

# SCIENTIFIC REPORTS



OPEN

## Nitrogen fixation and *nifH* diversity in human gut microbiota

Katsura Igai<sup>1,2,3,†</sup>, Manabu Itakura<sup>4</sup>, Suguru Nishijima<sup>5</sup>, Hirohito Tsurumaru<sup>4</sup>, Wataru Suda<sup>6</sup>, Takumi Tsutaya<sup>7</sup>, Eriko Tomitsuka<sup>2</sup>, Kiyoshi Tadokoro<sup>2</sup>, Jun Baba<sup>8</sup>, Shingo Odani<sup>9</sup>, Kazumi Natsuhara<sup>10</sup>, Ayako Morita<sup>2</sup>, Minoru Yoneda<sup>7</sup>, Andrew R. Greenhill<sup>11,12</sup>, Paul F. Horwood<sup>11</sup>, Jun-ichi Inoue<sup>13</sup>, Moriya Ohkuma<sup>13</sup>, Yuichi Hongoh<sup>14</sup>, Taro Yamamoto<sup>1,3</sup>, Peter M. Siba<sup>11</sup>, Masahira Hattori<sup>6</sup>, Kiwamu Minamisawa<sup>4</sup> & Masahiro Umezaki<sup>2</sup>

Received: 23 December 2015

Accepted: 27 July 2016

Published: 24 August 2016

It has been hypothesized that nitrogen fixation occurs in the human gut. However, whether the gut microbiota truly has this potential remains unclear. We investigated the nitrogen-fixing activity and diversity of the nitrogenase reductase (NifH) genes in the faecal microbiota of humans, focusing on Papua New Guinean and Japanese individuals with low to high habitual nitrogen intake. A <sup>15</sup>N<sub>2</sub> incorporation assay showed significant enrichment of <sup>15</sup>N in all faecal samples, irrespective of the host nitrogen intake, which was also supported by an acetylene reduction assay. The fixed nitrogen corresponded to 0.01% of the standard nitrogen requirement for humans, although our data implied that the contribution in the gut *in vivo* might be higher than this value. The *nifH* genes recovered in cloning and metagenomic analyses were classified in two clusters: one comprising sequences almost identical to *Klebsiella* sequences and the other related to sequences of Clostridiales members. These results are consistent with an analysis of databases of faecal metagenomes from other human populations. Collectively, the human gut microbiota has a potential for nitrogen fixation, which may be attributable to *Klebsiella* and Clostridiales strains, although no evidence was found that the nitrogen-fixing activity substantially contributes to the host nitrogen balance.

Nitrogen fixation, the reduction of atmospheric dinitrogen to ammonium, is a process that changes a biologically inactive form of nitrogen to a compound that can be readily used by organisms<sup>1</sup>. Diverse microbial communities use this process to fulfil their nitrogen demands. For example, the microbiota associated with invertebrates living in nitrogen-poor environments occasionally include diazotrophs (nitrogen-fixing prokaryotes). Nitrogen fixation in these animals, including termites<sup>2</sup>, wood-boring beetles<sup>3</sup>, shipworms<sup>4</sup>, sponges<sup>5</sup>, and corals<sup>6</sup>, has been tested by the measurement of <sup>15</sup>N<sub>2</sub> incorporation and of reduction of acetylene, an analogue of dinitrogen, and recently also by nanoscale secondary ion mass spectrometry (NanoSIMS). The phylogenetic diversity of the nitrogenase reductase (NifH) gene, a molecular marker of nitrogen fixation, has also been extensively examined in the symbiotic microbiota of various invertebrates<sup>2,3,5,7</sup>.

In contrast, there have been few studies of nitrogen fixation in vertebrates. Several reports have been published, but almost exclusively in the 1970s. No or only slight acetylene-reducing activity was detected in the

<sup>1</sup>Graduate School of Biomedical Sciences, Nagasaki University, Nagasaki, Japan. <sup>2</sup>Department of Human Ecology, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan. <sup>3</sup>Department of International Health, Institute of Tropical Research, Nagasaki University, Nagasaki, Japan. <sup>4</sup>Department of Environmental Life Sciences, Graduate School of Life Sciences, Tohoku University, Miyagi, Japan. <sup>5</sup>Department of Computational Biology and Medical Sciences, Graduate School of Frontier Sciences, The University of Tokyo, Chiba, Japan. <sup>6</sup>Center for Omics and Bioinformatics, Graduate School of Frontier Sciences, The University of Tokyo, Chiba, Japan. <sup>7</sup>Department of Integrated Biosciences, Graduate School of Frontier Sciences, The University of Tokyo, Chiba, Japan. <sup>8</sup>Department of Sociological Studies, Faculty of Human Sciences, Wako University, Tokyo, Japan. <sup>9</sup>Faculty of Letters, Chiba University, Chiba, Japan. <sup>10</sup>The Japanese Red Cross Akita College of Nursing, Akita, Japan. <sup>11</sup>Papua New Guinea Institute of Medical Research, Eastern Highlands Province, Papua New Guinea. <sup>12</sup>School of Applied and Biomedical Sciences, Federation University Australia, Victoria, Australia. <sup>13</sup>Japan Collection of Microorganisms, RIKEN BioResource Center, Ibaraki, Japan. <sup>14</sup>Department of Life Science and Technology, Tokyo Institute of Technology, Tokyo, Japan. <sup>†</sup>Present address: Department of Life Science and Technology, Tokyo Institute of Technology, 2-12-1, W3-48 Ookayama, Meguro, Tokyo, Japan 152-8550. Correspondence and requests for materials should be addressed to K.I. (email: kigai@bio.titech.ac.jp)

| Sampling area and year | ID   | Sex <sup>a</sup> | Age | BMI <sup>b</sup> | N intake <sup>c</sup> | Sample type <sup>d</sup> | Analyses <sup>e</sup> |             |
|------------------------|------|------------------|-----|------------------|-----------------------|--------------------------|-----------------------|-------------|
|                        |      |                  |     |                  |                       |                          | Activity              | <i>nifH</i> |
| PNG                    |      |                  |     |                  |                       |                          |                       |             |
| Levani (2012)          | L005 | M                | 23  | 22.9             | 85.2                  | R                        |                       | ✓           |
| Levani (2012)          | L006 | M                | 35  | 22.8             | 89.6                  | R                        |                       | ✓           |
| Levani (2012)          | L015 | M                | 20  | 21.2             | 69.6                  | R                        |                       | ✓           |
| Levani (2013)          | V001 | M                | 39  | 22.9             | 103.4                 | Fro, R                   | ✓                     |             |
| Levani (2013)          | V004 | M                | 22  | 22.7             | 109.6                 | Fro, R                   | ✓                     | ✓           |
| Levani (2013)          | V009 | M                | 35  | 24.3             | 117.2                 | Fro, R                   | ✓                     |             |
| Levani (2013)          | V014 | M                | 22  | 26.2             | 67.3                  | Fro, R                   | ✓                     | ✓           |
| Levani (2013)          | V022 | F                | 35  | 27.2             | 87.3                  | Fro, R                   | ✓                     |             |
| Levani (2013)          | V035 | M                | 40  | 22.9             | 148.0                 | Fro, R                   | ✓                     | ✓           |
| Japanese               |      |                  |     |                  |                       |                          |                       |             |
| Tokyo (2013)           | J001 | M                | 44  | 25.1             | 209.1                 | Fro, R                   | ✓                     | ✓           |
| Tokyo (2013)           | J002 | M                | 40  | 21.6             | 183.2                 | Fro, R                   | ✓                     | ✓           |
| Tokyo (2014)           | J004 | M                | 31  | 22.9             | 273.3                 | R                        |                       | ✓           |
| Tokyo (2014)           | J005 | M                | 24  | 22.2             | 210.3                 | R                        |                       | ✓           |
| Tokyo (2016)           | N001 | M                | 25  | 24.4             | 170.9                 | Unfrozen                 | ✓                     |             |
| Tokyo (2016)           | N002 | M                | 23  | 22.7             | 176.4                 | Unfrozen                 | ✓                     |             |
| Tokyo (2016)           | N003 | F                | 41  | 19.2             | 250.0                 | Unfrozen                 | ✓                     |             |
| Tokyo (2016)           | N004 | M                | 39  | 24.6             | 177.1                 | Unfrozen                 | ✓                     |             |
| Tokyo (2016)           | N005 | F                | 38  | 19.7             | 197.6                 | Unfrozen                 | ✓                     |             |

**Table 1. PNG and Japanese sample information.** <sup>a</sup>M, male; F, female. <sup>b</sup>BMI, body mass index, calculated as body weight (kg) per body height (m)<sup>2</sup>. <sup>c</sup>Values are given as total N (mg) per body weight (kg) per day. <sup>d</sup>R, preserved in RNAlater; Fro, frozen. <sup>e</sup>Activity, <sup>15</sup>N<sub>2</sub>-incorporation and/or acetylene reduction assays; *nifH*, cloning analysis.

rumen contents or faeces of herbivorous mammals, such as sheep<sup>8–10</sup>, goats<sup>8,10</sup>, cows<sup>11,12</sup>, rabbits<sup>8</sup>, and reindeer<sup>8</sup>. *In situ* examinations of nitrogen fixation were conducted in a sheep<sup>13</sup> and a guinea pig<sup>14</sup>. In the sheep, acetylene gas was injected into the rumen of a live animal, and acetylene-reduction rates of 0–5.1 nmol ethylene/ml rumen gas/h were detected. The live guinea pig was maintained in an atmosphere containing acetylene, and 75 nmol of ethylene was detected in the rearing chamber after 22 h.

The human distal gut is an anaerobic environment, like those of other mammals, where dinitrogen accounts for 60–80% of the gaseous content<sup>15</sup>. In a study of people with a low-nitrogen diet in Papua New Guinea (PNG), Bergersen and Hipsley (1970) hypothesized that nitrogen fixation occurs in the human distal gut<sup>14</sup>. The habitual nitrogen intake of the study subjects was estimated to be 49–74 mg/kg body weight/day in 1967<sup>16</sup>, which is much lower than the estimated average requirement for humans (105 mg/kg body weight/day)<sup>17</sup>. Nevertheless, these individuals displayed no disorders associated with nitrogen (protein) deficiency<sup>18</sup>. Bergersen and Hipsley collected faeces from an unspecified number of PNG individuals and five Europeans, and analysed them with an acetylene reduction assay. They detected 0.75 nmol/g/h acetylene-reducing activity in one of the European faecal samples and ‘slight but significant’ activities in all the PNG samples. However, the details were not described, and the samples were supplemented with glucose, which may have altered their acetylene-reducing activity. Because there have been no subsequent studies, whether the human gut microbiota truly has a capacity for nitrogen fixation, and if it does, whether it is specific to individuals with low nitrogen intake, remain unclear.

The diet of certain PNG individuals is still composed mostly of sweet potatoes today, resulting in a below-standard nitrogen intake<sup>19</sup>, even though their average nitrogen intake has been increasing since the 1970s. Therefore, we collected faecal samples from PNG individuals with low to sufficient nitrogen intake, and from Japanese individuals, with much higher nitrogen intake. To detect the nitrogen-fixing activity of the faecal microbiota and evaluate the effects of the habitual nitrogen intake, the <sup>15</sup>N<sub>2</sub> incorporation and acetylene reduction by the faecal samples were tested. We also examined the phylogenetic diversity of the *nifH* gene in the faecal samples by analysing clone libraries of *nifH* genes and transcripts and with metagenomic sequencing. In addition to *nifH*, the nitrogenase (NifDK) genes were identified in the PNG and Japanese metagenomes and in publicly available databases of human faecal metagenomes from other populations. This study demonstrates the nitrogen-fixing activity of the human faecal microbiota and presents a list of candidate nitrogen-fixing bacteria in the human gut.

## Results

**Nitrogen-fixing activity.** Table 1 lists the faecal samples used in this study, with information on the PNG and Japanese host individuals, including their habitual nitrogen intake. Most samples were collected from males, because it has been suggested that nitrogen requirements fluctuate during the menstrual cycle and pregnancy in females<sup>20,21</sup>. To measure nitrogen-fixing activity, we selected six PNG (V001–035) and two Japanese (J001 and 002) individuals, so that the host nitrogen intake ranged from below to above (67.3–209.1 mg/kg body weight/day)

| Sample ID       | $\delta^{15}\text{N}$ value <sup>a</sup> (‰) |                           | $\Delta\delta^{15}\text{N}^b$ | $p$ value <sup>c</sup> |
|-----------------|--|---------------------------|-------------------------------|------------------------|
|                 | $^{15}\text{N}_2$ gas (+)                    | $^{15}\text{N}_2$ gas (-) |                               |                        |
| PNG             |  |                           |                               |                        |
| V001            | 5.90 ± 0.32                                  | 4.83 ± 0.10               | <b>1.07</b>                   | 0.0025                 |
| V004            | 4.26 ± 0.22                                  | 3.09 ± 0.38               | <b>1.17</b>                   | 0.0049                 |
| V009            | 6.17 ± 0.26                                  | 5.33 ± 0.04               | <b>0.84</b>                   | 0.0139                 |
| V014            | 6.44 ± 0.60                                  | 5.32 ± 0.05               | <b>1.11</b>                   | 0.0429                 |
| V022            | 6.32 ± 0.60                                  | 5.64 ± 0.05               | <b>0.68</b>                   | 0.0283                 |
| V035            | 3.74   | 2.66                      | <b>1.08</b>                   | N. A.                  |
| Japanese        |  |                           |                               |                        |
| J001            | 7.48   | 6.22 ± 0.33               | <b>1.26</b>                   | N. A.                  |
| J002            | 5.58 ± 0.18                                  | 4.78 ± 0.21               | <b>0.80</b>                   | 0.0041                 |
| Sterile control |  |                           |                               |                        |
| V004            | 3.22   | 3.09 ± 0.38               | <b>0.13</b>                   | N. A.                  |
| V009            | 4.86   | 5.33 ± 0.04               | - <b>0.47</b>                 | N. A.                  |
| J002            | 4.66 ± 0.08                                  | 4.78 ± 0.21               | - <b>0.12</b>                 | 0.2066                 |

**Table 2. Incorporation of  $^{15}\text{N}_2$  gas by PNG and Japanese faecal samples.** <sup>a</sup>Averages and SD of triplicate samples are shown. For duplicate samples, only average values are shown. <sup>b</sup> $\Delta\delta^{15}\text{N} = \delta^{15}\text{N}[^{15}\text{N}_2 (+)] - \delta^{15}\text{N}[^{15}\text{N}_2 (-)]$ . <sup>c</sup>Student's *t* test for V001 and V004; Welch's *t* test for the other samples. N. A., not applicable.

| Sample ID           | Time between sampling and assay <sup>a</sup> (h) | Acetylene reduction rates (nmol/g/h) |               |
|---------------------|--|--------------------------------------|---------------|
|                     |  | Unfrozen                             | Frozen        |
| N001-1 <sup>b</sup> | 8  | 1.390 ± 0.237                        | 0.472 ± 0.109 |
| N001-2 <sup>b</sup> | 6  | 1.608 ± 0.422                        | 0.954 ± 0.010 |
| N002                | 8  | 0.660 ± 0.403                        | N. A.         |
| N003                | 15.5   | 0.466 ± 0.117                        | 0.049 ± 0.034 |
| N004                | 16.5   | 0.331 ± 0.127                        | 0.082 ± 0.055 |
| N005                | 27   | 0.055 ± 0.010                        | N. A.         |

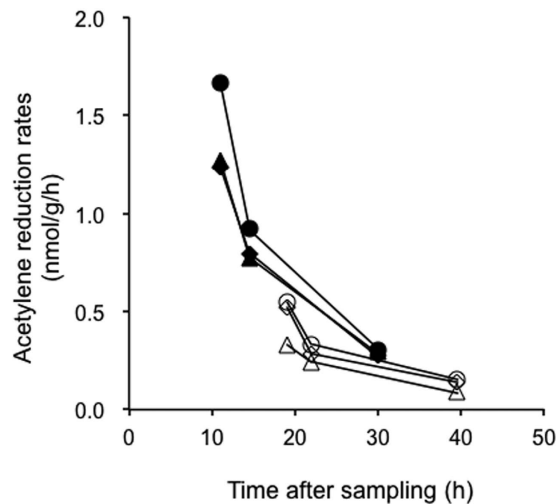
**Table 3. Acetylene-reduction rates in Japanese faecal samples.** <sup>a</sup>Time between defecation and the addition of acetylene to the unfrozen samples or freezing (the frozen samples). <sup>b</sup>N001-1 and N001-2 were sampled from the same individual (N001) on different days.

the standard requirement (105 mg/kg body weight/day)<sup>17</sup> (Table 1). Frozen faecal samples from these individuals were subjected to a  $^{15}\text{N}_2$  incorporation assay.

All the faecal samples from both the PNG and Japanese individuals incorporated  $^{15}\text{N}_2$  gas, whereas faecal samples sterilized by autoclaving did not (Table 2). The increases in the  $\delta^{15}\text{N}$  values ( $\Delta\delta^{15}\text{N}$ ) ranged from 0.68‰ to 1.26‰, and were statistically significant (Student's *t* test or Welch's *t* test:  $p < 0.05$ ,  $n = 3$ ). These values corresponded to nitrogen fixation rates of 0.065–0.147 nmol/g wet faeces/h (Supplementary Table S1). There was no correlation between  $\Delta\delta^{15}\text{N}$  and the nitrogen intake of the host (Pearson's product-moment correlation:  $r = 0.22$ ; Spearman's rank correlation:  $\rho = 0.21$ ). The acetylene reduction assay supported the occurrence of nitrogen fixation in the two Japanese samples with high nitrogen intake (209.1 and 183.2 mg/kg body weight/day). The emission of ethylene was observed in the presence of acetylene (Supplementary Figure S1), and the nitrogen fixation rates were estimated to be 0.008 and 0.143 nmol/g/h, when the theoretical reduction ratio  $\text{C}_2\text{H}_2:\text{N}_2 \approx 3^{22}$  was used.

To further confirm the occurrence of nitrogen fixation in the faecal samples of human individuals with high nitrogen intake, we additionally examined the acetylene-reducing activities in six faecal samples from five Japanese individuals (N001–005) (Table 1). In this assay, each faecal sample was divided into two: one was analysed directly ('unfrozen sample') and the other was frozen before analysis ('frozen sample'). By comparing these two types of sample, we also evaluated the effect of freezing on the nitrogen-fixing activity. All the unfrozen samples showed acetylene-reducing activities, ranging 0.055–1.608 nmol/g/h, which were 1.6–9.5 times higher than the values for the corresponding frozen samples (Table 3). The differences were statistically significant for samples N001-1 and N003 (Welch's *t* test:  $p < 0.01$ ,  $n = 3$ ). We also observed that the acetylene-reducing activity declined within 24 h of sampling (defecation) (Fig. 1). Furthermore, the acetylene-reduction rates correlated negatively with the time after defecation in the unfrozen samples (Pearson's product-moment correlation:  $r = -0.86$ ), although the host difference might have affected the data (Table 3).

**PCR and reverse transcription (RT)-PCR amplification of *nifH*.** DNA and RNA were extracted from each of the RNAlater<sup>®</sup>-preserved faecal samples from six PNG individuals and four Japanese individuals (Table 1). PCR and RT-PCR were performed with universal primers for *nifH*<sup>23</sup>. Genes and transcripts of the predicted size, around 360 base pairs (bp), were detected in all the samples. A total of 259 genes and 265 transcripts encoding



**Figure 1.** Decline in the acetylene-reduction rates in unfrozen Japanese faecal samples, N001-1 and N003. The experiments were performed in triplicate. Closed symbols, N001-1; open symbols, N003.

NifH homologues were detected and sorted to 44 operational taxonomic units (OTUs) (Supplementary Table S2). These were phylogenetically affiliated with *nifH* Cluster I (3 OTUs, 56 clones), Cluster III (2 OTUs, 24 clones), or Cluster IV (39 OTUs, 444 clones) (Fig. 2), according to the classification of *nifH* sequences by Zehr *et al.*<sup>24</sup>

Cluster I is composed of *nifH* sequences from diverse proteobacteria and cyanobacteria<sup>24</sup>. The OTU-24 sequence was recovered as gene clones from the PNG and Japanese samples (Fig. 2a). The nucleotide sequences were nearly identical (99%) to the *nifH* sequences of *Klebsiella* species, such as *Klebsiella pneumoniae*. The OTU-02 and OTU-41 sequences were almost identical to known contaminants frequently detected in PCR reagents<sup>25</sup>, and were not detected in the metagenomic datasets described below. Therefore, it is likely that these were not true constituents of the *nifH* community in the faecal samples.

Cluster III is known to contain *nifH* sequences of anaerobes from various environments. OTU-08 and OTU-34 showed high amino acid sequence similarities to NifH of anaerobic bacteria isolated from mammalian intestinal tracts (Fig. 2a). OTU-08 was recovered as both genes and transcripts from a PNG faecal sample. This OTU shared 98% amino acid sequence identity with NifH of *Lachnospira multipara* and also 97% identity with those of *Butyrivibrio* species and Ruminococcaceae bacterium AE2021. These bacteria have been isolated from ruminant foreguts, according to the description in the sequence database. OTU-34 was recovered as gene clones from two PNG samples. This OTU shared more than 97% amino acid sequence identity with NifH of Lachnospiraceae bacterium 3-1 isolated from a mouse cecum and Ruminococcaceae bacterium AE2021 and *Prevotella bryantii* isolated from ruminant foreguts.

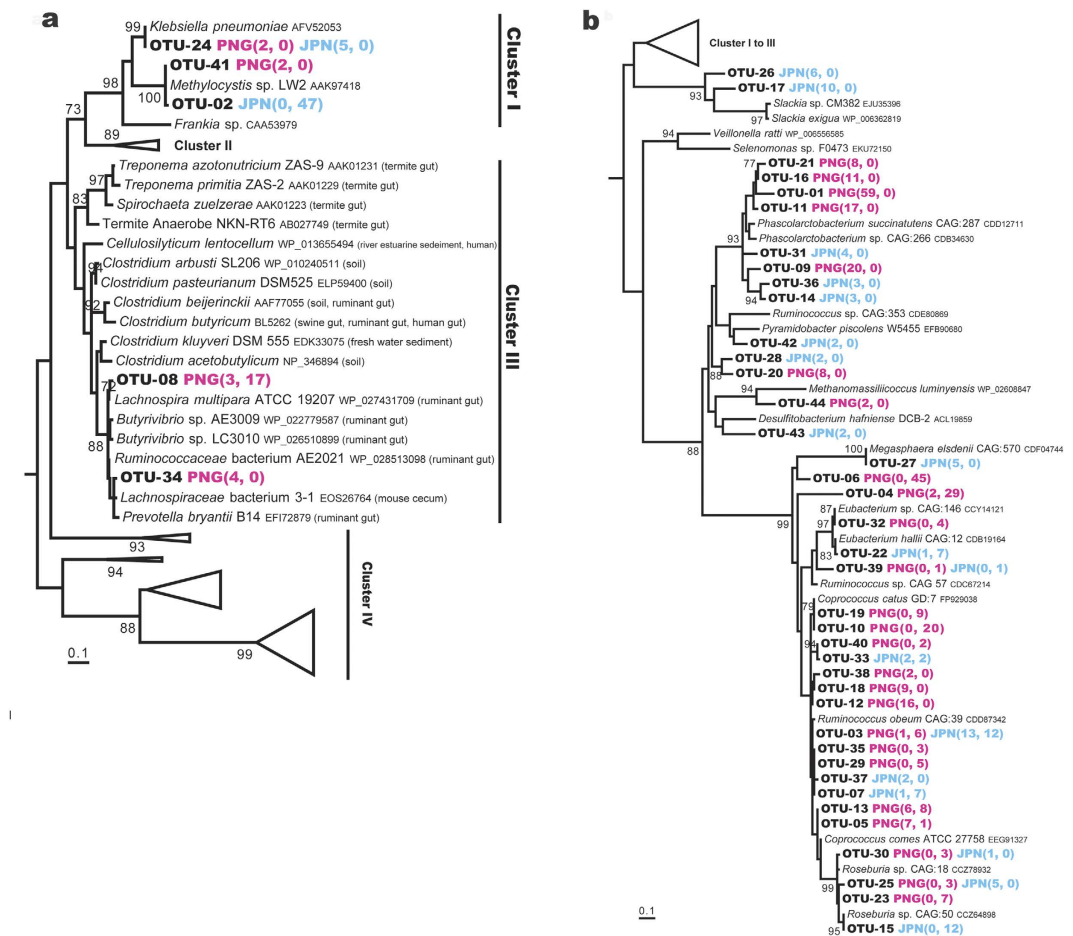
Diverse and abundant sequences were assigned to Cluster IV, which is also called the ‘NifH-like cluster’. The function of Cluster IV NifH is largely unknown, but it has been suggested to play a role other than in nitrogen fixation<sup>24,26</sup>. Most of the 39 OTUs showed high amino acid sequence similarities to NifH-like sequences of members of the order Clostridiales, belonging to genera such as *Ruminococcus*, *Coprococcus*, and *Eubacterium* (Fig. 2b).

**Detection of *nifH* Cluster I and III sequences with quantitative PCR (qPCR).** Because the abundant Cluster IV *nifH* sequences probably interfered with the detection of Cluster I and III sequences in the clone analysis discussed above, we also performed a qPCR analysis of nine PNG and four Japanese faecal samples (Table 4) to specifically detect sequences of OTU-24, OTU-08, and OTU-34. The detection limit was  $1.25 \times 10^4$  copies per one gram of faecal sample. Of the 13 samples tested, 10 showed the presence of at least one of these OTUs.

OTU-24 of Cluster I was detected in samples V004 and J002, as in the clone analysis, and also in another two PNG samples, L005 and V001. The copy number of the OTU-24 sequence ranged from  $10^5$  to  $10^6$  per gram wet faeces in the three PNG samples, whereas the Japanese sample J002 contained  $10^9$  copies per gram wet faeces (Table 4). OTU-08 of Cluster III was detected in five PNG samples, including L005, in which the OTU was found with the clone analysis. OTU-34 was detected more widely, in six PNG and three Japanese samples, including samples L005 and V014, in both of which the OTU was also found in the clone analysis. The copy numbers of the OTU-08 and OTU-34 sequences ranged from  $10^6$  to  $10^8$  per gram wet faeces in the PNG samples, while  $10^5$  to  $10^6$  in the Japanese samples (Table 4).

**Metagenomic analysis.** The faecal DNA samples from 16 PNG and two Japanese individuals, including the six PNG and two Japanese individuals tested in the  $^{15}\text{N}_2$  incorporation assay, were subjected to whole-genome shotgun sequencing (Supplementary Table S3). To facilitate homology searches, non-redundant gene sets were prepared from the metagenomic contigs from the PNG and Japanese samples. A published metagenomic dataset<sup>27</sup> was incorporated into the Japanese set.

In the PNG non-redundant gene set, four *nifH* homologues were identified and designated PNG-1–4 (Supplementary Table S4). These were affiliated with *nifH* Cluster III, sharing 94–98% amino acid sequence identities with NifH of *Lachnospira* and *Butyrivibrio* species. PNG-1 and PNG-3 showed 100% amino acid sequence



**Figure 2.** Phylogenetic positions of *nifH* recovered from Papua New Guinean and Japanese faecal samples, based on the deduced amino acid sequences. (a) Sequences affiliated with *nifH* Clusters I and III. (b) Sequences affiliated with *nifH* Cluster IV. Numbers in parentheses indicate the numbers of clones of genes and transcripts affiliated with the OTUs. A total of 113 amino acid positions were used, corresponding to positions 45–157 of the *Klebsiella pneumoniae* NifH sequence (AFV52053). 100 bootstrap resamplings were performed. Chlorophyllide reductase subunit BchX of *Rhodobacter sphaeroides* (CAB38747) was used as an outgroup. Only bootstrap confidence levels  $\geq 70\%$  are shown. OTU-02 and OTU-41 are most probably sequences contaminating in reagents.

similarity to OTU-08 and OTU-34, respectively (Fig. 3). The PNG-2 and PNG-4 sequence regions did not overlap most regions of the OTU sequences, and are therefore not shown in Fig. 3. In the Japanese gene set, two NifH homologues, designated JPN-1 and JPN-2, were identified and affiliated with *nifH* Clusters I and III, respectively (Fig. 3). JPN-1 shared 98% amino acid sequence identity with OTU-24, and JPN-2 shared 99% with OTU-34.

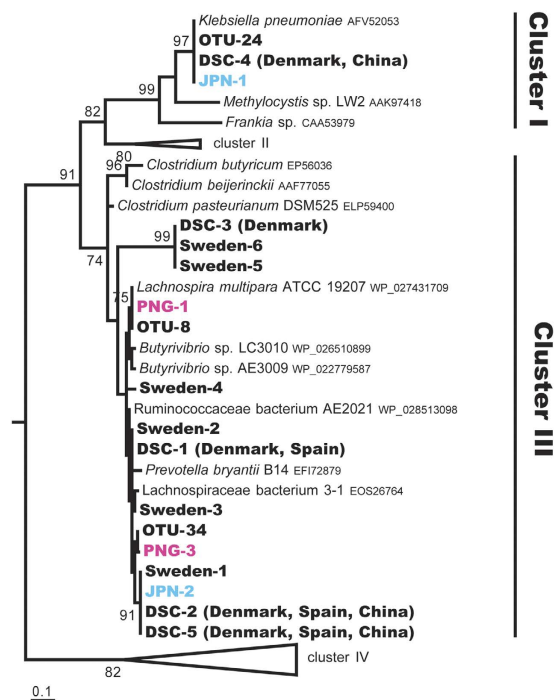
We also searched for *nifH* sequences in non-redundant gene sets constructed from publicly available databases of faecal metagenomes from other human populations, each consisting of 291 Danish<sup>28,29</sup>, 39 Spanish<sup>28</sup>, 145 Swedish<sup>30</sup>, or 363 Chinese<sup>31</sup> individuals. In total, 11 NifH homologues were identified. Ten were assigned to *nifH* Cluster III, sharing 83–100% amino acid sequence identity with the amino acid sequences of Clostridiales species (*Lachnospira* and *Butyrivibrio* species) and also of OTU-08 and OTU-34 (Fig. 3 and Supplementary Table S4). The remaining homologue shared an identical amino acid sequence with NifH of *Klebsiella* species and OTU-24 in *nifH* Cluster I. Cluster III sequences were detected in 31.7%, 21.6%, 17.9%, and 10.7% of the Swedish, Danish, Spanish, and Chinese individuals, respectively, when  $10^6$  sequencing reads were randomly chosen from each population. *Klebsiella*-related Cluster I sequences were detected in 0.7% and 6.6% of Danish and Chinese individuals, respectively.

The sequencing reads of the faecal metagenomes of the PNG and Japanese individuals who were tested with the  $^{15}\text{N}_2$  incorporation assay were further searched for *nifH*, using MG-RAST v3.2<sup>32</sup>. The results were consistent with the data presented above (Supplementary Table S5). We also searched for genes encoding the nitrogenase Mo-Fe protein subunits NifD and NifK in all the non-redundant gene sets as well as the sequencing reads of the PNG and Japanese, used in the search for NifH. NifD and NifK were encoded in all the datasets and shared high amino acid sequence identities with those of *Klebsiella* and members of the Clostridiales (Supplementary Tables S6 and S7).



| ID       | Target copy number (copies/wet faeces g) |                                     |                                     |
|----------|--|-------------------------------------|-------------------------------------|
|          | OTU-24                                   | OTU-08                              | OTU-34                              |
| PNG      |  |                                     |                                     |
| L005     | $8.4 \times 10^4$                        | $2.4 \times 10^6$                   | <u><math>3.1 \times 10^7</math></u> |
| L006     | UD                                       | UD                                  | $1.4 \times 10^6$                   |
| L015     | UD                                       | <u><math>6.1 \times 10^7</math></u> | $1.4 \times 10^8$                   |
| V001     | $1.1 \times 10^5$                        | UD                                  | UD                                  |
| V004     | <u><math>5.1 \times 10^6</math></u>      | DNQ                                 | UD                                  |
| V009     | UD                                       | UD                                  | UD                                  |
| V014     | UD                                       | $4.0 \times 10^6$                   | <u><math>3.9 \times 10^7</math></u> |
| V022     | UD                                       | DNQ                                 | $1.4 \times 10^6$                   |
| V035     | UD                                       | UD                                  | $9.6 \times 10^7$                   |
| Japanese |  |                                     |                                     |
| J001     | UD                                       | UD                                  | DNQ                                 |
| J002     | <u><math>1.6 \times 10^9</math></u>      | UD                                  | $1.4 \times 10^6$                   |
| J004     | UD                                       | UD                                  | UD                                  |
| J005     | UD                                       | UD                                  | $1.8 \times 10^5$                   |

**Table 4.** Detection of Cluster I and Cluster III *nifH* genes with qPCR. Underlining indicates OTUs detected in the cloning analysis of the corresponding samples. Each reaction was conducted in triplicate, and standards ( $10^1$ – $10^7$  plasmid copies) and negative controls were measured in duplicate. Quantification was performed when the target sequences were detected at  $\geq 5$  copies in at least two of the triplicate reactions. Abbreviations: UD, under detection limit; DNQ, detected but not quantified.



**Figure 3.** Phylogenetic positions of *nifH* from the human faecal metagenomes from six countries, based on deduced amino acid sequences. A total of 112 amino acid positions were used, corresponding to positions 45–157 of the *Klebsiella pneumoniae* NifH sequence (AFV52053). Of the four *nifH* genes from Papua New Guinean individuals, PNG-2 and -4 were not included because they corresponded to positions 138–283 and 124–264, respectively. ‘DSC’ indicates a combined non-redundant gene set of Danish, Spanish, and Chinese populations. See also the legend to Fig. 2.

## Discussion

We have demonstrated that the human gut microbiota truly has a capacity for nitrogen fixation, using  $^{15}\text{N}_2$  incorporation and acetylene reduction assays. The nitrogen-fixing activities were detected irrespective of the host habitual nitrogen intake; even the samples from Japanese individuals with nitrogen intake much higher than the

standard requirement exhibited the activity. In addition, *nifHDK* genes were detected in the faecal metagenomes of PNG, Japanese, Chinese, and European individuals. Collectively, we suggest that a capacity for nitrogen fixation is maintained in various human populations.

The acetylene-reduction rates in the human faecal samples were up to 1.6 nmol/g/h, which is comparable to those detected in herbivorous mammals, such as sheep (0–1.4 nmol/g/h in the rumen)<sup>8–10</sup>, a rabbit (0.38 nmol/g/h in the cecum)<sup>8</sup>, and a reindeer (0–0.77 nmol/g/h in faeces)<sup>8</sup>. The nitrogen fixed in the human faecal samples corresponded to 0.36 mg/kg gut content/day, and the contribution to the host nitrogen balance was estimated to be 0.01%, at most. Thus, the nutritional impact of nitrogen fixation should be low in humans. However, caution is needed. Our data suggest that the nitrogen fixation by the human faecal microbiota is sensitive to the time after sampling (Fig. 1 and Table 3), as seen in termites. The nitrogen-fixing activity in termites declines within several hours of sampling<sup>33–36</sup>. Therefore, the nitrogen-fixing activity in the human gut *in vivo* might be higher than the values estimated in this study. In addition, freezing considerably reduced the nitrogen-fixing activity; this activity of the PNG faecal samples was possibly underestimated, because the sampling required a long-distance transportation and freezing.

It is generally believed that nitrogen fixation is suppressed when available nitrogen sources, such as ammonia and nitrate, are sufficient in the environment. Therefore, it has been assumed that the concentration of nitrogen compounds in the human gut is too high to allow biological nitrogen fixation<sup>26</sup>. However, previous studies have reported plasticity in the regulation of nitrogen-fixing activity in natural environments. In aquatic environments, such as marine sediments and salt marshes, nitrogen fixation (acetylene reduction) has been detected under high concentrations of nitrogen compounds (e.g., millimolar levels of ammonia)<sup>37,38</sup>. A high demand for nitrogen and balance between nitrogen fixation and nitrification have been suggested to explain these examples. In another case, certain plant-associated diazotrophs are less sensitive to ammonia in the symbiotic phase than in the free-living phase, and contribute to the nutrition of the host<sup>1</sup>. Our data suggest that the human gut microbiota also contains members capable of fixing nitrogen, even when the nitrogen input is sufficient.

In both of the PNG and Japanese faecal samples, the *nifH* genes identified (except the Cluster IV *nifH*-like genes) were affiliated with two clades: those in one set were almost identical to the *nifH* genes of *Klebsiella* species, and these genes in the other set were closely related to those of Clostridiales members. This result is consistent between the cloning results and metagenomic analyses, and the *nifH* genes found in the European and Chinese faecal metagenomes also fell within these two phylogenetic groups. Thus, *Klebsiella* and Clostridiales species are the candidate nitrogen-fixers in the human gut, although the possibility that other bacterial lineages possess the *nif*-genes via horizontal gene transfer cannot be excluded. The total copy numbers of these *Klebsiella*- and Clostridiales-related *nifH* genes were 10<sup>5</sup>–10<sup>9</sup> copies per gram wet faeces (Table 4) and not correlated with the host nitrogen intake (Table 1) or with the nitrogen-fixing activity (Table 2). These *nifH* copy numbers are comparable to those in other diazotrophic environments, such as corals (10<sup>7</sup> copies per gram tissue)<sup>7</sup> and the rhizosphere (10<sup>5</sup>–10<sup>7</sup> copies per gram soil)<sup>39,40</sup>.

Many *Klebsiella* strains have been identified as active diazotrophs in various environments, including soil, water, and plants<sup>41</sup>. *Klebsiella* species are generally not predominant, but are widely distributed in the gut microbiota of humans<sup>28,42,43</sup>. Bergersen and Hipsley (1970) predicted that *Klebsiella* species are one of the nitrogen fixers present in the human gut microbiota<sup>14</sup>. They isolated three bacterial strains, provisionally identified as *Klebsiella aerogenes*, on a nitrogen-free medium, from human faecal samples and determined their <sup>15</sup>N<sub>2</sub>-incorporating activities.

The nitrogen-fixing activities of various Clostridiales strains containing Cluster III NifH were demonstrated with an acetylene reduction or <sup>15</sup>N<sub>2</sub> incorporation assay<sup>44</sup>. However, the bacterial strains, such as *Lachnospira* and *Butyrivibrio* species, of which NifH sequences constitute a clade with those from the human faecal microbiota, were not tested for nitrogen-fixing activity. A computational study suggested that certain bacteria expressing Cluster III NifH are likely to reduce acetylene and azide, but not dinitrogen, because they lack a histidine residue in the NifD protein (His422 in *Azotobacter vinelandii*)<sup>45</sup>. However, <sup>15</sup>N<sub>2</sub> fixation has recently been reported in *Endomicrobium proavitum*, which lacks His422 but retains all cysteine residues in NifD<sup>46</sup>. Among the relevant members of the Clostridiales and Bacteroidales (Figs 2a and 3), *L. multipara* ATCC 19207, Ruminococcaceae bacterium AE2021, and *P. bryantii* B14 lack His422 in NifD, but all the essential cysteine residues are conserved, as in *E. proavitum*. The other members of Clostridiales in this clade retain both the histidine and cysteine residues. In the NifD homologues recovered from the human faecal metagenomes, both of those with and without the histidine residue were identified. None of the NifD homologues lacked the cysteine residues. Therefore, there is currently no reason to consider that these bacteria cannot fix dinitrogen.

This study has shown that the human faecal microbiota has a potential capacity for nitrogen fixation, and that this capacity may exist in a wide range of human populations. However, evidence that this capacity substantially contributes to the host nitrogen balance was not obtained; the ecological meaning of nitrogen fixation in the human gut remains to be clarified.

## Materials and Methods

**Ethics statement.** The study, involving human participants, was approved by the Institutional Review Board at the Papua New Guinea Institute of Medical Research (1025), the Papua New Guinea Medical Research Advisory Committee (11.25), and the Research Ethics Committee in the Graduate School of Medicine, The University of Tokyo (3391). All the participants provided their written informed consent. All experiments were performed in accordance with the approved study protocols.

**Study population.** PNG participants were recruited in the Levani area (Levani Valley, Hela Province, Papua New Guinea) in March 2012 and March 2013, and in the Maprik area (East Maprik, East Sepik Province) in August 2012. Dietary data were obtained with a semi-quantitative food frequency questionnaire developed in

our previous study<sup>19</sup>, including demographic and anthropometric data. Nineteen healthy PNG individuals, aged between 15 and 40 years, with body mass indices (BMIs) ranging from 19.2 to 27.2 kg/m<sup>2</sup>, were selected for the study (Table 1 and Supplementary Table S3). Their nitrogen intake was calculated based on their protein intake using a nitrogen-to-protein conversion factor of 6.25 (nitrogen weight = protein weight/6.25)<sup>47</sup>.

The Japanese participants were recruited in Tokyo, Japan, between June 2013 and April 2016. Their dietary information was obtained with a brief self-administered diet history questionnaire<sup>48</sup>, and their age, height, and weight were self-reported. Nine individuals, with ages and BMIs similar to those of the PNG participants, were selected (Table 1).

**Sample collection in PNG.** In Levani in 2012, faecal samples were collected by each participant in a plastic container, and a portion of approximately 500 mg of the faeces was immediately suspended in 2 ml of RNAlater<sup>®</sup> (Ambion) (Table 1). In Maprik in 2012 and in Levani in 2013, faecal samples were collected and immediately enclosed in an AnaeroPack<sup>™</sup> (Mitsubishi Gas Chemical), placed in a cooler container, and brought to our research base in town (Table 1 and Supplementary Table S3), which took 7–8 h to drive and/or walk from the rural communities in Maprik and Levani. At the base, a portion of the faeces was suspended in RNAlater<sup>®</sup> as described above, and another portion of approximately 2.0 g was frozen in a liquid nitrogen dry vapor shipper. These samples were transported to our laboratory at The University of Tokyo, Japan, and stored at –80 °C until analysis.

**Sample collection in Japan.** Faecal samples were collected in 2012 and 2013 in Tokyo (Table 1), and suspended in RNAlater<sup>®</sup> immediately after defecation by each participant. Frozen samples from two individuals were also prepared. The samples were enclosed anaerobically in an AnaeroPack<sup>™</sup> by each participant, brought to the laboratory at the ambient temperature (around 15 °C), and stored at –80 °C until analysis. In 2016, additional faecal samples were collected from five individuals without freezing or treatment with RNAlater<sup>®</sup> (Table 1). The samples were enclosed in an AnaeroPack<sup>™</sup> by each participant and were stored at 4 °C until analysis.

**<sup>15</sup>N<sub>2</sub> incorporation assay.** The frozen faeces were divided (cracked) into pieces without thawing. The faecal pieces were placed in a vacuum desiccator (7 l), as duplicate or triplicate samples, at the ambient temperature. The gas phase was immediately replaced with argon (Ar) and subsequently adjusted to (60% [v/v] <sup>15</sup>N<sub>2</sub> [99.7% atoms] and 40% [v/v] Ar) (SI Science, Saitama, Japan). The samples were incubated at 37 °C for 48 h. Control experiments without <sup>15</sup>N<sub>2</sub> gas were also performed in duplicate or triplicate. Sterile controls were prepared by autoclaving the samples at 120 °C for 20 min and incubating them with <sup>15</sup>N<sub>2</sub> gas under the same conditions. The <sup>15</sup>N abundance ( $\delta^{15}\text{N}$ ) was measured with the DELTA V Advantage ConFlo IV system (Thermo Fisher Scientific) at SI Science. The total nitrogen mass concentration (%N) was measured with a Flash 2000 elemental analyser (Thermo Fisher Scientific) in the Isotope Ecology Laboratory at The University of Tokyo. The analytical standard deviation (SD) was approximately <0.2‰ for  $\delta^{15}\text{N}$  and <0.77% for %N. For samples analysed in triplicate, Student's *t* test or Welch's *t* test was performed based on the results of an *F*-test. The correlation between  $\Delta\delta^{15}\text{N}$  and the host nitrogen intake was determined with Pearson's product-moment correlation coefficient (*r*) and Spearman's rank correlation coefficient ( $\rho$ ). The calculation of the <sup>15</sup>N mass is shown in Supplementary Table S1.

**Acetylene reduction assay.** Frozen faeces from samples J001 and J002 were cracked into approximately 1.0–2.0 g pieces and divided into glass vials (net capacity 8.7 ml). The headspace was replaced with nitrogen gas followed by the injection of acetylene gas at a final concentration of 15% (v/v). The vials were placed at 37 °C. Faecal samples incubated without acetylene gas and empty vials with only acetylene gas were also prepared as controls. An aliquot (0.2 ml) of the headspace gas was analysed to measure the ethylene concentrations from day 0 to day 3 at 40 °C using a GC7A gas chromatograph (Shimadzu) equipped with a Propack N column (80/100 mesh, 2 m in length, 2.2 mm in diameter; Shimadzu) and a flame ionization detector.

Unfrozen faeces from samples N001 to N005 were divided into approximately 1.0–2.0 g portions in glass vials in an anaerobic chamber. Half the vials were frozen at –80 °C for 10–12 h for comparisons with the corresponding unfrozen samples. Aliquots (0.125 ml) of the headspace gas from both the unfrozen and frozen samples were analysed at 3, 6, and 21 h after the addition of acetylene gas for samples N001-1 and N003, and at 1 and 3 h after the addition of acetylene gas for the other samples.

**DNA extraction.** DNA was extracted from the faecal samples preserved in RNAlater<sup>®</sup> or from frozen samples, according to the previously described methods<sup>49,50</sup>, with modifications. The faecal samples in RNAlater<sup>®</sup> were diluted and washed twice with phosphate-buffered saline (PBS). Faecal aliquots of 4 mg were prepared and suspended in 300  $\mu$ l of Tris-SDS (250  $\mu$ l of 200 mM Tris-HCl, 80 mM EDTA, pH 9.0, and 50  $\mu$ l of 10% SDS). To this suspension, 300 mg of glass beads (0.1 mm diameter) and 500  $\mu$ l of Tris-EDTA (TE)-saturated phenol (Wako, Japan) were added, and the cells disrupted with a Shakemaster Auto (Biomedical Science, Tokyo, Japan) for 15 min. After centrifugation, the supernatant was subjected to phenol/chloroform/isoamyl alcohol extraction and isopropanol precipitation. The extracted DNA was suspended in 100  $\mu$ l of TE.

A frozen faecal sample was treated with 15 mg/ml lysozyme at 37 °C for 1 h, with 60 units/ml purified achromopeptidase (Wako 015-09951, Japan) at 37 °C for 30 min, and then with 1% SDS and 1 mg/ml proteinase K at 55 °C for 1 h. The sample was subjected to phenol/chloroform/isoamyl alcohol extraction and isopropanol precipitation. The extracted DNA was treated with RNase A, and purified with polyethylene glycol precipitation to remove any residual protein. The purified DNA was suspended in 300  $\mu$ l of TE.

**RNA extraction and cDNA synthesis.** RNA was extracted from faecal samples preserved in RNAlater<sup>®</sup>, using the RNeasy<sup>®</sup> Mini Kit (Qiagen). The samples were washed in PBS, and 20 mg aliquots of faeces were



combined with a mixture of 600  $\mu$ l of the RLT buffer contained in the kit, 7  $\mu$ l of  $\beta$ -mercaptoethanol, 100  $\mu$ l of TE buffer, and 300 mg of glass beads, and disrupted with the Shakemaster Auto for 5 min. The subsequent procedure was performed according to the manufacturer's guidelines. DNA was removed from the extracted RNA with the Turbo DNA-free Kit (Ambion) at 37 °C for 1 h. Reverse transcription was performed at 55 °C using the SuperScript III First Strand Synthesis System (Invitrogen) with 0.5  $\mu$ M nifH3 primer<sup>23</sup>. Negative controls without reverse transcription were prepared for each sample.

**PCR and RT-PCR amplification, cloning, and sequencing.** Genes and transcripts encoding NifH were amplified with PCR and RT-PCR, respectively, using primers specific for *nifH* under previously described conditions<sup>31</sup>, with modifications (Supplementary Methods). The PCR products were separated by agarose gel electrophoresis and those of approximately 360 bp in length were purified with the QIAquick Gel Extraction Kit (Qiagen), cloned into the pGEM<sup>®</sup>-T Easy Vector System (Promega), and used to transform *Escherichia coli* DH5 $\alpha$  cells. Randomly chosen clones were sequenced with the T7 primer using the BigDye Terminator Cycle Sequencing Kit V3.1 (Applied Biosystems) on an ABI 3730 Genetic Analyzer (Applied Biosystems).

**Identification and phylogenetic analysis of *nifH*.** The sequences obtained with PCR and RT-PCR were examined with BLASTX searches of the National Center for Biotechnology Information (NCBI) non-redundant database. The sequences that showed high similarities (E-value  $\leq 10^{-50}$ ) to known *nifH* sequences and encoded the amino acid residues conserved in NifH<sup>23</sup> were subjected to subsequent screening. The selected sequences were aligned with ClustalX<sup>52</sup>, and a distance matrix was calculated with the DNADIST program in the PHYLIP package<sup>53</sup>. The nucleotide sequences were sorted to OTUs with a 96% similarity cut-off, using the mothur program<sup>54</sup>. A phylogenetic analysis was performed based on the deduced amino acid sequences, using a maximum likelihood method with the Le and Gascuel substitution model in MEGA 6.0<sup>55</sup>.

**Quantitative PCR.** Specific PCR primers and TaqMan<sup>®</sup> probes were designed for the *nifH* phylotypes OTU-08, OTU-24, and OTU-34 (Supplementary Table S8). Standard curves were constructed in triplicate using serial dilutions of linearized plasmids containing each target sequence. Quantitative PCR amplification was performed with the FastStart Essential DNA Probes Master (Roche Diagnostics) on the LightCycler Nano System (Roche) under cycling conditions: 95 °C for 10 min, and 40 cycles of 95 °C for 10 s and 60 °C for 30 s. Specific amplification was verified by cloning and sequencing the PCR products.

**Metagenomic analysis.** Whole-genome shotgun sequencing was performed on an Ion PGM<sup>™</sup> System with the Ion 318<sup>™</sup> Chip Kit and Ion PGM<sup>™</sup> 400 Sequencing Kit (Life Technologies). A total of 51,633,321 and 7,099,635 reads (average 3,227,083 and 3,549,818) were obtained for the 16 PNG and two Japanese samples (Supplementary Table S3), respectively, and a non-redundant gene set was constructed for each population. The method is described in detail in the Supplementary Methods.

Faecal metagenomic sequences for Danish<sup>28,29</sup>, Spanish<sup>28</sup>, Swedish<sup>30</sup>, and Chinese<sup>31</sup> populations were retrieved from the NCBI archives. A combined non-redundant gene set for the Danish, Spanish, and Chinese populations was obtained at the GigaDB database (<http://gigadb.org>), and the non-redundant gene set for the Swedish population was provided by Fredrik Karlsson (Chalmers University of Technology, Sweden).

The amino acid sequences deduced from the non-redundant gene sets were searched for NifH homologues using BLASTP with a cut-off level of E-value  $\leq 10^{-5}$ , against *nifH* (K02588) genes from the Kyoto Encyclopedia of Genes and Genomes (KEGG) database. A supercomputer at the Human Genome Center, The University of Tokyo (<http://sc.hgc.jp/shirokane.html>) was used for the analysis. Sequences of more than 120 amino acids were incorporated into the *nifH* database in the ARB format created by Zehr's laboratory ([www.jzehrlab.com/#!/nifh-database/c1coj](http://www.jzehrlab.com/#!/nifh-database/c1coj)) to classify the sequences into *nifH* Clusters<sup>24</sup>. The BLASTX search results for all the amino acid sequences containing  $\leq 120$  amino acids were also manually checked to eliminate Cluster IV NifH. The frequencies of the retrieved NifH homologues in each metagenomic dataset were calculated by mapping 10<sup>6</sup> randomly chosen reads to the corresponding non-redundant gene sets, using the Bowtie2 tool<sup>56</sup>, with a criterion of 95% nucleotide sequence identity. The NifD and NifK sequences were identified with BLASTP using a cut-off of E-value  $\leq 10^{-5}$  against NifD (K02586) and NifK (K02591) in the KEGG database, respectively. NifD and NifK sequences with more than 80% amino acid identity to those of bacteria containing Cluster I or Cluster III NifH were selected.

Additional searches for *nifH*, *nifD*, and *nifK* were made in the metagenomic sequence reads from the samples used for the <sup>15</sup>N incorporation assay, using MG-RAST<sup>31</sup>. The method is described in detail in the Supplementary Methods.

## References

- Dixon, R. & Kahn, D. Genetic regulation of biological nitrogen fixation. *Nat Rev Microbiol* **2**, 621–631 (2004).
- Brune, A. & Ohkuma, M. Role of the termite gut microbiota in symbiotic digestion in *Biology of Termites: a Modern Synthesis* (eds Bignell, D. E., Roisin, Y., & Lo, N.) 439–475 (Springer, Dordrecht, 2011).
- Ceja-Navarro, J. A. *et al.* Compartmentalized microbial composition, oxygen gradients and nitrogen fixation in the gut of *Odontotaenium disjunctus*. *ISME J* **8**, 6–18 (2014).
- Lechene, C. P., Luyten, Y., McMahon, G. & Distel, D. L. Quantitative imaging of nitrogen fixation by individual bacteria within animal cells. *Science* **317**, 1563–1566 (2007).
- Mohamed, N. M., Colman, A. S., Tal, Y. & Hill, R. T. Diversity and expression of nitrogen fixation genes in bacterial symbionts of marine sponges. *Environ Microbiol* **10**, 2910–2921 (2008).
- Lesser, M. P. *et al.* Nitrogen fixation by symbiotic cyanobacteria provides a source of nitrogen for the scleractinian coral *Montastraea cavernosa*. *Mar Ecol Prog Ser* **346**, 143–152 (2007).
- Olson, N. D., Ainsworth, T. D., Gates, R. D. & Takabayashi, M. Diazotrophic bacteria associated with Hawaiian *Montipora* corals: Diversity and abundance in correlation with symbiotic dinoflagellates. *J Exp Mar Biol Ecol* **371**, 140–146 (2009).

8. Granhall, U. & Ciszuk, P. Nitrogen fixation in rumen contents indicated by the acetylene reduction test. *J Gen Microbiol* **65**, 91–93 (1971).
9. Jones, K. & Thomas, J. G. Nitrogen fixation by the rumen contents of sheep. *J Gen Microbiol* **85**, 97–101 (1974).
10. Li Pun, H. H. & Satter, L. D. Nitrogen fixation in ruminants. *J Anim Sci* **41**, 1161–1163 (1975).
11. Hardy, R. W., Holsten, R. D., Jackson, E. K. & Burns, R. C. The acetylene-ethylene assay for N<sub>2</sub> fixation: laboratory and field evaluation. *Plant Physiol* **43**, 1185–1207 (1968).
12. Elleway, R. F., Sabine, J. R. & Nicholas, D. J. Acetylene reduction by rumen microflora. *Arch Mikrobiol* **76**, 277–291 (1971).
13. Hobson, P. N., Summers, R., Postgate, J. R. & Ware, D. A. Nitrogen fixation in the rumen of a living sheep. *J Gen Microbiol* **77**, 225–226 (1973).
14. Bergersen, F. J. & Hipsley, E. H. The presence of N<sub>2</sub>-fixing bacteria in the intestines of man and animals. *J Gen Microbiol* **60**, 61–65 (1970).
15. Levitt, M. D., Bond, John H., Levitt & David G. Gastrointestinal gas in *Physiology of the Gastrointestinal Tract* (ed. Johnson, L. R.) 1301–1316 (Raven Press, 1981).
16. Oomen, H. A. Interrelationship of the human intestinal flora and protein utilization. *Proc Nutr Soc* **29**, 197–206 (1970).
17. Rand, W. M., Pellett, P. L. & Young, V. R. Meta-analysis of nitrogen balance studies for estimating protein requirements in healthy adults. *Am J Clin Nutr* **77**, 109–127 (2003).
18. Luyken, R., Luyken-Koning, F. W. & Pikaar, N. A. Nutrition studies in New Guinea. *Am J Clin Nutr* **14**, 13–27 (1964).
19. Morita, A. *et al.* Development, validation, and use of a semi-quantitative food frequency questionnaire for assessing protein intake in Papua New Guinean Highlanders. *Am J Hum Biol* **27**, 349–357 (2015).
20. Stephens, T. V., Payne, M., Ball, R. O., Pencharz, P. B. & Elango, R. Protein requirements of healthy pregnant women during early and late gestation are higher than current recommendations. *J Nutr* **145**, 73–78 (2015).
21. Egun, G. N. & Atinmo, T. Protein requirement of young adult Nigerian females on habitual Nigerian diet at the usual level of energy intake. *Br J Nutr* **70**, 439–448 (1993).
22. Montoya, J. P., Voss, M., Kahler, P. & Capone, D. G. A simple, high-precision, high-sensitivity tracer assay for N<sub>2</sub> fixation. *Appl Environ Microbiol* **62**, 986–993 (1996).
23. Zehr, J. P. & McReynolds, L. A. Use of degenerate oligonucleotides for amplification of the *nifH* gene from the marine cyanobacterium *Trichodesmium thiebautii*. *Appl Environ Microbiol* **55**, 2522–2526 (1989).
24. Zehr, J. P., Jenkins, B. D., Short, S. M. & Steward, G. F. Nitrogenase gene diversity and microbial community structure: a cross-system comparison. *Environ Microbiol* **5**, 539–554 (2003).
25. Goto, M., Ando, S., Hachisuka, Y. & Yoneyama, T. Contamination of diverse *nifH* and *nifH*-like DNA into commercial PCR primers. *FEMS Microbiol Lett* **246**, 33–38 (2005).
26. Raymond, J., Siefert, J. L., Staples, C. R. & Blankenship, R. E. The natural history of nitrogen fixation. *Mol Biol Evol* **21**, 541–554 (2003).
27. Nishijima, S. *et al.* The gut microbiome of healthy Japanese and its microbial and functional uniqueness. *DNA Res* **2**, 125–133 (2016).
28. Qin, J. *et al.* A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* **464**, 59–65 (2010).
29. Le Chatelier, E. *et al.* Richness of human gut microbiome correlates with metabolic markers. *Nature* **500**, 541–546 (2013).
30. Karlsson, F. H. *et al.* Gut metagenome in European women with normal, impaired and diabetic glucose control. *Nature* **498**, 99–103 (2013).
31. Qin, J. *et al.* A metagenome-wide association study of gut microbiota in type 2 diabetes. *Nature* **490**, 55–60 (2012).
32. Meyer, F. *et al.* The metagenomics RAST server - a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC Bioinformatics* **9**, 386 (2008).
33. Prestwich, G. D., Bentley, B. L. & Carpenter, E. J. Nitrogen-sources for neotropical Nasute termites - fixation and selective foraging. *Oecologia* **46**, 397–401 (1980).
34. Lovelock, M., O'Brien, R. W. & Slaytor, M. Effect of laboratory containment on the nitrogen-metabolism of termites. *Insect Biochem* **15**, 503–509 (1985).
35. Yamada, A. *et al.* Nitrogen fixation by termites in tropical forests, Thailand. *Ecosystems* **9**, 75–83 (2006).
36. Pandey, S., Waller, D. A. & Gordon, A. S. Variation in acetylene-reductoin (nitrogen-fixation) rates in *Reticulitermes* spp. (Isoptera: Rhinotermitidae). *Virginia J Sci* **43**, 333–338 (1992).
37. Dekas, A. E., Chadwick, G. L., Bowles, M. W., Joye, S. B. & Orphan, V. J. Spatial distribution of nitrogen fixation in methane seep sediment and the role of the ANME archaea. *Environ Microbiol* **16**, 2012–2029 (2013).
38. Dicker, H. J. & Smith, D. W. Physiological ecology of acetylene reduction (nitrogen fixation) in a Delaware salt marsh. *Microb Ecol* **6**, 161–171 (1980).
39. Hai, B. *et al.* Quantification of key genes steering the microbial nitrogen cycle in the rhizosphere of sorghum cultivars in tropical agroecosystems. *Appl Environ Microbiol* **75**, 4993–5000 (2009).
40. Yousuf, B., Kumar, R., Mishra, A. & Jha, B. Differential distribution and abundance of diazotrophic bacterial communities across different soil niches using a gene-targeted clone library approach. *FEMS Microbiol Lett* **360**, 117–125 (2014).
41. Brisse, S. G. F. & Grimont, P. A. D. The Genus *Klebsiella* in *Proteobacteria: Gamma Subclass* (eds Dworkin, M., Falkow, S., Rosenberg, E., Schleifer, K. & Stackebrandt, E.) 159–196 (Springer, 2006).
42. Martinez, I. *et al.* The gut microbiota of rural Papua New Guineans: composition, diversity patterns, and ecological processes. *Cell Rep* **11**, 527–538 (2015).
43. Zhang, J. *et al.* A phylo-functional core of gut microbiota in healthy young Chinese cohorts across lifestyles, geography and ethnicities. *ISME J* **9**, 1979–1990 (2015).
44. Rosenblum, E. D. & Wilson, P. W. Fixation of isotopic nitrogen by *Clostridium*. *J Bacteriol* **57**, 413 (1949).
45. Dos Santos, P. C., Fang, Z., Mason, S. W., Setubal, J. C. & Dixon, R. Distribution of nitrogen fixation and nitrogenase-like sequences amongst microbial genomes. *BMC Genomics* **13**, 162–174 (2012).
46. Zheng, H., Dietrich, C., Radek, R. & Brune, A. *Endomicrobium proavitum*, the first isolate of *Endomicrobia* class. nov. (phylum *Elusimicrobia*) - an ultramicrobacterium with an unusual cell cycle that fixes nitrogen with a Group IV nitrogenase. *Environ Microbiol* **18** (2015).
47. Jones, D. B. Factors for converting percentages of nitrogen in foods and feeds into percentages of protein. Report No. 183, (US Dept. Agric. Cir., Washington, D.C., 1931).
48. Kobayashi, S. *et al.* Comparison of relative validity of food group intakes estimated by comprehensive and brief-type self-administered diet history questionnaires against 16 d dietary records in Japanese adults. *Public Health Nutr* **14**, 1200–1211 (2011).
49. Matsuki, T. *et al.* Quantitative PCR with 16S rRNA-gene-targeted species-specific primers for analysis of human intestinal bifidobacteria. *Appl Environ Microbiol* **70**, 167–173 (2004).
50. Ueno, M. *et al.* Assessment and Improvement of Methods for Microbial DNA Preparation from Fecal Samples in *Handbook of Molecular Microbial Ecology II* (ed. Bruijn, F. J. D.) 191–198 (John Wiley & Sons, Inc., 2011).
51. Moisaner, P. H., Beinart, R. A., Voss, M. & Zehr, J. P. Diversity and abundance of diazotrophic microorganisms in the South China Sea during intermonsoon. *ISME J* **2**, 954–967 (2008).
52. Larkin, M. A. *et al.* Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947–2948 (2007).
53. Felsenstein, J. PHYLIP - Phylogeny Inference Package (Version 3.2). *Cladistics* **5**, 164–166 (1989).

54. Schloss, P. D. *et al.* Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl Environ Microbiol* **75**, 7537–7541 (2009).
55. Tamura, K., Stecher, G., Peterson, D., Filipiński, A. & Kumar, S. MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. *Mol Biol Evol* **30**, 2725–2729 (2013).
56. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**, 357–359 (2012).

### Acknowledgements

We thank Hirokazu Tsuji and Koji Nomoto of Yakult Honsha Co., Ltd. for their helpful advice on sampling and preparing nucleic acids. We also thank Erika Iioka, Misa Kiuchi, Rina Kurokawa, Keiko Komiyama, Naoko Yamashita, Chie Shindo, and Yasue Hattori for technical assistance with DNA sequencing, and Takayuki Wada for advising on data analysis. This study was supported by the Next Generation World-Leading Researchers from Japan Society for the Promotion of Science LS024 to M.U. and partially a Grant-in-Aid for Scientific Research 26292047 to M.O. and 23117003 to Y.H., and also supported by an incubation study supported by the Research Institute for Humanity and Nature, Japan, to T.Y. This work was partly conducted at the Joint Usage/Research Center on Tropical Disease, Institute of Tropical Medicine, Nagasaki University.

### Author Contributions

M.U. conceived the study. K.I. and K.M. designed the experiments. M.Y., M.O., Y.H., T.Y., M.H., K.M. and M.U. provided the reagents and equipment. K.T., J.B., S.O., K.N., A.M., A.R.G., P.F.H., P.M.S. and M.U. conducted field studies. K.I., M.I., H.T., T.T., E.T., J.I. and K.M. performed experiments and data analyses. K.I., S.N., W.S. and M.H. performed metagenomic analyses. K.I. and Y.H. wrote the manuscript with significant contributions from A.R.G., P.F.H., M.O., T.Y., K.M. and M.U. and input from all the other authors.

### Additional Information

**Accession numbers:** The *nifH* sequences of the 41 OTUs were deposited at GenBank/EMBL/DDBJ under accession numbers, LC097013–56.

**Supplementary information** accompanies this paper at <http://www.nature.com/srep>

**Competing financial interests:** The authors declare no competing financial interests.

**How to cite this article:** Igai, K. *et al.* Nitrogen fixation and *nifH* diversity in human gut microbiota. *Sci. Rep.* **6**, 31942; doi: 10.1038/srep31942 (2016).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2016

## Supplementary Information

### Nitrogen fixation and *nifH* diversity in human gut microbiota

Katsura Igai, Manabu Itakura, Suguru Nishijima, Hirohito Tsurumaru, Wataru Suda, Takumi Tsutaya, Eriko Tomitsuka, Kiyoshi Tadokoro, Jun Baba, Shingo Odani, Kazumi Natsuhara, Ayako Morita, Minoru Yoneda, Andrew R. Greenhill, Paul F. Horwood, Jun-ichi Inoue, Moriya Ohkuma, Yuichi Hongoh, Taro Yamamoto, Peter M. Siba, Masahira Hattori, Kiwamu Minamisawa, Masahiro Umezaki

Supplementary Methods

Supplementary Table S1

Supplementary Table S2

Supplementary Table S3

Supplementary Table S4

Supplementary Table S5

Supplementary Table S6

Supplementary Table S7

Supplementary Table S8

Supplementary Figure S1

## Supplementary Methods

### PCR and RT-PCR conditions

The first-round amplification of nested PCR was conducted in 25  $\mu$ l volume with 1  $\times$  *Ex Taq*<sup>®</sup> buffer, 0.2 mM dNTPs, 1.0  $\mu$ M primers (*nifH3* and *nifH4*)<sup>1</sup>, 0.06 U *Ex Taq*<sup>®</sup> (Takara, Shiga, Japan), and 1  $\mu$ l of DNA extract was added as a template for PCR and 2  $\mu$ l of cDNA for RT-PCR. Amplification was performed under the conditions: 2 min of initial denaturation at 95°C, followed by 30 cycles of denaturation (30 s at 95°C), annealing (30 s at 57°C), and extension (1 min at 72°C), and final extension at 72°C for 7 min. The second-round amplification was conducted with 1  $\mu$ l of the first-round PCR products as template and primers (*nifH1* and *nifH2*)<sup>1</sup> under the same conditions.

### Construction of non-redundant gene sets

The metagenomic reads were quality-filtered using a pipeline<sup>2</sup> and assembled into contigs using Newbler v2.7. The contigs  $\geq$  500 bp were subjected to gene prediction using MetaGeneAnnotator<sup>3</sup>. Non-redundant gene sets were constructed for the PNG and Japanese samples, respectively, with CD-HIT<sup>4</sup> using a 95% sequence similarity and a 90% length coverage as thresholds. The Japanese non-redundant gene set was combined with the set prepared from a published Japanese fecal metagenome<sup>5</sup>.

### Search for *nifH* in metagenomic sequencing reads

The metagenome sequencing reads from the samples used for the <sup>15</sup>N<sub>2</sub> incorporation assay were further searched for *nifH*, *nifD*, and *nifK*, using MG-RAST (v3.2)<sup>6</sup> with the integrated microbial genome (IMG) database (<https://img.jgi.doe.gov/cgi-bin/er/main.cgi>). The following criteria were used: amino acid sequence identity greater than or equal to 80 %, E-value less than or equal to 10<sup>-5</sup>, and 15 or more amino acids in length. The candidate reads were further screened by BLASTx searches: reads sharing 80% or greater sequence identity and positive score 90% or greater with known NifH sequences were selected.



## Supplementary References

- 1 Zehr J P & McReynolds L A. Use of degenerate oligonucleotides for amplification of the *nifH* gene from the marine cyanobacterium *Trichodesmium thiebautii*. *Appl Environ Microbiol* **55**, 2522-2526 (1989).
- 2 Gomez-Alvarez V, Teal T K & Schmidt T M. Systematic artifacts in metagenomes from complex microbial communities. *ISME J* **3**, 1314-1317 (2009).
- 3 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).
- 4 Li W & Godzik A. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* **22**, 1658-1659 (2006).
- 5 Nishijima S *et al.* The gut microbiome of healthy Japanese and its microbial and functional uniqueness. *DNA Res* (2016).
- 6 Meyer F *et al.* The metagenomics RAST server - a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC Bioinformatics* **9**, 386 (2008).

Table S1. Calculation of  $^{15}\text{N}$  absolute mass and weight in faecal samples.

| ID   | Sample N ratio (%) | $^{15}\text{N}_2$ exposure (-)  |                                |                                 | $^{15}\text{N}_2$ exposure (+)  |                                |                                 | Nitrogen fixation rates<br>$^{15}\text{N}$ nmoles/<br>sample g/ h |
|------|--------------------|---------------------------------|--------------------------------|---------------------------------|---------------------------------|--------------------------------|---------------------------------|---|
|      |                    | $\delta^{15}\text{N}(\text{‰})$ | $^{15}\text{N}(\text{atom}\%)$ | $^{15}\text{N}$ mg/<br>sample g | $\delta^{15}\text{N}(\text{‰})$ | $^{15}\text{N}(\text{atom}\%)$ | $^{15}\text{N}$ mg/<br>sample g |   |
| V001 | 5.158              | 4.8277                          | 0.3680                         | 0.1898                          | 5.7415                          | 0.3683                         | 0.1900                          | 0.128   |
| V004 | 1.846              | 3.0860                          | 0.3674                         | 0.0678                          | 4.3825                          | 0.3679                         | 0.0679                          | 0.065   |
| V009 | 6.330              | 5.3287                          | 0.3682                         | 0.2331                          | 6.1710                          | 0.3685                         | 0.2333                          | 0.145   |
| V014 | 4.900              | 5.3273                          | 0.3682                         | 0.1804                          | 6.4357                          | 0.3686                         | 0.1806                          | 0.147   |
| V022 | 7.620              | 5.6390                          | 0.3683                         | 0.2807                          | 6.3203                          | 0.3686                         | 0.2808                          | 0.124   |
| V035 | 5.650              | 2.6545                          | 0.3672                         | 0.2075                          | 3.7380                          | 0.3676                         | 0.2077                          | 0.146   |
| J001 | 4.068              | 6.2177                          | 0.3685                         | 0.1499                          | 7.4790                          | 0.3690                         | 0.1501                          | 0.139   |
| J002 | 5.340              | 4.7787                          | 0.3680                         | 0.1965                          | 5.5830                          | 0.3683                         | 0.1967                          | 0.117   |

To evaluate  $^{15}\text{N}$  mass,  $^{15}\text{N}$  absolute ratios (atom %) were calculated using the equation:  $A$  (atom %) =  $100 \times [(\delta^{15}\text{N} \times 10^{-3} + 1)(^{15}\text{N}/^{14}\text{N})_{\text{atmosphere}} / 1 + (\delta^{15}\text{N} \times 10^{-3} + 1)(^{15}\text{N}/^{14}\text{N})_{\text{atmosphere}}]$ , where  $(^{15}\text{N}/^{14}\text{N})_{\text{atmosphere}} = 0.03676$  and  $\delta^{15}\text{N}$  = average value in replicates. Then,  $^{15}\text{N}$  mass ( $^{15}\text{N}$  mg/sample g) was obtained by multiplying ‘Sample N ratio’ by ‘ $^{15}\text{N}$  absolute ratio’. ‘Nitrogen fixation rates’ (nmoles/sample g/h) were estimated based on subtraction of ‘ $^{15}\text{N}$  mg/sample g’ values between the conditions ‘ $^{15}\text{N}_2$  exposure (+)’ and ‘ $^{15}\text{N}_2$  exposure (-)’.

Table S2. *nifH* DNA and cDNA sequences identified from PNG and Japanese faecal samples.

|             | <i>nifH</i> DNA sequences |      |      |      |      |          |      |      |      |      | <i>nifH</i> cDNA sequences |      |      |      |      |          |      |      |      |      |
|-------------|---------------------------|------|------|------|------|----------|------|------|------|------|----------------------------|------|------|------|------|----------|------|------|------|------|
|             | PNG                       |      |      |      |      | Japanese |      |      |      |      | PNG                        |      |      |      |      | Japanese |      |      |      |      |
|             | V004                      | V014 | V035 | L005 | L006 | L015     | J001 | J002 | J004 | J005 | V004                       | V014 | V035 | L005 | L006 | L015     | J001 | J002 | J004 | J005 |
| OTU-01      |                           | 29   | 9    | 4    |      |          |      |      |      |      |                            |      |      |      |      |          |      |      |      |      |
| OTU-02      |                           |      |      |      |      |          |      |      |      |      |                            |      |      |      |      |          | 14   | 14   | 19   |      |
| OTU-03      |                           |      | 1    |      |      |          |      | 13   |      |      |                            | 6    |      |      |      | 6        |      | 6    |      |      |
| OTU-04      |                           |      | 3    |      |      |          |      |      |      |      |                            |      | 29   |      |      |          |      |      |      |      |
| OTU-05      | 7                         |      |      |      |      |          |      |      |      |      |                            | 1    |      |      |      |          |      |      |      |      |
| OTU-06      |                           |      |      |      |      |          |      |      |      |      | 22                         | 23   |      |      |      |          |      |      |      |      |
| OTU-07      |                           |      |      |      |      |          |      |      |      |      |                            |      |      |      |      | 22       |      |      |      |      |
| OTU-08      |                           |      |      |      |      | 3        |      |      |      |      |                            |      |      |      | 17   |          |      |      |      |      |
| OTU-09      |                           |      |      |      | 5    | 15       |      |      |      |      |                            |      |      |      |      |          |      |      |      |      |
| OTU-10      |                           |      |      |      |      |          |      |      |      |      |                            |      |      |      | 20   |          |      |      |      |      |
| OTU-11      |                           |      |      |      |      |          |      | 17   |      |      |                            |      |      |      |      |          |      |      |      |      |
| OTU-12      | 8                         |      |      |      | 1    | 7        |      |      |      |      |                            |      |      |      |      |          |      |      |      |      |
| OTU-13      | 6                         |      |      |      |      |          |      |      |      |      |                            | 8    |      |      |      |          |      |      |      |      |
| OTU-14      |                           | 17   |      |      |      |          |      |      | 13   |      |                            |      |      |      |      |          |      |      |      |      |
| OTU-15      |                           |      |      |      |      |          |      |      |      |      |                            |      |      |      |      |          | 12   |      |      |      |
| OTU-16      |                           | 8    | 1    |      | 1    | 1        |      |      |      |      |                            |      |      |      |      |          |      |      |      |      |
| OTU-17      |                           |      |      |      |      |          |      |      | 10   |      |                            |      |      |      |      |          |      |      |      |      |
| OTU-18      |                           |      | 4    | 3    | 2    |          |      |      |      |      |                            |      |      |      |      |          |      |      |      |      |
| OTU-19      |                           |      |      |      |      |          |      |      |      |      |                            | 5    |      | 4    |      |          |      |      |      |      |
| OTU-20      |                           |      |      |      | 8    |          |      |      |      |      |                            |      |      |      |      |          |      |      |      |      |
| OTU-21      |                           | 4    |      |      | 4    |          |      |      |      |      |                            |      |      |      |      |          |      |      |      |      |
| OTU-22      |                           |      |      |      |      |          |      | 1    |      |      |                            |      |      |      |      | 3        |      | 4    |      |      |
| OTU-23      |                           |      |      |      |      |          |      |      |      |      |                            | 7    |      |      |      |          |      |      |      |      |
| OTU-24      | 2                         |      |      |      |      |          |      | 5    |      |      |                            |      |      |      |      |          |      |      |      |      |
| OTU-25      |                           |      |      |      |      |          |      | 5    |      |      |                            |      |      |      | 2    |          |      |      |      |      |
| OTU-26      |                           |      |      |      |      | 6        |      |      |      |      |                            |      |      |      |      |          |      |      |      |      |
| OTU-27      |                           |      |      |      |      |          |      | 5    |      |      |                            |      |      |      |      |          |      |      |      |      |
| OTU-28      |                           |      |      |      |      |          |      | 5    |      |      |                            |      |      |      |      |          |      |      |      |      |
| OTU-29      |                           |      |      |      |      |          |      |      |      |      |                            |      |      |      | 5    |          |      |      |      |      |
| OTU-30      |                           |      |      |      |      |          |      | 1    |      |      | 3                          |      |      |      |      |          |      |      |      |      |
| OTU-31      |                           |      |      |      |      |          |      | 4    |      |      |                            |      |      |      |      |          |      |      |      |      |
| OTU-32      |                           |      |      |      |      |          |      |      |      |      |                            | 4    |      |      |      |          |      |      |      |      |
| OTU-33      |                           |      |      |      |      |          |      |      | 2    |      |                            |      |      |      |      |          |      |      |      | 2    |
| OTU-34      |                           | 1    |      | 3    |      |          |      |      |      |      |                            |      |      |      |      |          |      |      |      |      |
| OTU-35      |                           |      |      |      |      |          |      |      |      |      |                            |      |      | 3    |      |          |      |      |      |      |
| OTU-36      |                           |      |      |      |      |          |      |      | 3    |      |                            |      |      |      |      |          |      |      |      |      |
| OTU-37      |                           |      | 2    |      |      |          |      | 1    | 1    |      |                            |      |      |      |      |          |      |      |      |      |
| OTU-38      |                           |      |      |      |      |          |      |      |      |      |                            |      |      |      |      |          |      |      |      |      |
| OTU-39      |                           |      |      |      |      |          |      |      |      |      | 1                          |      |      |      |      |          |      |      |      | 1    |
| OTU-40      |                           |      |      |      |      |          |      |      |      |      |                            |      |      |      |      | 2        |      |      |      |      |
| OTU-41      | 2                         |      |      |      |      |          |      |      |      |      |                            |      |      |      |      |          |      |      |      |      |
| OTU-42      |                           |      |      |      |      |          |      | 2    |      |      |                            |      |      |      |      |          |      |      |      |      |
| OTU-43      |                           |      |      |      |      |          |      | 2    |      |      |                            |      |      |      |      |          |      |      |      |      |
| OTU-44      |                           |      | 2    |      |      |          |      |      |      |      |                            |      |      |      |      |          |      |      |      |      |
| Total count | 25                        | 30   | 30   | 24   | 26   | 28       | 23   | 25   | 23   | 25   | 26                         | 30   | 24   | 29   | 27   | 24       | 33   | 26   | 24   | 22   |

Table S3. PNG and Japanese individuals for metagenomic sequencing.

| Sampling area and year | ID   | Sex <sup>a</sup> | Age | BMI <sup>b</sup> | N intake <sup>c</sup> |
|------------------------|------|------------------|-----|------------------|-----------------------|
| Levani (2013)          | V001 | M                | 39  | 22.9             | 103.4                 |
| Levani (2013)          | V004 | M                | 22  | 22.7             | 109.6                 |
| Levani (2013)          | V005 | M                | 15  | 20.8             | 117.7                 |
| Levani (2013)          | V009 | M                | 35  | 24.3             | 117.2                 |
| Levani (2013)          | V010 | M                | 32  | 25.6             | 55.4                  |
| Levani (2013)          | V014 | M                | 22  | 26.2             | 67.3                  |
| Levani (2013)          | V018 | F                | 23  | 19.2             | 41.8                  |
| Levani (2013)          | V020 | F                | 24  | 24.9             | 46.2                  |
| Levani (2013)          | V022 | F                | 35  | 27.2             | 87.3                  |
| Levani (2013)          | V031 | F                | 27  | 23.4             | 74.7                  |
| Levani (2013)          | V035 | M                | 40  | 22.9             | 148.0                 |
| Levani (2013)          | V036 | F                | 36  | 19.4             | 81.1                  |
| Maprik (2012)          | V038 | F                | 25  | 23.2             | 65.5                  |
| Maprik (2012)          | M025 | M                | 33  | 23.3             | 68.7                  |
| Maprik (2012)          | M031 | F                | 29  | 22.6             | 115.5                 |
| Maprik (2012)          | M101 | F                | 22  | 19.2             | 66.3                  |
| Tokyo (2013)           | J001 | M                | 44  | 25.1             | 209.1                 |
| Tokyo (2013)           | J002 | M                | 40  | 21.6             | 183.2                 |

<sup>a</sup> M, male; F, female.

<sup>b</sup> BMI, body mass index, given by body weight (kg) per body height(m)<sup>2</sup>.

<sup>c</sup> Values were given by total N (mg) per body weight (kg) per day.

Table S5. *nifH* homologues identified from PNG and Japanese faecal metagenome sequencing reads.

| Sample ID | Query sequence             | Closest reference sequence          |                  | Identity % | Alignment length (aa) | E-value  |
|-----------|----------------------------|-------------------------------------|------------------|------------|-----------------------|----------|
|           |                            | Species                             | Accession number |            |                       |          |
| PNG_V004  | 6INH6:01185:00221_1_153_+  | <i>Clostridium beijerinckii</i>     | WP_026886176     | 87.8       | 49                    | 4.00E-21 |
| PNG_V014  | XK7CQ:00442:01839_1_278_+  | Lachnospiraceae bacterium 3-1       | WP_016286008     | 96.7       | 90                    | 5.68E-55 |
| PNG_V014  | XK7CQ:01216:00792_1_240_-  | Lachnospiraceae bacterium 3-1       | WP_016286008     | 100        | 73                    | 5.33E-41 |
| PNG_V014  | XK7CQ:00625:02773_1_346_-  | Lachnospiraceae bacterium 3-1       | WP_016286008     | 91.9       | 111                   | 7.99E-62 |
| PNG_V014  | XK7CQ:01216:00792_1_240_-  | Lachnospiraceae bacterium 3-1       | WP_016286008     | 100        | 73                    | 5.33E-41 |
| PNG_V014  | XK7CQ:03187:01948_1_161_+  | <i>Clostridium</i> sp. CAG:75*      | WP_004203553     | 98.0       | 49                    | 4.91E-24 |
| PNG_V022  | Z28L8:00332:02004_1_324_+  | Lachnospiraceae bacterium 3-1       | WP_016286008     | 97.2       | 107                   | 1.00E-66 |
| PNG_V035  | EU8EW:00090:01025_1_224_-  | Lachnospiraceae bacterium 3-1       | WP_016286008     | 98.4       | 61                    | 3.00E-32 |
| PNG_V035  | SEU8EW:00502:02449_1_365_+ | Lachnospiraceae bacterium 3-1       | WP_016286008     | 98.2       | 108                   | 1.00E-67 |
| PNG_V035  | EU8EW:00795:02054_1_240_-  | Lachnospiraceae bacterium 3-1       | WP_016286008     | 96.2       | 78                    | 8.00E-45 |
| PNG_V035  | EU8EW:01823:00842_1_316_+  | Lachnospiraceae bacterium 3-1       | WP_016286008     | 97.1       | 104                   | 3.00E-64 |
| PNG_V035  | EU8EW:02502:00267_80_316_- | Lachnospiraceae bacterium 3-1       | WP_016286008     | 98.7       | 77                    | 2.00E-47 |
| PNG_V035  | EU8EW:02796:01992_1_200_+  | Lachnospiraceae bacterium 3-1       | WP_016286008     | 96.9       | 65                    | 7.00E-34 |
| PNG_V035  | EU8EW:03345:01041_1_232_-  | Lachnospiraceae bacterium 3-1       | WP_016286008     | 97.3       | 75                    | 4.00E-41 |
| PNG_V035  | EU8EW:03452:01767_1_275_+  | Lachnospiraceae bacterium 3-1       | WP_016286008     | 97.8       | 91                    | 1.00E-54 |
| PNG_V035  | EU8EW:01384:03160_1_194_-  | <i>Butyrivibrio proteoclasticus</i> | WP_026651776     | 100        | 62                    | 9.00E-33 |
| PNG_V035  | EU8EW:00304:01641_1_221_-  | <i>Lachnospira multipara</i>        | WP_031556311     | 96.2       | 52                    | 5.00E-26 |
| PNG_V035  | EU8EW:00850:02075_1_355_-  | <i>Clostridium</i> sp. CAG:75       | WP_022024781     | 96.1       | 77                    | 2.00E-44 |
| PNG_V035  | EU8EW:02766:01481_1_321_+  | <i>Clostridium</i> sp. CAG:75       | WP_022024781     | 97.1       | 104                   | 1.00E-65 |
| PNG_V035  | EU8EW:01843:03240_1_149_+  | <i>Clostridium</i> sp. CAG:75       | WP_022024781     | 100        | 37                    | 2.00E-15 |
| PNG_V035  | EU8EW:03175:03042_1_214_+  | <i>Clostridium</i> sp. CAG:75       | WP_022024781     | 100        | 69                    | 8.00E-40 |
| JPN_J002  | 41HRQ:01239:00544_1_229_+  | <i>Klebsiella</i> spp.              | WP_004203553     | 80.0       | 60                    | 2.00E-19 |
| JPN_J002  | 41HRQ:01239:00544_1_229_+  | <i>Klebsiella</i> spp.              | WP_004203553     | 98.7       | 75                    | 9.00E-45 |
| JPN_J002  | 41HRQ:02314:03199_1_181_-  | <i>Enterobacter</i> spp.            | WP_039078970     | 97.1       | 35                    | 3.00E-13 |

\* The data of *Clostridium* sp. CAG:75 is derived from human gut metagenome (BioSample: SAMEA3138515).



Table S4. *nifH* homologues identified in human faecal metagenome gene sets.

| Gene set            | Identified homologues | Closest reference sequence       |                  | Identity % | Alignment length (aa) | E-value   |
|---------------------|-----------------------|----------------------------------|------------------|------------|-----------------------|-----------|
|                     |                       | Species                          | Accession number |            |                       |           |
| PNG                 | PNG-1                 | <i>Lachnospira multipara</i>     | WP_031556311     | 94.4       | 213                   | 8.12E-140 |
|                     | PNG-2                 | Lachnospiraceae bacterium 3-1    | EOS26764         | 97.6       | 123                   | 6.65E-78  |
|                     | PNG-3                 | <i>Clostridium</i> sp. CAG:75*   | CCZ52034         | 97.9       | 146                   | 7.47E-98  |
|                     | PNG-4                 | <i>Clostridium</i> sp. CAG:75    | CCZ52034         | 97.9       | 141                   | 2.45E-94  |
| Japanese            | JPN-1                 | <i>Klebsiella pneumoniae</i> 342 | AFV52053         | 99.4       | 179                   | 9.17E-123 |
|                     | JPN-2                 | <i>Clostridium</i> sp. CAG:75    | CCZ52034         | 100.0      | 282                   | 0.00E+00  |
| Denmark_Spain_China | DSC-1                 | Lachnospiraceae bacterium 3-1    | EOS26764         | 97.2       | 282                   | 0.00E+00  |
|                     | DSC-2                 | <i>Clostridium</i> sp. CAG:75    | CCZ52034         | 100.0      | 282                   | 0.00E+00  |
|                     | DSC-3                 | Lachnospiraceae bacterium MC2017 | WP_044906710     | 89.0       | 173                   | 8.88E-105 |
|                     | DSC-4                 | Enterobacteriaceae spp.          | WP_004203553     | 100.0      | 293                   | 0.00E+00  |
|                     | DSC-5                 | <i>Clostridium</i> sp. CAG:75    | CCZ52034         | 99.5       | 201                   | 5.69E-139 |
| Sweden              | Sweden-1              | <i>Clostridium</i> sp. CAG:75    | CCZ52034         | 99.6       | 282                   | 0.00E+00  |
|                     | Sweden-2              | Lachnospiraceae bacterium 3-1    | EOS26764         | 97.2       | 282                   | 0.00E+00  |
|                     | Sweden-3              | Lachnospiraceae bacterium 3-1    | EOS26764         | 99.3       | 282                   | 0.00E+00  |
|                     | Sweden-4              | <i>Butyrivibrio</i> spp.         | WP_026655912     | 96.1       | 282                   | 0.00E+00  |
|                     | Sweden-5              | <i>Butyrivibrio</i> spp.         | WP_026655912     | 83.4       | 271                   | 3.30E-166 |
|                     | Sweden-6              | Lachnospiraceae bacterium MC2017 | WP_044906710     | 87.2       | 203                   | 1.64E-123 |

\* The data of *Clostridium* sp. CAG:75 is derived from human gut metagenome (BioSample: SAMEA3138515).

Table S6. *nifD* and *nifK* homologues identified from human faecal metagenome gene sets.

| Gene set   | Query sequence   | Closest reference sequence                             |                                    |                               | Identity %   | Alignment length (aa) | E-value   |
|--|--|--|------------------------------------|-------------------------------|--------------|-----------------------|-----------|
|  |  | Definition   | Species                            | Accession number              |              |                       |           |
| PNG  | V036_contig00879_gene_6_2543_2968_minus_0_11_39.5395_s           | NifD   | Ruminococcaceae bacterium AE2021   | WP_028513096                  | 96.2         | 133                   | 6.79E-88  |
|  | V036_contig00879_gene_7_2946_3293_minus_0_11_18.0651_s           | NifD   | Lachnospira multipara              | WP_027436974                  | 83.5         | 97                    | 4.74E-50  |
|  | V036_contig00879_gene_8_3290_3841_minus_0_11_43.5437_s           | NifD   | Lachnospira multipara              | WP_031556307                  | 86.7         | 181                   | 4.23E-111 |
|  | V036_contig00879_gene_9_3790_3972_minus_0_11_6.7026_s            | NifD   | Ruminococcaceae bacterium AE2021   | WP_028513096                  | 93.5         | 46                    | 2.13E-21  |
|  | V035_contig00800_gene_10_6894_7547_plus_0_11_36.3085_s           | NifK   | Butyrivibrio sp. AE3003            | WP_026654166                  | 81.5         | 211                   | 1.61E-123 |
|  | V035_contig00800_gene_11_7604_8182_plus_0_11_50.996_s            | NifK   | <i>Clostridium</i> sp. CAG:75*     | CCZ52039                      | 80.2         | 192                   | 6.69E-106 |
| Japan  | FTagent01_contig08661_gene_2_845_2053_minus_0_11_80.4959_b       | NifD   | <i>Clostridium</i> sp. CAG:75      | CCZ52038                      | 100.0        | 402                   | 0.00E+00  |
|  | FTagent12_contig72540_gene_2_332_475_plus_0_10_8.24811_b         | NifD   | Lachnospiraceae bacterium 3-1      | WP_016286010                  | 91.7         | 48                    | 1.94E-21  |
|  | Miseqsingleton_scaffold_193445_gene_1_1_319_plus_2_00_60.7152_b  | NifD   | <i>Klebsiella variicola</i>        | WP_049165071                  | 100.0        | 105                   | 1.56E-68  |
|  | PGMsingleton_contig300576_gene_2_213_706_plus_0_10_44.649_b      | NifD   | <i>Klebsiella variicola</i>        | WP_044613654                  | 87.2         | 156                   | 5.23E-91  |
|  | re-singleton_contig113700_gene_1_1_228_minus_0_10_15.4132_p      | NifD   | <i>Klebsiella pneumoniae</i>       | WP_046882228                  | 93.4         | 76                    | 5.01E-41  |
|  | FTagent01_contig08661_gene_1_8_826_minus_0_11_43.0035_b          | NifK   | <i>Clostridium</i> sp. CAG:75      | CCZ52039                      | 97.5         | 238                   | 8.54E-166 |
|  | FTagent01_contig15875_gene_1_238_363_plus_0_11_4.57206_b         | NifK   | <i>Clostridium</i> sp. CAG:75      | CCZ52039                      | 100.0        | 41                    | 3.80E-19  |
|  | Miseqsingleton_scaffold_199724_gene_1_1_181_minus_0_10_35.3352_p | NifK   | <i>Klebsiella pneumoniae</i>       | WP_061153956                  | 100.0        | 60                    | 1.05E-32  |
|  | re-singleton_contig101838_gene_1_1_557_plus_2_01_88.5988_b       | NifK   | Enterobacteriaceae                 | WP_008804133                  | 100.0        | 184                   | 1.78E-131 |
|  | singleton_HQWINQD02GQM01_gene_1_1_597_minus_0_10_83.428_b        | NifK   | <i>Clostridium butyricum</i> 60E.3 | ENZ32193                      | 100.0        | 199                   | 3.06E-143 |
|  | singleton_HQWINQD02GQM01_gene_1_1_597_minus_0_10_83.428_b        | NifK   | <i>Klebsiella pneumoniae</i>       | CAA29588                      | 100.0        | 199                   | 3.06E-143 |
| Denmark_Spain_China                                    | 35118  | NifD   | <i>Clostridium</i> sp. CAG:75      | CCZ52038                      | 99.8         | 481                   | 0.00E+00  |
|  | 1818158  | NifD   | Lachnospiraceae bacterium 3-1      | WP_016286010                  | 88.7         | 467                   | 0.00E+00  |
|  | 35118  | NifD   | <i>Clostridium</i> sp. CAG:75      | CCZ52038                      | 88.8         | 480                   | 0.00E+00  |
|  | 3277371  | NifD   | <i>Klebsiella variicola</i>        | WP_049165071                  | 99.2         | 131                   | 2.20E-89  |
|  | 3280711  | NifD   | <i>Klebsiella variicola</i>        | WP_044613654                  | 99.7         | 335                   | 0.00E+00  |
|  | 3708105  | NifD   | <i>Clostridium</i> sp. CAG:75      | CCZ52038                      | 100.0        | 273                   | 0.00E+00  |
|  | 764122   | NifK   | Butyrivibrio sp. AE3003            | WP_026654166                  | 83.6         | 134                   | 3.82E-75  |
|  | 1797041  | NifK   | <i>Clostridium</i> sp. CAG:75      | CCZ52039                      | 90.0         | 50                    | 1.04E-22  |
|  | 1818155  | NifK   | Butyrivibrio sp. LB2008            | WP_019553696                  | 80.9         | 424                   | 0.00E+00  |
|  | 3273820  | NifK   | <i>Klebsiella pneumoniae</i>       | WP_032740067                  | 99.6         | 451                   | 0.00E+00  |
|  | 3323959  | NifK   | <i>Clostridium</i> sp. CAG:75      | CCZ52039                      | 98.7         | 394                   | 0.00E+00  |
|  | 3691959  | NifK   | Lachnospiraceae bacterium MA202    | WP_044922962                  | 82.0         | 61                    | 3.26E-27  |
|  | 3708106  | NifK   | <i>Clostridium</i> sp. CAG:75      | CCZ52039                      | 99.0         | 96                    | 9.74E-58  |
|  | Sweden   | gene_15804495_484_NODE_77899_length_17005_cov_6.235048 | NifD                               | Lachnospiraceae bacterium 3-1 | WP_016286010 | 88.7                  | 486       |
| gene_14996018_543_NODE_29503_length_19686_cov_7.427614 |  | NifD   | <i>Clostridium</i> sp. CAG:75      | CCZ52038                      | 99.8         | 481                   | 0.00E+00  |
| gene_677624_317_NODE_122348_length_1171_cov_4.495303   |  | NifD   | Lachnospiraceae bacterium 3-1      | WP_016286010                  | 95.5         | 403                   | 0.00E+00  |
| gene_18672520_88_NODE_16775_length_1066_cov_4.862101   |  | NifD   | <i>Clostridium</i> sp. CAG:75      | CCZ52038                      | 89.3         | 367                   | 0.00E+00  |
| gene_15595724_243_NODE_77119_length_1084_cov_3.226937  |  | NifD   | <i>Clostridium</i> sp. CAG:75      | CCZ52038                      | 87.9         | 340                   | 0.00E+00  |
| gene_10179291_195_NODE_43917_length_1157_cov_4.478825  |  | NifD   | <i>Clostridium</i> sp. CAG:75      | CCZ52038                      | 89.4         | 292                   | 0.00E+00  |
| gene_9808769_552_NODE_116611_length_662_cov_2.649547   |  | NifD   | Butyrivibrio sp. MC2021            | WP_026517309                  | 89.8         | 49                    | 2.14E-22  |
| gene_416326_NODE_2750038_length_11457_cov_23.302784    |  | NifD   | Lachnospiraceae bacterium 3-1      | WP_016286010                  | 95.5         | 490                   | 0.00E+00  |
| gene_328325_NODE_2098381_length_3771_cov_9.653142      |  | NifD   | <i>Clostridium</i> sp. CAG:75      | CCZ52038                      | 90.2         | 256                   | 4.73E-169 |
| gene_588445_NODE_4031947_length_6943_cov_22.383984     |  | NifD   | <i>Clostridium</i> sp. CAG:75      | CCZ52038                      | 88.5         | 480                   | 0.00E+00  |
| gene_1581485_NODE_9865591_length_8092_cov_22.446737    |  | NifD   | <i>Clostridium</i> sp. CAG:75      | CCZ52038                      | 89.2         | 480                   | 0.00E+00  |
| gene_10206795_195_NODE_77711_length_2199_cov_4.713506  |  | NifK   | Butyrivibrio sp. AE2005            | WP_026668199                  | 82.5         | 93                    | 0.00E+00  |
| gene_18706167_88_NODE_48280_length_577_cov_3.233969    |  | NifK   | <i>Clostridium</i> sp. CAG:75      | CCZ52039                      | 82.4         | 125                   | 9.82E-62  |
| gene_15801722_484_NODE_75170_length_6822_cov_6.351363  |  | NifK   | <i>Treponema</i> sp. C6A8          | WP_027728473                  | 89.3         | 112                   | 9.70E-67  |
| gene_14139807_432_NODE_147117_length_1403_cov_3.354954 |  | NifK   | <i>Clostridium</i> sp. CAG:75      | CCZ52039                      | 99.0         | 101                   | 3.21E-65  |
| gene_4311480_305_NODE_127346_length_1476_cov_4.300136  |  | NifK   | <i>Clostridium</i> sp. CAG:75      | CCZ52039                      | 100.0        | 60                    | 1.78E-31  |
| gene_8870726_547_NODE_67519_length_13382_cov_21.034674 |  | NifK   | <i>Clostridium</i> sp. CAG:75      | CCZ52039                      | 90.0         | 50                    | 1.46E-22  |
| gene_18022002_225_NODE_3657_length_13385_cov_22.584087 |  | NifK   | Lachnospiraceae bacterium MA2020   | WP_044922962                  | 82.0         | 61                    | 3.07E-27  |

\* The data of *Clostridium* sp. CAG:75 is derived from human gut metagenome (BioSample: SAMEA3138515).

Table S7. *nifD* and *nifK* homologues identified using MG-RAST from PNG and Japaense faecal metagenome sequencing reads.

| Sample ID | Query sequence              | Closest reference sequence |                                  |                  | Identity % | Alignment length (aa) | E-value  |
|-----------|-----------------------------|----------------------------|----------------------------------|------------------|------------|-----------------------|----------|
|           |                             | Definition                 | Species                          | Accession number |            |                       |          |
| PNG_V004  | 6INH6:02571:02398_1_353_-   | NifD                       | <i>Raoultella terrigena</i>      | AFV36358         | 99.0       | 115                   | 3.00E-75 |
| PNG_V004  | 6INH6:02764:01137_1_150_+   | NifK                       | <i>Klebsiella</i> spp.           | WP_016161448     | 98         | 47                    | 3.00E-23 |
| PNG_V014  | XK7CQ:00908:00644_1_289_+   | NifK                       | <i>Clostridium</i> sp. CAG:75*   | CCZ52039         | 87.0       | 85                    | 4.00E-52 |
| PNG_V035  | EU8EW:01327:00961_1_243_+   | NifK                       | <i>Clostridium</i> sp. CAG:75    | CCZ52039         | 81.0       | 74                    | 5.00E-30 |
| PNG_V035  | EU8EW:01620:00672_1_337_-   | NifK                       | Lachnospiraceae bacterium MA2020 | WP_044922962     | 94.0       | 104                   | 7.00E-41 |
| PNG_V035  | EU8EW:02451:00193_1_280_-   | NifK                       | <i>Clostridium</i> sp. CAG:75    | CCZ52039         | 91.0       | 88.0                  | 2.00E-42 |
| PNG_V035  | EU8EW:02954:00789_1_350_-   | NifK                       | <i>Clostridium</i> sp. CAG:75    | CCZ52039         | 88.0       | 92.0                  | 7.00E-44 |
| JPN_J002  | 41HRQ:01232:02393_1_337_-   | NifD                       | <i>Klebsiella pneumoniae</i>     | WP_060590496     | 100.0      | 47                    | 5.00E-23 |
| JPN_J002  | 41HRQ:02346:01381_1_343_-   | NifD                       | <i>Klebsiella variicola</i>      | WP_044613654     | 96.0       | 107                   | 3.00E-67 |
| JPN_J002  | 41HRQ:03302:02188_1_285_+   | NifD                       | <i>Klebsiella variicola</i>      | WP_044613654     | 97.1       | 93                    | 2.00E-59 |
| JPN_J002  | 41HRQ:00900:00564_1_210_+   | NifK                       | <i>Klebsiella</i> spp.           | WP_004203556     | 99         | 68                    | 5.00E-39 |
| JPN_J002  | 41HRQ:01033:02206_105_354_- | NifK                       | <i>Klebsiella</i> spp.           | WP_016161448     | 95         | 79                    | 6.00E-45 |

\* The data of *Clostridium* sp. CAG:75 is derived from human gut metagenome(BioSample: SAMEA3138515).

Table S8. Quantitative PCR primers and probes (5'–3')

| Primers and probes | Sequence (5'–3')             |
|--------------------|------------------------------|
| OTU-08 forward     | ACCGTATTCTGAAAGAGGGATATGG    |
| OTU-08 reverse     | GGATGTGATAATTCCTCGGCCAG      |
| OTU-08 probe       | CATCCTACGCCTGGTTCTGGTCCTCCTG |
| OTU-24 forward     | GGACTCCACCCGTCTGATC          |
| OTU-24 reverse     | CAACGGACATCGCCATAGC          |
| OTU-24 probe       | ACGTCTTCGAGCTCCAGATCCTCGAC   |
| OTU-34 forward     | TCGTCAGAGAAGGCTTTGGC         |
| OTU-34 reverse     | GCGTCCTGCACATCCAACA          |
| OTU-34 probe       | CGGTCCGCCGGATTCCACACATCTG    |

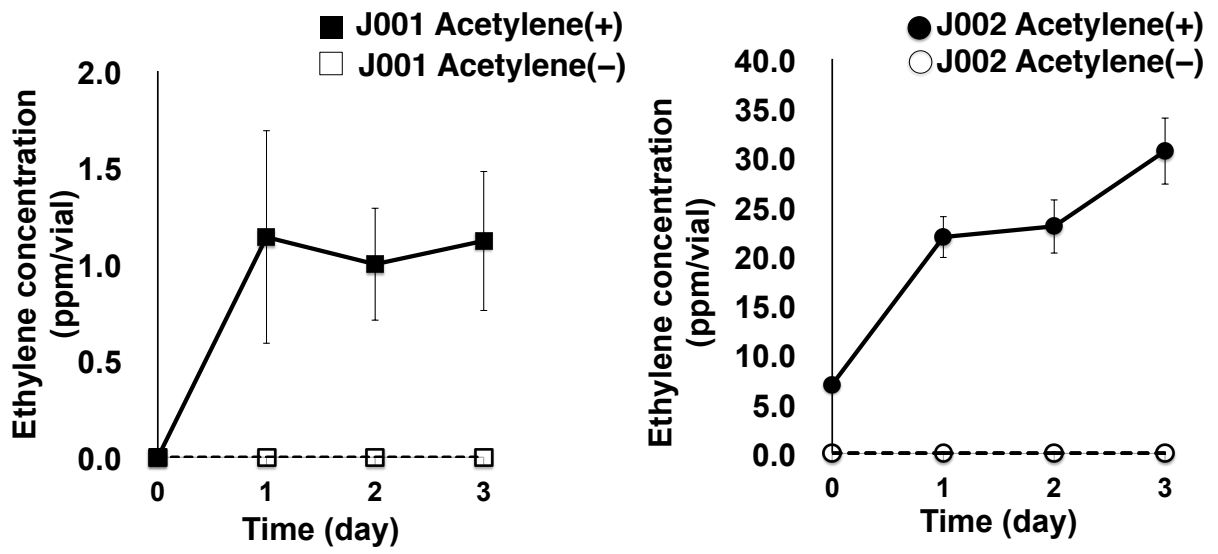


Figure S1.

Acetylene reduction assays in Japanese frozen fecal samples (J001 and J002).