OPEN

The ISME Journal (2015) 9, 629–642

www.nature.com/ismej

ORIGINAL ARTICLE

Comparative genome analysis and identification of competitive and cooperative interactions in a polymicrobial disease

Akiko Endo[1], Takayasu Watanabe[2], Nachiko Ogata[2], Takashi Nozawa[2,3], Chihiro Aikawa[2,3], Shinichi Arakawa[4], Fumito Maruyama[2,3,5], Yuichi Izumi[1] and Ichiro Nakagawa[2,3]

[1]*Department of Periodontology, Graduate School of Medical and Dental Sciences, Tokyo Medical and Dental University, Tokyo, Japan;* [2]*Section of Bacterial Pathogenesis, Graduate School of Medical and Dental Sciences, Tokyo Medical and Dental University, Tokyo, Japan;* [3]*Department of Microbiology, Graduate School of Medicine, Kyoto University, Kyoto, Japan;* [4]*Department of Lifetime Oral Health Care Sciences, Graduate School of Medical and Dental Sciences, Tokyo Medical and Dental University, Tokyo, Japan and* [5]*Section of Microbial Genomics and Ecology, Graduate School of Medical and Dental Sciences, Tokyo Medical and Dental University, Tokyo, Japan*

Polymicrobial diseases are caused by combinations of multiple bacteria, which can lead to not only mild but also life-threatening illnesses. Periodontitis represents a polymicrobial disease; *Porphyromonas gingivalis*, *Treponema denticola* and *Tannerella forsythia*, called 'the red complex', have been recognized as the causative agents of periodontitis. Although molecular interactions among the three species could be responsible for progression of periodontitis, the relevant genetic mechanisms are unknown. In this study, we uncovered novel interactions in comparative genome analysis among the red complex species. Clustered regularly interspaced short palindromic repeats (CRISPRs) of *T. forsythia* might attack the restriction modification system of *P. gingivalis*, and possibly work as a defense system against DNA invasion from *P. gingivalis.* On the other hand, gene deficiencies were mutually compensated in metabolic pathways when the genes of all the three species were taken into account, suggesting that there are cooperative relationships among the three species. This notion was supported by the observation that each of the three species had its own virulence factors, which might facilitate persistence and manifestations of virulence of the three species. Here, we propose new mechanisms of bacterial symbiosis in periodontitis; these mechanisms consist of competitive and cooperative interactions. Our results might shed light on the pathogenesis of periodontitis and of other polymicrobial diseases.
The ISME Journal (2015) **9,** 629–642; doi:10.1038/ismej.2014.155; published online 29 August 2014

Introduction

Infectious diseases are the second leading cause of death in the world (http://www.who.int/whr/2000/en/). A polymicrobial disease is defined as an infectious disease caused by the mixed infection of several microbes. It is difficult to elucidate onset/progress mechanisms of a polymicrobial disease because this nosological entity is not eligible for the basic concept established by Robert Koch. Periodontitis (Ingham and Sisson, 1984) and respiratory illness (Hament *et al.*, 1999) are representative examples of polymicrobial diseases.

Periodontitis, which is known as a main cause of tooth loss (Papapanou *et al.*, 1999; Irfan *et al.*, 2001), leads to the destruction of periodontal ligaments and adjacent supportive alveolar bone. According to a report by the World Health Organization, 15–30% of 35- to 44-year-old human beings are afflicted by periodontitis (http://www.who.int/en/), and the prevalence of periodontitis is the highest of all the infectious diseases in the world. More than 300 existing bacterial species are implicated in the initiation and progression of periodontitis, especially the following three bacterial species: *Porphyromonas gingivalis*, *Treponema denticola* and *Tannerella forsythia*, which are classified as 'the red complex'. They are detected together at a high rate in deep periodontal pockets (Socransky *et al.*, 1998; Darveau, 2010).

In bacterial communities, one bacterial species tends to obtain different functions from the other species (Scott *et al.*, 2008). The red complex species might interact with one another to adapt to the niche and thus may cause periodontitis in a cooperative manner. Many studies revealed the virulence factors

Correspondence: F Maruyama, Department of Microbiology, Graduate School of Medicine, Kyoto University, Yoshida-Konoe-cho, Sakyo-ku, Kyoto 606-8501, Japan.
E-mail: fumito-m.bac@tmd.ac.jp
Received 26 February 2014; revised 11 July 2014; accepted 17 July 2014; published online 29 August 2014

of *T. denticola* and *P. gingivalis*, whereas the reports of *T. forsythia* virulence factors are rarer (Ishihara, 2010; Grenier and La, 2011; Dashper *et al.*, 2011a; Bostanci and Belibasakis, 2012). Some reports are suggestive of synergistic effects of virulence factors between *T. denticola* and *P. gingivalis*; *T. denticola* does not form biofilms by itself *in vitro* (Vesey and Kuramitsu, 2004), but coculture of *P. gingivalis* and *T. denticola* results in biofilm formation and coaggregation (Kuramitsu *et al.*, 2005). On the other hand, *T. forsythia* grows slowly and requires special precise conditions for its growth (Honma *et al.*, 2001; Sakakibara *et al.*, 2007). The information about *T. forsythia* is limited, and there is no genome-wide analysis to identify any interactions among the red complex species.

Major research projects on the diagnosis of polymicrobial disease are mainly based on the metagenomic analyses, which can identify a wide range of bacteria from various environments (Brogden *et al.*, 2005). Nonetheless, to the best of our knowledge, no data in the literature describe simultaneous comparative genome analysis and focuses on the specific bacterial species that are strongly related to polymicrobial diseases. In this study, we performed comparative genome analyses among the red complex species and demonstrated reciprocal relationships among the three species: competitive and cooperative relationships. Our results are suggestive of significance of competitive and cooperative interactions in the periodontitis as a novel mechanism of bacterial symbiosis.

Materials and methods

Detailed information is described in Supplementary Materials and methods.

Bacterial strains and culture conditions
We used 18 *T. forsythia* strains that were isolated at Tokyo Medical and Dental University, Tokyo, Japan.

Determination of T. forsythia complete genome sequences
Complete genome sequences of two *T. forsythia* strains (KS16 and 3313) were determined using a combination of the GS Junior Titanium instrument (454 Life Sciences a Roche Company, Branford, CT, USA) for shotgun/paired-end sequencing, Genome Analyzer IIx (GAIIx; Illumina, San Diego, CA, USA) for 101-bp paired-end sequencing and traditional Sanger sequencing (ABI 3730; Applied Biosystems, Foster City, CA, USA). Then, these reads were assembled in Newbler v.2.6 to obtain scaffolds.

Determination of T. forsythia draft genome sequences
Draft genome sequences of 16 *T. forsythia* strains (sit3, 15, 20, 2612, MH6, L7, 273, KM4, 3322, 291, 3114, TR1, 2444, 1224, 2442 and K8) were determined using GAIIx paired-end sequencing. Then, the draft genomes were annotated using Rapid Annotation using Subsystem Technology server v.4.0 (http://rast.nmpdr.org/; Aziz *et al.*, 2008).

Characterization of T. forsythia species by genome information
We characterized 19 *T. forsythia* strains by estimating the pan-genome size, by dotplot analysis and via construction of phylogenetic tree, as described below in the methods for the three species of the red complex. The findings of these experiments are described in Supplementary Results.

Estimation of pan-genome size
The gene sequences were clustered among the strains of each species by PGAP v.1.11 (Zhao *et al.*, 2012). Following the clustering of the genes, we estimated the number of new genes when sequencing more genomes and the pan-genome size for each species. Regression analysis for new genes and the pan-genome was conducted as described previously (Tettelin *et al.*, 2008).

Multiple dotplot analysis
Multiple dotplot analysis was conducted as described previously (Watanabe *et al.*, 2013).

Construction of phylogenetic tree and single-nucleotide polymorphism analysis
In each species in the red complex, the core genes that were single copy in all the strains were tested to see whether rearrangement in a gene was significant, by using the Phi test (Huson and Bryant, 2006). Then, the concatenated amino-acid sequences from all the single-copy core genes were used for construction of phylogenetic trees. For all the gene clusters in each species where a gene was located in more than two genomes, a dN/dS ratio was calculated by means of codeml in the PAML v.4.6 (Yang, 1997).

CRISPR analysis
Clustered regularly interspaced short palindromic repeat (CRISPR) regions and CRISPR-associated (*cas*) genes were identified, and a target sequence of each spacer was characterized as described previously (Watanabe *et al.*, 2013).

Metabolism analysis
In each species, the core genes were mapped on metabolic pathways using KAAS (http://www.genome.jp/kaas-bin/; Moriya *et al.*, 2007), and were visualized using the iPath v.2 (Yamada *et al.*, 2011).

Virulence gene
To test for the presence of known virulence genes in all the genomes of strains in each species, BLASTP was used for all the genes of 38 strains (i.e., the sum of the three species strains) against the virulence genes described in the previous reports (Holt *et al.*, 1999; Sharma, 2010; Dashper *et al.*, 2011a). In addition, we tried to find any novel virulence gene candidates using BLASTP against the virulence factor database (last accessed 25 Novemer 2013; Chen *et al.*, 2012) and the microbial database of protein toxins, virulence factors and antibiotic resistance genes for biodefense applications (MvirDB; last accessed 21 March 2007; Zhou *et al.*, 2007).

Results and discussion

Comparative genomics of the red complex species
The red complex species are reported to persist together in the periodontal pocket (Socransky *et al.*, 1998; Paster *et al.*, 2001; Kumar *et al.*, 2003; Griffen *et al.*, 2012; Casarin *et al.*, 2013; Ge *et al.*, 2013), and other oral bacteria are not always detected (Supplementary Table S1). Nevertheless, there are no reports in the literature about the precise relationship among the red complex species. Comparative genome analysis among the red complex species may be a significant clue not only to the characteristic features of each species but also to the interactions among them. For the comparative genome analyses, we applied the same annotation criteria (see Materials and methods) to all the strains of the red complex species including the publicly available genome sequences (Supplementary Tables S2 and S3).

Gene content
An average genome size of *T. forsythia*, *T. denticola* and *P. gingivalis* was 3.38, 2.83 and 2.34 Mbp, respectively; the genome size of *T. forsythia* was the largest in the red complex species (Supplementary Table S3). In the complete genome sequences, the average number of coding DNA sequences (CDSs) of *T. forsythia*, *T. denticola* and *P. gingivalis* were 3005, 2693 and 2175, respectively. Within each species, all the strains had the same numbers of rRNA and tRNA genes; the number of rRNA operons of *T. forsythia*, *T. denticola* and *P. gingivalis* were 2, 2 and 4, respectively, and the number of tRNA genes of *T. forsythia*, *T. denticola* and *P. gingivalis* were 46, 45 and 54, respectively. The core genes (shared within all the strains in the species) and the species-specific genes of *T. forsythia*, *T. denticola* and *P. gingivalis* were 1733/1130, 1966/1899 and 1610/1005, respectively (Supplementary Figure S1 and Supplementary Tables S4–S6). The number of core genes accounted for 60% (1733/2912), 73% (1966/2693) and 74% (1610/2175) of average number of CDSs in 19 *T. forsythia* strains, 5 *T. denticola* strains and 3 *P. gingivalis* strains, respectively. Forty-five genes were shared among the red complex species, including housekeeping genes and molecular chaperones such as ClpB (Hsp104) and DnaK (Hsp70; Supplementary Table S7). A previous report indicated that these chaperones regulate overproduction of reactive oxygen species that is associated with the inflammatory response (Steeves *et al.*, 2011). When comparing the genes within each pair of species, the number of gene homologs was 548, 10 and 12 between *T. forsythia* and *P. gingivalis*, between *T. forsythia* and *T. denticola*, and between *T. denticola* and *P. gingivalis*, respectively (Supplementary Figure S1). The homologs shared between *T. forsythia* and *P. gingivalis* were 19% (548/2912) of all *T. forsythia* genes and 25% (548/2175) of all *P. gingivalis* genes.

Next, we estimated pan-genome size based on the 19 *T. forsythia* genomes, 14 *T. denticola* genomes and 5 *P. gingivalis* genomes to determine the genetic landscape (Supplementary Figure S2). First, we estimated the number of new genes that would be discoveredby using power-law regression analysis, $n = \kappa \cdot N - \alpha$ (Tettelin *et al.*, 2008). We then estimated the number of pan-genomes that would be accumulated by using the same analysis, $n = \varepsilon \cdot N\gamma$ (Tettelin *et al.*, 2008). The values were $\alpha = 0.494$ and $\gamma = 0.254$ in *T. forsythia*, $\alpha = 0.733$ and $\gamma = 0.256$ in *T. denticola* and $\alpha = 0.834$ and $\gamma = 0.246$ in *P. gingivalis*, respectively. These values in red complex species were indicative of an open pan-genome because the exponent α was <1.0 and γ was >0. Recent studies reported that some bacterial species also show an open pan-genome, including *Propionibacterium acnes* and *Corynebacterium pseudotuberculosis* (Soares *et al.*, 2013; Tomida *et al.*, 2013). Although the pan-genomes of the red complex species tend to be open, these results suggested that almost 20% genes were homologs shared between *T. forsythia* and *P. gingivalis*. These results may be the consequence of the fact that *P. gingivalis* and *T. forsythia* belong to the same phylum Bacteroidetes.

Genomic collinearity
Genomic rearrangements have a significant effect on bacterial evolution, including the genome reductive processes and the creation of new DNA regions (Sun *et al.*, 2012). We observed only one or two continuous lines in the dot plots of *T. forsythia* and *T. denticola.* On the contrary, the dot plots of *P. gingivalis* contained many parallel and numerous discontinuous lines in its plot (Figure 1). Similar to the previous report (Naito *et al.*, 2008; Watanabe *et al.*, 2013), we identified frequent genome rearrangements in *P. gingivalis*, which contained inversions, translocations, simple insertions, deletions and replacements. Simple insertions, deletions and replacements were observed two times more frequently than inversions or translocations.

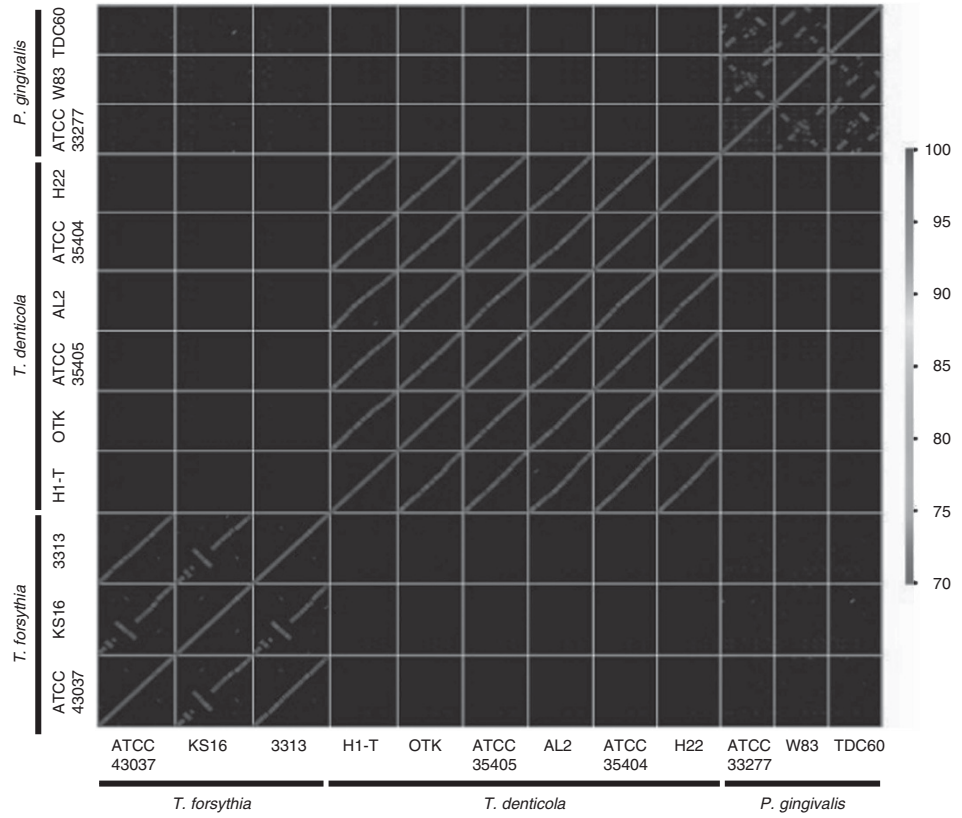

Figure 1 Dotplot analysis of *T. forsythia*, *T. denticola* and *P. gingivalis*. A dot is plotted when the 20-bp sequences exhibited similarity between two genomes. The dots are colored according to the similarity score, with gradation from 70% (blue) to 100% (red). A color version of this figure is available on *The ISME Journal* online.

Because mobile genetic elements (MGEs) are intimately involved in the extensive genome rearrangement of *P. gingivalis* (Watanabe *et al.*, 2013), we next counted the number of MGEs: insertion sequences, conjugative transposons and miniature inverted-repeat transposable elements (Supplementary Figure S3). The number of each MGE per CDS was the lowest in *T. denticola* and highest in *P. gingivalis* among the red complex species. The ratio of insertion sequences/CDS of *T. denticola* was lower than that of the other two species, with statistical significance ($P<0.01$ for *T. denticola* and *P. gingivalis*; $P<0.05$ for *T. denticola* and *T. forsythia*; Student's *t*-test).

These results suggest that the number of MGEs is high when extensive genome rearrangements are frequently observed. On the other hand, the number of MGEs is low when genome rearrangements are rare. The red complex species might be classified into two types, based on their genome structure: one possible type, including *T. forsythia* and *T. denticola*, has maintained stable genome structure, and the other type (represented only by *P. gingivalis*) had undergone frequent changes in the genome structure as a result of complicated genomic rearrangements mediated by transposition of MGEs, demonstrated in the previous reports (Naito *et al.*, 2008; Watanabe *et al.*, 2013).

The ratio of nonsynonymous substitutions to synonymous substitutions

To assess the dynamics of molecular evolution and to detect proteins undergoing adaptation to environmental changes in each species (Kimura, 1983; Gillespie, 1991; Ohta and Ina, 1995), we estimated the ratio of the non-synonymous substitution rates (d*N*) to the synonymous substitution rates (d*S*; the d*N*/d*S* ratio) for the core genes of each species. In *T. forsythia*, *T. denticola* and *P. gingivalis*, 5.4% (54/995), 0.7% (10/1 320) and 2.6% (11/425) of the core genes exhibited d*N*/d*S*>1, respectively, indicating positive selection (Figure 2 and Supplementary Table S8). The 54 genes in *T. forsythia* contained 26 hypothetical genes; among these hypothetical genes, no protein motif was found in 16 genes, whereas the other 10 genes encoded several functional domains including those in peptidase and outer membrane proteins. In *T. denticola*, all 10 genes with d*N*/d*S*>1 were hypothetical genes, including four genes that contained the domains such as the RHS repeat-associated core domains and cysteine protease domains. In *P. gingivalis*, 11 genes with d*N*/d*S*>1 included seven hypothetical genes, the nitroimidazole resistance gene and the SOS mutagenesis/repair gene. Among the hypothetical genes, six genes encoded several domains including TlpA-like family related to the cytochrome maturation and

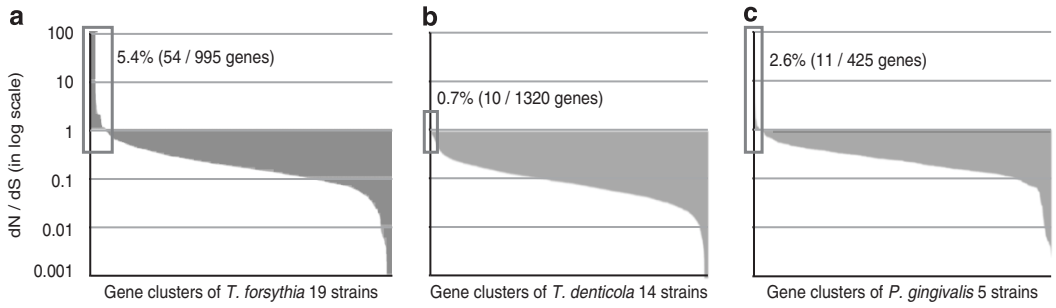

Figure 2 The d*N*/d*S* ratio of *T. forsythia*, *T. denticola* and *P. gingivalis*. The d*N*/d*S* ratios of all the gene clusters in this analysis are shown in an ascending order for (**b**) 19 *T. forsythia*, (**b**) 14 *T. denticola* and (**c**) 5 *P. gingivalis* strains, respectively. Red squares indicate the genes that correspond to the condition d*N*/d*N*>1. A color version of this figure is available on *The ISME Journal* online.

phosphopantetheinyl transferase superfamily related to the biosynthesis of butyric acid.

As a conclusion from the above three analyses, we propose different evolutionary strategies among the red complex species; intraspecies genetic diversity is low in *T. forsythia* reflecting stable genome structure, and amino-acid sequence alterations may be a major driving force of diversification in *T. forsythia. T. denticola* shows the most stable genome structure and the lowest selective pressure among the three species. With respect to *P. gingivalis*, the main evolutionary strategy may be based on frequent genome rearrangement by transposing MGEs rather than amino-acid sequence variation.

Competitive interactions between T. forsythia and the other red complex species via CRISPR/Cas systems

Previous reports showed that the CRISPR/Cas system prevents acquisition of conjugative plasmids (Marraffini and Sontheimer, 2008), integrative conjugative elements (Lopez-Sanchez *et al.*, 2012) and environmental DNA via natural transformation (Bikard *et al.*, 2012; Zhang *et al.*, 2013b). Similarly, the CRISPR/Cas system might affect the molecular interactions and the survival strategies among the red complex species.

The red complex species had different sets of CRISPR-associated genes (*cas* genes) and CRISPR repeat sequences (Supplementary Table S9). When we investigated the complete genome sequences of three *T. forsythia* strains, according to the most recent classification by Makarova *et al.* (2011), all three strains carried a type I-B CRISPR/Cas locus with six *cas* genes (*cas*6, *cas*7, *cas*3, *cas*4, *cas*1 and *cas*2; Supplementary Figure S4). Strain KS16 had another CRISPR/Cas locus, which was a type II CRISPR/Cas system with three *cas* genes (*cas*2, *cas*1 and *cas*9). We next identified the CRISPR/Cas loci in the 19 *T. forsythia* genomes including draft genome sequences. The CRISPR/Cas loci were detected in all 19 strains. Five strains (KS16, 2612, 291, 3114 and 3322) had both type I-B and II CRISPR/Cas systems, and strain TR1 had a type I-B and an unclassified type of CRISPR/Cas locus, whereas the other 13 strains possessed only a type I-B CRISPR/Cas system (Supplementary Table S9). In the type I-B CRISPR/Cas locus, a gene deletion was observed in six strains without differences in clinical parameters (sit3, KS16, 3322, 2612, 3114 and 2444; Supplementary Figure S4). Furthermore, a 3.5-kbp insertion was found between the *csm4* and *csm5* genes in 3313. The insertion region contained hypothetical genes and phage-encoded genes, and hereafter we call this region a 'phage-like structure'. Such structures were found in five strains (3313, 2444, 1224, 2442 and TR1) and were not found in any regions other than this CRISPR region. In the phylogenetic trees (see Supplementary Materials and methods and Supplementary Results for the construction procedure), strain 3313 was phylogenetically distant from the four strains (2444, 1224, 2442 and TR1; Supplementary Figure S5), suggesting that strain 3313 obtained the phage-like structure via a different mechanism compared with the above-mentioned four strains.

On the other hand, the *cas* gene sets of *T. denticola* were highly conserved among the six draft genomes. The CRISPR/Cas locus in *T. denticola* was classified as a type II-A CRISPR/Cas system because of the genes *csn*2, *cas*2, *cas*1 and *cas*9 (Supplementary Figure S6). Strain H1-T had only one repeat sequence, whereas we identified no repeat sequences in strain H22. The structure of the *cas* gene in *P. gingivalis* genomes was also conserved at two CRISPR loci, including type I-C and III-B (Watanabe *et al.*, 2013).

Polymorphisms of the repeat sequences were observed among the strains in each species (Supplementary Figures S7 and S8 and Supplementary Tables S10 and S11), in line with a previous report (Watanabe *et al.*, 2013). We identified a few mutations at the end and in the middle of the repeat sequences of *T. forsythia* and *T. denticola*, respectively. Three types of *T. forsythia* CRISPR repeats had different secondary structures and were grouped into different categories based on superclasses, classification of the phylum, *cas* gene types and sequence families (Supplementary Figure S9). Protospacer-adjacent motifs were not clearly detected in all CRISPR types of *T. forsythia* and

T. denticola, which may be because of the lesser amount of the spacers exhibiting high nucleotide similarity to the sequences in the databases (Supplementary Figure S10).

The total number of spacers and unique spacers in each species (*T. forsythia*, *T. denticola* and *P. gingivalis*) were 1631/1149, 78/66 and 267/244, respectively (Supplementary Table S12). Among these spacers, none was common for the three species. Dendrogram-based analysis of spacers showed that strains 3322 and 3313 were clustered as the same group, and this finding was consistent with the structure of phylogenetic trees of the core genes (Supplementary Figure S11). In *P. gingivalis*, CRISPR typing was shown to be capable of clustering the strains beyond the fimbrial gene types (Watanabe *et al.*, 2013); in this study, we demonstrated that the CRISPR typing was applicable to *T. forsythia* to show strain clusters that can be formed in the core gene-based phylogenetic tree.

A previous study reported that the CRISPR spacers in *P. gingivalis* possibly inhibit both genomic rearrangements and intercellular recombination among *P. gingivalis* strains (Watanabe *et al.*, 2013). Nonetheless, there are no studies focused on CRISPR targets in *T. forsythia* and *T. denticola*. To identify such targets of the CRISPR/Cas system in *T. forsythia* and *T. denticola*, we analyzed the spacers in these species by using nucleotide similarity searches against the seven databases (see Materials and methods). In *T. forsythia* and *T. denticola*, 106/1631 (6.5%) and 7/78 (9.0%) non-redundant spacers exhibited high nucleotide similarity to the sequences in the databases, respectively (Figure 3, Table 1, Supplementary Figure S12 and Supplementary Table S13).

In *T. denticola*, 6/7 spacers showed significant nucleotide similarity to hypothetical genes in the genome of *T. denticola* (Table 1, Supplementary Figure S12 and Supplementary Table S13). Four out of those six spacers were located only in the ATCC 35404 genome, and their targets were the genes in the ATCC 35405 genome that perfectly matched the nucleotide sequence. Meanwhile, the remaining two of the six spacers were present only in ATCC 35405 and targeted the genes on the same chromosome, with one single-nucleotide polymorphism present in their protospacer regions. This phenomenon can be called 'self-targeting', in the sense that the CRISPR spacer targets the same chromosome endogenously. One study reported that 100 of 23 550 spacers (0.4%) are self-targeting spacers; these spacers are widely distributed over diverse phylogenetic lineages (Stern *et al.*, 2010). The self-targeting spacers may have influenced genome-scale evolution by participating in regulation of gene expression (Aklujkar and Lovley, 2010) or by being lethal for the strain harboring such spacers (Westra *et al.*, 2014). Generally, a strict nucleotide match (no more than 1-bp mismatch) between a CRISPR spacer and its target site is required for CRISPR interference (Sorek *et al.*, 2013), although it was recently reported that a maximal 5-bp mismatch is tolerated in type I if these nucleotides are outside the seed region in a protospacer (Semenova *et al.*, 2011). Because strain ATCC 35405 is culturable even harboring the self-targeting spacer, it is possible that ATCC 35405 may evade lethality via unknown mechanisms that do not involve the mismatch being in the target site and may use the CRISPR/Cas system for its evolution.

In *T. forsythia*, 55/106 nonredundant spacers showed high nucleotide similarity to the annotated sequences in the databases (Figure 3, Table 1 and Supplementary Table S13). Of these 55 spacers, 22 exhibited significant nucleotide similarity to *T. forsythia*'s own genome, which contained outer membrane protein genes and bacterial group 2 immunoglobulin-like protein genes. Meanwhile, 25/55 spacers showed similarity to other bacterial genomes; 16/25 and 3/25 spacers exhibited significant nucleotide similarity to genomes of *P. gingivalis* and *T. denticola*, respectively. These results suggest that the CRISPR/Cas system of *T. forsythia* might target the genomes of other red complex species. In particular, the spacers of strains 2444, 1224, 2442 and TR1 showed significant nucleotide similarity to the methyltransferase gene of *P. gingivalis*. According to an investigation of the adjacent regions of the methyltransferase gene on *P. gingivalis* genomes, this gene was not likely to be associated with any transposable elements (data not shown). In addition, it is notable that the spacers of strains 3313, 3322 and 15 exhibited significant nucleotide similarity to several MGEs (conjugative transposons and insertion sequences) in *P. gingivalis*. In *P. gingivalis*, DNA transfer is reported via conjugation or natural transformation (Naito *et al.*, 2011; Tribble *et al.*, 2012). Competitive interactions between *T. forsythia* and the other two species in the red complex are possible; *P. gingivalis* delivers its DNA into the cells of other species to assert its dominance in the niche, whereas *T. forsythia* attacks the methyltransferase gene of *P. gingivalis* by delivering the spacer and related Cas proteins of *T. forsythia* CRISPR/Cas system into *P. gingivalis* cells, to impede their persistence.

Bacterial interactions may be important as a competitive interaction among the species that causes growth inhibition (Rao *et al.*, 2005), metabolic inactivation (Stringfellow and Aitken, 1995) and downregulation of genes (Xie *et al.*, 2007). We demonstrated a possibility of competitive interactions among the red complex species via the CRISPR/Cas system. The CRISPR/Cas system in *T. forsythia* may work as both a weapon toward the other members in the community and a barrier against them. Further research will be needed to prove our hypothesis, for example, by characterizing targets of the CRISPR spacers that are not characterized in this study or by demonstrating DNA/protein transfer among the red complex species.

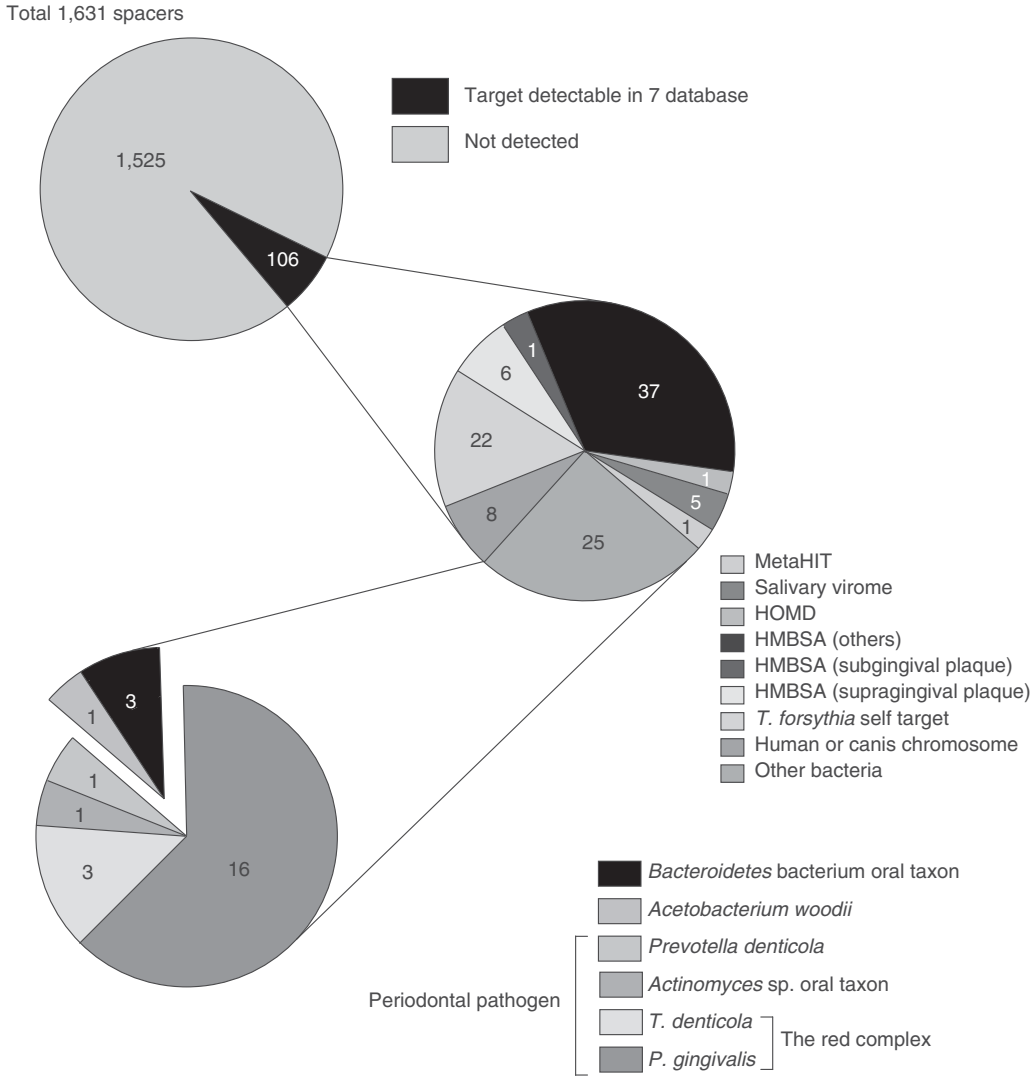

Figure 3 Similarity searches of CRISPR spacers in *T. forsythia*. In the *T. forsythia* CRISPR spacers, the presence of the targets is shown as a pie chart, where the targets of the 106 spacers are shown as other charts. The targets of the 25 spacers against the bacteria other than the three red complex species are indicated by the species names. A color version of this figure is available on *The ISME Journal* online.

On the other hand, we also demonstrated that evolutionary strategies were different among the three species living in the same niche; this observation is consistent with that in other environments (Venter *et al*., 2004; Eckburg *et al*., 2005) because environmental conditions may affect the evolutionary strategy of bacteria (Roszak and Colwell, 1987). The periodontal pocket is an environment that can be changing continuously; the environment is altered by medical treatments with the removal of subgingival calculus and biofilm to reduce pocket depth (Hellström *et al*., 1996), although the red complex species tend to form a biofilm to protect themselves from deleterious conditions. The different evolutionary strategies of the red complex species may be essential for adaptation to this variable environment. One study suggested that *P. gingivalis* produces a number of surface molecules and structures to interact with *T. denticola* and *T. forsythia* (Holt and Ebersole, 2005). We propose different roles for *P. gingivalis* and its two partner species to persist in the periodontal pocket; *P. gingivalis* might establish an ecological niche in the pocket, followed by participation of *T. denticola* and *T. forsythia* carrying biofilm-related genes to form a community in which the three species are the main members. It is still an open question as to why they dominate other kinds of oral bacteria.

Cooperative interactions among the red complex species

Several relations between the metabolism and niche habitat of organisms have been shown by both experimental and *in silico* analyses (Kastenmüller *et al*., 2008). Metabolic pathways are important for both viability of bacterial cells and expression of virulence, which is achieved by production of

Table 1 Sequences in the databases exhibiting high nucleotide similarity to unique spacers

Spacer ID					*Species*	*Gene name*
T. forsythia						
546					*T. forsythia* ATCC 43037	Putative membrane protein
744					*T. forsythia* ATCC 43037	Hypothetical protein
750					*T. forsythia* ATCC 43037	Putative membrane protein
975					*T. forsythia* ATCC 43037	Phenylalanine and histidine ammonia lyase
1090					*T. forsythia* ATCC 43037	Hypothetical protein
1256					*T. forsythia* ATCC 43037	Recombination protein O
1416	1558	1683	1806		*T. forsythia* ATCC 43037	Hypothetical protein
1434	1436	1448			*T. forsythia* ATCC 43037	Bacterial group 2 Ig-like protein
1435	1437	1438	1440	1442	*T. forsythia* ATCC 43037	Bacterial group 2 Ig-like protein
1439					*T. forsythia* ATCC 43037	Bacterial group 2 Ig-like protein
1820					*T. forsythia* ATCC 43037	Putative nuclease sbccd, subunit C
1859					*T. forsythia* ATCC 43037	Hypothetical protein
157	1032				*P. gingivalis* TDC60	IS (12 bp terminal inverted repeat)
545					*P. gingivalis* TDC60	Conjugative transposon protein traq
578	756	1258			*P. gingivalis* TDC60	Hypothetical protein
765					*P. gingivalis* TDC60	50S ribosomal protein L10
769					*P. gingivalis* ATCC 33277	30S ribosomal protein S16
1165					*P. gingivalis* ATCC 33277	Hypothetical protein
1171					*P. gingivalis ATCC 33277*	Hemagglutinin protein haga
1182					*P. gingivalis* ATCC 33277	DNA-binding protein HU
1275					*P. gingivalis* ATCC 33277	Intergenic region
1389	1531	1656	1779		*P. gingivalis* ATCC 33277	Decarboxylating precorrin-6Y C5,15-methyltransferase
265					*T. denticola* ATCC 35405	Hypothetical protein
267	268				*T. denticola* ATCC 35405	Hypothetical protein
547					*P. denticola* F0289	Hypothetical protein
199					*Actinomyces* sp. oral taxon 178 str. F0338 contig00144	
118	1146	1294			*Bacteroidetes* bacterium oral taxon 274 F0058	
795	867				*Canis familiaris* chromosome 9	
1309	1451	1576			Human chromosome 14 DNA sequence BAC R-662J14 of library RPCI-11	
1310	1452	1577			*Homo sapiens* chromosome	
1917					*Acetobacterium woodii* DSM 1030	
T. denticola						
229					*T. denticola* ATCC 35405	Hypothetical protein
251					*T. denticola* ATCC 35405	Hypothetical protein
265	266	267	268		*T. denticola* ATCC 35405	Hypothetical protein

Abbreviations: Ig, immunoglobulin; IS, insertion sequences.

virulence substances via multiple metabolic reactions (Panos and Ajl, 1963). We mapped the core genes in each species in the red complex onto the metabolic pathways (Supplementary Table S14, see Materials and methods), and obtained the complementation ratio of the number of mapped genes to the total number of genes involved in the respective pathways.

All 29 main pathways exhibited an abundance pattern of the mean complementation ratio when the number of species that were included in the mapping analysis was increased (Supplementary Figure S13 and Supplementary Table S15). Notably, the complementation ratio of the fatty acid biosynthesis pathway was the highest among the 29 pathways for each number of the species (one species: $68.84 \pm 6.99\%$; two species: $76.81 \pm 2.51\%$; and three species: 78.26%, respectively). Fatty acids are known to stimulate bacterial growth (Grenier, 1992a) and to be associated with inflammatory conditions (via increased concentration), such as periodontitis (Kurita-Ochiai *et al.*, 1995; Niederman *et al.*, 1997) and anaerobic infections (Ladas *et al.*, 1979; Mills *et al.*, 2006). We examined the complementation patterns in the fatty acid biosynthesis pathway, and found that only the butyric acid metabolism pathway was the case where all essential genes were complemented in the three species (Figure 4 and Supplementary Figure S13). Moreover, the fold increase was the highest in the butyric acid pathway among the 29 main pathways described above. Our results are suggestive of a possibility that the biosynthesis of fatty acids is an essential metabolic pathway for the red complex species, and the fatty acid produced via cooperative interactions among the red complex species might stimulate growth of these species and development of periodontitis. We also noted that the pathways except for the butyric acid biosynthesis were incomplete even if the red complex species coexisted.

Butyric acid is reported to be responsible for progression of periodontitis because this substance affects the host immune system in the periodontal pocket (Kurita-Ochiai *et al.*, 1995; Kurita-Ochiai and Ochiai, 2010). Previous reports have shown that

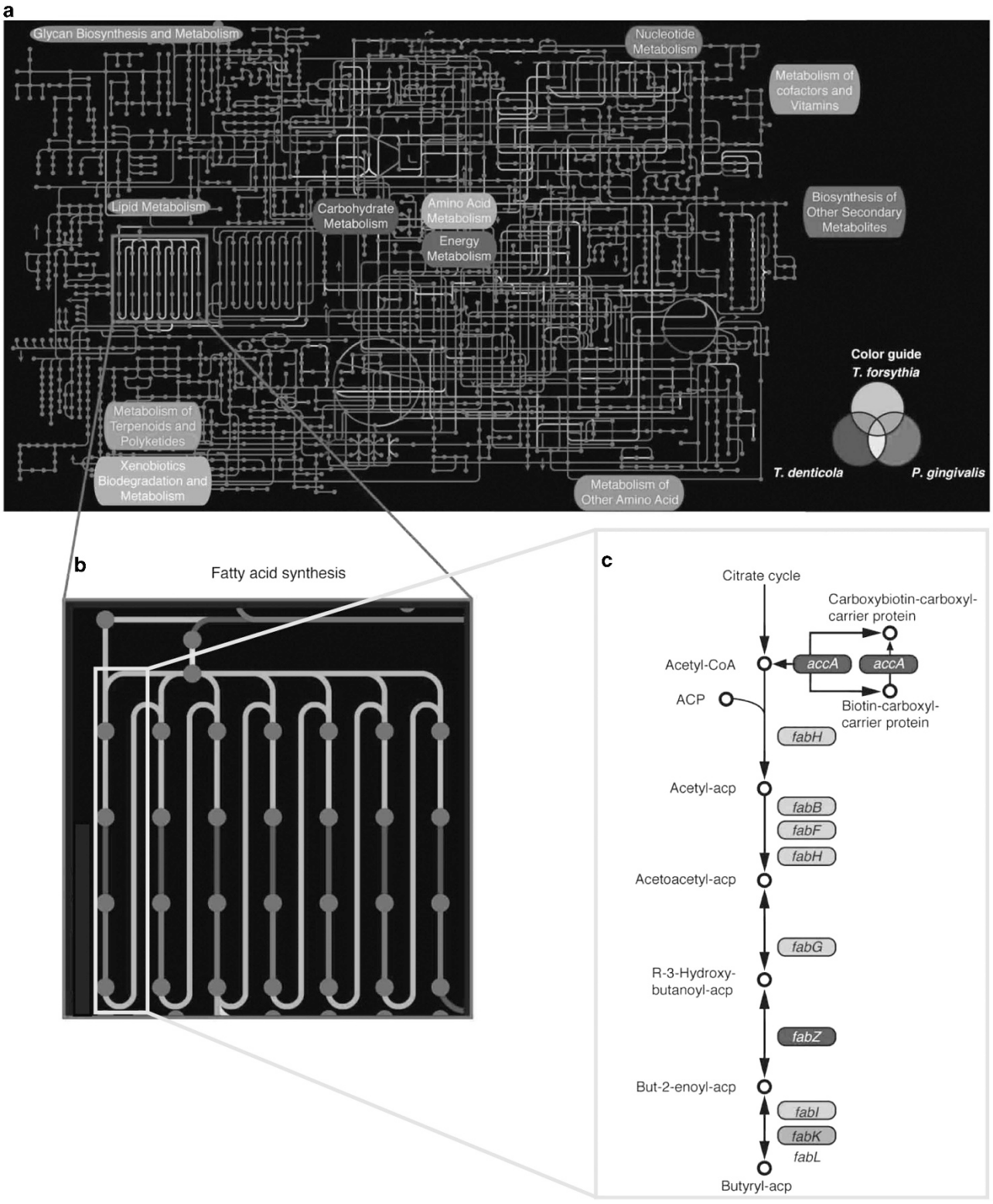

Figure 4 Metabolic analyses of *T. forsythia*, *T. denticola* and *P. gingivalis*. (**a**) The stand-alone pathways in the small pathway groups are colored if they are mapped by the core genes of either one, two or three species, as follows: gray (a known metabolic pathway), green (*T. forsythia*), blue (*T. denticola*), red (*P. gingivalis*), purple (*T. forsythia* and *T. denticola*), yellow (*T. denticola* and *P. gingivalis*), pink (*P. gingivalis* and *T. forsythia*), light blue (*T. forsythia*, *T. denticola* and *P. gingivalis*), respectively. (**b**) A fatty acid synthesis pathway group, which includes a butyric acid pathway, is shown by enlarging the corresponding region in (**a**). (**c**) In the butyric acid pathway, the gene required for each change of substance is shown, with the colors described in (**a**). A color version of this figure is available on *The ISME Journal* online.

there were several pathways for the synthesis of the butyric acids. Metabolic pathways vary among *Fusobacterium* species with distinct pathways (Gharbia and Shah, 1991) and metabolic shift from butyrate formation to lactate and acetate formation in the fermentation is found to be associated with changes in the activities of several key enzymes (Zhu and Yang, 2004). However, the pathway for the synthesis of the butyric acids among the red complex species might not apply to any of them. Furthermore, the fatty acid cross-feeding has previously been demonstrated in *P. gingivalis* and *T. denticola* (Grenier, 1992b), which is consistent with the downregulation of three genes (*fabG*, *fabF* and *acpP*) encoding enzymes participating in the initial stage of fatty acid synthesis (Tan *et al.*, 2014). These results suggest that the presence of *T. denticola* helps *P. gingivalis* reduces energy-consuming processes (Tan *et al.*, 2014). Complementation of essential genes in the butyric acid metabolism pathway may be important for the three species to facilitate their persistence in the niche. Because of the hypothetical protein transfer among the three species as described the above section, proteins encoded by those essential genes may be secreted to extracellular space and may participate in sequential reactions with the intermediates. Alternatively, it is also possible that the intermediates are transferred from the extracellular space into bacterial cells, and vice versa, and undergo reactions inside a bacterial cell possessing a relevant protein. The intermediates may be transferred via endocytosis or exocytosis; an endocytosis-like process was reported in *Gemmata obscuriglobus* (Fuerst and Sagulenko, 2010), whereas an outer membrane vesicle can be an example of exocytic transfer (Haurat *et al.*, 2011).

Furthermore, one study showed that cell-to-cell bacterial communication affected their growth (Toyofuku *et al.*, 2010). Despite these scenarios, there is still a possibility that one chromosome harbors all essential genes in a metabolic pathway, which can be shown via functional assignment of hypothetical genes. To find them, we searched the genes that are not associated with metabolic pathways for the presence of any protein motifs from the pathways, in NCBI's Conserved Domain Database (CDD; last accessed 21 March 2013; Marchler-Bauer *et al.*, 2011). There were no genes with significant probability of participating in metabolic pathways. Our findings are based on the analyses focused on the three red complex species; a biofilm in a periodontal pocket harbors hundreds of other bacterial species, which can compensate the metabolic pathways described above. Nevertheless, because the species other than the red complex are not always detected in periodontal pockets (Supplementary Table 1), they may not be capable of compensation of the metabolic pathways.

No study has shown that mixed cultures of the red complex species produce higher amounts of butyric acid compared with sole cultures. However, several studies have shown that the red complex species showed a clear increase in abundance and prevalence with increase in the severity of periodontal diseases (Loozen *et al.*, 2014) and the concentration of butyric acid in periodontal pockets was correlated with the severity of the diseases (Margolis *et al.*, 1988; Botta *et al.*, 1985). Furthermore, one study has shown that the isobutyric acid was found to promote the growth of *P. gingivalis* and *T. denticola* when both were inoculated together (Grenier, 1992a). The possible biological significance of our findings is that the mutual compensation of the butyric acid metabolism genes could facilitate persistence and virulence of the three red complex species.

In addition to the analyses of metabolic pathways, we examined the presence of virulence genes in the genome of each species. Most core genes of *T. forsythia* species were found to be housekeeping genes, and more than 90% of known virulence genes were included in the core genes (Supplementary Information and Supplementary Tables S4 and S16). In addition, by searching for amino-acid sequence similarities between the core genes and the virulence factors in the database, we identified four novel virulence gene candidates, which encode L-asparaginase, a hypothetical protein, an ATP-binding protein, and a translocase (Supplementary Table S17). We searched the genes of *T. denticola* and *P. gingivalis* strains to find such candidates too; there was no candidate in both species. Then we searched for the known virulence genes in each strain genome in the two species. In *T. denticola* and *P. gingivalis*, all the known virulence genes were conserved among all the strains; however, one strain of *T. forsythia* species lacked one virulence gene, which was the forsythia cell-detaching factor (Supplementary figure S14). It is notable that all the virulence genes were species specific except for a novel dipeptidyl peptidase *dpp*, whose mechanism of virulence in *T. forsythia* and *T. denticola* remains unknown (Ohara-Nemoto *et al.*, 2014), and that there was little redundancy in function of the virulence genes among the species (Table 2). Aside from compensation of metabolic pathways, it is likely that each red complex species may provide its virulence factors to the other species in which such factors are absent: this may be an example of synergistic interactions among the three species.

Bacterial interactions may enhance pathogenicity in natural host systems (Raymond *et al.*, 2012), improve sharing of a nutritious substance (Mashburn *et al.*, 2005), and promote metabolism (Shoaie *et al.*, 2013). Moreover, one study showed that the rats infected with the red complex species exhibit a significant increase in the level of serum immunoglobulin G (specific antibodies) and suggested that the interaction of the red complex species might promote periodontal tissue destruction (Kesavalu *et al.*, 2007). Nonetheless, genetic mechanisms of the actual molecular interactions

Table 2 Function of virulence genes of the red complex

Species	*Virulence gene*	*Virulece function*
T. forsythia	*Exo alpha sialidase, trypsin-like proteases*	Degradation of host proteins
	PrtH protease, forsythia cell-detaching factor protein	Inducing detachment of adherent cells and interleukin-8 production in human fibroblasts
	Surface layer protein A	Promoting epithelial cell adherence, invasion and biofilm formation
	BNR Asp-box repeat protein	Bacterial adherence and invasion of epithelial cells
	Surface antigen BspA	Binding to fibrinogen and fibronectin and the induction of proinflammatory cytokine expression in host cells by activating Toll-like receptor 2
	Methylglyoxalsynthase	Enolization of DHAP to produce methylglyoxal
Novel virulence gene candidates of *T. forsythia*	*L-asparaginase I*	Inhibiting T-cell responses of host
	Hypothetical protein	Protein of unknown function*
	ATP-binding protein	ABC transporters involved in nutrient or metal ion uptake may result indirectly from attachment of the bacteria to host cells
	Protein translocase subunit secA 1	Secretion of factors associated with pathogenesis or stress response
T. denticola	*Leucine-rich repeat protein*	Binding to *T. forsythia*
	Dentilisin	Disrupting or modulating intercellular host signaling pathways and degradation of intercellular adhesion proteins
	Major outer sheath protein Msp	Mediating colonization of host tissue
	Lipoprotein OppA, FhbB	Epithelial cell binding and invasion, subversion of the complement cascade or tissue invasion
	Chemotaxis CheR, CheB	Damaging host tissue
P. gingivalis	*Novel dipeptidylpeptidase dpp*	Abscess formation and lethality
	Hemagglutinin HagA	Periodontal tissue/cell invasion
	Hemagglutinin HagB	Stimulate strong IgG and Th immune responses; induce immune protection
	Gingipains RgpA, RgpB, Kgp	Tissue destruction and alter cytokine/chemokine and Igs bioactivity (i.e., IL-12,TNFa, C3 and C5, IgG/A)
	Fimbriae FimA	Bacterial colonization, induce host IgA, IgG and Th1 immune responses
	Capsular polysaccharide	Stimulating immune protection in mouse oral challenge model (via IgG activity)

Abbreviations: ABC, ATP-binding cassette; DHAP, dihydroxyacetone phosphate; IgG, immunoglobulin G; IL, interleukin; Th, T helper; TNF, tumor necrosis factor.
*Structurally this family is part of the six hairpin glycosidase superfamily, suggesting a glycosyl hydrolase, which is one of the virulence gene (Niu *et al.*, 2013).

among the red complex species are unknown at present. In this study, we demonstrated complementary relationships between the red complex species both in the metabolic pathways and in the virulence genes. Such relationships may have an important role not only in persistence in the deep periodontal pocket, where nutrient flexibility is limited, but also in the manifestation of virulence toward the host, as a progression mechanism of periodontitis. We showed the diversity of the symbiotic strategies among the red complex species and an interesting role of competition among the three species, although they persist together. During the initial stages of periodontitis, we believe that the interaction between the red complex species is important. Despite our findings, the precise mechanism of onset is yet to be elucidated; thus, further studies are needed for the validation of the interactions using well-established animal experimental periodontitis models and applying meta-analysis, which includes metatranscriptomic and metagenomic approaches.

In conclusion, this is the first piece of genome-wide evidence that the red complex species coexist in an oral environment based on the cooperative and competitive interactions among bacterial species. The interactions may allow bacteria to use different evolutionary strategies to survive in the variable environment. Our results will shed light on the mechanisms of interaction among not only the red complex species in periodontitis but also the bacteria involved in other polymicrobial diseases. Because this study is focused exclusively on the red complex species, behavior of the orange complex, which is a bacterial group believed to persist at high rates in the niches of the red complex (Socransky *et al.*, 1998), will be examined in future studies.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgements

We thank Dr M Umeda (Department of Periodontology, Osaka Dental University) for providing the *T. forsythia* isolates. This research was supported in part by grants-in-aid from the Japan Society for the Promotion of Science (JSPS); the Japanese Ministry of Education, Global Center of Excellence (GCOE) Program, International Research Center for Molecular Science in Tooth and Bone Diseases, MEXT (Ministry of Education, Culture, Sports, Science, and Technology), Japan; Grant-in-Aid for Scientific Research (22592032, 25293370, 25305013, 25670776 and 25713060); Grant-in-Aid for Scientific Research in Innovative Areas (24117508 and 70294113); Grant-in-Aid for Daiichi Sankyo Foundation of Life Science and the Funding Program for Next Generation World-Leading Researchers (LS041), JSPS.

References

Aklujkar M, Lovley DR. (2010). Interference with histidyl-tRNA synthetase by a CRISPR spacer sequence as a factor in the evolution of *Pelobacter carbinolicus*. *BMC Evol Biol* **10**: 230.

Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA *et al.* (2008). The RAST Server: rapid annotations using subsystems technology. *BMC Genom* **9**: 75.

Bikard D, Hatoum-Aslan A, Mucida D, Marraffini LA. (2012). CRISPR interference can prevent natural transformation and virulence acquisition during *in vivo* bacterial infection. *Cell Host Microbe* **12**: 177–186.

Bostanci N, Belibasakis GN. (2012). *Porphyromonas gingivalis*: an invasive and evasive opportunistic oral pathogen. *FEMS Microbiol Lett* **333**: 1–9.

Botta GA, Radin L, Costa A, Schito G, Blasi G. (1985). Gas–liquid chromatography of the gingival fluid as an aid in periodontal diagnosis. *J Periodontal Res* **20**: 450–457.

Brogden KA, Guthmiller JM, Taylor CE. (2005). Human polymicrobial infections. *Lancet* **365**: 253–255.

Casarin R, Barbagallo A, Meulman T, Santos V, Sallum E, Nociti F *et al.* (2013). Subgingival biodiversity in subjects with uncontrolled type-2 diabetes and chronic periodontitis. *J Periodontal Res* **48**: 30–36.

Chen L, Xiong Z, Sun L, Yang J, Jin Q. (2012). VFDB 2012 update: toward the genetic diversity and molecular evolution of bacterial virulence factors. *Nucleic Acids Res* **40**: D641–D645.

Darveau RP. (2010). Periodontitis: a polymicrobial disruption of host homeostasis. *Nat Rev Microbiol* **8**: 481–490.

Dashper SG, Seers CA, Tan KH, Reynolds EC. (2011a). Virulence factors of the oral spirochete *Treponema denticola*. *J Dent Res* **90**: 691–703.

Eckburg PB, Bik EM, Bernstein CN, Purdom E, Dethlefsen L, Sargent M *et al.* (2005). Diversity of the human intestinal microbial flora. *Science* **308**: 1635–1638.

Fuerst JA, Sagulenko E. (2010). Protein uptake by bacteria: an endocytosis-like process in the planctomycete *Gemmata obscuriglobus*. *Commun Integr Biol* **3**: 572–575.

Ge X, Rodriguez R, Trinh M, Gunsolley J, Xu P. (2013). Oral microbiome of deep and shallow dental pockets in chronic periodontitis. *PLoS One* **8**: e65520.

Gharbia SE, Shah HN. (1991). Pathways of glutamate catabolism among *Fusobacterium* species. *J Gen Microbiol* **137**: 1201–1206.

Gillespie JH. (1991). The burden of genetic load. *Science* **254**: 1049.

Grenier D. (1992a). Nutritional interactions between two suspected periodontopathogens, *Treponema denticola* and *Porphyromonas gingivalis*. *Infect Immun* **60**: 5298–5301.

Grenier D. (1992b). Demonstration of a bimodal coaggregation reaction between *Porphyromonas gingivalis* and *Treponema denticola*. *Oral Microbiol Immunol* **7**: 280–284.

Grenier D, La VD. (2011). Proteases of *Porphyromonas gingivalis* as important virulence factors in periodontal disease and potential targets for plant-derived compounds: a review article. *Curr Drug Targets* **12**: 322–331.

Griffen A, Beall C, Campbell J, Firestone N, Kumar P, Yang Z *et al.* (2012). Distinct and complex bacterial profiles in human periodontitis and health revealed by 16S pyrosequencing. *ISME J* **6**: 1176–1185.

Hament JM, Kimpen JL, Fleer A, Wolfs TF. (1999). Respiratory viral infection predisposing for bacterial disease: a concise review. *FEMS Immunol Med Microbiol* **26**: 189–195.

Haurat MF, Aduse-Opoku J, Rangarajan M, Dorobantu L, Gray MR, Curtis MA *et al.* (2011). Selective sorting of cargo proteins into bacterial membrane vesicles. *J Biol Chem* **286**: 1269–1276.

Hellström MK, Ramberg P, Krok L, Lindhe J. (1996). The effect of supragingival plaque control on the subgingival microflora in human periodontitis. *J Clin Periodontol* **23**: 934–940.

Holt SC, Kesavalu L, Walker S, Genco CA. (1999). Virulence factors of *Porphyromonas gingivalis*. *Periodontol 2000* **20**: 168–238.

Holt SC, Ebersole JL. (2005). *Porphyromonas gingivalis*, *Treponema denticola*, and Tannerella forsythia: the 'red complex,' a prototype polybacterial pathogenic consortium in periodontitis. *Periodontol 2000* **38**: 72–122.

Honma K, Kuramitsu HK, Genco RJ, Sharma A. (2001). Development of a gene inactivation system for *Bacteroides forsythus*: construction and characterization of a BspA mutant. *Infect Immun* **69**: 4686–4690.

Huson DH, Bryant D. (2006). Application of phylogenetic networks in evolutionary studies. *Mol Biol Evol* **23**: 254–267.

Ingham HR, Sisson PR. (1984). Pathogenic synergism. *Microbiol Sci* **1**: 206–208.

Irfan UM, Dawson DV, Bissada NF. (2001). Epidemiology of periodontal disease: a review and clinical perspectives. *J Int Acad Periodontol* **3**: 14–21.

Ishihara K. (2010). Virulence factors of *Treponema denticola*. *Periodontol 2000* **54**: 117–135.

Kastenmüller G, Gasteiger J, Mewes HW. (2008). An environmental perspective on large-scale genome clustering based on metabolic capabilities. *Bioinformatics* **24**: i56–i62.

Kesavalu L, Sathishkumar S, Bakthavatchalu V, Matthews C, Dawson D, Steffen M *et al.* (2007). Rat model of polymicrobial infection, immunity, and alveolar bone resorption in periodontal disease. *Infect Immun* **75**: 1704–1712.

Kimura M. (1983). Diffusion model of intergroup selection, with special reference to evolution of an altruistic character. *Proc Natl Acad Sci USA* **80**: 6317–6321.

Kumar P, Griffen A, Barton J, Paster B, Moeschberger M, Leys E. (2003). New bacterial species associated with chronic periodontitis. *J Dent Res* **82**: 338–344.

Kuramitsu HK, Chen W, Ikegami A. (2005). Biofilm formation by the periodontopathic bacteria *Treponema denticola* and *Porphyromonas gingivalis*. *J Periodontol* **76**: 2047–2051.

Kurita-Ochiai T, Fukushima K, Ochiai K. (1995). Volatile fatty acids, metabolic by-products of periodontopathic bacteria, inhibit lymphocyte proliferation and cytokine production. *J Dent Res* **74**: 1367–1373.

Kurita-Ochiai T, Ochiai K. (2010). Butyric acid induces apoptosis via oxidative stress in Jurkat T-cells. *J Dent Res* **89**: 689–694.

Ladas S, Arapakis G, Malamou-Ladas H, Palikaris G, Arseni A. (1979). Rapid diagnosis of anaerobic infections by gas-liquid chromatography. *J Clin Pathol* **32**: 1163–1167.

Loozen G, Ozcelik O, Boon N, De Mol A, Schoen C, Quirynen M *et al.* (2014). Inter-bacterial correlations in subgingival biofilms: a large-scale survey. *J Clin Periodontol* **41**: 1–10.

Lopez-Sanchez MJ, Sauvage E, Da Cunha V, Clermont D, Ratsima Hariniaina E, Gonzalez-Zorn B *et al.* (2012). The highly dynamic CRISPR1 system of *Streptococcus agalactiae* controls the diversity of its mobilome. *Mol Microbiol* **85**: 1057–1071.

Makarova KS, Haft DH, Barrangou R, Brouns SJ, Charpentier E, Horvath P *et al.* (2011). Evolution and classification of the CRISPR-Cas systems. *Nat Rev Microbiol* **9**: 467–477.

Margolis HC, Duckworth JH, Moreno EC. (1988). Composition and buffer capacity of pooled starved plaque fluid from caries-free and caries-susceptible individuals. *J Dent Res* **67**: 1476–1482.

Marraffini LA, Sontheimer EJ. (2008). CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. *Science* **322**: 1843–1845.

Mashburn LM, Jett AM, Akins DR, Whiteley M. (2005). *Staphylococcus aureus* serves as an iron source for *Pseudomonas aeruginosa* during *in vivo* coculture. *J Bacteriol* **187**: 554–566.

Marchler-Bauer A, Lu S, Anderson JB, Chitsaz F, Derbyshire MK, DeWeese-Scott C *et al.* (2011). CDD: a Conserved Domain Database for the functional annotation of proteins. *Nucleic Acids Res* **39**: D225–D229.

Mills SW, Montgomery SH, Morck DW. (2006). Evaluation of the effects of short-chain fatty acids and extracellular pH on bovine neutrophil function *in vitro*. *Am J Vet Res* **67**: 1901–1907.

Moriya Y, Itoh M, Okuda S, Yoshizawa AC, Kanehisa M. (2007). KAAS: an automatic genome annotation and pathway reconstruction server. *Nucleic Acids Res* **35**: W182–W185.

Naito M, Hirakawa H, Yamashita A, Ohara N, Shoji M, Yukitake H *et al.* (2008). Determination of the genome sequence of *Porphyromonas gingivalis* strain ATCC 33277 and genomic comparison with strain W83 revealed extensive genome rearrangements in *P. gingivalis*. *DNA Res* **15**: 215–225.

Naito M, Sato K, Shoji M, Yukitake H, Ogura Y, Hayashi T *et al.* (2011). Characterization of the *Porphyromonas gingivalis* conjugative transposon CTnPg1: determination of the integration site and the genes essential for conjugal transfer. *Microbiology* **157**: 2022–2032.

Niederman R, Buyle-Bodin Y, Lu BY, Robinson P, Naleway C. (1997). Short-chain carboxylic acid concentration in human gingival crevicular fluid. *J Dent Res* **76**: 575–579.

Niu S, Luo M, Tang J, Zhou H, Zhang Y, Min X *et al.* (2013). Structural basis of the novel S. pneumoniae virulence factor, GHIP, a glycosyl hydrolase 25 participating in host-cell invasion. *PLoS One* **8**: e68647.

Ohara-Nemoto Y, Rouf SM, Naito M, Yanase A, Tetsuo F, Ono T *et al.* (2014). Identification and characterization of prokaryotic dipeptidyl-peptidase 5 from *Porphyromonas gingivalis*. *J Biol Chem* **289**: 5436–5448.

Ohta T, Ina Y. (1995). Variation in synonymous substitution rates among mammalian genes and the correlation between synonymous and non-synonymous divergences. *J Mol Evol* **41**: 717–720.

Panos C, Ajl Sj. (1963). Metabolism of microorganisms as related to their pathogenicity. *Annu Rev Microbiol* **17**: 297–328.

Papapanou PN, Engebretson SP, Lamster IB. (1999). Current and future approaches for diagnosis of periodontal diseases. *N Y State Dent J* **65**: 32–37.

Paster B, Boches S, Galvin J, Ericson R, Lau C, Levanos V *et al.* (2001). Bacterial diversity in human subgingival plaque. *J Bacteriol* **183**: 3770–3783.

Rao D, Webb JS, Kjelleberg S. (2005). Competitive interactions in mixed-species biofilms containing the marine bacterium *Pseudoalteromonas tunicata*. *Appl Environ Microbiol* **71**: 1729–1736.

Raymond B, West SA, Griffin AS, Bonsall MB. (2012). The dynamics of cooperative bacterial virulence in the field. *Science* **337**: 85–88.

Roszak DB, Colwell RR. (1987). Survival strategies of bacteria in the natural environment. *Microbiol Rev* **51**: 365–379.

Sakakibara J, Nagano K, Murakami Y, Higuchi N, Nakamura H, Shimozato K *et al.* (2007). Loss of adherence ability to human gingival epithelial cells in S-layer protein-deficient mutants of *Tannerella forsythensis*. *Microbiology* **153**: 866–876.

Scott JJ, Oh DC, Yuceer MC, Klepzig KD, Clardy J, Currie CR. (2008). Bacterial protection of beetle–fungus mutualism. *Science* **322**: 63.

Semenova E, Jore MM, Datsenko KA, Semenova A, Westra ER, Wanner B *et al.* (2011). Interference by clustered regularly interspaced short palindromic repeat (CRISPR) RNA is governed by a seed sequence. *Proc Natl Acad Sci USA* **108**: 10098–10103.

Sharma A. (2010). Virulence mechanisms of *Tannerella forsythia*. *Periodontol 2000* **54**: 106–116.

Shoaie S, Karlsson F, Mardinoglu A, Nookaew I, Bordel S, Nielsen J. (2013). Understanding the interactions between bacteria in the human gut through metabolic modeling. *Sci Rep* **3**: 2532.

Soares SC, Silva A, Trost E, Blom J, Ramos R, Carneiro A *et al.* (2013). The pan-genome of the animal pathogen *Corynebacterium pseudotuberculosis* reveals differences in genome plasticity between the biovar ovis and equi strains. *PLoS One* **8**: e53818.

Socransky S, Haffajee A, Cugini M, Smith C, Kent R. (1998). Microbial complexes in subgingival plaque. *J Clin Periodontol* **25**: 134–144.

Sorek R, Lawrence CM, Wiedenheft B. (2013). CRISPR-mediated adaptive immune systems in bacteria and archaea. *Annu Rev Biochem* **82**: 237–266.

Steeves CH, Potrykus J, Barnett DA, Bearne SL. (2011). Oxidative stress response in the opportunistic oral pathogen *Fusobacterium nucleatum*. *Proteomics* **11**: 2027–2037.

Stern A, Keren L, Wurtzel O, Amitai G, Sorek R. (2010). Self-targeting by CRISPR: gene regulation or autoimmunity? *Trends Genet* **26**: 335–340.

Stringfellow WT, Aitken MD. (1995). Competitive metabolism of naphthalene, methylnaphthalenes, and fluorene by phenanthrene-degrading pseudomonads. *Appl Environ Microbiol* **61**: 357–362.

Sun S, Ke R, Hughes D, Nilsson M, Andersson DI. (2012). Genome-wide detection of spontaneous chromosomal rearrangements in bacteria. *PLoS One* **7**: e42639.

Tan KH, Seers CA, Dashper SG, Mitchell HL, Pyke JS, Meuric V *et al.* (2014). *Porphyromonas gingivalis* and *Treponema denticola* exhibit metabolic symbioses. *PLoS Pathogen* **10**: e1003955.

Tettelin H, Riley D, Cattuto C, Medini D. (2008). Comparative genomics: the bacterial pan-genome. *Curr Opin Microbiol* **11**: 472–477.

Tomida S, Nguyen L, Chiu BH, Liu J, Sodergren E, Weinstock GM *et al.* (2013). Pan-genome and comparative genome analyses of propionibacterium acnes reveal its genomic diversity in the healthy and diseased human skin microbiome. *MBio* **4**: e00003–e00013.

Toyofuku M, Nakajima-Kambe T, Uchiyama H, Nomura N. (2010). The effect of a cell-to-cell communication molecule, *Pseudomonas* quinolone signal (PQS), produced by *P. aeruginosa* on other bacterial species. *Microbes Environ* **25**: 1–7.

Tribble GD, Rigney TW, Dao DH, Wong CT, Kerr JE, Taylor BE *et al.* (2012). Natural competence is a major mechanism for horizontal DNA transfer in the oral pathogen *Porphyromonas gingivalis*. *MBio* **3**: e00231–11.

Venter JC, Remington K, Heidelberg JF, Halpern AL, Rusch D, Eisen JA *et al.* (2004). Environmental genome shotgun sequencing of the Sargasso Sea. *Science* **304**: 66–74.

Vesey PM, Kuramitsu HK. (2004). Genetic analysis of *Treponema denticola* ATCC 35405 biofilm formation. *Microbiology* **150**: 2401–2407.

Watanabe T, Nozawa T, Aikawa C, Amano A, Maruyama F, Nakagawa I. (2013). CRISPR regulation of intraspecies diversification by limiting IS transposition and intercellular recombination. *Genome Biol Evol* **5**: 1099–1114.

Westra ER, Buckling A, Fineran PC. (2014). CRISPR-Cas systems: beyond adaptive immunity. *Nat Rev Microbiol* **12**: 317–326.

Xie H, Lin X, Wang BY, Wu J, Lamont RJ. (2007). Identification of a signalling molecule involved in bacterial intergeneric communication. *Microbiology* **153**: 3228–3234.

Yamada T, Letunic I, Okuda S, Kanehisa M, Bork P. (2011). iPath2.0: interactive pathway explorer. *Nucleic Acids Res* **39**: W412–W415.

Yang Z. (1997). PAML: a program package for phylogenetic analysis by maximum likelihood. *Comput Appl Biosci* **13**: 555–556.

Zhang Y, Heidrich N, Ampattu BJ, Gunderson CW, Seifert HS, Schoen C *et al.* (2013b). Processing-independent CRISPR RNAs limit natural transformation in *Neisseria meningitidis*. *Mol Cell* **50**: 488–503.

Zhao Y, Wu J, Yang J, Sun S, Xiao J, Yu J. (2012). PGAP: pan-genomes analysis pipeline. *Bioinformatics* **28**: 416–418.

Zhou CE, Smith J, Lam M, Zemla A, Dyer MD, Slezak T. (2007). MvirDB—a microbial database of protein toxins, virulence factors and antibiotic resistance genes for bio-defence applications. *Nucleic Acids Res* **35**: D391–D394.

Zhu Y, Yang ST. (2004). Effect of pH on metabolic pathway shift in fermentation of xylose by *Clostridium tyrobutyricum*. *J Biotechnol* **110**: 143–157.

Supplementary Information accompanies this paper on The ISME Journal website (http://www.nature.com/ismej)