Molecular and clinical analyses of the gene encoding the collagen-binding adhesin of *Streptococcus mutans*

Ryota Nomura,1 Kazuhiko Nakano,1 Naho Taniguchi,1 Jinthana Lapirattanakul,1 Hirotoshi Nemoto,1 Lisa Grönoos,2 Satu Alaluusua2,3 and Takashi Ooshima1

1Department of Pediatric Dentistry, Osaka University Graduate School of Dentistry, Suita, Osaka, Japan
2Department of Oral and Maxillofacial Diseases, Helsinki University Central Hospital, Helsinki, Finland
3Department of Pediatric and Preventive Dentistry, Institute of Dentistry, University of Helsinki, Helsinki, Finland

*Streptococcus mutans* is a known pathogen of dental caries and its major cell surface antigens have been widely investigated. Recently, an approximately 120 kDa Cnm protein with binding properties to type I collagen was identified, and its encoding gene (*cnm*) cloned and sequenced. In the present study, we sequenced *cnm* from 47 different clinical *S. mutans* strains and found that the nucleotide alignment of the collagen-binding domain was well conserved. We devised a PCR method for identifying the *cnm* gene, examined the prevalence of *cnm*-positive *S. mutans* strains in various mother–child groups, and assessed the significance of such strains for transmission and dental caries. The detection rate of *cnm*-positive strains was significantly lower in strains isolated from Japanese children in the 2000s (8.0 %) as compared to those isolated in the 1980s (15.8 %) (P<0.05). Furthermore, the presence of *S. mutans* possessing *cnm* in salivary specimens collected from 55 *S. mutans*-positive mother–child pairs was 40 and 32.7 % in the mothers and children, respectively. The frequency of *cnm*-positive children whose mothers were also positive was 72 %, which was significantly higher than that of *cnm*-positive children with negative mothers (P<0.0001, odds ratio 17.5). In addition, clinical parameters indicating dental caries were significantly increased in children with *cnm*-positive *S. mutans* in saliva (n=13), as compared to those with *cnm*-negative *S. mutans* (n=15) and *S. mutans*-negative children (n=20) (P<0.01). These results indicate that *cnm*-positive *S. mutans* strains are closely correlated with dental caries, while vertical transmission in *cnm*-positive mother–child pairs was also demonstrated.

INTRODUCTION

*Streptococcus mutans*, known to be a major pathogen of dental caries, is classified into four serotypes (c, e, f and k) based on the chemical composition of the serotype-specific rhamnose glucose polymers (Hamada & Slade, 1980; Nakano et al., 2004a). The distribution frequency of those serotypes among clinical oral isolates has been investigated, which showed that the majority, approximately 70–80 %, were classified as serotype c, followed by e (approx. 20 %), with less than 5 % of the strains classified as serotype f or k (Hirasawa & Takada, 2003; Shibata et al., 2003; Nakano et al., 2004b).

The major cell surface protein antigens of *S. mutans*, which include three types of glucosyltransferases (Aoki et al., 1986; Pucci et al., 1987; Hanada & Kuramitsu, 1989), a 190 kDa protein antigen (PA) (Okahashi et al., 1989) and four types of glucan-binding proteins (Gbps) (Banas et al., 1997; Mattos-Graner et al., 2001; Sato et al., 1997; Shah & Russell, 2004), have been widely investigated in studies related to its virulence for dental caries. Recently, the 120 kDa Cnm protein encoding the collagen-binding adhesin of *S. mutans* was characterized, and its encoding gene was cloned and sequenced (Sato et al., 2004). The distribution frequency of the *cnm* gene in 102 *S. mutans* strains in our previous study was shown to be 21.4 %, with *cnm*-positive strains showing a predominant distribution among strains with the minor serotypes f and k, which...
were also demonstrated to have unique features (Nakano et al., 2007).

The development of dental caries results in the destruction of enamel on tooth surfaces, leading to dentin exposure (Selwitz et al., 2007). The Cnm protein has been shown to possess binding activity to type I collagen (Sato et al., 2004), a major organic component of dentin, while approximately 20% of S. mutans strains studied were also shown to possess collagen-binding activity, which is regarded as an advantage for binding to exposed dentin (Switalski et al., 1993). In the present study, the distribution of strains with the cnm gene was analysed using saliva specimens from children and their mothers. Furthermore, we investigated the correlation between subjects with cnm-positive strains and clinical parameters for dental caries.

**METHODS**

**Bacterial strains.** From 2002 to 2006, we isolated 213 oral strains of S. mutans from samples collected from 213 Japanese children who came to our clinic and 60 strains from 60 of their mothers who participated in oral hygiene lectures held at our clinic. In addition, 95 strains isolated from Japanese children who visited our clinic between 1982 and 1990, and 110 strains isolated in Finland in the early 1990s were investigated. All strains were confirmed to be S. mutans based on biochemical properties, such as positive sugar fermentation profiles for mannitol, sorbitol, raffinose and melibiose (1% each) in phenol red broth (Difco), and a rough colony morphology on Mitis Salivarius agar (Difco) plates containing bacitracin (0.2 U ml⁻¹; Sigma Chemical) and 15% (w/v) sucrose. Furthermore, confirmation of S. mutans and serotype determination were carried out by PCR using S. mutans-specific and serotype-specific sets of primers, which are listed in Table 1.

**Determination of cnm sequence.** The entire length of the cnm gene in all of the strains was amplified by PCR using TaKaRa Ex Taq polymerase (Takara Bio) with appropriate primers (cnm-1F and cnm-1R; Table 1, Fig. 1), which were constructed based on the cnm nucleotide alignment of strain Z1 in the GenBank database (GenBank accession no. AB102689). PCR amplification was performed in a thermal cycler (iCycler; Bio-Rad) with the following cycling parameters: initial denaturation at 95°C for 4 min, and then 30 cycles consisting 95°C for 30 s, 60°C for 30 s and 72°C for 2 min, with a final extension at 72°C for 7 min. The amplified fragments

![Fig. 1. Proposed structure of Cnm in S. mutans clinical isolates. (a) The putative structure of Cnm in representative strains (MT4289 and MT4363) in comparison to that in strain Z1 is illustrated. Horizontal lines indicate the entire length of the cnm gene. Arrows indicate the primer positions, designed based on the nucleotide alignment of the cnm gene of strain Z1. Filled rectangles show the ORFs, CBD, Collagen-binding domain. (b) Number of repeats in the B repeats region for 47 clinical isolates. The bold vertical bar indicates the mean value.](image-url)
mRNA expression of the Cnm-encoding gene. To confirm mRNA expression of the collagen-binding domain of the cnm gene, an RT-PCR method was used with cnm-CF and cnm-CR primers (Table 1), as described previously (Nomura et al., 2005). In order to confirm that cDNA was successfully extracted, amplification of 16S rRNA was performed using the following sets of primers: 5′-GGT GGA CCC AAG GAA ACA CAC TGT GC-3′ and 5′-CGT CGG CTT GGT AAG CTC TTA CCT TAC C-3′ (Matsumoto-Nakano & Kuramitsu, 2006).

Collagen-binding properties of clinical strains. Collagen-binding properties were evaluated according to the method described by Waterhouse & Russell (2006), with some modifications. Type I collagen [collagen (type I) in 0.25 M acetic acid; Sigma] was coated onto 96-well tissue culture plates (Beckton Dickinson) and incubated overnight at 4 °C, then the plates were washed three times with PBS and blocked for 1.5 h with 5 % BSA in PBS at 37 °C. Next, the wells were washed again with PBS containing 0.01 % Tween 20. Cells from overnight cultures of S. mutans grown in brain heart infusion broth (Difco) were collected by centrifugation and added to the wells, after the bacterial numbers were adjusted with PBS. After 3 h incubation at 37 °C, adherent cells were washed three times with PBS, then fixed with 200 μl 25 % formaldehyde at room temperature for 30 min. After another three washes with PBS, the adherent cells were stained with 200 μl 0.05 % crystal violet (Wako) in water for 1 min and washed three times with PBS, then the dye was dissolved by adding 7 % acetic acid (200 μl) before determining the Abs492. First, the correlations of the amount of collagen (0.002–2 mg) and the number of bacteria (1 × 108–1 × 1010 c.f.u.) with the collagen-binding activities of the TW871 (cnm-positive) and MT8148 (cnm-negative) strains were evaluated. Then, the collagen-binding activities of 99 clinical strains (47 cnm-positive and 52 cnm-negative strains) were analysed under the fixed condition of 2 mg collagen and 1 × 1010 bacterial cells. The results for each strain were expressed as a percentage compared to the binding property of TW871, which was defined as 100 %.

PCR methods for identification of strains with the cnm gene. Based on the nucleotide alignment determined for the 47 strains, primer sets designed to specifically detect cnm-positive strains were constructed (cnm-BF and cnm-1R; Table 1, Fig. 1). PCR amplification was performed using TaKaRa Ex Taq polymerase (Takara Bio) with the following cycling parameters: initial denaturation at 95 °C for 4 min, and then 30 cycles consisting of 95 °C for 30 s, 60 °C for 30 s and 72 °C for 1 min, with a final extension at 72 °C for 7 min. First, the specificity of our method was confirmed using 102 strains in which the presence of the cnm gene had been confirmed by Southern blot hybridization in our previous study (Nakano et al., 2007). In addition, PCR analysis was carried out for all bacterial strains used in this study whose complete sequence had been determined, with those results also used to confirm specificity. Finally, the sensitivity of the method was determined using diluted cultures of S. mutans NN2115 (serotype f). When analysed using simultaneous PCR assays with known numbers of bacterial cells diluted in sterile distilled water, the detection limit of our method was revealed to range from 10 to 100 cells.

Detection of S. mutans with the cnm gene in saliva specimens from children and their mothers. All of the procedures in the present study were approved by the Ethical Committee of the Osaka University Graduate School of Dentistry. Expectorated saliva specimens were collected from 101 mother–child pairs (mothers, 24–47 years of age; children, 47 boys and 54 girls, 2–15 years old) who participated in oral hygiene lectures at our clinic from 2005 to 2007. The specimens were processed for the PCR assay using a method reported previously (Hoshino et al., 2004). PCR analyses were performed to determine the presence of S. mutans DNA and the cnm gene using the method described in the section above.

Correlation of the presence of S. mutans species with the cnm gene and dental caries. In order to compare the incidence of dental caries in primary teeth, 48 children who came to our clinic from 2005 to 2007 (22 boys and 26 girls, 3–8 years of age) were selected based on the following criteria: precise clinical records were available regarding dental caries when the saliva specimens were taken, and the subject possessed more than 14 primary teeth and less than 8 permanent teeth, which corresponded to children under the age of 8. Clinical examinations were carried out with a mirror and explorer under a dental operation light by a single skilled examiner according to criteria established by the WHO (1987). The subjects were divided into three groups: those with cnm-positive S. mutans (group A), those with cnm-negative S. mutans (group B) and those without S. mutans detected (group C). The condition of the primary teeth was expressed as numbers of teeth that were decayed (d), extracted due to dental caries (e) and had filled surfaces (f) (dfs). The results are expressed as the ‘dfs rate’, which was calculated by dividing (d + e + f) by the total number of surfaces (s).

Statistical analyses. Statistical analyses were carried out using the computational software packages StatView 5.0 (SAS Institute) and Prism 4 (GraphPad Software). Collagen-binding activities of the cnm-positive and cnm-negative strains were compared using Student’s t-test. Fisher’s protected least-significant difference test was used to compare detection rates of S. mutans with the cnm gene in children whose mothers possessed cnm-positive S. mutans as compared to children whose mothers did not. The relationship between the intensity of the mRNA expression of the cnm gene and collagen-binding properties was analysed using regression analysis. The values for the odds ratio and 95 % confidence interval were calculated to determine the significance of the association between children with S. mutans with the cnm gene and their mothers possessing the same. Comparisons among scores for dental caries in subjects possessing cnm-positive S. mutans, those possessing cnm-negative S. mutans, and those without S. mutans were performed using one-way ANOVA (Bonferroni multiple comparison test). P values less than 0.05 were considered to be significant.

RESULTS AND DISCUSSION

Prevalence of cnm-positive S. mutans clinical isolates

A total of 47 of the 478 tested strains were shown to possess the cnm gene, thus the prevalence of S. mutans with cnm among oral isolates was 9.8 %. Table 2 lists the prevalence of cnm-positive strains based on serotype classification. The rate of detection for the cnm gene in serotype f strains was the highest, followed by serotype k, while 25 of 359 serotype c strains were found to be positive. A comparison of frequencies in Japanese children in different decades showed that strains isolated in 2002–2006 had a significantly lower
variation (Sato et al. 2004), whereas the numbers in the species analysed in the present study varied from 6 (strain MT4363) to 25 (strain MT4289), with the mean number for all 47 clinical strains being 19.7, and most strains contained 15–25 repeats (Fig. 1b). The number of repeats and the length of the repeating units were shown to be dependent on species.

Collagen-binding properties of clinical strains

A number of studies have noted the cellular adhesion of S. mutans to collagen, and strains with and without collagen-binding properties have been reported (Love et al., 1997; Petersen et al., 2002). Recently, the Cnm protein was shown to possess collagen-binding activity (Sato et al., 2004) and its distribution frequency was reported to be approximately 20% (Nakano et al., 2007), which is quite similar to that shown for S. mutans strains with collagen-binding properties (Switalski et al., 1993). Thus, the collagen-binding properties of S. mutans was investigated in the present study by focusing on the Cnm protein.

The collagen-binding properties of TW871 were shown to increase with corresponding increases in the amounts of collagen and bacteria (Fig. 2a, b), while strain MT8148 showed no collagen-binding properties. Furthermore, the assay with 99 clinical strains using 2 mg collagen and 10^{10} bacterial cells showed that the collagen-binding properties of strains with the cnm gene were significantly higher than those of strains without the gene (P<0.01) (Fig. 2c). With the properties of TW871 defined as 100%, the binding intensity of the strains with the cnm gene ranged from 0.0 to 217.9%. Approximately half of the strains had high binding properties, for 20% strains though the binding properties were extremely weak. RT-PCR results for the strains with extremely weak binding showed that the mRNA expression of the cnm gene in those was also extremely weak. In addition, RT-PCR analysis revealed that the mRNA expression of the cnm gene was extremely weak in 11 of the 47 strains, half of which were isolated from Japanese children in the 1980s to early 1990s.

### Table 2. Prevalence of cnm gene in S. mutans oral isolates from various populations

<table>
<thead>
<tr>
<th>Serotype</th>
<th>Proportion of strains isolated in the 1980s and early 1990s</th>
<th>Proportion of strains isolated in the 2000s</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>c</td>
<td>10/72 (13.9%)</td>
<td>0/18 (0%)</td>
<td>1/63 (1.6%)</td>
</tr>
<tr>
<td>e</td>
<td>3/20 (15.0%)</td>
<td>0/3 (0%)</td>
<td>2/17 (11.8%)</td>
</tr>
<tr>
<td>f</td>
<td>2/3 (66.7%)</td>
<td>0/2 (0%)</td>
<td>3/3 (100%)</td>
</tr>
<tr>
<td>k</td>
<td>No strains</td>
<td>1/1 (100%)</td>
<td>0/3 (0%)</td>
</tr>
<tr>
<td>Total</td>
<td>15/95 (15.8%)*</td>
<td>1/24 (4.2%)</td>
<td>6/86 (7.0%)</td>
</tr>
</tbody>
</table>

*P<0.05 between the two groups by Fisher’s exact probability test.

cnm gene in S. mutans is exogenous and derived from multiple species. In addition, the mRNA expression of the cnm gene was extremely weak in 11 of the 47 strains, half of which were isolated from Japanese children in the 1980s to early 1990s.

### Molecular characterization of cnm in S. mutans clinical isolates

Fig. 1(a) shows an illustration of the presumed Cnm protein structure in the clinical isolates analysed in the present study. The putative amino acid sequence of the cnm gene of strain Z1 (GenBank accession no. AB102689) was reported to contain 538 amino acids, which includes the collagen-binding domain (CBD; residues 152–316), putative B repeats domain (residues 328–455) and an LPXTG motif (residues 507–511) (Sato et al., 2004). This is the first study to perform molecular analyses of the cnm gene using a large number of S. mutans clinical isolates.

Our analysis of the nucleotide alignment of cnm in 47 clinical isolates showed that the number of amino acids ranged from 448 (strain MT4363) to 564 (strain MT4289). The amino acid sequence of the CBD in all of the strains was nearly the same and the LPXTG motif was confirmed to be conserved among all. As for the number of B repeats, strain Z1 was reported to contain 21 repeats (Sato et al., 2004), whereas the numbers in the species analysed in the present study varied from 6 (strain MT4363) to 25 (strain MT4289), with the mean number for all 47 clinical strains shown to be 19.7, and most strains contained 15–25 repeats (Fig. 1b). The number of repeats and the length of the repeating units were shown to be dependent on species variation (Sato et al., 2004), thus we speculated that the
Vertical transmission of *S. mutans* with *cnm* analysed using saliva specimens

Interfamilial transmission of *S. mutans* has been widely investigated using serotype classification (Berkowitz *et al.*, 1975), bacteriocin activity profiles (Davey & Rogers, 1984), chromosomal DNA fingerprinting (Emanuelsson & Thornqvist, 2000), ribotyping (Köhler *et al.*, 2003), arbitrarily primed PCR method (Li & Caufield, 1998) and multilocus sequence typing (Nakano *et al.*, 2007), with the results of most of those indicating mother-to-child transmission. However, few investigations of mother-to-child transmission have focused on specific virulent genes. Li *et al.* (2005) analysed the transmitted and non-transmitted strains focusing on the genes encoding mutacin, which showed that the strains with mutAI were found frequently in the non-transmitted group. In the present study, we focused on the *cnm* gene in order to determine whether *S. mutans* strains with or without the *cnm* gene were highly transmissible. Saliva specimens from 55 pairs of mothers and their children were positive for *S. mutans*, which were divided into 4 groups based on the absence or presence of the bacterium and the *cnm* gene (Fig. 3a). The *cnm* gene was detected in 32.7% of the *S. mutans*-positive children.

**Fig. 2.** Collagen-binding properties of *S. mutans* cells. (a) *S. mutans* MT8148 and TW871 cells (1×10⁹ c.f.u.) were added separately to wells coated with various amounts of collagen. (b) Various amounts of *S. mutans* MT8148 and TW871 cells (1×10⁸–1×10¹⁰ c.f.u.) were added separately to wells coated with 2 mg collagen. (c) *S. mutans* strains (1×10¹⁰ c.f.u.) with or without the *cnm* gene were added to plates coated with 2 mg collagen. Statistically significant differences were found between the two groups (Student's *t*-test **P<0.01). Horizontal bars indicate the mean values. □, MT8148: ■, TW871; ○, *cnm*-negative strain; ●, *cnm*-positive strain. The collagen binding of TW871 was defined as 100%.

**Fig. 3.** Detection of *S. mutans* with the *cnm* gene in saliva specimens from mother–child pairs. (a) The occurrence of strains with the *cnm* gene was analysed in children, and the subjects were categorized into those with mothers who did or did not possess *cnm*-positive strains. A total of 55 mother–child pairs shown to harbour *S. mutans* in saliva specimens were classified into four categories based on identification of the *cnm* gene in the bacteria found in the children and their mothers. (b) Detection rates of the *cnm* gene in children categorized as those with mothers who did or did not possess strains with the *cnm* gene. Statistically significant differences were found between the two groups (Fisher's protected least-significant difference test ***P<0.0001).
and 40% of the *S. mutans*-positive mothers. Furthermore, the rate of children with *cnm*-positive *S. mutans* whose mothers also possessed *cnm*-positive *S. mutans* was 72%, which was significantly higher than that of whose mothers did not (9.1%) (P<0.0001, odds ratio 17.5, 95% confidence interval 1.2–250.4) (Fig. 3b). The rate of children with *cnm*-positive *S. mutans* whose mothers also possessed *cnm*-positive *S. mutans* was high, which indicates that vertical transmission occurred in the *cnm*-positive mother–child pairs.

### Dental caries experience of children with *cnm*-positive saliva specimens

Once dental caries progresses, demineralization of the tooth surface occurs and the dentin becomes exposed (Selwitz et al., 2007). It is reasonable to speculate that strains with collagen-binding properties have an advantage with respect to attaching to dentin, in which the main organic component is type I collagen. In this study, we focused on detection of strains with the *cnm* gene encoding collagen-binding adhesin in order to consider the clinical association between the presence of strains with the *cnm* gene and dental caries. It has been reported that multiple genotypes of *S. mutans* are present in the oral cavity (Grönnroos & Alaluusua, 2000; Kozai et al., 1999; Saarela et al., 1996). Thus, saliva specimens, which carry microbes representative of the whole oral cavity, were considered appropriate for the identification of *cnm*-positive *S. mutans* in the oral cavity as compared to analyses of isolated strains. In addition, saliva specimen analysis requires much less time than isolation of strains and extraction of their genomic DNA.

Forty-eight children were selected based on the availability of precise clinical records regarding their dental caries and number of teeth emerged in the oral cavity (more than 14 primary teeth and less than 8 permanent teeth). A total of 28 were positive for *S. mutans*, of whom 13 possessed *cnm*-positive *S. mutans* (group A) and 15 did not (group B) (Fig. 4a). As for children with *S. mutans* either positive or negative for *cnm*, those in group A had a significantly higher defs rate than those in group B (P<0.01) (Fig. 4b). In addition, children without *S. mutans* detected (group C) had a significantly lower defs rate as compared to group A (P<0.01), whereas there was no significant difference between groups B and C. This finding implies that the presence of the Cnm protein of *S. mutans* is closely correlated with experiencing dental caries. On the other hand, there were several subjects positive for *S. mutans* with *cnm* who had low caries scores, which indicates that other factors also play important roles in the development of dental caries. Taken together, our results suggest that individuals with *cnm*-positive *S. mutans* are at possible risk for the development of dental caries.

### ACKNOWLEDGEMENTS

This study was supported by the 21st Century Center of Excellence (COE) programme entitled ‘Origination of Frontier BioDentistry’ at Osaka University Graduate School of Dentistry supported by the Ministry of Education, Culture, Sports, Science and Technology of Japan, Grants-in-Aid for Scientific Research (A) 19209063 and (B) 16390605, Grants-in-Aid for Exploratory Research 17659647 and 19659538 from Japan Society for Promotion of Science, and Grants-in-Aid for Young Scientists (A) 18689050 and (B) 19791572 from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

### REFERENCES


