Journal of Microbiological Methods 165 (2019) 105718

Contents lists available at ScienceDirect

Journal of Microbiological Methods

journal homepage: www.elsevier.com/locate/jmicmeth

Method for absolute quantification of microbial communities by using both microarrays and competitive PCR

Ayaka Yazawa[a], Shigeki Kamitani[a], Naoyuki Togawa[b,*,1]

[a] College of Health and Human Sciences, Osaka Prefecture University, 3-7-30 Habikino, Habikino-City, Osaka 583-8555, Japan
[b] Bio-Device Group, Tsurumi R&D Center, Mitsubishi Chemical Co., Ltd, Yokohama-City, Japan

ARTICLE INFO

Keywords:
Microarray
Competitive PCR
Absolute quantification
Hybridization affinity ratio

ABSTRACT

Methods for the robust quantification of bacterial communities are still under development. In this context, the present study aimed to evaluate a method combining competitive PCR (cPCR) and microarray assays for the determination of absolute content of total bacteria and individual bacterial species in samples. For this, a competitor DNA for cPCR and microarrays containing three types of DNA probes was prepared. A calibration curve was generated with genomic DNA samples as standards, which was then utilized for cPCR-based determination of the total number (in moles) of 16S rRNA genes in other bacterial samples. Moreover, scatter plots of species-specific probes versus total bacteria probe for each genomic DNA of known concentration was fit to the regression model, and the obtained slope value was defined as the hybridization affinity ratio.

The cPCR assay was performed for both a commercially available mixed genomic DNA sample and human oral bacterial DNA samples, and the total number of moles of 16S rRNA genes was determined. These values were distributed among each species on the basis of the signal intensities of species-specific probes and the hybridization affinity ratio. The total number of bacterial genomes and those of individual species were determined by dividing the copy number of 16S rRNA genes per genome. The obtained results were confirmed by quantitative real-time PCR (qPCR). For values of $> 1 \times 10^2$ copies determined by qPCR, the ratio of the values measured by DNA chips to by qPCR was 1.53-fold on average and < 2.6-fold for all data. These results show that the combined method of cPCR and microarray is useful to quantify the absolute numbers of several types of bacteria in a sample at one time.

1. Introduction

Owing to developments in next-generation sequencing (NGS) techniques, the Human Microbiome Project has progressed (Arumugam et al., 2011) and this has led to the discovery of new groups of bacteria. The bacteria belonging to the first group are found in specific parts of the body, such as the intestine, oral cavity, and skin, while the bacteria belonging to the second group appear to be associated with intestinal diseases and periodontal disease progression (Teles et al., 2006; Miyake et al., 2015; Tsuruya et al., 2016; Kageyama et al., 2017). These studies are beginning to contribute toward the diagnosis and prevention of diseases, and NGS is expected to be used for bacterial testing; however, the relative abundance data estimated by using NGS will not be enough for accomplishing this purpose.

Recently, the American Type Culture Collection (Manassas, VA, USA) launched the sale of bacterial genomic DNA mixtures and the mixture of strains as a reference material for comparison of quantification results. Although many studies have reported the relative abundance of bacteria by using NGS (D'Amore et al., 2016; Fouhy et al., 2016), several research groups have tried to estimate the total number of bacteria and the number of each species in samples by combining them with spike-in bacteria or standards; however, the differences between the values measured by using NGS and quantitative real-time PCR (qPCR) have been > 10-fold (Stämmler et al., 2016; Tourlousse et al., 2017).

Even though NGS is a powerful tool for comprehensive sequence analysis, when 10–100 bacterial species whose complete genome sequences have been published are analyzed to detect the types of bacteria associated with a disease, microarray analysis is more suitable because it is cost-effective. Many researchers have measured bacterial populations with high-throughput methods by using microarrays, but

Abbreviations: cPCR, competitive polymerase chain reaction; NGS, next-generation sequencing; qPCR, quantitative real-time PCR

* Corresponding author.
E-mail address: togawa.naoyuki.ms@m-chemical.co.jp (N. Togawa).
[1] Present address: 10-1, Daikokucho, Tsurumi-ku, Yokohama, Kanagawa 230-0053, Japan.

https://doi.org/10.1016/j.mimet.2019.105718
Received 23 May 2019; Received in revised form 7 September 2019; Accepted 7 September 2019
Available online 10 September 2019
0167-7012/

only qualitative or relative quantitative results have been obtained to date (Roth et al., 2004; Palmer et al., 2006). Specific problems are still associated with NGS- and microarray-based quantification methods.

Competitive PCR (cPCR) is one of the most effective methods to overcome the issues with quantitative analysis. (Zentilin and Giacca, 2007). Mangin et al. (2006) performed cPCR by adding serially diluted solutions of competitor DNA to a constant amount of sample DNA. To quantify the number of Bifidobacterium 16S rRNA genes in a sample, they constructed a calibration curve based on the fluorescence intensity values from electrophoresis bands of cPCR products. They further calculated the total number of Bifidobacterium genomes by using the 16S rRNA gene copy number per genome, but could not quantify the number of each species according to the relative value obtained from the fluorescence intensities of the polymerase chain reaction temperature gradient gel electrophoresis (PCR-TTGE) bands.

Treimo et al. (2006) also conducted cPCR to quantify the total number of bacteria and number of each species by using microarrays equipped with species-specific probes. Additionally, they attempted to quantify the number of each species based on species composition determined from the detection results of 16S rRNA genes. Their microarrays contained the following three different types of probes: the first probe hybridized competitor DNA in cPCR products, the second probe consisted of a 16S rRNA gene conserved sequence region that hybridized all 16S rRNA amplification products obtained through cPCR, and the third probe hybridized species-specific 16S rRNA region.

Treimo et al. hybridized the cPCR products to this microarray and calculated the total amount of bacterial DNA based on the ratio of the fluorescence intensity between the cPCR competitor and 16S rRNA gene probes and the dilution factor of the competitor. Moreover, the DNA composition ratio of each bacterium was calculated from the fluorescence intensity ratios of the probes for 16S rRNA genes and those for species-specific 16S rRNA. They confused the total amount of bacterial DNA with the total number of 16S rRNA genes, and the absolute number of each bacterium was not determined by their method.

The present study aimed to propose a method that combine cPCR and microarray technology to accurately determine the total number of bacteria and number of each species in samples.

In this study, cPCR along with a microarray comprising the abovementioned three different types of probes were used; however, unlike previous studies, calibration curves were constructed based on cPCR and microarray analyses, to quantify the total number of moles of 16S rRNA genes before measuring each sample. The hybridization affinity ratios between species-specific and total bacteria probe were also determined before measuring each sample because the 16S rRNA gene amplification product could hybridize the probes for both 16S rRNA genes and species-specific 16S rRNA.

In our method, the equivalent values of total bacteria probe for each species-specific probe were calculated according to the hybridization affinity ratios determined prior to measurement and fluorescence intensities of the species-specific probes obtained during measurement. The composition ratio of each bacterium was determined by calculating the ratio between equivalent value and signal intensity derived from fluorescence intensity of total bacteria probe. The number of genomes and the weight of the bacterial DNA were then calculated by using the genome size and number of 16S rRNA gene copies per genome. To determine the accuracy of our method, quantitative evaluations were performed by using toothbrush-derived samples containing 12 bacterial species closely related to periodontal disease as indicators. In addition, the results of the cPCR and DNA chip analyses were verified by using qPCR.

2. Materials and methods

2.1. DNA chip design

In the present study, a three-dimensional gel-type Genopal™ DNA chip ORAL (Mitsubishi Chemical Co., Ltd. Tokyo, Japan) was used for microarray testing. Three different types of probes were spotted on this DNA chip: One used to measure the total number of bacteria, a second one used as competitor DNA, and a third one useful to quantify individual species (species-specific) in a sample (Fig. S1).

To design probes from target sequences, including 16S ribosomal subunit markers, about 20,000 DNA sequences were collected from both the 16S Ribosomal RNA Project (https://www.ncbi.nlm.nih.gov/refseq/targetedloci/) and the PopSet database (https://www.ncbi.nlm.nih.gov/popset/).

Twelve species-specific probes were designed at the V3 site of the 16S rRNA gene by ProbeQuest™ (https://www.dynacom.co.jp/english/product/). The “total bacteria” probe for detecting 16S rRNA genes in almost all bacteria (Wang and Qian, 2009) and the “competitor DNA” probe for competitive DNA, shown in Fig. S1(b) and (c), respectively, were also designed by using this software. Sequence specificity was confirmed by ClustalX (Thompson et al., 1997) and BLAST (Altschul et al., 1990). The sequences of the species-specific probes, the competitor DNA probe, and the total bacteria probe are shown in Table S1; whereas the location of the probes on the DNA chip is shown in Fig. S2. As there are two major strains of *Prevotella intermedia* and *Aggregatibacter actinomycetemcomitans* and we were interested in identifying all of them, specific probes were designed for each one. For each species, equal amounts of the probes were mixed and immobilized on the chip, on spot No. 8 for *P. intermedia* and on spot No. 10 for *A. actinomycetemcomitans*. Our DNA chip platform data were deposited in the Gene Expression Omnibus database, hosted by the NCBI (www.ncbi.nlm.nih.gov/geo/, platform registry number GPL25612).

2.2. Genomic DNA for quantitative verification tests

To test the specificity of the probes, genomic DNA of 12 bacterial strains and genomic DNA mixtures were used (Table S2). All 12 genomic DNAs were purchased from ATCC (Manassas, VA, USA) and RIKEN BioResource Center (Tsukuba, Ibaraki, Japan), and two DNA mixture solutions (ATCC MSA-1002™ and MSA-1003™) were also purchased from ATCC. The concentration of the DNA solutions was determined by measuring absorbance at 260 nm with a Nanodrop™ 2000c (Thermo Fisher Scientific, San Jose, CA). Besides, one DNA mixture solution was manually prepared by mixing eight of the genomic DNA solutions. The final composition of the three different DNA mixtures is shown in Tables S3, S4, and S5.

DNA mixtures were diluted as follows: 9 serial 4-fold dilutions of the 20 Strain Even Mix Genomic Material (ATCC MSA-1002™), 11 serial 4-fold dilutions of the 20 Strain Staggered Mix Genomic Material (ATCC MSA-1003™), and 6 serial 16-fold dilutions of the genomic DNA mixture of eight strains. Both the genomic DNA mixture of eight strains and the 20 Strain Staggered Mix Genomic Material were used for the preparation of calibration curves. As a quantitative assessment sample, the 20 Strain Even Mix Genomic Material was used.

2.3. Oral bacterial samples

Five human dental plaque samples were collected from the erupted teeth of five different persons (ages: under 5 years, $n = 1$; 20–30 years, $n = 2$; over 60 years old, $n = 2$) by brushing for 1 min with sterile toothbrushes. The plaques attached to the brushes were removed by washing each of them several times in 5 mL sterile distilled water in 15 mL-test tubes and centrifuged at 1600*g* for 20 min. Supernatants were discarded and pellets stored at −20 °C until extraction of genomic DNA (Okada et al., 2000). Genomic DNA was extracted from samples using a Wizard Genome DNA Purification Kit (Promega, Madison, WI, USA), the DNA concentration in the final solutions was measured, and samples were stored at −20 °C until use.

This experiment was performed after obtaining the consent from the subjects in accordance with the Declaration of Helsinki and was

conducted with the approval of the Ethical Review Board of the Graduate School of Comprehensive Rehabilitation of the Osaka Prefecture University (Approval No.: 2015–311).

2.4. Primers and competitor DNA for competitive PCR (cPCR) assays

For cPCR assays, the forward primer, labeled with Cy5™ (5′-Cy5-TACGGGAGGCAGCAG-3′), was designed in the conserved region between the V2 and V3 regions of the 16S rRNA gene; whereas the reverse primer, containing an inosine base (5′-TACCIGGGTATCTAATCC-3′), was designed in the conserved region between the V4 and V5 regions of the 16S rRNA gene (Fig. S1(b)). These primers enable the amplification of about 90% of the 16S rRNA gene (Wang and Qian, 2009). For the construction of the competitor DNA, an artificial sequence of a size close to that of the amplification product that would be amplified by the already mentioned primers was designed not to hybridize with bacterial 16S rRNA genes according to a BLAST search. The forward primer sequence and the complementary sequence of the reverse primer were attached to the corresponding end of the artificial sequence (Fig. S1(c)). The sequence of the competitor DNA is shown in Table S6. The competitor DNA fragment was cloned into the *SmaI* site of a pUC19 vector (2686 bp), checked by DNA sequencing at Fasmac (Kanagawa, Japan), and used for cPCR assays.

2.5. Competitive PCR assays

As shown in Fig. S3(a), each sample was added to a fixed amount of competitor DNA and a cPCR reaction was performed. cPCR was optimized with 0.5 amol (attomol, 10^{-18} mol) of competitor DNA per reaction mixture. DNA fragments were amplified by using TaKaRa Premix Ex Taq® HotStart version (TaKaRa Bio Inc., Otsu, Japan) in a 20 μL reaction mixture consisting of 10 μL of 2× Premix solution, 1 μL of each primer 20 μM, 2 μL of competitor DNA solution 0.25 amol μL^{-1}, 5 or 1 μL of the sample solution, and 1 or 5 μL of water. cPCR was performed under the following amplification conditions: a 60 s initial denaturation step at 94 °C followed by 40 cycles at 98 °C for 10 s, 55 °C for 30 s, and 72 °C for 20 s by using a ProFlex™ PCR System (Thermo Fisher Scientific, San Jose, CA). After completion of the reaction, cPCR mixtures were cooled to 4 °C and stored at −30 °C until use.

2.6. Hybridization and detection of signals

Following the reaction, each cPCR solution was mixed with 48 μL of Tris/HCl (pH 7.5) 1 M, 48 μL of NaCl 1 M, 20 μL of a Tween® 20 solution 0.5%, and 64 μL of water for a total volume of 200 μL of hybridization solution. Hybridization and washing were performed by using an Automatic hybridization and washing instrument (Mitsubishi Chemical Co., Ltd. Tokyo, Japan; https://www.m-chemical.co.jp/genome/products/auto_hybrid.html) for 16 h. After treatment, DNA chips were further washed with wash solution (0.24 M Tris/HCl pH 7.5 and 0.24 M NaCl) at about 25 °C for 5 min.

Subsequently, hybridization signals were examined by using a Genopal Reader™ (Mitsubishi Chemical Co., Ltd. Tokyo, Japan). Fluorescence intensities were detected at different exposure times: 0.1, 1, 4, and 40 s. Those spot images with the longest exposure times that were not saturated (showed no pixel saturation) were selected and quantified to obtain the fluorescence intensity per second.

Next, the probe signal intensity was defined by the fluorescence intensity. The mean and the standard deviation of the fluorescence intensity of the background spot without the probe were calculated, and the sum of the mean and three times the standard deviation was defined as the background signal intensity. Probe signal intensity was calculated by subtracting the background signal intensity from the median of the fluorescence intensity of the quintuple spots for each probe. DNA chip data were deposited in the Gene Expression Omnibus database, hosted by the NCBI (www.ncbi.nlm.nih.gov/geo/, data series number GSE125085).

2.7. Specificity of species-specific probes and affinity ratios to total bacteria probe

The concentrations of the 16S rRNA genes of the 12 strains of genomic DNA (Table S2) were adjusted to 1, 0.1, 0.01, and 0.001 amol μL^{-1}. These solutions were used to test the probes in duplicates. First, cPCR was carried out with 1 μL of each solution as template DNA as described in Section 2.5. The cPCR products were hybridized to the DNA chips, the signals were read, and the probe signal intensity was calculated. Probe specificity was verified through the obtained probe signal intensities.

The signal intensities of the species-specific probes were plotted against the signal intensities of the total bacteria probe (Fig. S1(a), lower). The slope and the slope coefficient were determined by the least-squares method passing through the origin. These slope values were defined as the hybridization affinity ratios and were used for the calibration of each species DNA content.

2.8. Preparation of calibration curves by cPCR and method for quantification

Scalar amounts of genomic DNA mixtures, 9 serial 4-fold dilutions of the 20 Strain Even Mix Genomic Material (ATCC MSA-1002™), 11 serial 4-fold dilutions of the 20 Strain Staggered Mix Genomic Material (ATCC MSA-1003™), and 6 serial 16-fold dilutions of the genomic DNA mixture of eight strains, were added to a fixed amount (0.5 amol) of competitor DNA. cPCR was performed either in duplicates or triplicates with 5 μL of each of the solutions as template solution (as described in Section 2.5.). cPCR products were then hybridized to the DNA chip and detected to calculate their signal intensities.

Before quantifying our samples, calibration curves were prepared by using the results from the cPCR assays done with the 20 Strain Staggered Mix Genomic Material and the genomic DNA mixture of eight strains. To confirm whether cPCR assays were well performed, the probe signal intensities corresponding to the competitor probe (C) and the total bacteria probe (T) were plotted against the log-transformed ratio of the total number of moles of 16S rRNA genes in the initial sample to the number of moles of competitor DNA in the initial sample. Then, to prepare the calibration curves, the log-transformed T/C ratios were also plotted against the log-transformed ratio of the total number of moles of 16S rRNA genes in the initial sample to the number of moles of competitor DNA in the initial sample (Fig. S1(a), upper). Two types of calibration curves by linear regression model or cubic regression model were prepared and the corresponding linear or cubic regression equations were created. The number of each species present in the initial samples were calculated based on these regression equations and evaluated by comparing them with the expected values.

The procedure for the quantification of unknown composition samples was as follows: First, using 1 or 5 μL of sample solution, cPCR was performed under the same conditions as those used for the construction of the calibration curves and the results of the DNA chip detection and the probe signal intensities were obtained. Then, the log-transformed T/C values were applied to the regression equations and 10^x values were calculated (Fig. S3(b)).

Then, the calculated value was multiplied by 0.5 (i.e., 0.5 corresponded to the number of attomoles of the competitor DNA) to obtain the total number of moles of 16S rRNA genes in the initial sample solution. For example, if $\log_{10}$ (T/C) = 1 and the linear regression equation was $y = 0.6x - 0.14$, x would be 1.9 and then the estimated total number of moles of 16S rRNA genes in the initial sample solution would be $0.5 \times 10^{1.9}$ amol 20 μL^{-1} (39.7×10^{-18} mol 20 μL^{-1}).

In addition, the number of moles of 16S rRNA genes for each species was calculated as follows (Fig. S3(c)): First, the signal intensities of species-specific probes were divided by each hybridization affinity

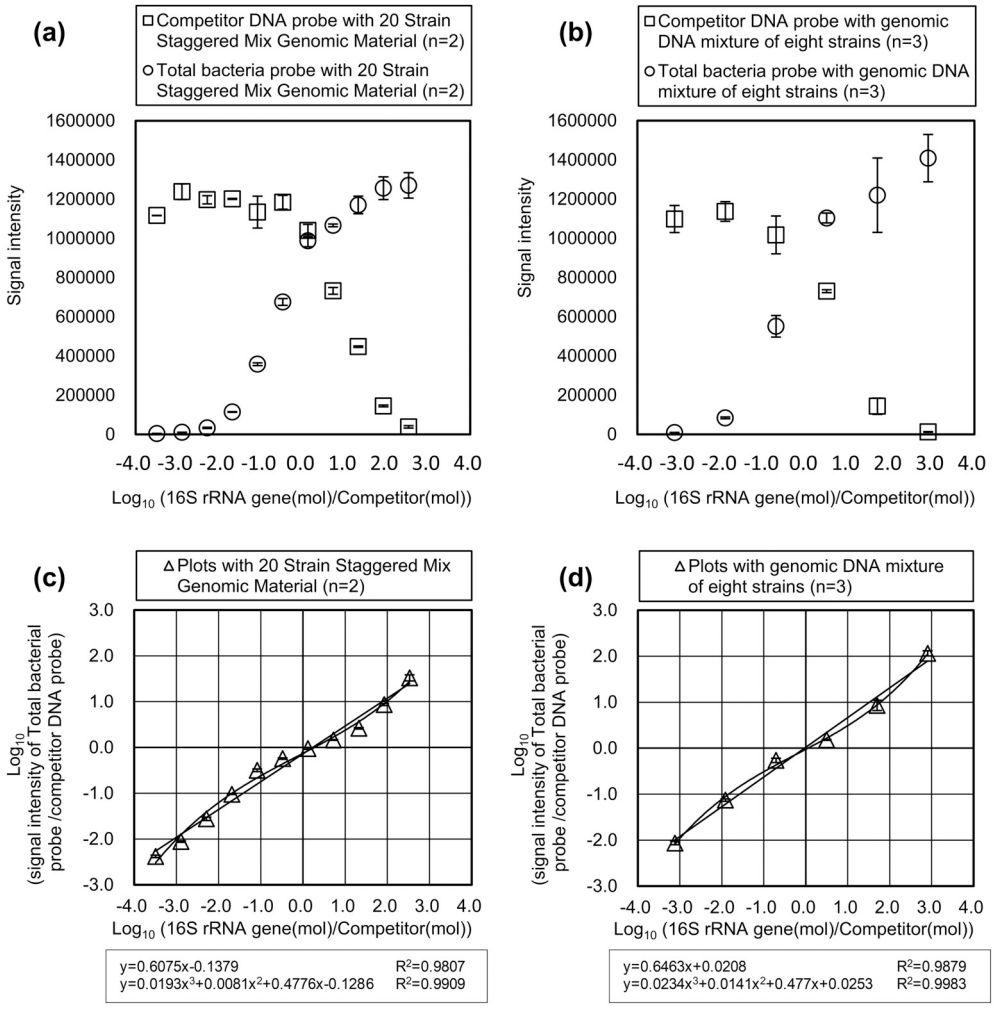

Fig. 1. Detection of mixed genomic DNA dilution series solutions of known concentration and creation of regression equation from the calibration curve.

Panels (a) and (b) represent scatter plots of signal intensities corresponding to the competitor probe (C: open squares) and the total bacteria probe (T: open circles) against the log-transformed molar ratio of total 16S rRNA genes to that of the competitor DNA in the initial samples. Data shown in panels (a) and (b) were obtained from cPCR by using 11 serial 4-fold dilutions of the 20 Strain Staggered Mix Genomic Material and 6 serial 16-fold dilutions of the genomic DNA mixture of eight strains, respectively. The x-axis of these scatter plots represents the log-transformed molar ratio of total 16S rRNA genes to that of competitor DNA in the initial samples. The range of x was −3.49–2.53 ($\log_{10}(0.00016/0.5)$–$\log_{10}(168.5/0.5)$) and −3.12–2.90 ($\log_{10}(0.00038/0.5)$–$\log_{10}(400/0.5)$), in the 20 Strain Staggered Mix Genomic Material, and the genomic DNA mixture of eight strains, respectively.

Subsequently, calibration curves were prepared by plotting the log-transformed T/C ratios against the log-transformed molar ratio of total 16S rRNA genes in the initial sample to that of competitor DNA in the initial sample (as shown in panels (c) and (d)). Two types of regression model were applied to the plotted data and linear or cubic regression equations were obtained. These equations and the corresponding regression coefficients obtained are represented below panels (c) and (d). The regression coefficient from using the linear regression model was > 0.98; whereas that from the cubic regression model was > 0.99. Error bars indicate standard deviation ($n = 2$ for panels (a) and (c), and $n = 3$ for panels (b) and (d)).

ratio, which had been determined prior to measurements, and converted to the equivalent of the signals of the total bacteria probe. Second, the composition ratios were calculated by the equivalent of each species divided by the signal intensity of the total bacteria probe. Finally, the composition ratios were multiplied by the total number of moles of 16S rRNA genes.

The number of genomes of each species per microliter in a sample was calculated by multiplying the number of moles of 16S rRNA genes for each species by Avogadro's number and then dividing by the 16S rRNA gene copy number per genome and the sample volume used for cPCR. The 16S rRNA gene copy numbers of each species are given in Table S2. To calculate the total number of genome copies, we used the average value of the 16S rRNA gene copy number: 4.8.

The expected values for the 16S rRNA gene and genome copy numbers were calculated; they are shown in Tables S3, S4, and S5. The expected values in the dilution series were also obtained by further dividing the calculation results. Using the result from the cPCR assays done with the 20 Strain Even Mix Genomic Material, the number of genome copies was calculated.

2.9. Measurement of the bacterial content of oral DNA samples

By using 1 μL of sample solution, a cPCR assay was performed under the same conditions as for calibration curves preparation. The amplified fragments were hybridized to the DNA chip and resulting fluorescence was measured. By using the cubic equation of the calibration curve prepared with the dilution series of the genomic DNA mixture of eight strains, quantitative results for 12 strains of bacteria and for total bacterial content on the DNA chip were estimated.

2.10. Validation by quantitative real-time PCR (qPCR)

To confirm the results from the DNA chips, an absolute quantitation was carried out by qPCR using an Applied Biosystems 7500 System (Applied Biosystems, Foster City, CA, USA). Absolute numbers of genome copies for *Porphyromonas gingivalis*, *Tannerella forsythia*, *Treponema denticola*, and the total bacteria were determined.

All specific primer sets and TaqMan™ probes were purchased from Thermo Fisher Scientific (San Jose, CA, USA) (shown in Table S7; Lyons et al., 2000; Suzuki et al., 2005). According to the manufacturer's recommended protocols, 2× TaqMan™ Fast Advanced Master Mix, primers, TaqMan™ probes, and 1 μL of the sample template solution were mixed to produce a 20 μL reaction mixture. qPCR amplification conditions included 50 cycles at 98 °C for 15 s and 60 °C for 60 s when measuring the number of gene copies for total bacteria, and 50 cycles at 98 °C for 3 s and 60 °C for 15 s when measuring the number of gene copies for each species. Calibration curves were prepared by 5 serial 16-fold dilutions of the genomic DNA mixture of eight strains. The quantification range of genome copies for *P. gingivalis*, *T. forsythia* and *T. denticola*, and total bacteria were 29–1,881,250, 11–752,500, 174–11,412,917 copies, respectively. The number of genome copies for total bacteria and each species was determined by dividing the number

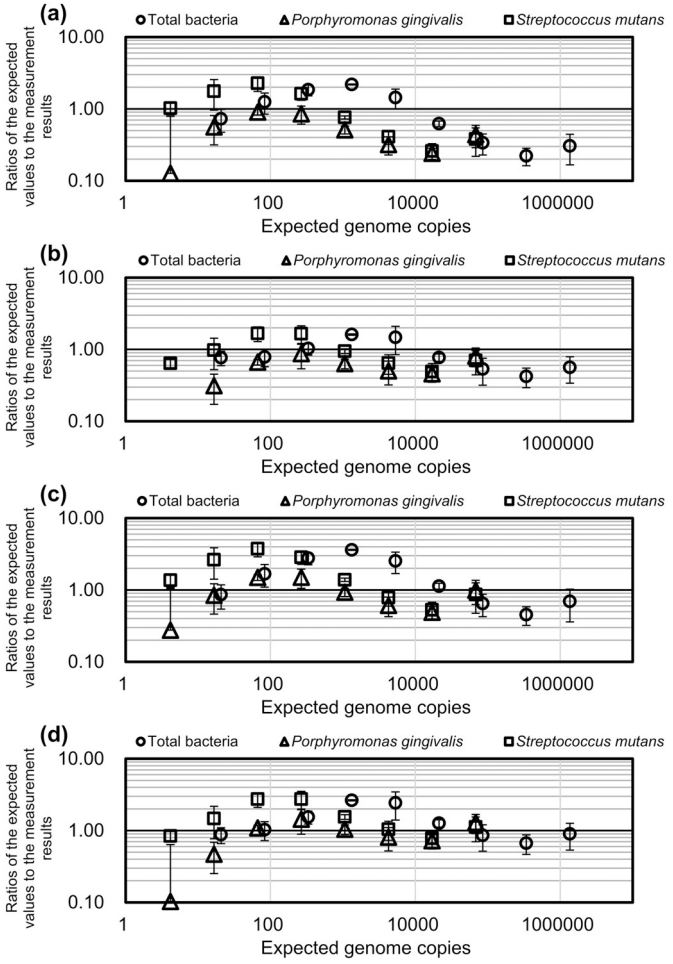

Fig. 2. Quantitative evaluation of results utilizing four different regression equations.

Nine serial 4-fold dilutions of the 20 Strain Even Mix Genomic Material were quantified in triplicates by using the method reported in the present study. Four different regression equations were used for the quantitation. For panels (a) and (b), data obtained from the genomic DNA mixture of eight strains were analyzed by linear and cubic regression equations, respectively. For panels (c) and (d), data obtained from the 20 Strain Staggered Mix Genomic Material were analyzed by linear and cubic regression equations, respectively.

The ratios of obtained quantification (y axis) to expected values (x-axis) are shown. In all four panels, open circles indicate quantitation of total bacteria; open triangles indicate that of *P. gingivalis*, and open squares that of *S. mutans*. The range of the number of expected genome copies of total bacteria was 21–1,361,150 copies, and that of *P. gingivalis* and *S. mutans* was 1–68,057 copies. Error bars represent associated standard deviations ($n = 3$).

of 16S rRNA gene copies obtained from the calibration curve by the 16S rRNA gene copy number per genome (Table S2).

Additionally, the total weight of the bacterial DNA in the sample DNA was estimated based on the total number of bacterial genome copies and the mean genome size of oral bacteria. The mean genome size, 2,426,000 bp, was previously calculated from 136 samples of subgingival and supragingival plaque forming bacteria (Nayfach and Pollard, 2015).

3. Results

3.1. Verification of probe specificity and determination of probe affinity ratio

Probe specificity was evaluated by hybridizing 12 different genomic DNAs. Resulting signal intensities are shown in Tables S8, S9, S10, and S11. The signal intensity other than that of the hybridized target probe was defined as a cross-hybridization signal intensity. The ratio of this cross-hybridization signal intensity to that of each hybridized target probe was calculated. The maximum ratio value obtained among the 12 different genomic DNAs tested was 0.163%.

Scatter plots of the signal intensity of the species-specific probes against the signal intensity of the total bacteria probe are presented in Fig. S4. The signal intensities increased proportionally to the concentration range tested. The slope values were calculated by a linear regression through the origin. The values of these slopes were determined as the hybridization affinity ratios of species-specific probes to the total bacteria probe and used for quantitation of each species. These affinity ratios are shown in Table S12.

3.2. Competitive PCR results and calibration curve creation

The concentration of the 20 Strain Even Mix Genomic Material was 5.0 ng μL^{-1}, that of the 20 Strain Staggered Mix Genomic Material was 14.1 ng μL^{-1}, and that of the genomic DNA mixture of eight strains was 29.9 ng μL^{-1}. Subsequently, the genome copy number of each bacterial species and the concentration of the 16S rRNA genes in the three genomic DNA mixtures were calculated. The concentration of the 16S rRNA genes were 11.87 amol μL^{-1}, 33.7 amol μL^{-1}, and 80.0 amol μL^{-1}, respectively (Tables S3–S5 at the lower right). According to the values calculated about the 20 Strain Staggered Mix Genomic Material and the genomic DNA mixture of eight strains, the range expected from the total bacterial genome copy number and the total number of moles of 16S rRNA genes in the initial sample was of 18–19,298,825 copies (0.00016–168.5 amol) and of 54–57,064,585 copies (0.00038–400 amol), respectively.

The signal intensities determined by cPCR for each of the two mixtures prepared in the dilution series are shown in Fig. 1(a) and (b). The scatterplots displayed in Fig. 1(c) and (d) were created from these results. These scatterplots formed slightly S-shaped curves; thus, two types of calibration curves were created to fit them: one by a linear regression model and another by a cubic regression model. The plots constructed by using the two mixtures almost fitted the same curve. Moreover, the cubic regression splines almost entirely overlapped, although there were differences near the upper and lower limits of the calibration curve because of the different concentrations of the two mixtures. When the linear regression model was used, the slopes of the calibration lines were slightly different, again because of differences in the concentration range of the two mixtures.

3.3. Quantitative evaluation

As shown in Fig. 2, the 20 Strain Even Mix Genomic Material was measured and compared with its expected values. The total genome copy number expected was in the range of 2.1×10^1–1.36×10^6 copies μL^{-1} (Table S3). The smallest genome copy numbers per microliter obtained, i.e., the lowest limit of quantitation per microliter, for total bacteria, *P. gingivalis*, and *Streptococcus mutans* were 2.1×10^1 copies μL^{-1}, 1.7×10^1 copies μL^{-1}, and 4×10^0 copies μL^{-1}, respectively.

The ratio between the expected and the measured values was < 10-fold, both when using the linear regression and the cubic regression models. Regarding measurement values above the lowest limit of quantitation per microliter, the ratio range between the values measured by DNA chips and those expected by calculation were 0.90–4.50-fold (Fig. 2(a)) by the linear regression model calculated with the genomic DNA mixture of eight strains, 1.50–3.76-fold (Fig. 2(c)) by the linear regression model calculated with the 20 Strain Staggered Mix Genomic Material, 0.87–3.20-fold (Fig. 2(b)) by the cubic regression model calculated with the genomic DNA mixture of eight strains, and 1.28–2.74-fold (Fig. 2(d)) by the cubic regression model calculated with the 20 Strain Staggered Mix Genomic Material.

It is known that, in terms of genome copies, the 20 Strain Even Mix Genomic Material has a 5% of DNA from *P. gingivalis* and a 5% from *S.*

Table 1
Genome copies measured by DNA chips in this study.

Species or total bacteria	No.47	No.48	No.409	No.12–1	No.412
Total bacteria	159,215 ± 7789	1,444,713 ± 81,364	360,671 ± 6606	269,683 ± 24,176	239,576 ± 10,378
Porphyromonas gingivalis	860 ± 47	1438 ± 339	170 ± 151	13,703 ± 1395	1273 ± 116
Tannerella forsythia	470 ± 20	580 ± 150	10 ± 17[a]	N.D.[b]	322 ± 19
Treponema denticola	187 ± 14	809 ± 130	830 ± 37	5 ± 4[a]	257 ± 15
Campylobacter rectus	1113 ± 84	5858 ± 630	18,149 ± 394	63 ± 21	2331 ± 127
Fusobacterium nucleatum subsp. *nucleatum*	2428 ± 126	47,853 ± 3617	8886 ± 197	838 ± 80	9617 ± 213
Prevotella intermedia	N.D.[b]	N.D.[b]	N.D.[b]	N.D.[b]	1768 ± 124
Prevotella nigrescens	661 ± 37	13,695 ± 1322	1384 ± 20	N.D.[b]	652 ± 15
Aggregatibacter actinomycetemcomitans	N.D.[b]	66 ± 57	184 ± 5	607 ± 67	565 ± 16
Capnocytophaga gingivalis	1150 ± 82	1800 ± 301	711 ± 14	4154 ± 488	1301 ± 46
Streptococcus gordonii	6780 ± 252	38,160 ± 2197	5518 ± 129	22,272 ± 1843	11,170 ± 286
Streptococcus intermedius	486 ± 14	1628 ± 175	475 ± 9	N.D.[b]	198 ± 1
Streptococcus mutans	909 ± 50	2145 ± 309	N.D.[b]	N.D.[b]	1219 ± 34

[a] *Tannerella forsythia* in sample No.409 and *Treponema denticola* in sample No.12-1 were below the limit of quantitation.
[b] N.D. means not detected. The error range was indicated by the value of standard deviation (n = 3).

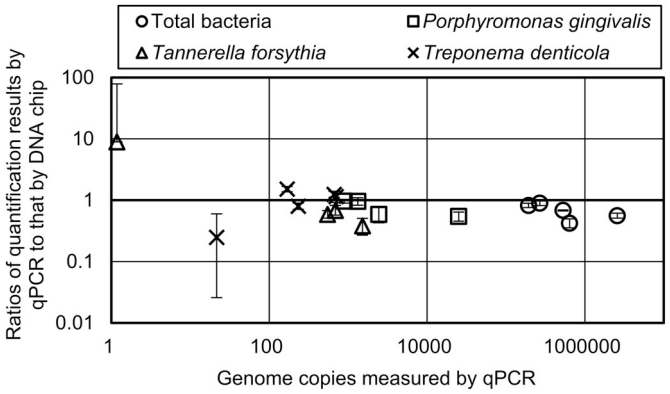

Fig. 3. Verification of DNA chip measurement results by using qPCR.
Five human dental plaque samples collected from teeth removed from five different persons (age: under five years, $n = 1$; 20–30 years, n = 2; over 60 years, n = 2) were quantified by using the DNA chip. For the same sample, absolute numbers of genome copies for *P. gingivalis* (open squares), *T. forsythia* (open triangles), *T. denticola* (crosses), and the total bacteria (open circles) were determined by qPCR.
The ratio (shown in the y-axis) for the values obtained by using qPCR to those determined by using the DNA chip was plotted against the number of copies measured by using qPCR (x-axis). Error bars represent standard deviation (n = 3). For values of $> 1.0 \times 10^2$ copies as determined by qPCR, the ratio was < 2.6-fold for the detected samples.

mutans. As shown in Fig. S5, the contents of *P. gingivalis* and *S. mutans* in each diluted sample were calculated based on total bacterial genome copies and of copies of each genome, giving average values of 3.6 ± 2.2% and 5.5 ± 0.7%, respectively.

3.4. Measurements of oral samples

The nucleic acid concentration found in the five oral samples, No. 47, No. 48, No. 409, No. 12–-1, and No. 412, were 3.6, 54.8, 3.9, 2.8, and 3.8 ng μL^{-1}, respectively. Among the 12 bacteria quantified with DNA chips, *S. gordonii* was present at relatively high levels in all five samples, as shown in Table 1.

The quantification results of three species known as the red complex (Teles et al., 2006) and the total number of bacteria were confirmed by qPCR. As shown in Fig. 3, for measurement values of $> 1.0 \times 10^2$ copies as determined by qPCR, the ratio between the values measured by DNA chips and by qPCR was < 2.6-fold.

The bacterial DNA content of the five oral samples, No. 47, No. 48, No. 409, No. 12–1, and No. 412, were estimated to be 12% (417 pg μL^{-1}), 7% (3784 pg μL^{-1}), 24% (945 pg μL^{-1}), 25% (706 pg μL^{-1}), and 17% (628 pg μL^{-1}).

4. Discussion

The results of the present study indicate that quantifying multiple bacteria by combining cPCR and microarray methods is as accurate as a qPCR. Treimo et al. (2006) also reported that a method combining cPCR and microarrays was useful. However, as they did not convert species-specific signal intensities to 16S rRNA gene equivalents, the relationship between the total number of 16S rRNA genes and the number of 16S rRNA genes belonging to each species was unclear. In contrast, the method we have proposed to calculate the number of each species was appropriate for this.

As shown in Fig. 1, when the total number of moles of 16S rRNA genes in the dilution series solution was almost equal to the number of moles of competitor DNA (0.5 amol), the signal intensities of the competitor DNA probe and the total bacteria probe were almost equal. When performing a typical cPCR method, a fixed amount of sample is mixed with a dilution series of the competitor, the reaction is performed, and the amplification products are quantified. Our results showed that even if a dilution series solution of reference DNA mixtures were added to a constant amount of competitor DNA, calibration curves could be created without any problem. The creation of calibration curves from the fluorescence intensity of DNA chip probes was also successful.

Intriguingly, even though the mixed genomic DNA solutions used to create the two calibration curves contained completely different bacterial species and composition ratios, the two calibration curves barely differed. These results indicate that the signal intensities of total bacteria probe depend only on the total number of moles of 16S rRNA genes.

To date, most researchers use linear regression models to generate calibration curves for cPCR, in the same way as Mangin et al. (2006) did. However, Suzuki and Giovannoni (1996) reported the simulation of a cPCR reaction model and noted that the relationship between the ratios among the amplification products and the ratios among the genetic materials present in the initial sample is sometimes not proportional, and that it depends on the concentration of the primers in the reaction mixture. Therefore, in the present research, cubic regression equations were created by a cubic regression model. An S-shaped plot was obtained for the calibration curve of cPCR and the approximation done by the cubic regression model showed to be a better fit than the linear regression model. The deviation from the expected values with the verification sample was smaller than the linear regression model, and the quantifiability was improved. It was considered useful to apply the cubic regression model to a calibration curve for cPCR.

According to Fig. 3, the average ratio between measured values and qPCR values was 1.53-fold; it was considered that accuracy was improved as compared with past reports on microarrays.

From the results shown in Figs. 2 and 3, it can be seen that our method had the same level of performance in terms of quantification when applied to reference samples and oral samples. The reproducibility of the performance achieved by our combined method was maintained even when the samples changed, demonstrating its reliability.

The proportion of bacterial DNA in the sample DNA was about 18% on average. Marotz et al. (2018) noted that oral samples contain significant amounts of human DNA and that their bacterial DNA content, as measured by NGS, was about 10%. Our own estimates are consistent with this report. Thus, the total amount of bacterial DNA in our samples was appropriately calculated.

Nevertheless, the limitations of our method are mainly due to information on the genome size and the number of 16S rRNA gene copies per genome. A similar sample to be examined must be examined in advance on a large scale with NGS to estimate the average genome size and 16S rRNA gene copy number per genome.

In recent years, bacterial genome sequencing data containing 16S rRNA gene sequences have been enriched with the development of NGS. At the same time, the information on the conserved regions of the 16S rRNA gene, bacterial species-specific sequences, and the 16S rRNA gene copy number per genome have been enriched. It is expected that more accurate quantification results around bacteria will be obtained by applying this information to microarray probe design and methods of calculation based on the copy number of 16S rRNA genes.

5. Conclusions

According to our results, the combination of cPCR and microarray analysis was useful to quantify the number of each species in samples at one time. The deviation between the value measured by using a DNA chip and that measured by using qPCR was < 10-fold. This assay will be applied to oral diagnostic tools that require quantification of microbial communities.

Declaration of Competing Interest

No conflict of interest declared.

Acknowledgments

The authors would like to thank Ai Nozawa, Eiji Sato and Kazuki Momose for technical discussion, and Nana Hoashi for technical assistance of the experiment. The authors are grateful to the Mitsubishi Chemical Bio-Devices Group's manufacturing team who produced DNA chips. The authors would like to thank Enago (www.enago.jp) for the English Proofreading. This work was supported by Women In Research Support Operation of Osaka Prefecture University.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.mimet.2019.105718.

References

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. J. Mol. Biol. 215, 403–410.
Arumugam, M., Raes, J., Pelletier, E., Le Paslier, D., Yamada, T., Mende, D.R., Fernandes, G.R., Tap, J., et al., 2011. Enterotypes of the human gut microbiome. Nature 473, 174–180.
D'Amore, R., Ijaz, U.Z., Schirmer, M., Kenny, J.G., Gregory, R., Darby, A.C., Shakya, M., Podar, M., Quince, C., Hall, N., 2016. A comprehensive benchmarking study of protocols and sequencing platforms for 16S rRNA community profiling. BMC Genomics 17, 55.
Fouhy, F., Clooney, A.G., Stanton, C., Claesson, M.J., Cotter, P.D., 2016. 16S rRNA gene sequencing of mock microbial populations-impact of DNA extraction method, primer choice and sequencing platform. BMC Microbiol. 16, 123.
Kageyama, S., Takeshita, T., Asakawa, M., Shibata, Y., Takeuchi, K., Yamanaka, W., Yamashita, Y., 2017. Relative abundance of total subgingival plaque-specific bacteria in salivary microbiota reflects the overall periodontal condition in patients with periodontitis. PLoS One 12, e0174782.
Lyons, S.R., Griffen, A.L., Leys, E.J., 2000. Quantitative real-time PCR for *Porphyromonas gingivalis* and total bacteria. J. Clin. Microbiol. 38, 2362–2365.
Mangin, I., Suau, A., Magne, F., Garrido, D., Gotteland, M., Neut, C., Pochart, P., 2006. Characterization of human intestinal bifidobacteria using competitive PCR and PCR-TTGE. FEMS Microbiol. Ecol. 55, 28–37.
Marotz, C.A., Sanders, J.G., Zuniga, C., Zaramela, L.S., Knight, R., Zengler, K., 2018. Improving saliva shotgun metagenomics by chemical host DNA depletion. Microbiome 6, 42.
Miyake, S., Kim, S., Suda, W., Oshima, K., Nakamura, M., Matsuoka, T., Chihara, N., Tomita, A., Sato, W., Kim, S.W., Morita, H., 2015. Dysbiosis in the gut microbiota of patients with multiple sclerosis, with a striking depletion of species belonging to Clostridia XIVa and IV clusters. PLoS One 10, e0137429.
Nayfach, S., Pollard, K.S., 2015. Average genome size estimation improves comparative metagenomics and sheds light on the functional ecology of the human microbiome. Genome Biol. 16, 51.
Okada, M., Hayashi, F., Nagasaka, N., 2000. Detection of *Actinobacillus actinomycetemcomitans* and *Porphyromonas gingivalis* in dental plaque samples from children 2 to 12 years of age. J. Clin. Periodontol. 27, 763–768.
Palmer, C., Bik, E.M., Eisen, M.B., Eckburg, P.B., Sana, T.R., Wolber, P.K., Relman, D.A., Brown, P.O., 2006. Rapid quantitative profiling of complex microbial populations. Nucleic Acids Res. 34, e5.
Roth, S.B., Jalava, J., Ruuskanen, O., Ruohola, A., Nikkari, S., 2004. Use of an oligonucleotide array for laboratory diagnosis of bacteria responsible for acute upper respiratory infections. J. Clin. Microbiol. 42, 4268–4274.
Stämmler, F., Gläsner, J., Hiergeist, A., Holler, E., Weber, D., Oefner, P.J., Gessner, A., Spang, R., 2016. Adjusting microbiome profiles for differences in microbial load by spike-in bacteria. Microbiome 4, 28.
Suzuki, M.T., Giovannoni, S.J., 1996. Bias caused by template annealing in the amplification of mixtures of 16S rRNA genes by PCR. Appl. Environ. Microbiol. 62, 625–630.
Suzuki, N., Yoshida, A., Nakano, Y., 2005. Quantitative analysis of multi-species oral biofilms by TaqMan real-time PCR. Clin. Med. Res. 3, 176–185.
Teles, R.P., Haffajee, A.D., Socransky, S.S., 2006. Microbiological goals of periodontal therapy. Periodontol. 42, 180–218.
Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., Higgins, D.G., 1997. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 25, 4876–4882.
Tourlousse, D.M., Yoshiike, S., Ohashi, A., Matsukura, S., Noda, N., Sekiguchi, Y., 2017. Synthetic spike-in standards for high-throughput 16S rRNA gene amplicon sequencing. Nucleic Acids Res. 45, e23.
Treimo, J., Vegarud, G., Langsrud, T., Marki, S., Rudi, K., 2006. Total bacterial and species-specific 16S rDNA micro-array quantification of complex samples. J. Appl. Microbiol. 100, 985–998.
Tsuruya, A., Kuwahara, A., Saito, Y., Yamaguchi, H., Tsubo, T., Suga, S., Inai, M., Aoki, Y., et al., 2016. Ecophysiological consequences of alcoholism on human gut microbiota: implications for ethanol-related pathogenesis of colon cancer. Sci. Rep. 6, 27923.
Wang, Y., Qian, P.Y., 2009. Conservative fragments in bacterial 16S rRNA genes and primer design for 16S ribosomal DNA amplicons in metagenomic studies. PLoS One 4, e7401.
Zentilin, L., Giacca, M., 2007. Competitive PCR for precise nucleic acid quantification. Nat. Protoc. 2, 2092–2104.