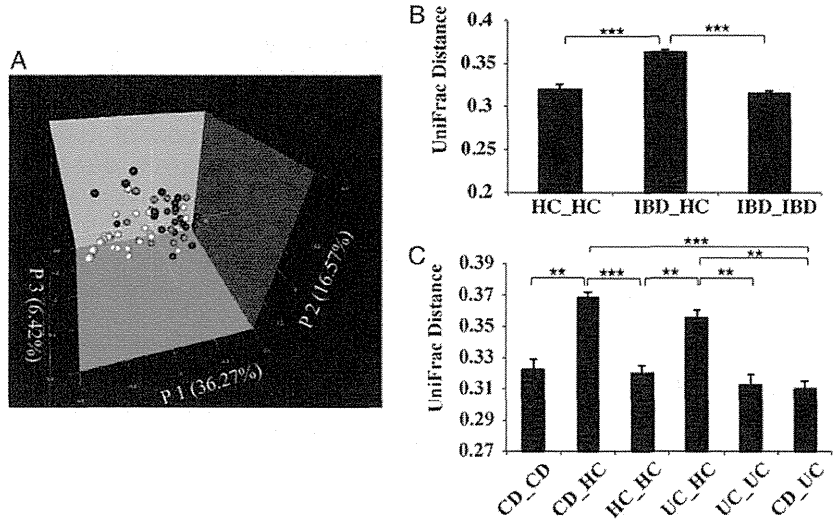


Figure 1. Analysis of the salivary microbiota of the HC, CD, and UC groups based on 16S data. (A) PCoA plot generated using weighted UniFrac metric. The three components explained 59.26% of the variance. White, grey, and black dots indicate HCs, UC, and CD samples, respectively. (B) Weighted UniFrac distance metric (a measure of differences in bacterial community structure) between HCs and the IBD (CD and UC) groups. (C) Weighted UniFrac distance metric between the HC, CD, and UC groups. Student's *t*-test was used; $*P < 0.01$, $**P < 10^{-5}$, and $***P < 10^{-10}$; mean ± S.E.M.

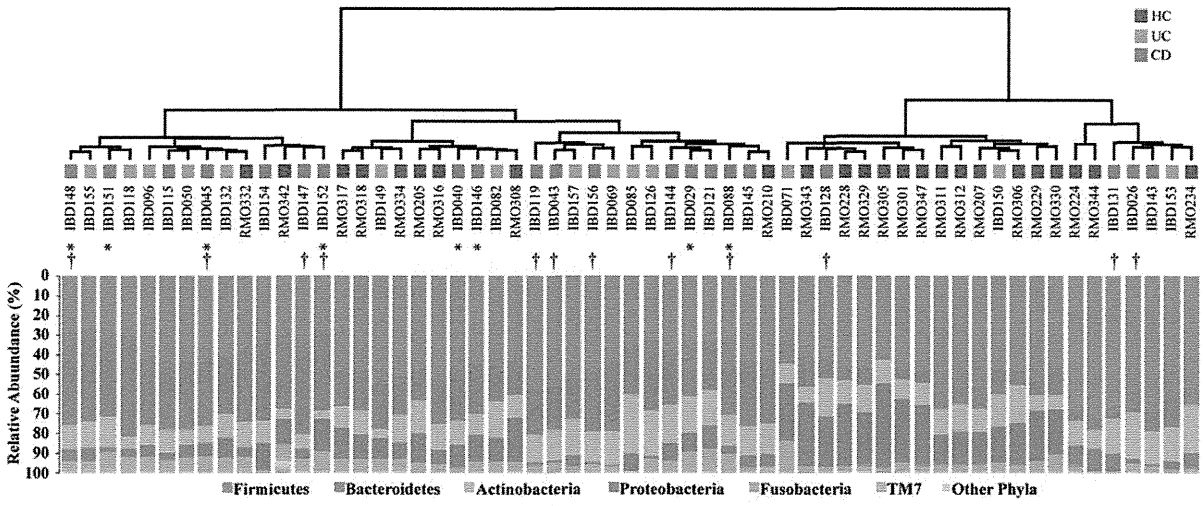


Figure 2. Cluster dendrogram generated using weighted UniFrac metric. Bar charts show the relative abundance of different phyla across the CD, UC and HC samples. Asterisks indicate samples taken during the active phase of CD. Dagger indicates anti-TNF-α antibody treated CD.

and UC groups as compared with HCs ($P < 0.01$), while that of Proteobacteria was significantly lower in both the CD and UC groups as compared with HCs ($P < 0.01$). No significant difference at the phylum level was observed between the UC and CD groups, which was consistent with the results of the UniFrac distance analysis.

In total, 107 bacterial genera were identified (at 95% identity), accounting for 97.8% of the total dataset. The remaining unclassified sequences (2.2%) were assigned to higher level taxa. Fourteen genera, including *Streptococcus*, *Prevotella*, *Rothia*, *Neisseria*, *Granulicatella*, *Actinomyces*, *Haemophilus*, *Veillonella*, *Gemella*, *Leptotrichia*, *Fusobacterium*, *Porphyromonas*, *Uncultured Lachnospiraceae*, and *Oribacterium*, predominated accounting for 92.7% of the total dataset. Other genera represented <0.5% each (Fig. 3; Supplementary Table S3). Two genera, *Prevotella* (phy. Bact.) and *Veillonella* (phy. Firm.), were significantly higher in both the CD and UC groups compared with HCs ($P < 0.01$). Two genera, *Streptococcus* (phy. Firm.) and *Haemophilus* (phy. Prot.), were significantly lower in both the CD and UC groups as compared with HCs ($P < 0.05$ and

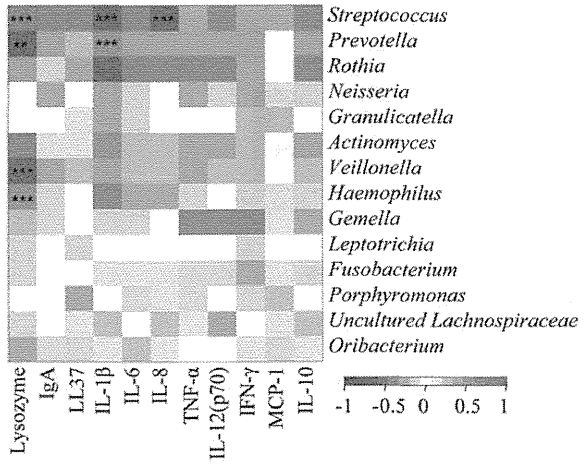


Figure 4. Correlation between the relative abundance of predominant genera and the level of immunological biomarkers in the saliva of IBD patients. Pearson product moment correlation coefficients are represented by colour ranging from blue, negative correlation (−1), to red, positive correlation (1). Normalized values of immunological biomarkers by total protein amount were used in this analysis. Significant correlations after *P*-value adjustment are marked by $*P < 0.05$, $**P < 0.01$, and $***P < 0.001$.

between HCs and IBD patients by 16S pyrosequencing analysis (Supplementary Table S4). Using these primers, we found strong correlations between 16S-based and qPCR data for the quantification of *P. melaninogenica* ($r = 0.87$, $P < 0.001$) and *H. parainfluenzae* ($r = 0.86$, $P < 0.001$), indicating the quantitative accuracy of our 16S pyrosequencing-based results (Fig. 5).

4. Discussion

4.1. Bacterial 16S rRNA-based pyrosequencing analysis

In this study, we used targeted amplicon sequencing of the 16S rRNA gene hypervariable V1–V2 region to evaluate bacterial composition at finer taxonomic levels. The use of primer 27Fmod enabled us to reduce underestimation of the relative abundance of *Bifidobacterium* species that predominate human microbiota, and thus the quantitative accuracy of the overall bacterial composition was greatly improved.[10,39] One limitation of clustering the 16S reads using the UCLUST program is selection of the representative sequence for each OTU. The quality of the representative sequence is not always the highest in the OTU, which affects the BLAST identity, *E*-value and score, sometimes providing inappropriate results for taxonomic assignment of the OTUs. We overcame this limitation by sorting the 16S reads by their average quality values prior to clustering, leading to 16S reads with the highest quality being selected as the representative sequence for each OTU. Our 16S-based results were also validated by strongly correlating with the qPCR data targeting bacterial species showing significant changes between HC and IBD samples (Fig. 5). In addition, clustering of the reads was performed with a 96% pairwise-identity cutoff to reduce overestimation of the number of bacterial species (or OTUs) largely due to 454 pyrosequencing errors.[10,40] Clustering with a 96% pairwise-identity cutoff should be applied for pyrosequencing reads obtained from other types of human microbiota.

4.2. Salivary microbiota composition in IBD patients

The abundant bacterial groups in the salivary microbiota detected in this study were similar to those previously reported,[41–44] but the compositions differed from those observed in plaque microbiota.[44] Our data clearly showed a significant difference in salivary microbiota composition between HCs and IBD patients. Shifts in oral microbiota composition were also observed in several oral manifestations such as dental caries,[45] periodontitis,[46] and oral squamous cell carcinoma.[47] Moreover, various components of the oral microbiota have been implicated in systemic diseases such as pancreatic disease including pancreatic cancer,[48] atherosclerosis,[49] bacteremia,[50] and endocarditis.[51]

Altered bacterial community structure in the gut microbiota of IBD patients is a common finding in comparison with that of healthy subjects. Previous studies showed overall structural changes as well as reduced species richness of the gut microbiota in IBD patients.[29–33] It is likely that the high microbial richness and diversity characterizing healthy microbiota may have a protective effect on humans. Unlike the gut microbiota of IBD patients, our estimates using several metrics revealed that microbial richness and diversity in the salivary microbiota of IBD patients was similar to that of HCs, despite significant changes in community structure (Fig. 1). These data suggest that the extent of the changes in the salivary microbiota is less than that in the gut microbiota of IBD patients.

Our data indicated a significant increase of the genus *Prevotella* in the salivary microbiota of IBD patients, in which its relative abundance was almost equivalent to that of reduced *Streptococcus*, which is most abundant in healthy salivary microbiota (Fig. 3). *Prevotella* is a Gram-negative, obligate anaerobe, and a member of the prevalent genera in the human microbiome.[52] Some *Prevotella* species were similarly increased, distinguishable from opportunistic infections, in bacterial vaginosis,[53] esophagitis,[54] antral gastritis,[55] and saliva of caries-active subjects.[45] These data suggest that the increase of *Prevotella*, with concurrently decreased *Streptococcus*, is clearly related with abnormal physiologies in IBD patients. The relative abundance of total Gram-positive and Gram-negative bacteria showed no significant difference between HCs and IBD patients

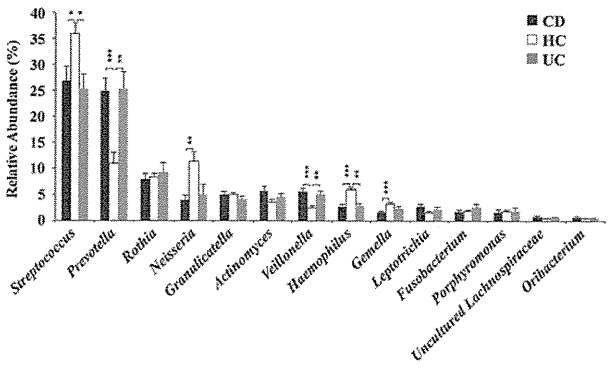


Figure 3. Mean genus abundance in the CD, UC and HC groups. Plotted values are the mean abundance of the 14 most abundant genera in each group. Welch's test with BH adjustment was used; $*P < 0.05$, $**P < 0.01$, and $***P < 0.001$; mean ± S.E.M.

0.01, respectively). Two other genera, *Neisseria* (phy. Prot.) and *Gemella* (phy. Firm.), were also found to be significantly lower only in the CD group as compared with HCs ($P < 0.01$ and 0.001, respectively). These results indicate that the relative increase of Bacteroidetes in IBD patients was mainly due to the increase of *Prevotella*, and the relative decrease of Proteobacteria in IBD patients was mainly due to the decrease of *Neisseria* and *Haemophilus*. No significant difference in the relative abundance of either Gram-positive or Gram-negative bacteria was observed among the three groups (Supplementary Table S3).

Clustering of all reads using a 96% pairwise-identity cutoff generated 1257 OTUs, of which only 40 OTUs represented 67.2% of the total reads analyzed. The remaining OTUs were present at relative abundance levels <0.5% of the total dataset (Supplementary Table S4). The relative abundance of several OTUs belonging to the genera *Streptococcus*, *Prevotella*, *Veillonella*, *Nesisseria*, *Haemophilus*, and *Gemella* showed significant differences in IBD patients as compared with HCs. These results were concordant with those detected at the genus level. Among the abundant OTUs, those most closely assigned to *Prevotella melaninogenica*, *Veillonella* sp. oral taxon 158, *Streptococcus mitis*, *Gemella sanguinis*, *Neisseria mucosa*, and *Haemophilus parainfluenzae* showed significant differences in relative abundance between the HC and IBD groups (Supplementary Table S4).

3.4. *Salivary immunological biomarkers in the HC, CD, and UC groups*

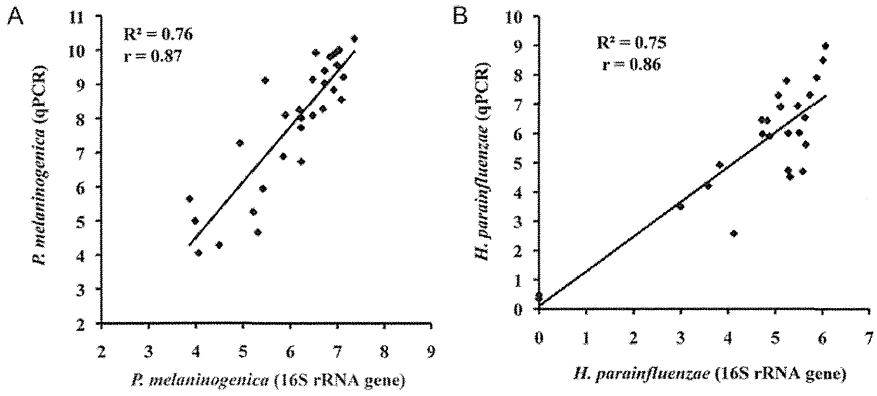
We evaluated the inflammatory state, considering its influence on shaping the salivary microbiota, in saliva of the CD and UC patients as compared with that of HCs. The analysis was performed by measuring secretory IgA, cytokines, and enzymes including lysozyme in unstimulated saliva of 15 HC, 14 CD, and 10 UC individuals (Supplementary Table S5 and Fig. S4). There was no significant difference in the total protein concentration in saliva of the CD and UC patients as compared with that of HCs ($P = 0.112$ and 0.192, respectively). The lysozyme level was significantly lower in saliva of both the CD and UC groups as compared with HCs ($P < 0.01$). On the other hand, the levels of IgA and LL37 in both CD and UC groups were higher than that of HCs with statistical significance. The use of Luminex technology was highly sensitive in measuring cytokines from small volumes of saliva samples. In saliva of the CD and UC groups, the level of IL-1β was significantly higher as compared with HCs ($P < 0.05$ and <0.01, respectively). The levels of IL-6, IL-8, and MCP-1 were significantly higher only in saliva of the UC group, while elevated TNF-α level was found only in the CD group with statistical significance. The levels of IgA and MCP-1 in the UC group were significantly higher than those in the CD group. These data indicate that the oral cavity of IBD patients is usually in the inflammatory state, and the levels tend to be slightly higher in the UC group than the CD group.

3.5. *Composition of the salivary microbiota in relation to immunological biomarkers*

We searched for correlations between the relative abundance of dominant bacterial genera and the measured biomarkers in the saliva of 39 subjects (Supplementary Table S5). The results are shown in Fig. 4. The relative abundance of *Streptococcus* negatively correlates with IL-1β and IL-8 ($r = -0.54$ and -0.51, respectively, $P < 0.001$), while it positively correlates with lysozyme ($r = 0.63$, $P < 0.001$). On the other hand, the abundance of *Prevotella* positively correlates with IL-1β ($r = 0.58$, $P < 0.001$) but negatively correlates with lysozyme ($r = -0.54$, $P < 0.01$). The relative abundance of *Veillonella* negatively correlates with lysozyme ($r = -0.54$, $P < 0.001$), while *Haemophilus* positively correlates with lysozyme ($r = 0.58$, $P < 0.001$). Linear regressions also validated correlations between the relative abundance of *Streptococcus* and *Prevotella* and the levels of lysozyme and IL-1β, and between the relative abundance of *Veillonella* and *Haemophilus* and the level of lysozyme (Supplementary Fig. S5). On the whole, *Prevotella*, *Actinomyces*, *Veillonella*, and *Lachnospiracea* tended to positively correlate, while *Streptococcus*, *Rothia*, *Neisseria*, *Haemophilus*, and *Gemella* tended to negatively correlate with elevated cytokines in saliva of IBD patients.

3.6. *Validation of 16S pyrosequencing data by targeted quantitative PCR*

We designed specific PCR primers for quantitative PCR (qPCR) targeting genomes of *P. melaninogenica* and *H. parainfluenzae*, which showed significant differences



C. Bacterial genome specific primers for quantitative PCR.

Targeted species	Primer	Sequence (5'-3')	Tm (°C)	Product size (bp)	Annealing temp (°C)
Prevotella melaninogenica	Fw	GCTTTTGGAGCAAACAGAGG	60	234	55
	Rv	GTTGTGAATACGCGGTCCTT	60		
Haemophilus parainfluenzae	Fw	ACCGTGGTCGTTTAGCAATC	60	252	55
	Rv	GTCCGGGTTTACGTTTAGCA	60		

Figure 5. Correlation between the 16S rRNA pyrosequencing and qPCR data. The results are shown in (A) for *P. melaninogenica* and (B) for *H. parainfleuenzae*. The *y*-axis represents the copy number per nanogram of bacterial DNA obtained from qPCR data, transformed by the inverse hyperbolic sine method. The *x*-axis represents the number of reads assigned as bacterial spp. obtained from the pyrosequencing data, transformed by inverse hyperbolic sine method. Pearson product moment correlation coefficient (*r*) on transformed data (using inverse hyperbolic sine transformation) is shown. (C) Primer sequences and PCR conditions used for qPCR experiments are shown.

(Supplementary Table S3). From these results, Gram-stain properties of bacterial surface structures may not be related with dysbiosis of IBD salivary microbiota, unlike the association of Gram-negative oral bacteria with dysbiosis observed in subgingival microbiota in periodontitis.[56]

4.3. *Salivary microbiota associated with immunological biomarkers*

Saliva contains a variety of components such as cytokines, immunoglobulins, and antimicrobial proteins involved in host defence mechanisms for maintaining oral and systemic health.[57] Alteration of the salivary microbiota in IBD patients suggests the occurrence of inflammatory immune responses in the oral cavity of IBD patients as intestinal inflammation associated with aberrant gut microbiota of IBD.[23–26] Our data showed that the levels of many salivary cytokines and IgA were significantly higher in both CD and UC patients than those observed in HCs, indicating that inflammatory responses are elicited in the oral cavity of the patients. Similarly, elevated salivary IL-1β, IL-6, and TNF-α levels in CD patients and an elevated IL-8 level in the saliva of patients with bowel disease were also reported.[58,59] Unexpectedly, the elevated level of inflammatory biomarkers in UC patients was similar to or slightly higher than that observed in CD patients, regardless of differences in disease states between IBD patients (Supplementary Fig. S4 and Table S5). Salivary IgA induction was observed in CD patients with oral symptoms but not in those without oral symptoms.[60] The elevated level of IgA in most IBD patients' saliva examined suggests that those patients may have oral manifestations, however, we did not have access to their oral health clinical records.

Salivary lysozyme levels were significantly reduced in both CD and UC patients as compared with that of HCs. Lysozyme is an antimicrobial protein, expressed by various cells including neutrophils, macrophages, and epithelial cells. It is abundant in saliva and plays an important role in the host constitutive defence system.[61] It has been reported that salivary lysozyme was significantly lower in patients with gingivitis and periodontitis as compared with healthy subjects.[62] In contrast, faecal lysozyme levels were significantly elevated in IBD patients.[63] Further analysis will be required to elucidate the difference in lysozyme levels between saliva and the intestine.

Lysozyme exclusively catalyses hydrolysis of Gram-positive bacterial cell wall. However, lysozyme can also be bactericidal for Gram-negative bacteria *in vivo* through synergistic action with salivary lactoferrin in the normal state.[64] Therefore, this *in vitro* specificity of lysozyme activity may not be largely involved in the dysbiosis of salivary microbiota in IBD patients, in which the abundance of Gram-positive bacteria was not significantly different as compared with HCs (Supplementary Table S3).

There were several subgroups of patients dependent on different medical treatments, and patients with different states of disease (Supplementary Tables S1 and S2). In addition, Infliximab (anti-TNF-α antibody) therapy is commonly used for IBD patients, but up to one-third of the patients have been shown not to respond.[65] Therefore, it was very difficult to precisely evaluate the differences in microbiota structure and biomarker levels between the subgroups. Nevertheless, phylogenetic analysis based on the weighted UniFrac distance metric did not show discrete clustering of particular subgroups, such as CD patients with or without Infliximab treatment and active CD, or CD in remission, suggesting limited contributions from the patients' disease state or medical treatment to the overall microbiota structure (Fig. 2).

Strong correlations between some inflammatory biomarkers and salivary microbiota compositions were revealed (Fig. 4). The lower lysozyme and elevated IL-1β, IL-8, IgA and several other biomarkers were likely to be synergistically or interactively associated with the abundance of the four dominant genera, *Streptococcus*, *Prevotella*, *Veillonella*, and *Haemophilus*. Interactions between these microbes and other species may also be involved in the dysbiosis of salivary microbiota of IBD patients.

Finally, it is still unknown whether the inflammatory state in the oral cavity of IBD patients is the cause or a consequence of imbalances in the salivary microbiota, and which local (the oral cavity) or systemic (the gut) immune response is more responsible for the observed dysbiosis of salivary microbiota. Our results strongly suggest the existence of certain defined mechanisms by which aberrant, but similar, salivary microbiota among IBD patients is formed. The human gut microbiota is gradually shaped to its matured assemblage in a few years after birth, with temporal changes in the diversity and rank of dominant species largely dependent on diet and host physiological state.[66] Salivary microbiota may also be established similar to gut microbiota. Since >1000 ml of saliva is produced per day in the average adult and it always flows into the gastrointestinal tract, bacteria in saliva also have many opportunities to reach the intestine. Therefore, it can be postulated that salivary microbiota affects the development of gut microbiota to some extent. To evaluate this hypothesis, it is necessary to investigate the progression of infant salivary microbiota and the oral inflammatory state. Additionally, further studies such as comparison of the salivary microbiota between IBD and other diseases will provide informative sources for discovering non-invasive salivary biomarkers specific to IBD.

Acknowledgements: All of the authors extend their deepest sympathy and condolences to the family of Dr Hiroshi Chinen, who sadly passed away in the middle of this study. We thank Dr Todd D. Taylor (RIKEN Yokohama Institute) for critical reading of the manuscript; K. Komiya, C. Shindo, H. Kuroyanagi, E. Iioka, Y. Takayama, E. Ohmori, M. Kiuchi, Y. Hattori (The University of Tokyo), and A. Nakano (Azabu University) for technical support; and Drs F. Kinjo and A. Hokama (University Hospital, Faculty of Medicine, University of the Ryukyus) for their kind help with the sampling and storing of the patient's saliva.

Supplementary data: Supplementary data are available at www.dnaresearch.oxfordjournals.org.

Funding

This work was supported in part by the global COE project of 'Genome Information Big Bang' from the Ministry of Education, Culture, Sports, Science, and Technology (MEXT) of Japan to M.H. and K.O., a research project grant from Azabu University to H.M., a grant from the Core Research for Evolutional Science and Technology (CREST) program of the Japan Science and Technology Agency (JST) to K.O., and Scientific Research B (No. 22370087) to H.I. and (Nos. 21370108 and 24370099) to H.O. from the Japan Society for the Promotion of Science (JSPS). H.S.S. acknowledges the fellowship from MEXT.

References

1. Metzker, M.L. 2010, Sequencing technologies — the next generation, *Nat. Rev. Genet.*, **11**, 31–46.
2. Kunin, V., Copeland, A., Lapidus, A., Mavromatis, K. and Hugenholtz, P. 2008, A bioinformatician's guide to metagenomics, *Microbiol. Mol. Biol. Rev.*, **72**, 557–78.
3. Hamady, M. and Knight, R. 2009, Microbial community profiling for human microbiome projects: tools, techniques, and challenges, *Genome Res.*, **19**, 1141–52.
4. Kuczynski, J., et al. 2012, Experimental and analytical tools for studying the human microbiome, *Nat. Rev. Genet.*, **13**, 47–58.
5. Song, S., Jarvie, T. and Hattori, M. 2013, Our second genome – human metagenome – how next generation sequencer changes our life through microbiology, *Adv. Microb. Physiol.*, **62**, 119–44.
6. Jumpstart Consortium Human Microbiome Project Data Generation Working Group. 2012, Evaluation of 16S rDNA-based community profiling for human microbiome research, *PLoS ONE*, **6**, e39315.
7. Andersson, A.F., Lindberg, M., Jakobsson, H., Bäckhed, F., Nyrén, P. and Engstrand, L. 2008, Comparative analysis of human gut microbiota by barcoded pyrosequencing, *PLoS ONE*, **3**, e2836.
8. Wang, Y. and Qian, P.Y. 2009, Conservative fragments in bacterial 16S rRNA genes and primer design for 16S ribosomal DNA amplicons in metagenomic studies, *PLoS ONE*, **4**, e7401.

9. Hamady, M., Walker, J.J., Harris, J.K., Gold, N.J. and Knight, R. 2008, Error-correcting barcoded primers for pyrosequencing hundreds of samples in multiplex, *Nat. Methods*, **5**, 235–7.
10. Kim, S.W., Suda, W., Kim, S., et al. 2013, Robustness of gut microbiota of healthy adults in response to probiotic intervention revealed by high-throughput pyrosequencing, *DNA Res.*, **20**, 241–53.
11. Starke, E.M., et al. 2006, Technology development to explore the relationship between oral health and the oral microbial community, *BMC Oral Health*, **6**(Suppl. 1), S10.
12. Avila, M., Ojcius, D.M. and Yilmaz, O. 2009, The oral microbiota: living with a permanent guest, *DNA Cell Biol.*, **28**, 405–11.
13. Dewhirst, F.E., Chen, T., Izard, J., et al. 2010, The human oral microbiome, *J. Bacteriol.*, **192**, 5002–17.
14. Curtis, M.A., Zenobia, C. and Darveau, R.P. 2011, The relationship of the oral microbiotia to periodontal health and disease, *Cell Host Microbe.*, **10**, 302–6.
15. Jose, F.A. and Heyman, M.B. 2008, Extraintestinal manifestations of inflammatory bowel disease, *J. Pediatr. Gastroenterol. Nutr.*, **46**, 124–33.
16. Femiano, F., Lanza, A., Buonaiuto, C., Perillo, L., Dell'Ermo, A. and Cirillo, N. 2009, Pyostomatitis vegetans: a review of the literature, *Med. Oral Patol. Oral Cir. Bucal.*, **14**, E114–117.
17. Rowland, M., Fleming, P. and Bourke, B. 2010, Looking in the mouth for Crohn's disease, *Inflamm. Bowel Dis.*, **16**, 332–7.
18. Veloso, F.T. 2011, Extraintestinal manifestations of inflammatory bowel disease: do they influence treatment and outcome? *World J. Gastroenterol.*, **17**, 2702–7.
19. Vavricka, S.R., Brun, L., Ballabeni, P., et al. 2011, Frequency and risk factors for extraintestinal manifestations in the Swiss inflammatory bowel disease cohort, *Am. J. Gastroenterol.*, **106**, 110–9.
20. Singhal, S., Dian, D., Keshavarzian, A., Fogg, L., Fields, J.Z. and Farhadi, A. 2011, The role of oral hygiene in inflammatory bowel disease, *Dig. Dis. Sci.*, **56**, 170–5.
21. Xavier, R.J. and Podolsky, D.K. 2007, Unravelling the pathogenesis of inflammatory bowel disease, *Nature*, **448**, 427–34.
22. Cho, J.H. 2008, The genetics and immunopathogenesis of inflammatory bowel disease, *Nat. Rev. Immunol.*, **8**, 458–66.
23. Peterson, D.A., Frank, D.N., Pace, N.R. and Gordon, J.I. 2008, Metagenomic approaches for defining the pathogenesis of inflammatory bowel diseases, *Cell Host Microbe*, **3**, 417–27.
24. Maloy, K.J. and Powrie, F. 2011, Intestinal homeostasis and its breakdown in inflammatory bowel disease, *Nature*, **474**, 298–306.
25. Khor, B., Gardet, A. and Xavier, R.J. 2011, Genetics and pathogenesis of inflammatory bowel disease, *Nature*, **474**, 307–17.
26. Maynard, C.L., Elson, C.O., Hatton, R.D. and Weaver, C.T. 2012, Reciprocal interactions of the intestinal microbiota and immune system, *Nature*, **489**, 231–41.
27. Greenberg, G.R. 2004, Antibiotics should be used as first-line therapy for Crohn's disease, *Inflamm. Bowel. Dis.*, **10**, 318–20.
28. Sellon, R.K., Tonkonogy, S., Schultz, M., et al. 1998, Resident enteric bacteria are necessary for development of spontaneous colitis and immune system activation in interleukin-10-deficient mice, *Infect. Immun.*, **66**, 5224–31.
29. Manichanh, C., Rigottier-Gois, L., Bonnaud, E., et al. 2006, Reduced diversity of faecal microbiota in Crohn's disease revealed by a metagenomic approach, *Gut*, **55**, 205–11.
30. Baumgart, M., Dogan, B., Rishniw, M., et al. 2007, Culture independent analysis of ileal mucosa reveals a selective increase in invasive Escherichia coli of novel phylogeny relative to depletion of Clostridiales in Crohn's disease involving the ileum, *ISME J.*, **1**, 403–18.
31. Frank, D.N., St Amand, A.L., Feldman, R.A., Boedeker, E.C., Harpaz, N. and Pace, N.R. 2007, Molecular-phylogenetic characterization of microbial community imbalances in human inflammatory bowel diseases, *Proc. Natl Acad. Sci. USA*, **104**, 13780–5.
32. Dicksved, J., Halfvarson, J., Rosenquist, M., et al. 2008, Molecular analysis of the gut microbiota of identical twins with Crohn's disease, *ISME J.*, **2**, 716–27.
33. Sokol, H., Pigneur, B., Watterlot, L., et al. 2008, Faecalibacterium prausnitzii is an anti-inflammatory commensal bacterium identified by gut microbiota analysis of Crohn disease patients, *Proc. Natl Acad. Sci. USA*, **105**, 16731–6.
34. Qin, J., Li, R., Raes, J., et al. 2010, A human gut microbial gene catalogue established by metagenomic sequencing, *Nature*, **464**, 59–65.
35. Morita, H., Kuwahara, T., Ohshima, K., et al. 2007, An improved DNA isolation method for metagenomic analysis of the microbial flora of the human intestine, *Microbes Environ.*, **22**, 214–22.
36. Lozupone, C., Hamady, M. and Knight, R. 2006, UniFrac–an online tool for comparing microbial community diversity in a phylogenetic context, *BMC Bioinformatics*, **7**, 371.
37. Gotelli, N.J. and Colwell, R.K. 2011, Estimating species richness, In: Magurran, A.E. and McGill, B.J. (eds.), *Frontiers in measuring biodiversity.* Oxford University Press: Oxford, UK, pp. 39–54.
38. Colwell, R.K. 2009, Biodiversity: concepts, patterns, and measurement, In: Levin, S.A., et al. (eds.), *The Princeton guide to ecology.* Princeton University Press: New Jersey, USA, pp. 257–63.
39. Hattori, M. and Taylor, T.D. 2009, The human intestinal microbiome: a new frontier of human biology, *DNA Res.*, **16**, 1–12.
40. Kunin, V., Engelbrektson, A., Ochman, H. and Hugenholtz, P. 2010, Wrinkles in the rare biosphere: pyrosequencing errors can lead to artificial inflation of diversity estimates, *Environ. Microbiol.*, **12**, 118–23.
41. Keijser, B.J., Zaura, E., Huse, S.M., et al. 2008, Pyrosequencing analysis of the oral microflora of healthy adults, *J. Dent. Res.*, **87**, 1016–20.
42. Nasidze, I., Li, J., Quinque, D., Tang, K. and Stoneking, M. 2009, Global diversity in the human salivary microbiome, *Genome Res.*, **19**, 636–43.

43. Diaz, P.I., Dupuy, A.K., Abusleme, L., et al. 2012, Using high throughput sequencing to explore the biodiversity in oral bacterial communities, *Mol. Oral Microbiol.*, **27**, 182–201.
44. Yamanaka, W., Takeshita, T., Shibata, Y., et al. 2012, Compositional stability of a salivary bacterial population against supragingival microbiota shift following periodontal therapy, *PLoS ONE*, **7**, e42806.
45. Yang, F., Zeng, X., Ning, K., et al. 2012, Saliva microbiomes distinguish caries-active from healthy human populations, *ISME J.*, **6**, 1–10.
46. Paju, S., Pussinen, P.J., Suominen-Taipale, L., Hyvönen, M., Knuuttila, M. and Könönen, E. 2009, Detection of multiple pathogenic species in saliva is associated with periodontal infection in adults, *J. Clin. Microbiol.*, **47**, 235–8.
47. Mager, D.L., Haffajee, A.D., Devlin, P.M., Norris, C.M., Posner, M.R. and Goodson, J.M. 2005, The salivary microbiota as a diagnostic indicator of oral cancer: a descriptive, non-randomized study of cancer-free and oral squamous cell carcinoma subjects, *J. Transl. Med.*, **3**, 27.
48. Farrell, J.J., Zhang, L., Zhou, H., et al. 2012, Variations of oral microbiota are associated with pancreatic diseases including pancreatic cancer, *Gut*, **61**, 582–8.
49. Koren, O., Spor, A., Felin, J., et al. 2011, Human oral, gut, and plaque microbiota in patients with atherosclerosis, *Proc. Natl Acad. Sci. USA*, **108**(Suppl. 1), 4592–8.
50. Poveda-Roda, R., Jiménez, Y., Carbonell, E., Gavaldá, C., Margaix-Muñoz, M.M. and Sarrión-Pérez, G. 2008, Bacteremia originating in the oral cavity. A review, *Med. Oral. Patol. Oral Cir. Bucal.*, **13**, E355–362.
51. Parahitiyawa, N.B., Jin, L.J., Leung, W.K., Yam, W.C. and Samaranayake, L.P. 2009, Microbiology of odontogenic bacteremia: beyond endocarditis, *Clin. Microbiol. Rev.*, **22**, 46–64, Table of Contents.
52. Alauze, C., Marchandin, H. and Lozniewski, A. 2010, New insights into Prevotella diversity and medical microbiology, *Future Microbiol.*, **5**, 1695–718.
53. Oakley, B.B., Fiedler, T.L., Marrazzo, J.M. and Fredricks, D.N. 2008, Diversity of human vaginal bacterial communities and associations with clinically defined bacterial vaginosis, *Appl. Environ. Microbiol.*, **74**, 4898–909.
54. Yang, L., Lu, X., Nossa, C.W., Francois, F., Peek, R.M. and Pei, Z. 2009, Inflammation and intestinal metaplasia of the distal esophagus are associated with alterations in the microbiome, *Gastroenterology*, **137**, 588–97.
55. Li, X.X., Wong, G.L., To, K.F., et al. 2009, Bacterial microbiota profiling in gastritis without Helicobacter pylori infection or non-steroidal anti-inflammatory drug use, *PLoS ONE*, **4**, e7985.
56. Darveau, R.P. 2010, Periodontitis: a polymicrobial disruption of host homeostasis, *Nat. Rev. Microbiol.*, **8**, 481–90.
57. Farnaud, S.J., Kosti, O., Getting, S.J. and Renshaw, D. 2010, Saliva: physiology and diagnostic potential in health and disease, *ScientificWorldJournal*, **10**, 434–56.
58. Szczeklik, K., Owczarek, D., Pytko-Polonczyk, J., Kesek, B. and Mach, T.H. 2012, Proinflammatory cytokines in the saliva of patients with active and non-active Crohn's disease, *Pol. Arch. Med. Wewn.*, **122**, 200–8.
59. Rathnayake, N., Akerman, S., Klinge, B., et al. 2013, Salivary biomarkers for detection of systemic diseases, *PLoS ONE*, **8**, e61356.
60. Savage, N.W., Barnard, K., Shirlaw, P.J., et al. 2004, Serum and salivary IgA antibody responses to *Saccharomyces cerevisiae*, *Candida albicans* and *Streptococcus mutans* in orofacial granulomatosis and Crohn's disease, *Clin. Exp. Immunol.*, **135**, 483–9.
61. Wiesner, J. and Vilcinskas, A. 2010, Antimicrobial peptides: the ancient arm of the human immune system, *Virulence*, **1**, 440–64.
62. Surna, A., Kubilius, R., Sakalauskiene, J., et al. 2009, Lysozyme and microbiota in relation to gingivitis and periodontitis, *Med. Sci. Monit.*, **15**, CR66–73.
63. Abraham, B.P. and Thirumurthi, S. 2009, Clinical significance of inflammatory markers, *Curr. Gastroenterol. Rep.*, **11**, 360–7.
64. Ellison, R.T. 3rd and Giehl, T.J. 1991, Killing of gram-negative bacteria by lactoferrin and lysozyme, *J. Clin. Invest.*, **88**, 1080–91.
65. Melmed, G.Y. and Targan, S.R. 2010, Future biologic targets for IBD: potentials and pitfalls, *Nat. Rev. Gastroenterol. Hepatol.*, **7**, 110–7.
66. Koenig, J.E., Spor, A., Scalfone, N., et al. 2011, Succession of microbial consortia in the developing infant gut microbiome, *Proc. Natl Acad. Sci. USA*, **108**(Suppl. 1), 4578–85.

LETTER

doi:10.1038/nature12331

T_{reg} induction by a rationally selected mixture of Clostridia strains from the human microbiota

Koji Atarashi[1,2,3]*, Takeshi Tanoue[1,2]*, Kenshiro Oshima[4,5]*, Wataru Suda[5], Yuji Nagano[1,2], Hiroyoshi Nishikawa[6], Shinji Fukuda[1,7], Takuro Saito[6], Seiko Narushima[1], Koji Hase[1,3], Sangwan Kim[5], Joëlle V. Fritz[8], Paul Wilmes[8], Satoshi Ueha[9], Kouji Matsushima[9], Hiroshi Ohno[1], Bernat Olle[10], Shimon Sakaguchi[6], Tadatsugu Taniguchi[2], Hidetoshi Morita[4,11], Masahira Hattori[5] & Kenya Honda[1,2,4]

Manipulation of the gut microbiota holds great promise for the treatment of inflammatory and allergic diseases[1,2]. Although numerous probiotic microorganisms have been identified[3], there remains a compelling need to discover organisms that elicit more robust therapeutic responses, are compatible with the host, and can affect a specific arm of the host immune system in a well-controlled, physiological manner. Here we use a rational approach to isolate $CD4^+FOXP3^+$ regulatory T (T_{reg})-cell-inducing bacterial strains from the human indigenous microbiota. Starting with a healthy human faecal sample, a sequence of selection steps was applied to obtain mice colonized with human microbiota enriched in T_{reg}-cell-inducing species. From these mice, we isolated and selected 17 strains of bacteria on the basis of their high potency in enhancing T_{reg} cell abundance and inducing important anti-inflammatory molecules—including interleukin-10 (IL-10) and inducible T-cell co-stimulator (ICOS)—in T_{reg} cells upon inoculation into germ-free mice. Genome sequencing revealed that the 17 strains fall within clusters IV, XIVa and XVIII of Clostridia, which lack prominent toxins and virulence factors. The 17 strains act as a community to provide bacterial antigens and a TGF-β-rich environment to help expansion and differentiation of T_{reg} cells. Oral administration of the combination of 17 strains to adult mice attenuated disease in models of colitis and allergic diarrhoea. Use of the isolated strains may allow for tailored therapeutic manipulation of human immune disorders.

$CD4^+FOXP3^+$ T_{reg} cells are present most abundantly in the intestinal mucosa at steady state, and contribute to intestinal and systemic immune homeostasis[4–7]. In germ-free mice, the frequency of colonic T_{reg} cells and levels of IL-10 expression by T_{reg} cells are markedly reduced[4–7]. We have shown previously that a combination of Clostridia strains isolated from conventionally reared mice potently affect the number and function of $CD4^+FOXP3^+$ T_{reg} cells in mouse colonic lamina propria[4]. In an attempt to enable clinical translation of our previous findings, we aimed to identify T_{reg}-cell-inducing bacterial strains derived from the human microbiota (see Supplementary Fig. 1 for a summary of the procedure).

We obtained a human stool sample from a healthy Japanese volunteer. Because we previously reported that the chloroform-resistant fraction of mouse gut microbiota was enriched in T_{reg}-cell-inducing species[4], the stool sample was either untreated or treated with chloroform and orally inoculated into IQI/Jic germ-free mice. Each group of ex-germ-free (exGF) mice was separately housed for 3–4 weeks in vinyl isolators to avoid further microbial contamination. Although a recent study showed that the human microbiota had no impact on the immune responses in the mouse small intestine[8], we observed a significant increase in the percentage of $FOXP3^+$ T_{reg} cells among $CD4^+$ T cells in the colons of exGF mice inoculated with untreated human faeces compared with germ-free mice (Fig. 1a and Supplementary Fig. 2). Notably, a more pronounced increase was observed in the colons of exGF mice inoculated with chloroform-treated human faeces (Fig. 1a). These findings suggest that the human intestinal microbiota contains T_{reg}-cell-inducing bacteria, and that they are enriched in the chloroform-resistant fraction. We also examined the effects of human faeces inoculation on colonic IL-17- and IFN-γ-expressing $CD4^+$ cells (T_H17 and T_H1 cells). In exGF mice inoculated with untreated or chloroform-treated human faeces, the frequency of T_H1 cells was unchanged compared with germ-free mice (Fig. 1b). By contrast, there

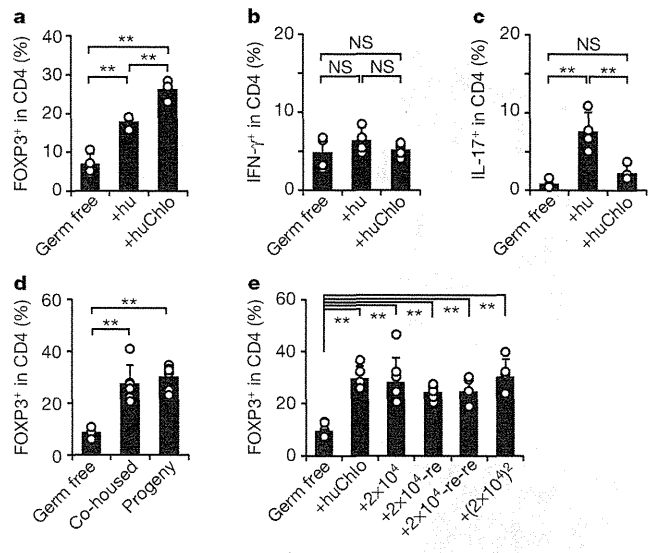


Figure 1 | T_{reg} cell accumulation in germ-free mice induced by inoculation with human microbiota. a–e, The percentages of $FOXP3^+$, $IL\text{-}17^+$ and $IFN\text{-}\gamma^+$ cells within the $CD4^+$ cell population in the colon lamina propria of the indicated mice are shown (see also Supplementary Fig. 2). Circles represent individual animals. The height of the black bars indicates the mean. All experiments were performed more than twice with similar results. Error bars indicate s.d. $^{**}P < 0.01$; NS, not significant. +hu, exGF mice inoculated with untreated human faeces; +huChlo, exGF mice inoculated with chloroform-treated human faeces. (See the main text for further definitions of *x*-axis labels.)

[1]RIKEN Center for Integrative Medical Sciences (IMS-RCAI), 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan. [2]Department of Immunology, Graduate School of Medicine, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan. [3]PRESTO, Japan Science and Technology Agency, 4-1-8 Honcho Kawaguchi, Saitama 332-0012, Japan. [4]CREST, Japan Science and Technology Agency, 4-1-8 Honcho Kawaguchi, Saitama 332-0012, Japan. [5]Graduate School of Frontier Sciences, The University of Tokyo, 5-1-5 Kashiwanoha, Kashiwa, Chiba 277-8561, Japan. [6]Experimental Immunology, Immunology Frontier Research Center, Osaka University, 3-1 Yamadaoka, Suita, Osaka 565-0871, Japan. [7]Institute for Advanced Biosciences, Keio University, 246-2 Mizukami, Tsuruoka, Yamagata 997-0052, Japan. [8]Luxembourg Centre for Systems Biomedicine, University of Luxembourg, Avenue des Hauts-Fourneaux, 7, Esch-sur-Alzette, L-4362, Luxembourg. [9]Department of Molecular Preventive Medicine, Graduate School of Medicine, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan. [10]PureTech Ventures, 500 Boylston Street, Suite 1600, Boston, Massachusetts 02116, USA. [11]School of Veterinary Medicine, Azabu University, 1-17-71 Fuchinobe, Sagamihara, Kanagawa 252-5201, Japan.
*These authors contributed equally to this work.

was a significant accumulation of T_H17 cells in the colons of exGF mice inoculated with untreated human faeces (Fig. 1c and Supplementary Fig. 2). Notably, the capacity of human faeces to induce T_H17 cells was greatly diminished after treatment with chloroform (Fig. 1c). These results indicate that the chloroform-sensitive bacterial fraction in the human stool tested contained T_H17-cell-inducing bacteria, whereas the chloroform-resistant bacteria preferentially promoted T_{reg} cell accumulation in the colon.

To investigate whether T_{reg} cell induction by the chloroform-resistant fraction of human intestinal bacteria is transmissible, adult germ-free mice were co-housed with exGF mice inoculated with chloroform-treated human faeces for 4 weeks. Co-housed mice showed a significant increase in the frequency of colonic T_{reg} cells (Fig. 1d). In addition, the progeny of exGF mice inoculated with chloroform-treated human faeces also showed increased numbers of T_{reg} cells (Fig. 1d). Therefore, T_{reg} cell induction by human intestinal bacteria is horizontally and vertically transmissible. Oral inoculation of germ-free mice with 2×10^4-fold diluted caecal samples from exGF mice inoculated with chloroform-treated human faeces fully induced the accumulation of T_{reg} cells in the colon lamina propria, suggesting that abundant rather than minor members of the intestinal microbiota in exGF mice inoculated with chloroform-treated human faeces drive the observed induction of T_{reg} cells (Fig. 1e). The T_{reg}-cell-inducing microbiota in mice inoculated with the 2×10^4-fold diluted sample ($+2 \times 10^4$ mice) was a stable community, because serial oral inoculation of caecal contents from these mice equally induced the accumulation of T_{reg} cells in secondary ($+2\times10^4$-re mice) and tertiary recipients ($+2\times10^4$-re-re mice) (Fig. 1e). To minimize nonessential components of the microbiota for T_{reg} cell induction, the caecal contents of $+2 \times 10^4$ mice were again diluted 2×10^4-fold and orally inoculated into another set of germ-free mice ($+(2 \times 10^4)^2$ mice). The $+(2 \times 10^4)^2$ mice had a marked accumulation of T_{reg} cells in the colon (Fig. 1e). These results suggested that we succeeded in obtaining mice colonized with a relatively restricted and stable community of bacterial species enriched for T_{reg} cell inducers.

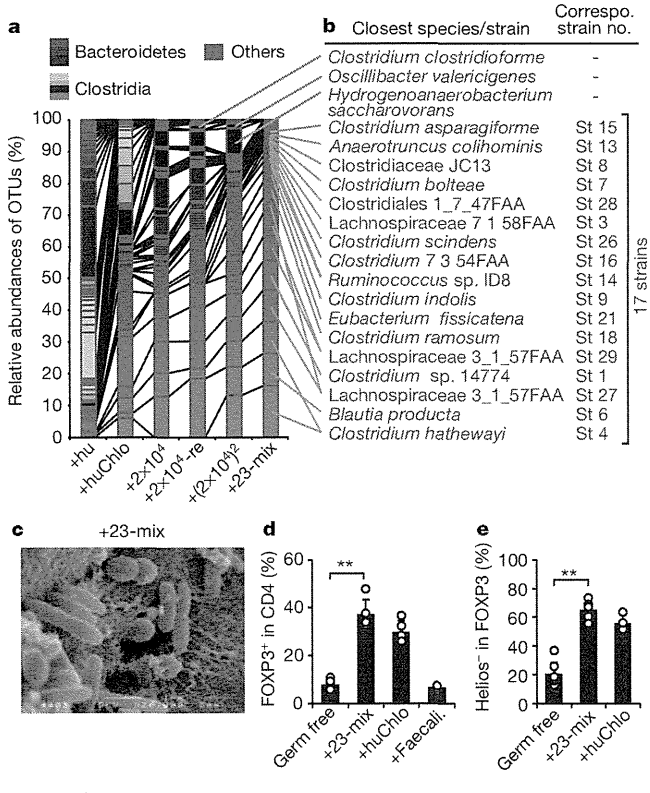


Figure 2 | Assessment of microbiota composition and isolation of T_{reg}-cell-inducing strains. **a**, **b**, Pyrosequencing of 16S rRNA genes was performed on caecal contents from the indicated mice. Relative abundance of OTUs (%) in the caecal bacterial community in each mouse (**a**), and the closest species/strain in the database and the corresponding isolated strain number for the indicated OTU (**b**) are shown. **c**, SEM showing the proximal colon of +23-mix mice. Original magnification, ~×20,000. **d**, **e**, The percentages of $FOXP3^+$ cells within the $CD4^+$ cell population (**d**) and $Helios^-$ cells in $CD4^+FOXP3^+$ cells (**e**) in the colon of the indicated mice. Circles represent individual animals. All experiments were performed more than twice with similar results. Error bars indicate s.d. $^{**}P < 0.01$.

Our composition of the gut microbiota in mice treated with human samples was analysed by 16S ribosomal RNA (rRNA) gene amplicon sequencing using a 454 sequencer. Quality filter-passed sequences (3,000 reads for each sample) were classified into operational taxonomic units (OTUs) based on sequence similarity (>96% identity). The numbers of detected reads and closest known species for each OTU are shown in Supplementary Table 1, and the relative abundance of OTUs in each caecal sample is shown in Fig. 2a. As expected, the OTU profiles of mice treated with human faeces were quite different from those of conventional specific pathogen-free (SPF) mice (Supplementary Fig. 3). In mice inoculated with untreated human faeces, OTUs belonging to Bacteroidetes accounted for about 50% of the caecal microbial community (Fig. 2a). By contrast, most OTUs in exGF mice inoculated with chloroform-treated human faeces were related to Clostridia species. Most bacteria in $+2 \times 10^4$, $+2 \times 10^4$-re and $+(2 \times 10^4)^2$ mice had 16S rRNA gene sequence similarities with about 20 species of Clostridia, listed in Fig. 2b.

To isolate bacterial strains with T_{reg}-cell-inducing capabilities, we cultured caecal contents from $+2 \times 10^4$, $+2 \times 10^4$-re and $+(2 \times 10^4)^2$ mice *in vitro* and picked 442 colonies. BLAST searches of 16S rRNA gene sequences of the isolated colonies revealed that 31 strains in total were present, all of which were Clostridia (Supplementary Fig. 4). Of the 31 strains, we selected 23 that had less than 99% 16S rRNA gene sequence identity to any of the other 30 strains (Supplementary Fig. 4). We then individually cultured the 23 strains, mixed them in equal amounts, and orally inoculated the mixture into germ-free IQI mice (+23-mix mice). Numerous rod- and round-shaped bacteria were observed by scanning electron microscopy (SEM) on the epithelial cell surface in +23-mix mice (Fig. 2c), and the size and appearance of the caeca were quite different from those in germ-free mice, indicating successful colonization (Supplementary Fig. 5a). Pyrosequencing of 16S rRNA genes revealed that the caecal microbiota composition in +23-mix mice was quite similar to that in $+(2 \times 10^4)^2$ mice (Fig. 2a). In +23-mix mice, we observed efficient induction of T_{reg} cells in the colonic lamina propria (Fig. 2d). The magnitude was comparable to that observed in exGF mice inoculated with chloroform-treated human faeces and much higher than that in mice colonized with *Faecalibacterium prausnitzii*, a human Clostridia strain well known for enhancing regulatory cell functions[9] (Fig. 2d). Most T_{reg} cells in +23-mix mice expressed low levels of Helios (also known as IKZF2), indicating antigen-experienced cells (Fig. 2e, Supplementary Fig. 5b and ref. 10).

Only 17 strains listed in Fig. 2b and Supplementary Fig. 4 were detected in +23-mix mice by 16S rRNA gene sequencing, indicating that these 17 strains may be sufficient to induce T_{reg} cells. Indeed, we found that the mixture of 17 strains (17-mix) induced $FOXP3^+$ T_{reg} cells to a similar extent as the 23-mix (Fig. 3a). The increase in T_{reg} cells induced by the 17-mix was reproducibly observed in exGF mice of different genetic backgrounds (IQI, BALB/c and C57BL/6) (Fig. 3a). Moreover, the mix was effective in other rodents: the frequency of colonic T_{reg} cells in exGF rats inoculated with 17-mix was significantly higher than that in germ-free rats and comparable to that in SPF rats (Fig. 3a). The colonization with 17-mix induced a significant increase in the frequency of IL-10^+ and/or $ICOS^+$ cells within the T_{reg} cell population, as revealed by analysis of exGF IL-10 reporter mice (*Il10*^{Venus} mice, ref. 4) colonized with the 17-mix (Fig. 3b). Furthermore, IL-10^+ T_{reg} cells in

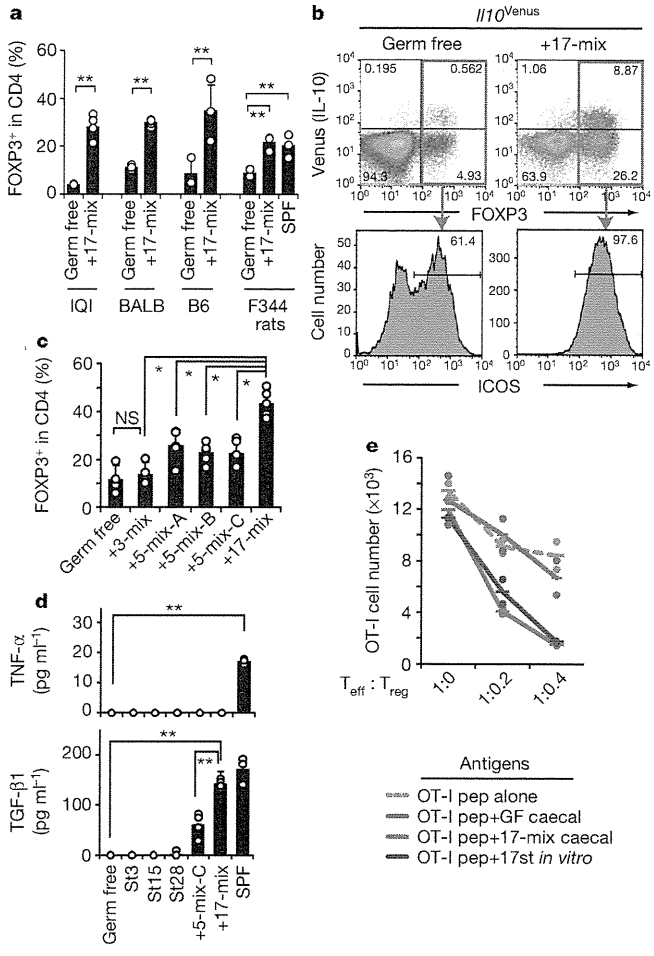


Figure 3 | Characterization of 17 T_{reg}-cell-inducing strains. a, The percentages of $FOXP3^+$ cells within the $CD4^+$ cell population in the colon lamina propria of the indicated mice and rats. **b**, The expression of Venus (IL-10) and FOXP3 by the gated colonic lamina propria $CD4^+$ cells, and ICOS expression by $CD4^+FOXP3^+$ cells in exGF *Il10*^{Venus} mice colonized with or without 17-mix. **c**, Percentages of $FOXP3^+$ cells within the $CD4^+$ cell population in IQI exGF mice colonized with the indicated mix. **d**, The production of TNF-α and TGF-β1 in HCT8 cells stimulated with caecal extracts from the indicated mice. **e**, $CD8^+$ T cells from OT-I mice (T_{eff}) and the indicated ratio of colon lamina propria $CD4^+CD25^+$ cells from +17-mix mice (T_{reg}) were incubated with $CD11c^+$ cells pulsed with OT-I peptide alone or in combination with autoclaved caecal contents from +17-mix mice (+17-mix caecal), germ-free mice (+GF caecal), or autoclaved 17 strains cultured *in vitro* (+17 st *in vitro*). Depicted data represent average of duplicates (see also Supplementary Fig. 9c). Circles in **a**, **c**–**e** represent samples from individual animals. All experiments were performed more than twice with similar results. Error bars indicate s.d. $^{**}P < 0.01$; $^{*}P < 0.05$; NS, not significant.

+17-mix mice expressed high levels of CTLA4 (Supplementary Fig. 5c). Because IL-10 and CTLA4 are essential for the immunosuppressive activity of T_{reg} cells[11,12], and ICOS is required for the T_{reg}-cell-mediated suppression of T_H2 responses[13], these results suggest that the mixture of 17 strains affects both the number and function of T_{reg} cells in the colon. Next, we monocolonized germ-free mice with one of each of the 17 individual strains to determine their individual T_{reg} cell induction capability. The monocolonized exGF mice exhibited low to intermediate levels of T_{reg} cell induction with inter-individual variability (Supplementary Fig. 6a). As expected, none of the strains induced T_H17 cells in the monocolonized mice (Supplementary Fig. 6b). We also examined T_{reg} cell induction by subsets of the 17-mix (randomly selected combinations of 3–5 strains: 3-mix, 5-mix-A, 5-mix-B, and 5-mix-C, see Supplementary Fig. 4). Although all tested combinations of 5-mix induced increases in the frequency of T_{reg} cells, the magnitude was substantially lower than that observed in +17-mix mice (Fig. 3c). Therefore, it is likely that the 17 strains act synergistically to amplify the induction of T_{reg} cells in a microbial-community-dependent fashion.

To investigate the mechanism for the T_{reg} cell induction by the community of 17 strains, we incubated various human and mouse intestinal epithelial cell lines and primary cells with aqueous extracts from caecal contents from the +17-mix mice, and assessed the production of the active form of TGF-β1, a key cytokine for the differentiation and expansion of T_{reg} cells. The caecal extracts from +17-mix mice routinely elicited TGF-β1, but not IL-6 and TNF-α, production, and the magnitude was significantly higher than that elicited by caecal extracts from single-strain or 5-mix-colonized mice (Fig. 3d and Supplementary Fig. 7). The induction of TGF-β1 was not inhibited by pre-treatment of the caecal extracts with a protease or nuclease (Supplementary Fig. 7c). Short-chain fatty acids (SCFAs) are protease- and nuclease-insensitive and have been associated with regulation of host immune homeostasis[14]. Quantitative analysis of SCFAs in caecal contents from +17-mix mice showed a significantly higher concentration of acetate, propionate, butyrate and isobutyrate than that in single-strain or 5-mix-colonized mice (Supplementary Fig. 8a). Furthermore, a mixture of sodium salts of these SCFAs elicited TGF-β1 production in epithelial cells to a level similar to that seen when the cells were stimulated with caecal extracts from +17-mix mice (Supplementary Fig. 8b). Therefore, the community of 17 strains cooperatively produces SCFAs that can elicit a TGF-β response, and this activity may contribute to the differentiation and expansion of T_{reg} cells. We also investigated whether the 17 strains provide bacterial antigens to T cells. To do this, we addressed the antigen specificity of T_{reg} cells accumulated in +17-mix mice using a cognate antigen-driven suppression assay. $CD4^+CD25^+$ lamina propria T cells from +17-mix mice substantially inhibited the OT-I ovalbumin (OVA) peptide-driven proliferation of OT-I CD8 T cells, and this suppression was markedly enhanced in the presence of autoclaved caecal content from +17-mix mice or autoclaved 17 strains cultured *in vitro*, but not in the presence of OT-II OVA peptide or caecal content from germ-free mice (Fig. 3e and Supplementary Fig. 9). These results are consistent with previous reports[15,16] and suggest that some fraction of colonic lamina propria T_{reg} cells in +17-mix mice is specific to the 17 strains of Clostridia. Next, we assessed the kinetics of T_{reg} cell accumulation and their expression of Ki67, a cell-cycle-associated nuclear protein, and gut-homing-associated molecules CD103 and β7 integrin. We observed a marked increase in the proportion of Ki67, CD103 and β7 expressing cells by 1 week after inoculation with the 17-mix (Supplementary Figs 10 and 11). Collectively, these observations indicate that the 17 strains provide SCFAs, bacterial antigens and probably other factors, which together contribute to differentiation, expansion and colonic homing of T_{reg} cells.

To define the identity of the 17 bacterial strains fully, we sequenced their genomes (Supplementary Fig. 12). Phylogenetic comparison of the 17 strains using ribosomal multi locus sequencing typing (rMLST) revealed that the 17 strains belong to bacterial species falling within clusters XIVa, IV and XVIII of Clostridia as defined previously[17] (in a recent taxonomy, members of cluster XVIII Clostridia were reclassified in the class Erysipelotrichi) (Supplementary Fig. 13). The genome sequencing also revealed that the 17 strains all lack strong virulence-related genes such as collagenase and phospholipase C, often identified in pathogenic Clostridia species (Supplementary Table 2). We then examined the relative abundance of the 17 strains in healthy and ulcerative colitis human subjects using draft genome sequences of the 17 strains and publicly available human microbiome genomes generated through the MetaHIT project[18]. Ulcerative colitis subjects showed a tendency towards a reduction of the 17 strains, and 5 out of the 17 strains were significantly reduced in ulcerative colitis patients (Supplementary Fig. 14).

To evaluate the potential benefits of supplementation with the 17 strains, 17-mix or control PBS was orally administered into adult SPF

mice every 2 or 3 days (SPF + 17-mix or SPF + ctrl, respectively). We confirmed a significant increase in the frequency of colonic T_{reg} cells in SPF + 17-mix mice compared with SPF + ctrl mice after 3 weeks of treatment (Fig. 4a). While being continuously treated with 17-mix or control, mice were subjected to the OVA-induced allergic diarrhoea model[19]. The occurrence and severity of diarrhoea and the OVA-specific serum IgE levels were significantly reduced in SPF + 17-mix mice relative to control mice (Fig. 4b–d). The protective effect of 17-mix was significantly attenuated by treatment of mice with a T_{reg}-cell-depleting anti-CD25 antibody (Supplementary Fig. 15). We also subjected mice to an experimental colitis model induced by trinitrobenzene sulphonic acid (TNBS)[20]. SPF + 17-mix mice showed less severe colon shortening and milder histological disease features, accompanied by lower mortality

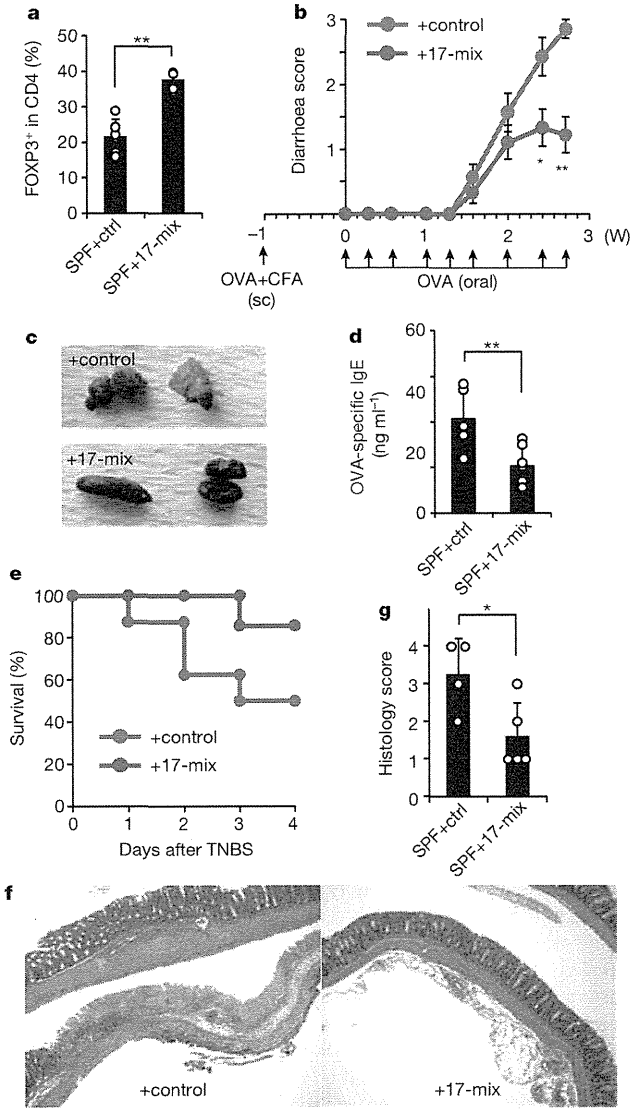


Figure 4 | Treatment with 17-mix suppresses experimental colitis models. **a**, The percentages of FOXP3$^+$ cells within the CD4$^+$ cell population in SPF + 17-mix or SPF + ctrl mice. **b–d**, SPF + 17-mix ($n = 9$) and SPF + ctrl ($n = 7$) mice were subjected to OVA-induced diarrhoea. The diarrhoea score (**b**; see Methods for definition), representative photographs of faeces (**c**), and OVA-specific IgE levels in the sera (**d**) are shown. sc, subcutaneous. **e–g**, SPF + 17-mix ($n = 8$) and SPF + ctrl ($n = 7$) were treated with TNBS. Animal survival (**e**), haematoxylin and eosin staining (original magnification, ×10) (**f**), and histology score of the distal colon (**g**) on day 4 after TNBS administration are shown. Data are representative of two independent experiments. Error bars indicate s.d. $^{**}P < 0.01$; $^{*}P < 0.05$.

than control mice (Fig. 4e–g and Supplementary Fig. 16a). In keeping with this clinical outcome, there was significantly increased expression of *Foxp3* and *Tgfb1* mRNA in SPF + 17-mix mice compared with control mice, as well as a tendency towards a reduction of inflammatory cytokine transcripts (Supplementary Fig. 16b). Identical suppression of colitis by 17-mix was also observed in an adoptive transfer model, in which germ-free SCID mice were orally inoculated with faeces from SPF mice together with or without 17-mix and then transferred with CD4$^+$CD45RB^{hi} T cells (Supplementary Fig. 17).

The clinical track record of efficacy of single-strain probiotics has been modest. It has been postulated that a collection of functionally distinct bacterial species rationally selected from the human gut microbiota may be more effective than single strains in preventing/treating disease[21]. In the present study, we isolated 17 strains within Clostridia clusters XIVa, IV and XVIII from a human faecal sample; these strains affect T_{reg} cell differentiation, accumulation and function in the mouse colon. It remains to be seen whether the 17 strains will have similar effects in the human intestine; however, a decreased prevalence of Clostridia clusters XIVa and IV in faecal samples from patients with inflammatory bowel disease and atopy[22–24] may suggest that supplementation with the 17-strain bacterial community might counterbalance dysbiosis, induce T_{reg} cells and aid in the management of allergic and inflammatory conditions.

METHODS SUMMARY

Experiments were performed with authorization from the Institutional Review Board for Human Research at RIKEN Yokohama Research Institute. Human stool from a healthy volunteer (Japanese, male, age 31 years) was obtained with informed consent. The sample was mixed with or without chloroform, and the aliquots were inoculated into germ-free IQI mice. Detailed procedures for lamina propria lymphocyte analysis, isolation of bacteria, extraction of bacterial DNA and sequencing are described in Methods. Statistical analysis was performed using the Student's *t*-test.

Full Methods and any associated references are available in the online version of the paper.

Received 14 September 2012; accepted 22 May 2013.
Published online 10 July 2013.

1. Round, J. L. & Mazmanian, S. K. The gut microbiota shapes intestinal immune responses during health and disease. *Nature Rev. Immunol.* **9,** 313–323 (2009).
2. Honda, K. & Littman, D. R. The microbiome in infectious disease and inflammation. *Annu. Rev. Immunol.* **30,** 759–795 (2012).
3. O'Toole, P. W. & Cooney, J. C. Probiotic bacteria influence the composition and function of the intestinal microbiota. *Interdiscip. Perspect. Infect. Dis.* **2008,** 175285 (2008).
4. Atarashi, K. *et al.* Induction of colonic regulatory T cells by indigenous *Clostridium* species. *Science* **331,** 337–341 (2011).
5. Geuking, M. B. *et al.* Intestinal bacterial colonization induces mutualistic regulatory T cell responses. *Immunity* **34,** 794–806 (2011).
6. Russell, S. L. *et al.* Early life antibiotic-driven changes in microbiota enhance susceptibility to allergic asthma. *EMBO Rep.* **13,** 440–447 (2012).
7. Round, J. L. & Mazmanian, S. K. Inducible Foxp3$^+$ regulatory T-cell development by a commensal bacterium of the intestinal microbiota. *Proc. Natl Acad. Sci. USA* **107,** 12204–12209 (2010).
8. Chung, H. *et al.* Gut immune maturation depends on colonization with a host-specific microbiota. *Cell* **149,** 1578–1593 (2012).
9. Sokol, H. *et al. Faecalibacterium prausnitzii* is an anti-inflammatory commensal bacterium identified by gut microbiota analysis of Crohn disease patients. *Proc. Natl Acad. Sci. USA* **105,** 16731–16736 (2008).
10. Thornton, A. M. *et al.* Expression of Helios, an Ikaros transcription factor family member, differentiates thymic-derived from peripherally induced Foxp3$^+$ T regulatory cells. *J. Immunol.* **184,** 3433–3441 (2010).
11. Rubtsov, Y. P. *et al.* Regulatory T cell-derived interleukin-10 limits inflammation at environmental interfaces. *Immunity* **28,** 546–558 (2008).
12. Wing, K. *et al.* CTLA-4 control over Foxp3$^+$ regulatory T cell function. *Science* **322,** 271–275 (2008).
13. Zheng, Y. *et al.* Regulatory T-cell suppressor program co-opts transcription factor IRF4 to control T_H2 responses. *Nature* **458,** 351–356 (2009).
14. Maslowski, K. M. & Mackay, C. R. Diet, gut microbiota and immune responses. *Nature Immunol.* **12,** 5–9 (2011).
15. Lathrop, S. K. *et al.* Peripheral education of the immune system by colonic commensal microbiota. *Nature* **478,** 250–254 (2011).
16. Cebula, A. *et al.* Thymus-derived regulatory T cells contribute to tolerance to commensal microbiota. *Nature* **497,** 258–262 (2013).

17. Collins, M. D. *et al.* The phylogeny of the genus *Clostridium*: proposal of five new genera and eleven new species combinations. *Int. J. Syst. Bacteriol.* **44,** 812–826 (1994).
18. Qin, J. *et al.* A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* **464,** 59–65 (2010).
19. Kweon, M. N., Yamamoto, M., Kajiki, M., Takahashi, I. & Kiyono, H. Systemically derived large intestinal $CD4^+$ Th2 cells play a central role in STAT6-mediated allergic diarrhea. *J. Clin. Invest.* **106,** 199–206 (2000).
20. Strober, W., Fuss, I. J. & Blumberg, R. S. The immunology of mucosal models of inflammation. *Annu. Rev. Immunol.* **20,** 495–549 (2002).
21. Lawley, T. D. *et al.* Targeted restoration of the intestinal microbiota with a simple, defined bacteriotherapy resolves relapsing *Clostridium difficile* disease in mice. *PLoS Pathog.* **8,** e1002995 (2012).
22. Frank, D. N. *et al.* Molecular-phylogenetic characterization of microbial community imbalances in human inflammatory bowel diseases. *Proc. Natl Acad. Sci. USA* **104,** 13780–13785 (2007).
23. Manichanh, C. *et al.* Reduced diversity of faecal microbiota in Crohn's disease revealed by a metagenomic approach. *Gut* **55,** 205–211 (2006).
24. Candela, M. *et al.* Unbalance of intestinal microbiota in atopic children. *BMC Microbiol.* **12,** 95 (2012).

Supplementary Information is available in the online version of the paper.

Acknowledgements This work was supported by JSPS NEXT program, Grant in Aid for Scientific Research on Innovative Areas 'Genome Science' from the Ministry of Education, Culture, Sports, Science and Technology of Japan (No.221S0002), the global COE project of 'Genome Information Big Bang' and the Waksman Foundation of Japan Inc. We thank M. Suyama, K. Furuya, C. Yoshino, H. Inaba, E. Iioka, Y. Takayama, M. Kiuchi, Y. Hattori, N. Fukuda and A. Nakano for technical assistance, and P. D. Burrows for review of the manuscript.

Author Contributions K.Ho. planned experiments, analysed data and wrote the paper together with B.O. and M.H.; K.A. and T.Tano. performed immunological analyses and bacterial cultures together with Y.N., S.N. and H.M.; W.S., K.O., S.K. and M.H. performed bacterial sequence analyses; K.M. and S.U. provided essential materials; H.N., T.S. and S.S. supervised the T_{reg} cell suppression assay; S.F., K.Ha., H.O., T.Tani., J.V.F. and P.W. were involved in data discussions.

Author Information All genome sequence data are deposited in DDBJ BioProject ID PRJDB521-543. Reprints and permissions information is available at www.nature.com/reprints. The authors declare competing financial interests: details are available in the online version of the paper. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to M.H. (hattori@k.u-tokyo.ac.jp) or K.Ho. (kenya@rcai.riken.jp).

METHODS

Mice and rats. C57BL/6, BALB/c, IQI/Jic and CB.17 SCID mice and F344 rats kept under SPF or germ-free conditions were purchased from Sankyo laboratories, Japan SLC, or CLEA Japan. IQI germ-free mice were used unless otherwise indicated. Germ-free and gnotobiotic mice were bred and maintained in vinyl isolators within the gnotobiotic facility of Sankyo laboratories. Germ-free *Il10*^{Venus} mice were generated as previously described[4]. OT-I and OT-II T-cell receptor transgenic mice were purchased from Taconic Farms. All animal experiments were approved by the Animal Research Committee of RIKEN Yokohama Institute and the University of Tokyo.

Chloroform treatment of human stool and generation of gnotobiotic mice. Human stool from a healthy volunteer (Japanese, male, age 31 years) was obtained with informed consent. Human stool and mouse caecal contents were directly frozen at −80 °C, or suspended in 4 times volume (w/v) of phosphate-buffered saline (PBS) + 20% glycerol solution, snap-frozen in liquid nitrogen and stored at −80 °C until use. The frozen stocks were thawed, suspended in 10 times volume (w/v) of PBS, and passed through a 70 μm cell strainer to eliminate clumps and debris. Then suspensions were mixed with chloroform (final concentration 3%), and incubated in a shaking water bath for 60 min. After evaporation of chloroform by bubbling with N_2 gas for 30 min, the aliquots containing the chloroform-resistant fraction of intestinal bacteria were inoculated into germ-free mice by intra-gastric administration (250 μl; per mouse). To generate a series of gnotobiotic mice inoculated with diluted samples, caecal contents from exGF mice were treated with chloroform, diluted with PBS, and inoculated into germ-free IQI mice. The caecal suspensions diluted 2×10^4-fold correspond to 2.5×10^4 bacterial cells per mouse. Each group of exGF mice was individually caged in the gnotobiotic isolator for 3–4 weeks at Sankyo Lab service.

Isolation of intestinal lamina propria lymphocytes and flow cytometry. The colons were collected and opened longitudinally, washed with PBS to remove all luminal contents and shaken in Hanks' balanced salt solution (HBSS) containing 5 mM EDTA for 20 min at 37 °C. After removing epithelial cells, muscle layers and fat tissue using forceps, the lamina propria layers were cut into small pieces and incubated with RPMI1640 containing 4% fetal bovine serum, 0.5 mg ml^{-1} collagenase D, 0.5 mg ml^{-1} dispase and 40 μg ml^{-1} DNase I (all Roche Diagnostics) for 1 h at 37 °C in a shaking water bath. The digested tissues were washed with HBSS containing 5 mM EDTA, resuspended in 5 ml of 40% Percoll (GE Healthcare) and overlaid on 2.5 ml of 80% Percoll in a 15-ml Falcon tube. Percoll gradient separation was performed by centrifugation at 800*g* for 20 min at 25 °C. The lamina propria lymphocytes were collected from the interface of the Percoll gradient and suspended in ice-cold PBS. For analysis of T_{reg} cells, isolated lymphocytes were labelled with the LIVE/DEAD fixable dead cell stain kit (Invitrogen) to exclude dead cells from the analysis. The cells were washed with staining buffer containing PBS, 2% FBS, 2 mM EDTA and 0.09% NaN_3 and surface staining was performed with PECy7- or Pacific blue-labelled anti-CD4 antibody (RM4-5, BD Biosciences), PE-labelled anti-ICOS antibody (C938.4A, BioLegend), Alexa488-labelled anti-CD103 antibody (2E7, BioLegend), and PerCP/Cy5.5-labelled anti-integrin-β7 antibody (FIB27, BioLegend). Intracellular staining of FOXP3, CTLA4, Helios and Ki67 was performed using the Alexa647-labelled anti-FOXP3 antibody (FJK-16 s, eBioscience), PE-labelled anti-CTLA4 antibody (UC10-4F10-11, BD Biosciences), PE-labelled anti-Helios antibody (22F6, BioLegend), PECy7-labelled anti-Ki67 antibody (B56, BD Biosciences) and FOXP3 staining buffer set (eBioscience). For analysis of T_H1 and T_H17 cells, isolated lymphocytes were stimulated for 4 h with 50 ng ml^{-1} phorbol 12-myristate 13-acetate (PMA, Sigma) and 1 μg ml^{-1} ionomycin (Sigma) in the presence of GolgiStop (BD Biosciences). After incubation for 4 h, cells were washed in PBS, labelled with the LIVE/DEAD fixable dead cell stain kit and surface CD4 was stained with PECy7-labelled anti-CD4 antibody. Cells were washed, fixed in Cytofix/Cytoperm, permeabilized with Perm/Wash buffer (BD Biosciences), and stained with the APC-labelled anti-IL-17 antibody (eBio17B7, eBioscience) and FITC-labelled anti-IFN-γ antibody (XMG1.2, BD Biosciences). The antibody-stained cells were analysed with LSR Fortessa or FACSAriaIII (BD Biosciences), and data were analysed using FlowJo software (Treestar).

Meta 16S rRNA gene sequencing. The caecal contents from exGF mice were suspended in 10 ml of Tris-EDTA containing 10 mM Tris-HCl and 1 mM EDTA (pH 8), and incubated with lysozyme (Sigma, 15 mg ml^{-1}) at 37 °C for 1 h with gentle mixing. A purified achromopeptidase (Wako) was added (final concentration 2,000 unit ml^{-1}) and further incubated at 37 °C for another 30 min. Then, sodium dodecyl sulphate (final concentration 1%) was added to the cell suspension and mixed well. Subsequently, proteinase K (Merck) was added (final concentration 1 mg ml^{-1}) to the suspension and the mixture was incubated at 55 °C for 1 h. High-molecular-mass DNA was isolated and purified by phenol/chloroform extraction, ethanol and finally polyethyleneglycol precipitation[25]. PCR was performed using Ex Taq (TAKARA) and (1) the 454 primer A (5′-CCATCTCA TCCCTGCGTGTCTCCGACTCAG (454 adaptor sequence) + barcode (10 bases) + AGRGTTTGATYMTGGCTCAG-3′ (27Fmod)) and (2) the 454 primer B (5′-CCTATCCCCTGTGTGCCTTGGCAGTCTCAG (454 adaptor sequence) + TGCTGCCTCCCGTAGGAGT-3′ (338R)) to the V1–V2 region of the 16S rRNA gene. Amplicons generated from each sample (~330 bp) were subsequently purified using AMPur XP (Beckman Coulter). The amount of DNA was quantified using Quant-iT Picogreen dsDNA assay kit (Invitrogen) and TBS-380mini fluorometer (Turner Biosystems). Then, the amplified DNA was used as template for 454 GS Junior (Roche) pyrosequencing using GS Junior Titanium emPCR Kit-Lib-L, GS Junior Titanium Sequencing Kit and GS Junior Titanium PicoTiterPlate Kit (all Roche) according to the manufacturer's instructions. Quality-filter-passed reads were obtained by removing reads that did not have both primer sequences, had the average quality value (QV) <25, and were possibly chimaeric[26]. Of the filter-passed reads, 3,000 reads trimming off both primer sequences for each sample were used and subjected to OTU analysis with the cutoff similarity of 96% identity. Representative sequences from each OTU were blasted to Ribosomal Database Project (RDP) of bacterial isolates, our genome database constructed from publically available genome sequences in NCBI and HMP databases, and 16S sequences of the 23 strains obtained in this study.

Isolation of bacterial strains. The frozen stocks of caecal contents from exGF mice were serially diluted with PBS and seeded onto non-selective agar plates (blood liver (BL) agar (Eiken Chemical) or Eggerth-Gagnon (EG) agar plates). EG agar plates contain the following components (quantities expressed per litre): Lab-Lemco Powder (2.8 g, Oxoid); proteose peptone no. 3 (10.0 g, Difco); yeast extract (5.0 g, Difco); Na_2HPO_4 (4.0 g); D(+)-glucose (1.5 g); soluble starch (0.5 g); L-cystine (0.2 g); L-cysteine-HCl-H_2O (0.5 g); Tween 80 (0.5 g); Bacto agar (16.0 g, Difco); and defibrinated horse blood (50 ml). After culture under aerobic conditions or strictly anaerobic conditions (80% N_2, 10% H_2, 10% CO_2) at 37 °C for 2 or 4 days, individual colonies were picked up and cultured for an additional 2 or 4 days at 37 °C in ABCM broth (Eiken Chemical) or EG agar plate. The isolated strains were collected into EG stock medium (10% glycerol) and stored at −80 °C. To identify the isolated strains, 16S rRNA gene sequences were determined. The 16S rRNA gene was amplified by colony-PCR using KOD FX (TOYOBO) and GeneAmp PCR System9700 (Applied Biosystems) using 16S rRNA gene-specific primer pairs: 8F (5′-AGAGTTTGATCMTGGCTCAG-3′) and 519R (5′-ATTACCGCGGCKGCTG-3′) or 1513R (5′-ACGGCTACCTTGTTACGACTT-3′). The amplification program consisted of one cycle at 98 °C for 2 min, followed by 40 cycles at 98 °C for 10 s, 57 °C for 30 s and 68 °C for 1 min 30 s. Each amplified DNA was purified from the reaction mixture using Illustra GFX PCR DNA and gel band purification kit (GE Healthcare). Sequence analysis was performed using BigDye Terminator V3.1 cycle sequencing kit (Applied Biosystems) and Applied Biosystems 3730xl DNA analyser (Applied Biosystems). The resulting sequences were compared with sequences in RDP database and genome database using BLAST to determine close species/strains.

Bacterial culture of isolated strains. The isolated strains of Clostridia and Erysipelotrichi were cultured in EG broth without horse blood under a strictly anaerobic condition (80% N_2, 10% H_2, 10% CO_2) at 37 °C in an anaerobic chamber (Coy Laboratory Products). To prepare the bacterial mixture, bacterial strains were individually grown in EG broth to confluence and mixed at equal amounts of media volume.

Scanning electron microscopy. Scanning electron microscopy was performed by Filgen, Inc., Japan. The proximal colon was removed from +23-mix mice, cut open longitudinally, prefixed with 2% glutaraldehyde in 0.1 M phosphate buffer (pH 7.4) for 24 h at 4 °C, and then postfixed with 2% osmium tetroxide for 1 h at 4 °C. Fixed samples were dehydrated for 5 min each in sequential baths of 50%, 70%, 90% and 100% ethanol, inserted into a critical point dryer until dry and coated with osmium in an OPC-80N osmium plasma coater (Filgen). Scanning electron micrographs were taken by a JEOL JSM-6320F instrument.

Measurement of organic acids. Organic acid concentrations in caecal contents were determined by gas chromatography-mass spectrometry (GC-MS). Caecal contents (10 mg) were disrupted using 3-mm zirconia/silica beads (BioSpec Products) and homogenized in extraction solution containing 100 μl of internal standard (100 μM crotonic acid), 50 μl of HCl and 200 μl of ether. After vigorous shaking using a Shakemaster neo (Bio Medical Science) at 1,500 r.p.m. for 10 min, homogenates were centrifuged at 1,000*g* for 10 min and then the top ether layer was collected and transferred into new glass vials. Aliquots (80 μl) of the ether extracts were mixed with 16 μl of *N*-tert-butyldimethylsilyl-*N*-methyltrifluoroacetamide (MTBSTFA). The vials were sealed tightly by screwing and heated at 80 °C for 20 min in a water bath, and left at room temperature for 48 h for derivatization. The derivatized samples were run through a 6890N Network GC System (Agilent Technologies) equipped with HP-5MS column (0.25 mm × 30 m × 0.25 μm) and 5973 Network Mass Selective Detector (Agilent Technologies). Pure helium (99.9999%) was used as a carrier gas and delivered at a flow rate of 1.2 ml min^{-1}.

The head pressure was set at 10 p.s.i. with split 10:1. The inlet and transfer line temperatures were 250 °C and 260 °C, respectively. The following temperature program was used: 60 °C (3 min), 60–120 °C (5 °C min^{-1}), 120–300 °C (20 °C min^{-1}). One microlitre quantity of each sample was injected with a run time of 30 min. Organic acid concentrations were quantified by comparing their peak areas with the standards.

Genome sequencing and gene prediction. The genome sequences of 17 T_{reg}-cell-inducing strains were determined by the whole-genome shotgun strategy using a 454GS FLX Ti or Ion PGM sequencer. Each 1–5 μg of the genomic DNA was sheared to obtain DNA fragments. Template DNA was prepared according to the supplier's protocol. The generated sequence data were assembled using Newbler v2.8 software to obtain the draft genome sequences. All genome sequence data were deposited in DDBJ BioProject ID: PRJDB521-543. Protein-encoding genes were predicted using MetaGeneAnnotator software[27]. Putative toxins and virulence factors were searched using the BLASTP program and virulence factor databases, VFDB (http://www.mgc.ac.cn/VFs/main.htm) and MvirDB (http://mvirdb.llnl.gov), with the *e*-value cutoff of 1.0×10^{-10}, the identity >30% and the length coverage >60%.

Phylogenetic tree. Sequences concatenated with genes encoding 26 ribosomal proteins (large subunit L10, L11, L14, L16, L17, L19, L20, L23, L24, L29, L31, L32, L35, L7/L12, and small subunit S10, S12, S13, S15, S16, S17, S20, S20, S21, S3, S4, S7, S8) predicted from the genomes of each strain were used to construct a phylogenetic tree. The sequences of other bacterial species used for the tree construction were obtained from the ribosomal multi-locus sequencing typing (MLST) database[28]. The calculation was performed using the MEGA v5.0 package and the neighbour-joining method with a bootstrap of 1,000 replicates.

Cognate antigen-driven T_{reg} cell suppression assay. Preparation of antigens in caecal contents was performed as previously reported[15]. Caecal contents from germ-free mice or +17-mix mice were collected and suspended in PBS (500 mg ml^{-1}); they were then filtered through a 70-μm mesh, and autoclaved at 121 °C for 15 min. To prepare antigens of bacterial components, the 17 strains of Clostridia were cultured *in vitro*, mixed, washed and suspended with 1 ml PBS, and autoclaved at 121 °C for 20 min. $CD11c^+$ cells were isolated by FACSAriaIII from spleens of SPF C57BL/6 mice and pulsed for 1 h with 0.5 μM SIINFEKL OT-I peptide alone or in combination with either of 5 μM ISQAVHAAHAEINEAGR OT-II peptide, autoclaved caecal contents from +17-mix mice or germ-free mice (diluted 1:200), or autoclaved 17 strains of bacteria cultured *in vitro* (diluted 1:200). The antigen-pulsed $CD11c^+$ cells were plated at 5×10^4 per well in 96-well round-bottomed plates. CD8 T cells (T_{eff} cells) were sorted from spleens of SPF OT-I mice by FACSAriaIII and added to the $CD11c^+$ cell-seeded plates at 5×10^4 per well. Then, $CD4^+CD25^+$ T cells (T_{reg} cells) sorted from colonic lamina propria of +17-mix mice or from spleens of SPF OT-II mice were added to the culture at the indicated ratio of T_{reg} to T_{eff} cells. After 3 days, all cells were harvested, stained with anti-CD4 and anti-CD8 antibodies, and analysed by FACSAriaIII to enumerate the number of CD8 OT-I T cells.

Intestinal epithelial cell stimulation with caecal extracts and SCFAs. To prepare caecal extracts, frozen caecal contents from germ-free, +17-mix or SPF mice were thawed and well suspended in 4 volumes of sterile water. After centrifugation (5,000 r.p.m. for 15 min), transparent supernatants were collected, filtered through 0.22 μm filter and used as caecal extracts. In some experiments, caecal extracts were treated with proteinase K (2 mg ml^{-1}, 55 °C for 1 h; Roche) or nuclease that degrades all forms of DNA and RNA (125 unit ml^{-1}, 37 °C for 4 h; Thermo), and subsequently heated at 95 °C for 5 min to inactivate the enzymes. Human intestinal epithelial cell lines (HCT8, HT29, Caco2, T84 and Colo205) and a mouse epithelial cell line (CMT93) were obtained from ATCC and maintained at 37 °C (5% CO2) in RPMI containing 10% heat-inactivated horse serum (Invitrogen). Cells were cultured at 1.5×10^5 cells in 150 μl medium in 48-well plates and stimulated with 4.5 μl caecal extract for 24 h. Human primary intestinal epithelial cells were obtained from Lonza and maintained at 33 °C (5% CO2) in SmGM-2 medium containing 10% FBS (Lonza) for 1–2 weeks (6×10^4 cells in 48-well plates). The medium was changed to 150 μl SmGM-2 containing 1% FBS before stimulation. Caecal extracts (4.5 μl) were added to the culture and incubated for 24 h. Culture supernatants were collected and the level of the active form of TGF-β1 (Promega), TNF-α (R&D) and IL-6 (R&D) was measured by ELISA. To stimulate epithelial cell lines with SCFAs, sodium salts of acetate, butyrate, propionate and isobutyrate were dissolved in PBS. SCFAs were added to the culture individually (final 0.5 mM) or in combination (final 0.5 mM each), and incubated for 24 h.

TNBS colitis. C57BL/6 SPF adult mice were orally inoculated with 17-mix or control PBS every 2 or 3 days for 3 weeks. 2,4,6-Trinitrobenzene sulphonic acid (TNBS)-induced colitis was induced by the intracolonic administration of 2.5 mg of TNBS (Sigma) in 50% ethanol into anaesthetized mice via a thin round-tip needle. The tip of the needle was inserted 4 cm proximal to the anal verge, and mice were held in a vertical position for 30 s after the injection. All the mice were observed daily and were killed on day 4 after TNBS administration. Colons were fixed with 4% paraformaldehyde, sectioned, and stained with haematoxylin and eosin. The degree of inflammation in the distal part of colon was graded from 0 to 4 as follows: 0, normal; 1, ulcer with cell infiltration limited to the mucosa; 2, ulcer with limited cell infiltration in the submucosa; 3, focal ulcer involving all layers of the colon; 4, multiple lesions involving all layers of the colon, or necrotizing ulcer larger than 1 mm in length.

Allergic diarrhoea. BALB/c SPF adult mice were primed by subcutaneous injection with 1 mg of OVA (Fraction V; Sigma) in 100 μl of Complete Freund Adjuvant (CFA, DIFCO). One week after priming, mice were given 50 mg of OVA dissolved in 200 μl of PBS by intra-gastric administration three times per week. 17-mix or control PBS was orally administered to mice every 2 or 3 days for the entire period of the experiments. Diarrhoea was monitored visually 1 h after each oral OVA challenge. Diarrhoea was scored as follows: 0, normal faeces (solid); 1, moist faeces (semi-solid); 2, mild diarrhoea (loose); and 3, severe diarrhoea (watery). Serum was collected from the cheek vein 1 h after the last OVA challenge and OVA-specific IgE levels were measured by ELISA (Chondrex).

Adoptive $CD4^+CD45RB^{hi}$ T-cell transfer model of colitis. Germ-free CB.17 SCID mice were orally inoculated with SPF faeces together with or without 17-mix of Clostridia. One week later, exGF SCID mice received 4×10^5 $CD4^+CD45RB^{hi}$ T cells by intraperitoneal injection. Naive $CD4^+CD45RB^{hi}$ T cells were isolated from spleens of SPF BALB/c mice by FACS sorting. All the mice were observed daily and were killed on day 14 after T-cell transfer.

25. Morita, H. *et al.* An improved isolation method for metagenomic analysis of the microbial flora of the human intestine. *Microbes Environ.* **22,** 214–222 (2007).
26. Kim, S. W. *et al.* Robustness of gut microbiota of healthy adults in response to probiotic intervention revealed by high-throughput pyrosequencing. *DNA Res.* **20,** 241–253 (2013).
27. Noguchi, H., Taniguchi, T. & Itoh, T. MetaGeneAnnotator: detecting species-specific patterns of ribosomal binding site for precise gene prediction in anonymous prokaryotic and phage genomes. *DNA Res.* **15,** 387–396 (2008).
28. Jolley, K. A. *et al.* Ribosomal multilocus sequence typing: universal characterization of bacteria from domain to strain. *Microbiology* **158,** 1005–1015 (2012).

LETTER

doi:10.1038/nature12347

Obesity-induced gut microbial metabolite promotes liver cancer through senescence secretome

Shin Yoshimoto[1,2]*, Tze Mun Loo[1,2,3], Koji Atarashi[4,5], Hiroaki Kanda[6], Seidai Sato[1,2], Seiichi Oyadomari[7], Yoichiro Iwakura[8], Kenshiro Oshima[9], Hidetoshi Morita[10], Masahira Hattori[9], Kenya Honda[2,4], Yuichi Ishikawa[6], Eiji Hara[1,2] & Naoko Ohtani[1,5]*

Obesity has become more prevalent in most developed countries over the past few decades, and is increasingly recognized as a major risk factor for several common types of cancer[1]. As the worldwide obesity epidemic has shown no signs of abating[2], better understanding of the mechanisms underlying obesity-associated cancer is urgently needed. Although several events were proposed to be involved in obesity-associated cancer[1,3], the exact molecular mechanisms that integrate these events have remained largely unclear. Here we show that senescence-associated secretory phenotype (SASP)[4,5] has crucial roles in promoting obesity-associated hepatocellular carcinoma (HCC) development in mice. Dietary or genetic obesity induces alterations of gut microbiota, thereby increasing the levels of deoxycholic acid (DCA), a gut bacterial metabolite known to cause DNA damage[6]. The enterohepatic circulation of DCA provokes SASP phenotype in hepatic stellate cells (HSCs)[7], which in turn secretes various inflammatory and tumour-promoting factors in the liver, thus facilitating HCC development in mice after exposure to chemical carcinogen. Notably, blocking DCA production or reducing gut bacteria efficiently prevents HCC development in obese mice. Similar results were also observed in mice lacking an SASP inducer[8] or depleted of senescent HSCs, indicating that the DCA–SASP axis in HSCs has key roles in obesity-associated HCC development. Moreover, signs of SASP were also observed in the HSCs in the area of HCC arising in patients with non-alcoholic steatohepatitis[3], indicating that a similar pathway may contribute to at least certain aspects of obesity-associated HCC development in humans as well. These findings provide valuable new insights into the development of obesity-associated cancer and open up new possibilities for its control.

Cellular senescence is a process occurring in normal cells in response to telomere erosion or oncogene activation, acting through checkpoint activation and stable cell-cycle arrest as a barrier to tumorigenesis[9,10]. Recent studies, however, reveal that senescent cells also develop a secretory profile composed mainly of inflammatory cytokines, chemokines and proteases, a typical signature termed the senescence-associated secretory phenotype (SASP)[4] or the senescence messaging secretome (SMS)[5], hereafter referred to as SASP. Some of the SASP factors have cell-autonomous activities that reinforce senescence cell-cycle arrest[5] and/or promote clearance of senescent cells[11,12], but other SASP factors have cell non-autonomous functions associated with inflammation and tumorigenesis promotion[4], indicating that SASP contributes positively and negatively to cancer development, depending on the biological context[4,5]. Because some of the SASP factors, such as IL-6 and PAI-1[4,5], are known to increase cancer risk in obesity[1,13], we propose that SASP may contribute to obesity-associated cancer.

To explore this possibility, we first set up a system to examine the impact of dietary obesity on tumorigenesis, using wild-type C57BL/6 mice. However, we were unable to detect a statistically significant difference in cancer development between obese mice fed a high-fat diet (HFD) and lean mice fed a normal diet (data not shown), implying that a certain level of oncogenic stimuli might be required for obesity-associated cancer, especially in wild-type mice maintained in a specific pathogen free (SPF) environment. Because the Ras-pathway is frequently activated in human cancers, including hepatocellular carcinoma (HCC)[14], we decided to use a treatment with DMBA (7,12-dimethylbenz(*a*)anthracene, a chemical carcinogen that causes an oncogenic Ras mutation) at the neonatal stage, a protocol known to generate a variety of tumours throughout the body[15]. In this setting, we also took advantage of using p21-*p*-luc mice, in which the expression of the $p21^{Waf1/Cip1}$ gene (a senescence inducer, also known as *Cdkn1a*) can be monitored noninvasively using a bioluminescence imaging (BLI) technique[16]. The neonatal p21-*p*-luc mice were therefore treated with a single application of DMBA, followed by feeding either HFD or normal diet for 30 weeks (Fig. 1a). Interestingly, a marked increase of the bioluminescent signal was observed in the abdomen of the obese mice, and it originated mainly from liver cancer (Supplementary Fig. 1). Notably, all HFD-fed mice developed HCC, whereas only 5% of -mice fed normal diet developed malignant tumours in lung, but not liver (Fig. 1b–e and Supplementary Fig. 2). Importantly, moreover, similar HCC development was also observed when genetically obese (*ob/ob*, also known as $Lep^{ob/ob}$) mice were treated with DMBA at the neonatal stage (Supplementary Fig. 3a–d), indicating that obesity, but not the HFD, promotes HCC development.

Because the induction of $p21^{Waf1/Cip1}$ expression was observed in liver, particularly in the area of liver cancer (Supplementary Fig. 1c), we speculated that senescent cells might be present in the vicinity of cancerous hepatocytes. Indeed, $p21^{Waf1/Cip1}$ expression was observed only in activated hepatic stellate cells (HSCs), which express α-smooth muscle actin (α-SMA) and desmin[7] (Fig. 1f). Furthermore, a number of other senescence markers[9,10], such as $p16^{INK4a}$ expression, signs of DNA damage (53BP1 foci and γH2AX foci) and inhibited cell proliferation (the absence of bromodeoxyuridine incorporation and Ki-67 expression), were also observed in activated HSCs despite absence of oncogenic *ras* mutation (Fig. 1f and Supplementary Figs 3e and 4). Interestingly, moreover, increased expression of IL-6, Gro-α, CXCL9 (major components of SASP)[4,5], but not HGF (a differentiation marker)[7], was observed in activated HSCs, but not in other types of liver cells (Fig. 1f and Supplementary Figs 3e and 5), indicating that these activated HSCs are senescing and may promote obesity-associated HCC development via SASP. It should be noted that, unlike the study using carbon tetrachloride (CCl_4)[11], fibrosis was not apparent in HFD-fed mice (Supplementary Fig. 6), precluding the possibility that the appearance of senescent HSCs was a by-product of liver fibrosis.

[1]Division of Cancer Biology, Cancer Institute, Japanese Foundation for Cancer Research, Koto-ku, Tokyo 135-8550, Japan. [2]CREST, Japan Science and Technology Agency, Kawaguchi, Saitama 332-0012, Japan. [3]Department of Applied Biological Science, Tokyo University of Science, Noda, Chiba 278-8510, Japan. [4]Research Center for Allergy and Immunology, RIKEN, Yokohama, Kanagawa 230-0045, Japan. [5]PRESTO, Japan Science and Technology Agency, Kawaguchi, Saitama 332-0012, Japan. [6]Division of Pathology, Cancer Institute, Japanese Foundation for Cancer Research, Koto-ku, Tokyo 135-8550, Japan. [7]Institute for Genome Research, University of Tokushima, Tokushima 770-8503, Japan. [8]Research Institute for Biological Science, Tokyo University of Science, Noda, Chiba 278-8510, Japan. [9]Graduate School of Frontier Sciences, University of Tokyo, Kashiwa, Chiba 277-8561, Japan. [10]School of Veterinary Medicine, Azabu University, Sagamihara, Kanagawa 229-8501, Japan.
*These authors contributed equally to this work.