Chromosomal inversion between *rrn* operons among *Streptococcus mutans* serotype c oral and blood isolates

Wen-Chuan Huang,1 Yi-Ywan M. Chen,2 Lee-Jene Teng,3 Huwei-Ting Lien,1 Jen-Yang Chen1,4 and Jean-San Chia1,5

1Graduate Institute of Microbiology, College of Medicine, National Taiwan University, Taipei, Taiwan, ROC
2Department of Microbiology and Immunology, College of Medicine, Chang Gung University, Taipei, Taiwan, ROC
3Department of Clinical Laboratory Sciences and Medical Biotechnology, Taipei, Taiwan, ROC
4National Health Research Institute, Taipei, Taiwan, ROC
5Department of Dentistry, National Taiwan University Hospital, Taipei, Taiwan, ROC

*Streptococcus mutans* causes dental caries and infective endocarditis. The aim of this study was to determine genomic diversity among serotype c *S. mutans* laboratory and clinical strains and to characterize the genetic events involved. A genome-based approach using PFGE coupled with Southern hybridization was employed to examine a total of 58 serotype c oral and blood isolates and seven laboratory strains and to compare them with *S. mutans* UA159. No significant differences were found in the phenotypic characteristics of the strains tested, except that some of the strains exhibited smooth rather than rough colony morphology. In contrast, PFGE profiles of clinical isolates, from either diseased or healthy subjects, exhibited diverse patterns, suggesting that recombination or point mutations occurred frequently in vivo. Diverse PFGE patterns, with various lengths of insertions and deletions, could be detected even within a localized chromosomal region between *rrn* operons. Comparative analysis using Southern hybridization with specific markers revealed that a large chromosomal inversion had also occurred between *rrn* operons in 25 strains.

**INTRODUCTION**

*Streptococcus mutans* together with members of mutans streptococci are the primary aetiologic agents of dental caries in humans (Loesche, 1986). *S. mutans*, and other viridans streptococci, such as *Streptococcus sanguinis* and *Streptococcus oralis*, are also common causes of infective endocarditis (Mylonakis & Calderwood, 2001). Based on the serological properties of cell-wall polysaccharide antigens, nine serotypes (Nomura *et al.*, 2005) and seven distinct species have been described in mutans streptococci, indicating a substantial heterogeneity among mutans streptococci. Of all serotypes, serotype c *S. mutans* is the most prevalent isolate found in humans, irrespective of age or race of subjects studied (Loesche, 1986). The whole genomic sequence of a serotype c *S. mutans* strain (UA159) was completed recently (Ajdic *et al.*, 2002). The unique pathogenetic potential of *S. mutans* in distinct tissue compartments, oral cavity or heart valves, is attributed in part to its ability to establish biofilms on teeth or damaged tissue matrix. The natural competence of *S. mutans* and the close contact with other bacteria in the dental biofilm have laid a perfect foundation for the promotion of horizontal gene transfer for the benefit of adaptation or survival in distinct environments (Davey & O’Toole, 2000).

Although serotyping is used broadly for identification of *S. mutans*, phenotypic or biochemical diversity is often observed within the same serotype. Certain serotype c clinical isolates are unable to ferment melibiose (Beighton *et al.*, 1991), and both rough and smooth colony formers on Mitis Salivarius (MS) agar are found within the serotype c isolates (Okahashi *et al.*, 1984). Clonal diversity of *S. mutans* depending on colonization niches has also been suggested (Gronroos & Alaluusua, 2000). Although it has been postulated that, in addition to interspecies differences among mutans streptococci, heterogeneity exists among closely related *S. mutans* isolates belonging to the same serotype c, genetic evidence is still lacking. In this study, a genome-based approach using PFGE coupled with Southern hybridization was conducted on a total of 58 serotype c oral and blood isolates as well as seven
laboratory strains. We have confirmed that genome diversity exists among serotype c clinical isolates and have also provided evidence to indicate that insertion, deletion and inversion between rrn operons contribute to the diversity.

METHODS

Bacterial strains and growth conditions. The seven serotype c laboratory strains included in this study were UA159 (ATCC 700610), GS-5 (Kuramitsu, 1975), MT8148 (Okahashi et al., 1984), JH1005 (Hillman et al., 1984), PC3370 (Crowley et al., 1999), Xc (Koga, et al., 1989) and NTU-5R. Twenty-five of the 50 clinical strains studied were isolated from the oral cavities of healthy individuals (caries-free) and the other 25 strains were from diseased (caries-active) individuals. In the caries-active group, samples were collected from dental plaque overlying untreated carious lesions from different sites (pooled plaque samples), whereas stimulated saliva samples were collected from caries-free individuals. S. mutans was isolated from the samples following serial dilution and plating on Mitis Salivarius (MS) agar plates (Difco). Selected colony-forming units from each donor were subsequently confirmed by biochemical reactions as well as by the serotyping analysis described below. Eight additional serotype c S. mutans strains that had been isolated previously from blood cultures of patients with infective endocarditis, were also included in the study. Approval for collection and investigation of human specimens was obtained from the Research Ethics Committee of the National Taiwan University Hospital.

All S. mutans strains used were grown routinely in Brain Heart Infusion (BHI) broth (Difco) and on MS agar at 37°C. The ability of biofilm formation was evaluated by using the method of Hamada et al. (1976). The serotypes of all the clinical isolates were determined using antisera specific for serotypes c, d, e, f and g. Over 90% of the isolates belonged to serotype c and a total of 58 serotype c isolates were selected for further analysis.

Biofilm formation and transformation of S. mutans. The degree of biofilm formation was an important virulence trait of S. mutans. The degree of biofilm formation was evaluated by using the method of O’Toole & Kolter (1998), except that the bacteria were grown in 3 ml BHI medium overnight at 37°C. All of the strains tested in this study produced biofilm, albeit at different degrees. The density of the biofilm was quantified from triplicate assays by measuring optical density at 570 nm after staining with crystal violet. Statistical analysis was conducted by using analysis of variance (ANOVA) and differences were considered significant when the error probability was less than 0.05. Most of the isolates formed biofilm at a density comparable to that observed for strain UA159.

Preparation of competent S. mutans cells and transformation were performed as described by Perry & Kuramitsu (1981). A shuttle-plasmid, pDL278 (Aspiras et al., 2000), was used to determine the transformation efficiency for each strain. All test strains were transformable, with variable transformation efficiencies ranging from 10 to 106 c.f.u. (μg plasmid DNA)–1.

DNA manipulation and PCR. Genomic DNA from S. mutans was isolated as described previously (Chia et al., 1991). The PCR primers used in this study are listed in Table 1. For PCR products greater than 15 kbp, 400 μM dNTPs and 2.5 U LA Taq polymerase (Takara Bio Inc.) were used in each reaction. All reactions were carried out according to the manufacturer’s instructions.

PFGE and Southern blots. Genomic DNA fragments from PFGE agarose gels was carried out at high stringency, as described previously (Chia et al., 1991). The hybridization probes were generated by PCR and were based on the genome sequence of strain UA159. All probes were labelled with [α-32P]dATP by random priming (NEBlot kit; New England Biolabs).

Statistical analysis. Analysis of the characteristics of the laboratory and clinical strains was performed using the chi-square test. P values of <0.0001 were regarded as significant. ANOVA was used to analyse the variance of biofilm production. P values of <0.05 were regarded as significant.

RESULTS AND DISCUSSION

Heterogeneity of serotype c S. mutans strains

The serotypes of all the clinical isolates were determined using antisera specific for serotypes c, d, e, f and g, as detailed in Methods, and 58 serotype c isolates were selected for further analysis. Phenotypic characterization showed that all test strains exhibited identical biochemical characteristics, with the exception of strains 16-1 and 16-7, which were negative for β-galactosidase. In addition, all but five strains (16-1, 16-7, 37, 41-2 and 8585) were able to ferment melibiose. Regardless of the minor differences in biochemical profiles and variable transformation
frequency, biofilm formation was quite compatible in all strains examined. To test whether genomic diversity is a general phenomenon within serotype c S. mutans strains, chromosomal DNA from all 58 clinical isolates and the seven laboratory strains was subjected to PFGE analysis. Total chromosomal DNA from all strains was digested initially with NotI or I-CeuI to completion and separated using PFGE. I-CeuI is an intron-encoded endonuclease that recognizes a unique 19 bp sequence present in the rrn locus for 23S rRNA. The reproducibility of PFGE was confirmed by using UA159 as the reference strain in each run. The PFGE profiles of all strains were confirmed in at least triplicate gels. The PFGE profiles of either NotI or I-CeuI digests revealed a substantial polymorphism, encompassing the whole genome present among all isolates, including the laboratory strains, and it was too complicated to clearly categorize the isolates into simple groups (data not shown).

For strain UA159, complete digestion with I-CeuI constantly revealed three consecutive bands of 189, 177 and 169 kbp. Albeit a substantial polymorphism was present in terms of the size of the I-CeuI fragments, all isolates, including the laboratory strains, exhibited the same number of I-CeuI fragments.

Characterization of the large chromosomal rearrangement in laboratory strains

Using PFGE analysis, Cappiello et al. (1999) constructed a genetic map of S. mutans GS-5. Based on the complete genome of strain UA159, we constructed a genetic map of UA159 using the Vector NTI Advance software package (data not shown). Comparative analysis of genetic maps of UA159 using the Vector NTI Advance software package (data not shown).

Previously, we characterized a S. mutans laboratory strain (NTU-5R) that did not contain the gbpA locus, and found that the deletion was located between rrrB and rrrC (Chia and others, unpublished data). Therefore, strains GS-5 and NTU-5R were selected for further genetic characterization, with an attempt to explore the genetic elements or events that contribute to the macrorestriction polymorphism observed among S. mutans strains.

Based on a computer-aided BLASTN search, a series of probes (Table 1) specific for each NotI fragment of S. mutans UA159 were constructed and were used to locate corresponding regions on the chromosomes of S. mutans

### Table 1. PCR primers used in this study

<table>
<thead>
<tr>
<th>Name</th>
<th>Forward*</th>
<th>Reverse*</th>
<th>Length of PCR product (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probe 1 (NotI fragment H-specific)</td>
<td>719893–719910</td>
<td>721316–721333</td>
<td>1440</td>
</tr>
<tr>
<td>Probe 2 (NotI fragment G-specific)</td>
<td>563789–563806</td>
<td>564568–564585</td>
<td>796</td>
</tr>
<tr>
<td>Probe 3 (NotI fragment A-specific)</td>
<td>514905–514922</td>
<td>516013–516030</td>
<td>1125</td>
</tr>
<tr>
<td>Probe 4 (NotI fragment A-specific)</td>
<td>183750–183767</td>
<td>185197–185214</td>
<td>1464</td>
</tr>
<tr>
<td>Probe 5 (NotI fragment F-specific)</td>
<td>1974287–1974304</td>
<td>1975250–1975267</td>
<td>980</td>
</tr>
<tr>
<td>Probe 6 (NotI fragment E-specific)</td>
<td>1938418–1938435</td>
<td>1940330–1940347</td>
<td>1929</td>
</tr>
<tr>
<td>Probe 7 (NotI fragment D-specific)</td>
<td>1701070–1701087</td>
<td>1702453–1702470</td>
<td>1400</td>
</tr>
<tr>
<td>Probe 8 (NotI fragment B-specific)</td>
<td>1324473–1324490</td>
<td>1325547–1325564</td>
<td>1091</td>
</tr>
<tr>
<td>Probe 9 (NotI fragment C-specific)</td>
<td>956792–956809</td>
<td>957478–957495</td>
<td>703</td>
</tr>
</tbody>
</table>

*The positions on the chromosome of S. mutans UA159 (GenBank accession no. AE014133) are listed.
GS-5 and NTU-5R. When chromosomal DNA was hybridized with probes specific for the B, C, G and H fragments (probes 8, 9, 2 and 1), signals with sizes predicted by the maps were observed in both GS-5 and NTU-5R (data not shown). These results indicate that the arrangement between NotI -B, -C, -H, and -G fragments is conserved among all three S. mutans strains. When chromosomal DNA was hybridized with two A-specific probes (probes 3 and 4 – one to each end of the A fragment), two signals, approximately 430 and 280 kbp, corresponding to GS-5 NotI-A and -D fragments, respectively, were observed in strain GS-5, but not in NTU-5R (Fig. 1b, probes 3 and 4). Furthermore, the same 430 kbp NotI fragment of strain GS-5 also hybridized to probe 6 (specific for fragment E), confirming inversion of S. mutans UA159 NotI-A and -E fragments in S. mutans GS-5. In contrast, with the NTU-5R chromosome, all probes (probes 3, 4, and 6) hybridized to a single NotI fragment of approximately 800 kbp (the size was confirmed in separate PFGE experiments under running conditions for resolution between 100 and 870 kbp) (Fig. 1b). Similarly, when the same blot was hybridized with probe 5, whose location was close to the juncture of fragments E and F on the UA159 chromosome, a signal of approximately 140 kbp was detected in GS-5, but not in NTU-5R (Fig. 1b, probe 5). This probe hybridized to the same NotI fragment that also hybridized with probes 3, 4 and 6 in GS-5, indicating that sequence alterations resulting in the loss of NotI recognition sequences had occurred in NTU-5R.

To further confirm the genomic arrangements in strains UA159 and GS-5, the hybridization patterns of I-CeuI-digested chromosomal DNA with probes 4, 6 and 7 were also determined (Fig. 1c). In agreement with the hybridization pattern of NotI-digested chromosomal DNA, signals with predicted sizes as found in UA159 were also observed in GS-5 with all three probes, indicating the presence of a genomic inversion that might occur through the combination of neighbouring rrn operons. In contrast, different hybridization patterns of I-CeuI fragments were observed in NTU-5R (Fig. 1c), suggesting that insertions within the I-CeuI-B and -C fragments (probes 4 and 6) and deletions within the I-CeuI-D fragment (probe 7) may have
occurred in the NTU-5R chromosome. Together, these data indicate that, compared with strain UA159, a large chromosomal inversion had occurred between \textit{rrn} operons in \textit{S. mutans} GS-5 and that deletion (other than \textit{gbpA}) could also occur in these regions, as found in NTU-5R.

**Chromosomal inversion in clinical isolates**

Based on the hybridization pattern of strain GS-5, we also investigated the \textit{NstI}-digested PFGE patterns of all isolates to predict whether recombination between \textit{rrnA} and \textit{rrnC} might have occurred. If inversion had occurred, probes 3 and 6 would hybridize to the same signals, whereas probe 4 would not. As shown in Fig. 2, in isolates NTU-1020, -3432 and -3930, the hybridization patterns obtained using probes 3 and 6 were identical, and probe 4 also hybridized to a signal (260 kbp) similar to that found in GS-5. These results indicated that chromosomal inversion might also occur in these three clinical isolates. Similar approaches indicated that a total of 24 clinical isolates (41.3\%) and one laboratory strain (14.3\%) exhibited \textit{NstI}-digested PFGE patterns similar to those of GS-5, suggesting that recombination might have occurred in both clinical and laboratory isolates.

**Characterization of the large chromosomal shuffling in strain NTU-5R**

Genomic analysis of gene positions in closely related micro-organisms suggests that replication forks are preferred sites for DNA recombination (Tillier & Collins, 2000). We noticed that the I-\textit{CeuI}-C fragments of strain UA159 contained the replication origin (\textit{ori}) and that an inversion had occurred in this region between strains UA159 and GS-5. Our analysis of the NTU-5R genome also indicated that a major deletion and/or sequence shuffling

\[\text{Fig. 2. PFGE and hybridization patterns of \textit{NstI}-digested chromosomes of } S. \textit{mutans} \text{ strains. (a) Digested chromosomal DNA from } S. \textit{mutans} \text{ NTU-1020, 1037, 3432, 3368 and 3930 were hybridized with various probes as indicated in Fig. 1(a). (b) PFGE of \textit{NstI}-digested DNA from laboratory } S. \textit{mutans} \text{ strains. S, R and U indicate chromosomes from strains GS-5, NTU-5R and UA159, respectively. 1020, 1037, 3432, 3368 and 3930 are the same as in (a). The sizes of restriction fragments of UA159 were used as a reference.}\]
had occurred within this region (Fig. 1). To further characterize the differences within this region between UA159, GS-5 and NTU-5R, an extensive series of PCR primers was designed based on the UA159 genome, and was used to amplify long stretches of DNA fragments within the I-Ceu I-C-fragment (Fig. 3a). Amplicons of the predicted sizes were observed only in UA159 and GS-5 with all pairs of primers used (Fig. 3b). PCRs with NTU-5R chromosomal DNA yielded a larger amplicon in the 5th region and a smaller one in the 7th region. The estimated insertion and deletion are 8 and 10 kbp, respectively. In addition, the deletion observed in the 7th region contained the gbpA locus (Fig. 3b). To further analyse the deleted region, chromosomal walking by PCR was performed and the nucleotide sequences of the PCR-amplified fragments were determined by direct sequencing. A series of 10 pairs of primers was included to encompass a total length of 16 kb flanking the 7th region of I-Ceu-C. The overlapping and contiguous PCR fragments obtained were about 1–4 kb, named segments a, b, c, d, e, f, g, h, i and j, as indicated in Fig. 3(a). Surprisingly, sequencing results confirmed that all PCR fragments, from a to j, derived from NTU-5R exhibited entities that were almost identical to that from UA159 (Fig. 4b). When PCR amplification of larger fragments was conducted by combining primers from segments c and h, a correct amplicon with a size of around 10 kbp was obtained (Fig. 4a, lane U). However, PCR amplicons obtained from NTU-5R consisted of two shorter fragments of 5 and 2.5 kbp (Fig. 4a, lane R). Further analysis using primers derived from segments e and g revealed similar PCR fragments (e–g) around 8 kbp from DNA from either NTU-5R or UA159. Therefore, a positive identity of fragments e–g found in NTU-5R suggested that this region was not deleted but might have located elsewhere in the chromosome. Taken together, these results suggest that sequence shuffling might have

Fig. 3. Analysis of the region between rnb and rnc using long-PCR. (a) Physical maps of S. mutans UA159 chromosomes digested with I-Ceu. The estimated size of each fragment is shown in kbp. Diagrammatic representations of the relative positions of the primers are indicated by arrows. The rns are indicated by open triangles. The relative location of gbpA is indicated. (b) Agarose gel electrophoresis of PCR products generated from S. mutans strains. Lanes 1, 2 and 3 show chromosomes from GS-5, NTU-5R and UA159, respectively.
In conclusion, we have demonstrated that genomic diversity observed at a significantly higher frequency in isolates from S. mutans exists among isolates from different colonizing niches in the volunteers tested. The most interesting findings were that genomic insertion, deletion and inversion occur frequently in S. mutans. Interestingly, in that report, the deletion of gbpA was also identified in several strains and the deletion was replaced by insertion of other genes as well as IS elements. The most recent microarray analysis indicated that a total of 385 (20%) of the UA159 open-reading frames were absent from one or more of the nine test S. mutans strains (Waterhouse et al., 2007). We found that IS199 (Macrina et al., 1996) was present in NTU-5R, but was absent from both GS-5 and UA159, by using PCR analysis (data not shown). However, chromosomal walking and sequence analysis of the 10 kb area of S. mutans UA159, we were able to detect a PCR product with gbpA at frequencies of 16, 4 and 12.5 %, respectively (Table 2). However, whether deletion in addition to gbpA or the underlying mechanism of chromosomal shuffling is involved awaits further investigation.

Deletion of the virulence gene gbpA among clinical isolates

The gbpA gene encodes one of the glucan-binding proteins (GBPA) and contributes to the adherence and biofilm formation of S. mutans on the surfaces of teeth (Banas & Vickerman, 2003). Using primers specific for gbpA of S. mutans UA159, we were able to detect a PCR product with the predicted size in chromosomal DNA from S. mutans GS-5 but not from NTU-5R (data not shown), confirming a deletion of gbpA in NTU-5R. The gbpA-deficient genotype was also observed in isolates from caries-free and caries-active subjects, and patients with endocarditis, at frequencies of 16, 4 and 12.5 %, respectively (Table 2). Based on this small-scale survey, deletion of gbpA was observed at a significantly higher frequency in isolates from caries-active than caries-free or endocarditis individuals (P<0.0001; chi-square test).

In conclusion, we have demonstrated that genomic diversity exists among S. mutans serotype c laboratory and clinical isolates from different colonizing niches in the volunteers tested. The most interesting findings were that genomic insertion, deletion and inversion occur frequently in S. mutans even within rather localized and defined regions between rrn operons. A large genomic rearrangement across the replication axis or recombination between rrn regions has been shown to occur in M3 and M5 strains (Manfredo et al., 1990). The underlying mechanism of chromosomal shuffling is involved awaits further investigation.

**Fig. 4.** Analysis of a 10 kbp chromosomal fragment of strain NTU-5R by PCR. The position and primers used to generate internal fragments of the 7th region were those depicted in Fig. 3(a). (a) Agarose gel electrophoresis of long-PCR products generated from S. mutans UA159 (U) and NTU-5R (R). M, 1 kb DNA ladder; N, negative control of PCR reaction. Bands of predicted sizes were found for UA159 with a ‘c–h’ PCR product, about 10 kbp, and an ‘e–g’ PCR product of about 8 kbp. In contrast, a shorter fragment of ‘c–h’ but not ‘e–g’ was obtained in NTU-5R. (b) Agarose gel electrophoresis of PCR products generated from subfragments in the 7th region from UA159 (U) or NTU-5R (R). ‘c, d, e, f, g and h’ PCR products amplified from either UA159 or NTU-5R are of identical length and with DNA sequence homology around 97 % (data not shown).
fragment in the 7th region at the I-CeuI-C region of NTU-5R (Fig. 3) did not reveal elements belonging to IS199 or others. Therefore, IS elements might not be involved in the gene shuffling observed in NTU-5R.

**ACKNOWLEDGEMENTS**

We thank Drs H. K. Kuramitsu, Y. Hamada, S. F. Lee, A. Bleiweis and Y. Yamashita for providing strains GS-5, MT8148, JH1005, PC3370 and XC strains. We thank Drs J. Y. Shew and S. C. Chang for assistance in PGFE and GelCompare software analysis. We thank S. Hamada for the serotype f specific antisera. This work was supported in part by the National Science Council (grants NSC-912320-B002-101, NSC-922320-B002-166, NSC-932320-B002-041, NSC-942320-B002-007, NSC-952320-B002-086-MY3-1 and NSC-963112-B002-031), National Health Research Institute (grants NHRI-EX94-9432SI, NHRI-EX95-9432SI and NHRI-EX96-9432SI), Chang Gung Memorial Hospital (grant CMRPD34001) and Center for Disease Control (grant DOH94-DC-1006).

**REFERENCES**


