

RESEARCH ARTICLE

# Association between functional lactase variants and a high abundance of *Bifidobacterium* in the gut of healthy Japanese people

Kumiko Kato<sup>1\*</sup>, Sachiko Ishida<sup>2</sup>, Masami Tanaka<sup>3</sup>, Eri Mitsuyama<sup>1</sup>, Jin-zhong Xiao<sup>1</sup>, Toshitaka Odamaki<sup>1</sup>

**1** Next Generation Science Institute, Morinaga Milk Industry Co., LTD., Kanagawa, Japan, **2** Business Planning Group, Next Generation Science Institute, DeNA Life Science, Inc., Tokyo, Japan, **3** R&D Group, DeNA Life Science, Inc., Tokyo, Japan

\* [ku-katou@morinagamilk.co.jp](mailto:ku-katou@morinagamilk.co.jp)



**OPEN ACCESS**

**Citation:** Kato K, Ishida S, Tanaka M, Mitsuyama E, Xiao J-z, Odamaki T (2018) Association between functional lactase variants and a high abundance of *Bifidobacterium* in the gut of healthy Japanese people. PLoS ONE 13(10): e0206189. <https://doi.org/10.1371/journal.pone.0206189>

**Editor:** Brenda A. Wilson, University of Illinois at Urbana-Champaign, UNITED STATES

**Received:** July 10, 2018

**Accepted:** October 7, 2018

**Published:** October 19, 2018

**Copyright:** © 2018 Kato et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the paper and its Supporting Information files.

**Funding:** Our work was funded by Morinaga Milk Industry Co., LTD., DeNA Life Science, Inc., and the Kanagawa Prefecture's model project related to the creation of ME-BYO industry in FY2016. Employees of Morinaga Milk Industry Co., LTD., KK, EM, JX and TO, and those of DeNA Life Science, Inc., SI and MT, were received salary from each company. The specific roles of these authors

## Abstract

Previous studies have shown that Japanese people exhibit a higher abundance of *Bifidobacterium* compared to people from other countries. Among the possible factors affecting the gut microbiota composition, an association of functional lactase gene variants with a higher abundance of *Bifidobacterium* in the gut has been proposed in some reports. However, no Japanese subjects were included in these studies. In this study, we investigated the possible contribution of functional lactase loci to the high abundance of *Bifidobacterium* in Japanese populations. Based on a data analysis assessing 1,068 healthy Japanese adults, a number of subjects is at least seven times greater than that reported in available online data. All subjects possessed CC genotype at rs4988235 and the GG at rs182549, which are associated with low lactase activity. We observed a positive correlation between dairy product intake and *Bifidobacterium* abundance in the gut. Considering previous reports, which revealed that four additional functional lactase loci, rs145946881, rs41380347, rs41525747 and rs869051967 (ss820486563), are also associated with low lactase activity in Japanese people, our findings imply the possible contribution of host genetic variation-associated low lactase activity to the high abundance of *Bifidobacterium* in the Japanese population.

## Introduction

The life expectancy of Japanese population is one of the longest life expectancies of any country in the world[1]. The underlying causes of long life in Japan have been debated, and the possibilities include good hygiene, a high level of health consciousness in Japan, and the Japanese diet[2]. In previous reports, Japanese healthy adults exhibited a higher abundance of *Bifidobacterium* compared to people from other countries[3–5]. Considering that *Bifidobacterium* naturally inhabits the human gastrointestinal tract (GIT) and is thought to play pivotal roles in

are articulated in the 'author contributions' section. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing interests:** I have read the journal's policy, and the authors of this manuscript have the following competing interests: Kumiko Kato, Eri Mitsuyama, Jin-zhong Xiao and Toshitaka Odamaki were employed by Morinaga Milk Industry Co., Ltd., and Sachiko Ishida and Masami Tanaka were employed by DeNA Life Science, Inc. This does not alter our adherence to PLOS ONE policies on sharing data and materials.

maintaining human health[6], the higher abundance of this genus seems to be another potential reason for long life in Japan. The proportion of *Bifidobacterium* in the gut microbiota is affected by many factors, such as host age[4], stress[7] and diet[8].

Host genetic variations related to fucosyltransferase 2[9] and lactase[10–13] have been reported to contribute to *Bifidobacterium* abundance. In particular, the functional lactase gene (*LCT*) variant rs4988235 has been associated with *Bifidobacterium* abundance in the gut in multiple reports[14–16]. The presence of the CC genotype at rs4988235 is related to low lactase activity and milk indigestibility in adulthood. Because *Bifidobacterium* assimilates lactose as a preferred carbon source for growth, it is reasonable that subjects with the CC genotype at this locus have a higher *Bifidobacterium* abundance in their gut. The locus rs4988235 is mainly reported as a characteristic locus related to lactase activity in Europeans. However, this single locus is insufficient to explain the frequency of the lactase phenotype present in various populations worldwide[11]. Another variant, the GG genotype at rs182549, has been reported to contribute to lactase persistence in Japanese-Brazilian and Chinese populations, whose genetic backgrounds are related more closely to those of the Japanese population than to those of the European population[11]. Therefore, the rs182549 locus might also be an important contributor to the lactase activity in the Japanese population.

Here, we investigated the associations between these two single nucleotide polymorphisms (SNPs), which have been reported as functional lactase variants, and the proportion of *Bifidobacterium* in the gut microbiota of 1,068 healthy Japanese adults. Furthermore, information for an additional four functional lactase variants that were reported previously[12,13] but were missing from our data set was also included using Japanese genome data obtained from two previous studies[13,17].

## Results/Discussion

Of 1,250 participants enrolled in this study, a total of 182 participants were removed from data analysis due to refusal to participate ( $n = 64$ ), pregnancy or lactation ( $n = 7$ ), medication use within the last two weeks ( $n = 96$ ), and mismatch with the criteria for quality control of genotypic data ( $n = 15$ , see [materials and methods](#)). Finally, a total of 1,068 healthy Japanese subjects were selected ([S1 Fig](#)).

The abundance of *Bifidobacterium* in the Japanese gut microbiota was confirmed by 16S ribosomal RNA gene sequencing. A total of 9,285,977 high-quality paired sequences were obtained from 1,068 samples, and  $8,695 \pm 2,255$  (average  $\pm$  standard deviation) reads per sample were generated. The average of *Bifidobacterium* abundance was  $6.3 \pm 7.8\%$ , which was lower than that in previous Japanese reports ( $17.9 \pm 15.2\%$ [3],  $10.5 \pm 11.0\%$ [4],  $13.6 \pm 17.7\%$ [5]). This discrepancy might be due to differences in faecal samplings, DNA extraction and analysis methods including the primers used[5,18,19], in addition to the distribution of the population in each study. Nevertheless, the percentage of *Bifidobacterium* abundance in this study is higher than those for data on subjects in many other countries[3,20–22]. A previous report showed that *Bifidobacterium* abundance was less than 5% in 9 of 11 countries, with an average of  $3.89 \pm 5.45\%$ [3]. The relatively higher abundance of *Bifidobacterium* in the Japanese gut microbiota was confirmed in our data.

We then investigated the genotype frequencies of two functional *LCT* variants, rs4988235 and rs182549, for lactase persistence. The genotyping results showed that these two variants were monomorphic: all subjects had a CC genotype at rs4988235 and GG genotype at rs182549 ([Table 1](#)), which are associated with low lactase activity. These genotypic frequencies for the two variants were confirmed in another 42 and 104 Japanese samples from the previous report [13] and the 1000 Genomes Project [17], respectively. In addition to the two variants,

**Table 1. Frequencies of LCT non-persistent genotypes in the Japanese population.**

| rsID                            | rs4988235                         | rs182549                          | rs145946881       | rs41380347        | rs41525747      | rs869051967<br>(ss820486563) |
|---------------------------------|-----------------------------------|-----------------------------------|-------------------|-------------------|-----------------|------------------------------|
|                                 | Chr.                              | 2                                 | 2                 | 2                 | 2               | 2                            |
| position                        | 136608646                         | 136616754                         | 136608746         | 136608651         | 136608643       | 136608745                    |
|                                 | LCT-13910 C>T                     | LCT-22018 G>A                     | LCT-14010 G>C     | LCT-13915 T>G     | LCT-13907 C>G   | LCT-14009 T>G                |
| Genotype of LCT non-persistence | CC                                | GG                                | GG                | TT                | CC              | TT                           |
| This study<br>(n = 1068)        | 100%<br>(1068/1068) <sup>*1</sup> | 100%<br>(1061/1061) <sup>*2</sup> | N/A <sup>*3</sup> | N/A               | N/A             | N/A                          |
| Reference [12]<br>(n = 42)      | 100%<br>(42/42)                   | N/A                               | 100%<br>(42/42)   | 100%<br>(42/42)   | 100%<br>(42/42) | 100%<br>(42/42)              |
| Reference [17]<br>(n = 104)     | 100%<br>(104/104)                 | 100%<br>(104/104)                 | 100%<br>(104/104) | 100%<br>(104/104) | N/A             | N/A                          |

\*1 Non-persistence/persistence subject number

\*2 No rs182549 data in seven subjects

\*3 N/A, not available

<https://doi.org/10.1371/journal.pone.0206189.t001>

four other *LCT* variants, rs145946881, rs41380347, rs41525747 and rs869051967 (ss820486563), have also been reported to affect lactase activity[12,13]. Our genotype dataset unfortunately lacked information concerning these four variants; however, we observed no variation at these SNPs in the Japanese population data from two previous studies[13,17] (Table 1). Table 2 shows the frequencies of LCT non-persistent genotypes at these six SNPs in different populations. Only the East Asian population, including Japanese population, has no variation at all of these SNPs. Based on these findings, all available data related to functional

**Table 2. Frequencies of LCT non-persistent genotypes in the different population based on reference [17].**

| rsID                            | rs4988235                           | rs182549              | rs145946881           | rs41380347            | rs41525747        | rs869051967<br>(ss820486563) |
|---------------------------------|-------------------------------------|-----------------------|-----------------------|-----------------------|-------------------|------------------------------|
|                                 | Chr.                                | 2                     | 2                     | 2                     | 2                 | 2                            |
| position                        | 136608646                           | 136616754             | 136608746             | 136608651             | 136608643         | 136608745                    |
|                                 | LCT-13910 C>T                       | LCT-22018 G>A         | LCT-14010 G>C         | LCT-13915 T>G         | LCT-13907 C>G     | LCT-14009 T>G                |
| Genotype of LCT non-persistence | CC                                  | GG                    | GG                    | TT                    | CC                | TT                           |
| All<br>(n = 2504)               | 75.60%<br>(1893/2504) <sup>*1</sup> | 75.40%<br>(1887/2504) | 99.40%<br>(2490/2504) | 99.90%<br>(2501/2504) | N/A <sup>*2</sup> | N/A                          |
| African<br>(n = 661)            | 95.20%<br>(629/661)                 | 95.20%<br>(629/661)   | 97.90%<br>(647/661)   | 100%<br>(661/661)     | N/A               | N/A                          |
| American<br>(n = 347)           | 63.10%<br>(219/347)                 | 62.80%<br>(218/347)   | 100%<br>(347/347)     | 99.10%<br>(344/347)   | N/A               | N/A                          |
| East Asian<br>(n = 504)         | 100%<br>(504/504)                   | 100%<br>(504/504)     | 100%<br>(504/504)     | 100%<br>(504/504)     | N/A               | N/A                          |
| European<br>(n = 503)           | 30.60%<br>(154/503)                 | 30.40%<br>(153/503)   | 100%<br>(503/503)     | 100%<br>(503/503)     | N/A               | N/A                          |
| South Asian<br>(n = 489)        | 79.10%<br>(387/489)                 | 78.30%<br>(383/489)   | 100%<br>(489/489)     | 100%<br>(489/489)     | N/A               | N/A                          |

\*1 Non-persistence/persistence subject number

\*2 N/A, not available

<https://doi.org/10.1371/journal.pone.0206189.t002>

lactase variants regarding the Japanese genome indicated that low lactase activity in the Japanese population seems to contribute to the higher abundance of *Bifidobacterium* in the Japanese population.

A previous study showed a positive correlation between *Bifidobacterium* abundance in the European population and the amount of dairy product consumption only in subjects with the GG genotypes at rs4988235 on the reverse strand, which is equivalent to the CC genotype on the forward strand [15]. This report indicates that the relatively higher amount of lactose contained in dairy products can directly reach the gut and be available to members of the genus *Bifidobacterium*, due to the low lactase activity. We therefore investigated the association between dairy product intake and *Bifidobacterium* abundance using 1,068 Japanese samples, all of which homogeneously presented the CC genotype. A brief-type diet history questionnaire (BDHQ) showed that the average dairy product consumption in the enrolled subjects was  $67.9 \pm 61.4$  g/1,000 kcal. A significant positive correlation was observed between the amount of dairy consumed and the relative abundance of *Bifidobacterium* in the gut ( $r = 0.164$ ,  $p < 0.01$ , Fig 1 and S1 Table), suggesting an association between the higher amount of lactose derived from dairy products and the higher abundance of *Bifidobacterium* in the gut. Considering that the influence of diet on the gut microbiota is very complex and the fact that lactose does not represent a selective growth substrate for bifidobacteria, the monophyletic lactase functional variants in Japanese people may partially contribute to the significant positive correlation between the amount of diet consumption of lactose and the abundance of *Bifidobacterium* in the gut. Nevertheless, our large scale survey of Japanese confirmed the observation, that subjects in the European population with the GG genotypes at rs4988235 could have increased *Bifidobacterium* abundance in the gut by consuming lactose [15].

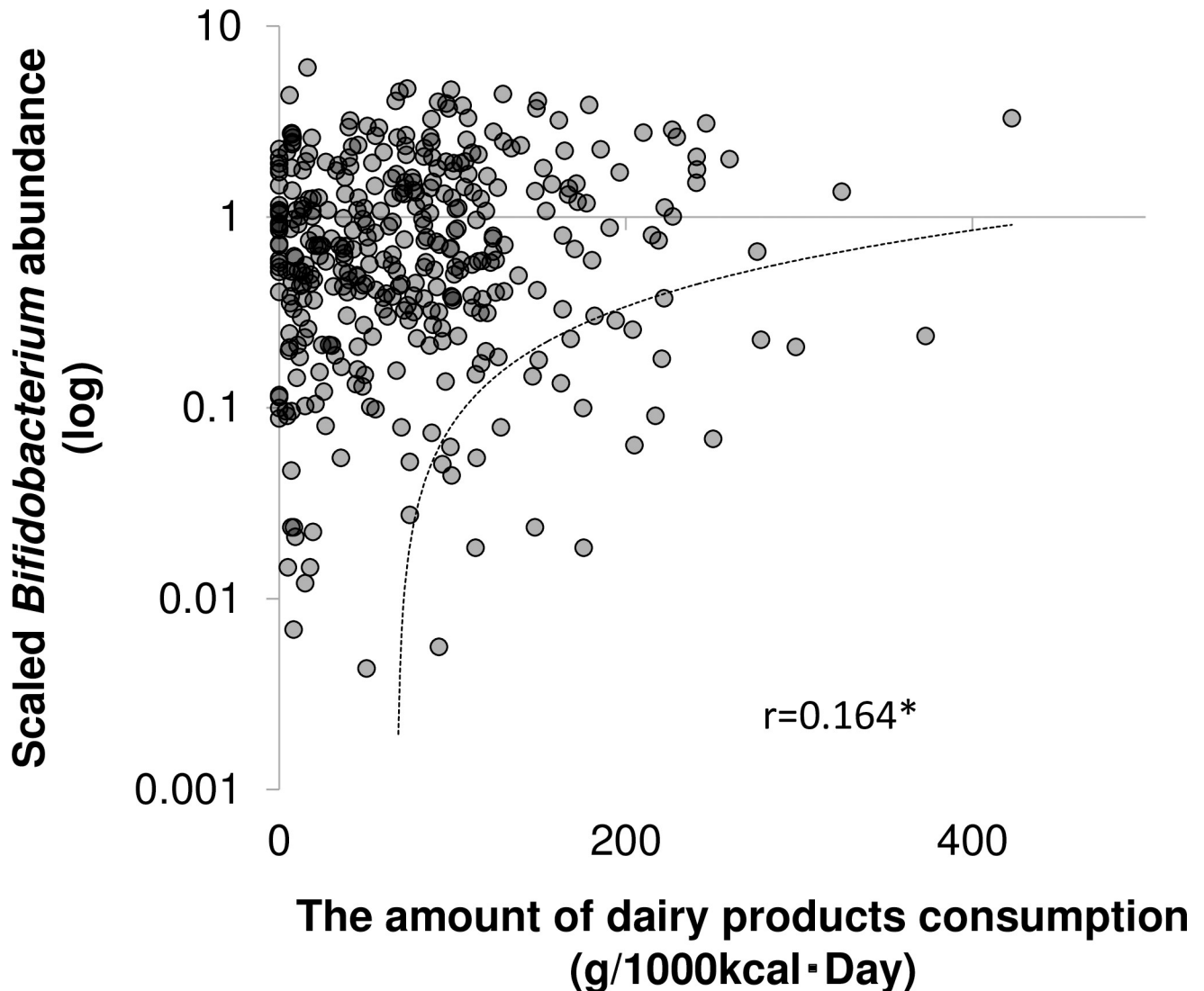
In conclusion, we demonstrated the association of six lactase functional variants with the relative abundance of *Bifidobacterium* in the Japanese gut using our large data set of 1,068 subjects as well as 146 Japanese genomes derived from two previous reports [13,17]. Our large scale data revealed that there was no variation in the SNPs associated with low lactase activity in the Japanese populations. Referring to a previous report, approximately 85% of Japanese people are reportedly lactose intolerant [23]. The underlying reason for the 15% population gap observed between the phenotype and genotype for lactase activity is unclear. A previous report has suggested that three *LCT* variants, rs41525747, rs41380347 and rs145946881, which differ from the variants that contributed to lactase activity in European and American populations, are effective functional variants in Africa and the Middle East [11]. Since the effective locus is thought to vary between population origins, there might be undiscovered lactase functional variants in Japanese populations.

Our results suggest a possible contribution of monophyletic lactase variants to the higher abundance of *Bifidobacterium* in the Japanese population (Fig 2). However, since lactose is not a selective growth substrate for *Bifidobacterium* in the gut, we should consider other possible contributing factors, such as the various types of soluble dietary fibres in Japanese diet. Further studies are needed to understand the mechanism by which the gut microbiota composition is established.

## Material and method

### Study subjects and sample collection

A total of 1,250 healthy Japanese adults were enrolled in this study as part of MYCODE Research, a research platform based on customers of MYCODE (DeNA Life Science Inc., Tokyo, Japan), a personal genome service in Japan. A total of 182 subjects were removed from the data analysis based on the following criteria (see also S1 Fig): 64 subjects declined participation, seven subjects were pregnant or lactating women, 96 subjects had taken medication



**Fig 1. Relationship between dairy product intake and *Bifidobacterium* abundance in the gut.** \* $p < 0.01$ , Spearman's correlation test. Dot line represents a linearized approximation to the data.

<https://doi.org/10.1371/journal.pone.0206189.g001>

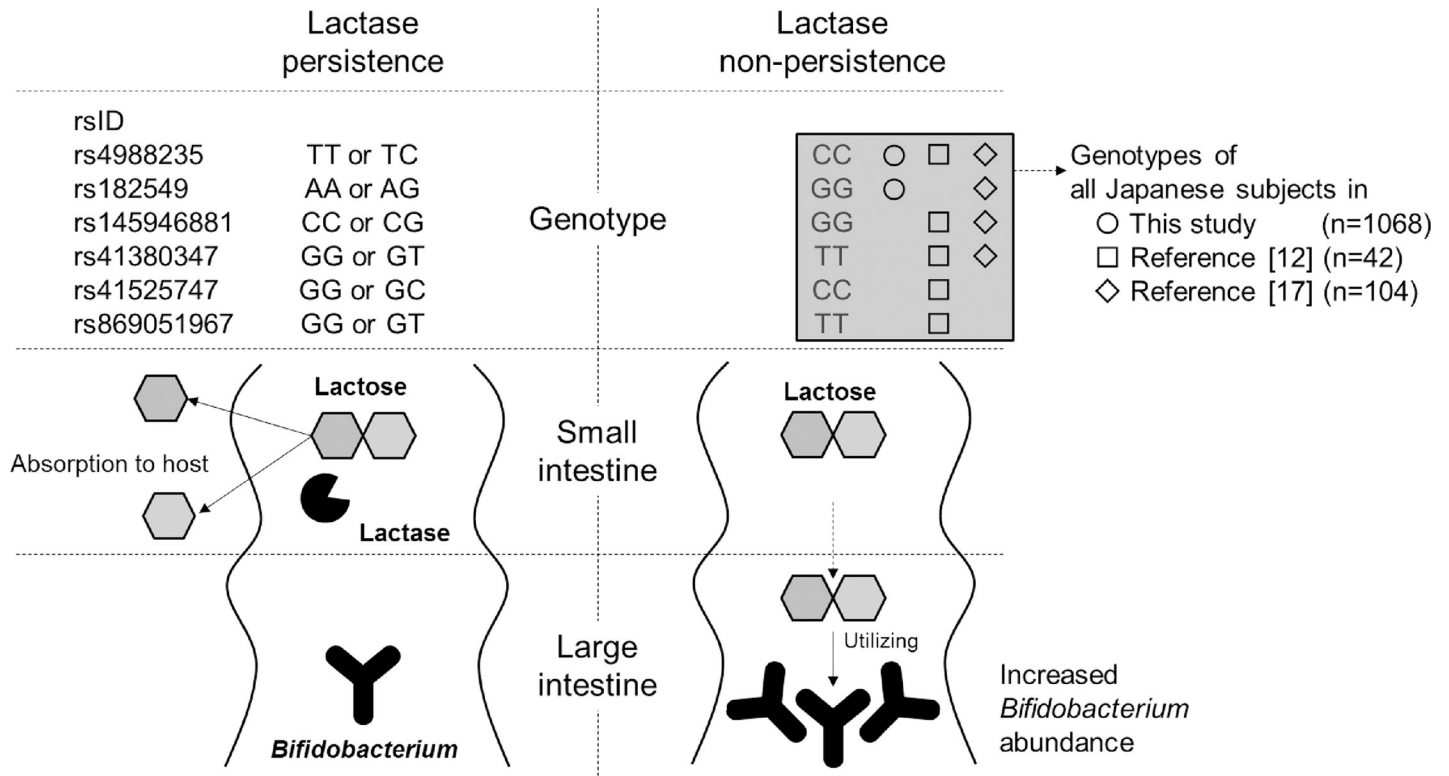
within the last two weeks, and 15 subjects did not meet the criteria for quality control of the genotypic data as described in detail below. Finally, the 1,068 participants consisted of 541 women and 527 men with a median age of 41 years (range 20–64 years) and median body mass index of 21.4 kg/m<sup>2</sup> (inter-quartile range 19.7–23.7 kg/m<sup>2</sup>).

The entire study was approved by both the ethics committee of DeNA Life Science Inc. (protocol #20160727\_1) and the Institute of Medical Science, University of Tokyo (protocol #28-29-1125) (Tokyo, Japan).

Written informed consent was initially obtained for MYCODE Research covering diverse genomic research comprehensively. Then, additional informed consent for this specific study was obtained from all subjects on the MYCODE website.

Saliva samples were collected for MYCODE genetic testing, and these genetic data were used for this study. All participants submitted their own stool samples for gut microbiota analyses. A stool sample aliquot was mixed with 1 ml of guanidine thiocyanate (GuSCN) solution





**Fig 2. Summary of LCT functional variants.** The large data set suggest a possible contribution of monophyletic lactase variants to the higher abundance of *Bifidobacterium* in the Japanese population.

<https://doi.org/10.1371/journal.pone.0206189.g002>

(TechnoSuruga Laboratory Co., Ltd, Shizuoka, Japan)[24] and was transported to the laboratory by postal mail at room temperature. Immediately upon receipt, the faecal samples were stored at  $-80^{\circ}\text{C}$  until the day of analysis.

### Genotyping and quality control

Genotyping of SNPs, was performed using either an Infinium OmniExpress-24+ BeadChip or a Human OmniExpress-24+ BeadChip (Illumina Inc., San Diego, CA, United States). Based on the genotypic data, a total of 15 subjects were removed in the process of quality controls using PLINK version 1.9[25] as follows: two subjects with call rates under 95% which indicated low reliability of their genotyping results; nine subjects who were one of a pair with a proportion of identical by descent (IBD)  $>0.185$ ; which indicates kinship; one subject showing a discordance between self-reported sex and genotyped sex; and three subjects determined to have non-Japanese ancestry by principal component analysis.

### DNA extraction from faecal samples

DNA extraction from human faecal samples was performed using the bead-beating method as previously described[4] with some modifications. Briefly, 500  $\mu\text{l}$  of faecal sample in GuSCN solution was vigorously vortexed with glass beads (300 mg; 0.1 mm in diameter) and 500  $\mu\text{l}$  of buffer-saturated phenol using a Multi-Beads Shocker (Yasui Kikai Co., Osaka, Japan) at a speed of 2,700 rpm for 180 s. After centrifugation at  $10,000\times g$  for 10 min, 400  $\mu\text{l}$  of the supernatant was extracted with phenol-chloroform, and 250  $\mu\text{l}$  of the supernatant was precipitated with isopropanol. The purified DNA was suspended in 1,000  $\mu\text{l}$  of Tris-EDTA buffer (pH 8.0).

## Sequencing and data processing of bacterial 16S rRNA sequences

16S rRNA gene sequencing was performed as previously described with minor modifications [4]. Briefly, the V3-V4 region of the bacterial 16S rRNA gene was amplified by PCR in triplicate using the TaKaRa Ex Taq HS Kit (TaKaRa Bio, Shiga, Japan) and the primer sets Tru357F (5'-CGCTCTTCCGATCTCTGTACGGRAGGCAGCA G-3') and Tru806R (5'-CGCTCTTCCGATCTGACG- GACTACHVGGGTWTCTAAT-3') with the following program: preheating at 94°C for 3 min; 30 cycles of denaturation at 94°C for 30 s, annealing at 50°C for 30 s and extension at 72°C for 30 s; and terminal extension at 72°C for 5 min. A 1- $\mu$ l sample of the combined PCR products was amplified with barcoded primers adapted for Illumina MiSeq sequencing: Fwd 5'-AATGATACGGCGACCACCGAGATCTACACXXXXXXXXXXCACTCTTTCCCTACACGACGCTCTTCCGATCTCTG-3' and Rev 5'-CAAGCAGAAGACGCATACGAGATXXXXXXXXXXGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGAC-3', where X represents a barcode base. Amplification was performed according to the program described above except only eight cycles were performed. The products were purified and quantified by a QIAquick PCR Purification Kit (Qiagen, Valencia, CA, United States) and Quant-iT PicoGreen dsDNA Assay Kit (Thermo Fisher Scientific, Waltham, MA, United States) according to the manufacturer's protocols. Equal amounts of amplicons were pooled and purified with the GeneRead Size Selection Kit (Qiagen) according to the manufacturer's protocol. The pooled libraries were sequenced with an Illumina MiSeq instrument and the MiSeq v3 Reagent Kit (Illumina Inc., San Diego, CA, United States).

After acquiring the Illumina paired-end reads, the Bowtie-2 program [26] (ver. 2-2.2.4) was used to remove reads mapped to the PhiX 174 sequence and the Genome Reference Consortium human build 38 (GRCh38). Thereafter, the 3' region of each read with a PHRED quality score of less than 17 was trimmed. Trimmed reads less than 150 bp in length with an average quality score of less than 25 or those lacking paired reads were also removed. The trimmed paired-end reads were combined by the fastq-join script in EA-Utils [27] (ver. 1.1.2-537). Potential chimeric sequences were removed by reference-based chimaera checking in USEARCH [28] (ver. 5.2.32) and the Genomes OnLine Database (GOLD) (<http://drive5.com/otupipe/gold.tz>).

Non-chimeric sequences were analysed via the QIIME software package version 1.8.0 [29,30]. For genus-level analysis, the sequences were assigned to operational taxonomic units (OTUs) by open-reference OTU picking [31] with a 97% pairwise identity threshold and the Greengenes reference database [32].

## The frequencies of *LCT* variants in different populations around the world

The frequencies of *LCT* variants in Japanese, African, American, East Asian, European, and South Asian populations were obtained from the Phase 3 1000 Genomes Project data (<http://phase3browser.1000genomes.org/index.html>) [17].

## Assessment of dairy product intake with a BDHQ

A BDHQ [33] was used to assess food intake habits during the month before sample collection.

To estimate the amount of lactose contained in an individual's diet, we used combined data with the categories "normal-fat milk" and "low-fat milk" (designated here as dairy products), after energy adjustment using the density method.

## Statistical analysis

All analyses were performed using the IBM SPSS Statistics, version 22.0, statistical software package (IBM Corp., Armonk, NY, USA). Intergroup differences in the composition of

*Bifidobacterium* were analysed using the Kruskal-Wallis test. Spearman's correlation coefficient was used to determine the relationship of *Bifidobacterium* abundance with the amount of dairy product intake. For all tests,  $p < 0.05$  was considered statistically significant.

## Supporting information

### S1 Fig. Flow chart of this study.

(TIF)

### S1 Table. Amount of dairy product consumption and *Bifidobacterium* abundance in the gut.

(DOCX)

## Acknowledgments

We would like to thank Dr. Chyn Boon Wong for her critical review of this manuscript.

## Author Contributions

**Conceptualization:** Toshitaka Odamaki.

**Data curation:** Kumiko Kato, Sachiko Ishida, Masami Tanaka.

**Formal analysis:** Kumiko Kato, Sachiko Ishida, Masami Tanaka.

**Investigation:** Kumiko Kato, Eri Mitsuyama.

**Project administration:** Toshitaka Odamaki.

**Supervision:** Jin-zhong Xiao, Toshitaka Odamaki.

**Writing – original draft:** Kumiko Kato.

**Writing – review & editing:** Kumiko Kato, Sachiko Ishida, Jin-zhong Xiao, Toshitaka Odamaki.

## References

1. Schwab K, editor. The Global Competitiveness Report 2017–2018. World Economic Forum. 2017.
2. Ikeda N, Saito E, Kondo N, Inoue M, Ikeda S, Satoh T, et al. What has made the population of Japan healthy? Lancet. Elsevier Ltd; 2011; 378: 1094–1105. [https://doi.org/10.1016/S0140-6736\(11\)61055-6](https://doi.org/10.1016/S0140-6736(11)61055-6)
3. Nishijima S, Suda W, Oshima K, Kim S-W, Hirose Y, Morita H, et al. The gut microbiome of healthy Japanese and its microbial and functional uniqueness. DNA Res. 2016; dsw002. <https://doi.org/10.1093/dnares/dsw002> PMID: 26951067
4. Odamaki T, Kato K, Sugahara H, Hashikura N, Takahashi S, Xiao J, et al. Age-related changes in gut microbiota composition from newborn to centenarian: a cross-sectional study. BMC Microbiol. BioMed Central; 2016; 16: 90. <https://doi.org/10.1186/s12866-016-0708-5> PMID: 27220822
5. Hosomi K, Ohno H, Murakami H, Natsume-Kitatani Y, Tanisawa K, Hirata S, et al. Method for preparing DNA from feces in guanidine thiocyanate solution affects 16S rRNA-based profiling of human microbiota diversity. Sci Rep. 2017; 7: 4339. <https://doi.org/10.1038/s41598-017-04511-0> PMID: 28659635
6. Di Gioia D, Aloisio I, Mazzola G, Biavati B. Bifidobacteria: their impact on gut microbiota composition and their applications as probiotics in infants. Appl Microbiol Biotechnol. Springer Berlin Heidelberg; 2014; 98: 563–577. <https://doi.org/10.1007/s00253-013-5405-9> PMID: 24287935
7. Logan AC, Katzman M. Major depressive disorder: Probiotics may be an adjuvant therapy. Med Hypotheses. 2005; 64: 533–538. <https://doi.org/10.1016/j.mehy.2004.08.019> PMID: 15617861
8. Scott KP, Gratz SW, Sheridan PO, Flint HJ, Duncan SH. The influence of diet on the gut microbiota. Pharmacol Res. Elsevier Ltd; 2013; 69: 52–60. <https://doi.org/10.1016/j.phrs.2012.10.020> PMID: 23147033



9. Wacklin P, Mäkiyuokko H, Alakulppi N, Nikkilä J, Tenkanen H, Räsänen J, et al. Secretor Genotype (FUT2 gene) Is Strongly Associated with the Composition of Bifidobacteria in the Human Intestine. *Otto M, editor. PLoS One. Public Library of Science*; 2011; 6: e20113. <https://doi.org/10.1371/journal.pone.0020113> PMID: 21625510
10. Blekhnman R, Goodrich JK, Huang K, Sun Q, Bukowski R, Bell JT, et al. Host genetic variation impacts microbiome composition across human body sites. *Genome Biol. Genome Biology*; 2015; 1–12. <https://doi.org/10.1186/s13059-014-0572-2>
11. Mattar R, Mazo Carrilho. Lactose intolerance: diagnosis, genetic, and clinical factors. *Clin Exp Gastroenterol.* 2012; 113. <https://doi.org/10.2147/CEG.S32368>
12. Itan Y, Jones BL, Ingram CJ, Swallow DM, Thomas MG. A worldwide correlation of lactase persistence phenotype and genotypes. *BMC Evol Biol.* 2010; 10. <https://doi.org/10.1186/1471-2148-10-36> PMID: 20144208
13. Liebert A, López S, Jones BL, Montalva N, Gerbault P, Lau W, et al. World-wide distributions of lactase persistence alleles and the complex effects of recombination and selection. *Hum Genet. Springer Berlin Heidelberg*; 2017; 136: 1445–1453. <https://doi.org/10.1007/s00439-017-1847-y> PMID: 29063188
14. Goodrich JK, Davenport ER, Beaumont M, Bell JT, Clark AG, Ley RE. Genetic Determinants of the Gut Microbiome in UK Twins Correspondence. *Cell Host Microbe. Elsevier Inc.*; 2016; 19: 731–743. <https://doi.org/10.1016/j.chom.2016.04.017> PMID: 27173935
15. Bonder MJ, Kurilshikov A, Tigchelaar EF, Mujagic Z, Imhann F, Vila AV, et al. The effect of host genetics on the gut microbiome. *Nat Genet. Nature Publishing Group*; 2016; 1–9. <https://doi.org/10.1038/ng.3663> PMID: 27694959
16. Rothschild D, Weissbrod O, Barkan E, Korem T, Zeevi D, Costea PI, et al. Environmental factors dominate over host genetics in shaping human gut microbiota composition. *Nature. Nature Publishing Group*; 2018; 555: 150540. <https://doi.org/10.1101/150540>
17. Auton A, Abecasis GR, Altshuler DM, Durbin RM, Bentley DR, Chakravarti A, et al. A global reference for human genetic variation. *Nature.* 2015; 526: 68–74. <https://doi.org/10.1038/nature15393> PMID: 26432245
18. Mackenzie BW, Waite DW, Taylor MW. Evaluating variation in human gut microbiota profiles due to DNA extraction method and inter-subject differences. *Front Microbiol.* 2015; 6: 1–11. <https://doi.org/10.3389/fmicb.2015.00001>
19. Jovel J, Patterson J, Wang W, Hotte N, O’Keefe S, Mitchel T, et al. Characterization of the gut microbiome using 16S or shotgun metagenomics. *Front Microbiol.* 2016; 7: 1–17. <https://doi.org/10.3389/fmicb.2016.00001>
20. Tandon D, Haque MM, Saravanan R, Shaikh S, Sriram P, Dubey AK, et al. A snapshot of gut microbiota of an adult urban population from Western region of India. *PLoS One.* 2018; 13: 1–20. <https://doi.org/10.1371/journal.pone.0195643> PMID: 29624599
21. Li J, Jia H, Cai X, Zhong H, Feng Q, Sunagawa S, et al. An integrated catalog of reference genes in the human gut microbiome. *Nat Biotechnol.* 2014; 32: 834–841. <https://doi.org/10.1038/nbt.2942> PMID: 24997786
22. Salamon D, Sroka-Oleksiak A, Kapusta P, Szopa M, Mrozińska S, Ludwig-Słomczyńska AH, et al. Characteristics of the gut microbiota in adult patients with type 1 and 2 diabetes based on the analysis of a fragment of 16S rRNA gene using next-generation sequencing. *Polish Arch Intern Med.* 2018; <https://doi.org/10.20452/pamw.4246> PMID: 29657308
23. de Vrese M, Stegelmann a, Richter B, Fenselau S, Laue C, Schrezenmeier J. Probiotics—compensation for lactase insufficiency. *Am J Clin Nutr.* 2001; 73: 421S–429S. <https://doi.org/10.1093/ajcn/73.2.421s> PMID: 11157352
24. Hisada T, Endoh K, Kuriki K. Inter- and intra-individual variations in seasonal and daily stabilities of the human gut microbiota in Japanese. *Arch Microbiol. Springer Berlin Heidelberg*; 2015; 197: 919–934. <https://doi.org/10.1007/s00203-015-1125-0> PMID: 26068535
25. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, et al. PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *Am J Hum Genet.* 2007; 81: 559–575. <https://doi.org/10.1086/519795> PMID: 17701901
26. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods.* 2012; 9: 357–359. <https://doi.org/10.1038/nmeth.1923> PMID: 22388286
27. Aronesty E. Comparison of Sequencing Utility Programs. *Open Bioinforma J.* 2013; 7: 1–8. <https://doi.org/10.2174/1875036201307010001>
28. Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R. UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics. Oxford University Press*; 2011; 27: 2194–2200. <https://doi.org/10.1093/bioinformatics/btr381> PMID: 21700674

29. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, et al. correspondence QIIME allows analysis of high-throughput community sequencing data Intensity normalization improves color calling in SOLiD sequencing. *Nat Publ Gr. Nature Publishing Group*; 2010; 7: 335–336. <https://doi.org/10.1038/nmeth0510-335>
30. Kuczynski J, Stombaugh J, Walters WA, González A, Caporaso JG, Knight R. Using QIIME to Analyze 16S rRNA Gene Sequences from Microbial Communities. *Current Protocols in Bioinformatics*. Hoboken, NJ, USA: John Wiley & Sons, Inc.; 2011. p. 10.7.1–10.7.20. <https://doi.org/10.1002/0471250953.bi1007s36> PMID: 22161565
31. Rideout JR, He Y, Navas-Molina JA, Walters WA, Ursell LK, Gibbons SM, et al. Subsampled open-reference clustering creates consistent, comprehensive OTU definitions and scales to billions of sequences. *PeerJ*. 2014; 2: e545. <https://doi.org/10.7717/peerj.545> PMID: 25177538
32. McDonald D, Price MN, Goodrich J, Nawrocki EP, DeSantis TZ, Probst A, et al. An improved GreenGenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. *ISME J*. 2011; 6: 610–618. <https://doi.org/10.1038/ismej.2011.139> PMID: 22134646
33. Kobayashi S, Murakami K, Sasaki S, Okubo H, Hirota N, Notsu A, 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*. 2011; 14: 1200–1211. <https://doi.org/10.1017/S1368980011000504> PMID: 21477414