Multilocus sequence typing analysis of *Streptococcus mutans* strains with the *cnm* gene encoding collagen-binding adhesin

Jinthana Lapirattanakul,¹ Kazuhiko Nakano,² Ryota Nomura,² Pattarawadee Leelataweewud,³ Narumon Chalermarp,⁴ Arthit Klaophimai,¹ Ratchapin Srisatjaluk,¹ Shigeyuki Hamada⁵ and Takashi Ooshima²

Correspondence

Jinthana Lapirattanakul
dtjlp@mahidol.ac.th

¹Department of Oral Microbiology, Faculty of Dentistry, Mahidol University, Bangkok 10400, Thailand

²Department of Pediatric Dentistry, Osaka University Graduate School of Dentistry, Osaka 565-0871, Japan

³Department of Pediatric Dentistry, Faculty of Dentistry, Mahidol University, Bangkok 10400, Thailand

⁴Department of Oral Medicine, Faculty of Dentistry, Mahidol University, Bangkok 10400, Thailand

⁵Research Institute for Microbial Diseases, Osaka University, Thailand–Japan Research Collaboration Center on Emerging and Re-emerging Infection, Nonthaburi 11000, Thailand

Streptococcus mutans is one of the oral pathogens associated with infective endocarditis (IE). With respect to bacterial binding ability to the extracellular matrix, the Cnm protein, a cell surface collagen-binding adhesin of *S. mutans*, is known as one of the possible virulence factors with regard to IE. In this study, we aimed to determine the distribution of the *cnm* gene, which encodes Cnm, in a large number of clinical isolates of *S. mutans* from Thai subjects. Then, the *cnm*-positive strains were classified using a multilocus sequence typing (MLST) scheme, which we constructed previously. In addition, the data were analysed together with our previous MLST data of *cnm*-positive strains from Japan and Finland in order to evaluate the clonal relationship among *S. mutans* strains harbouring the *cnm* gene. The *cnm* gene was detected in 12.4 % of all 750 Thai isolates, and serotype f showed the highest rate of detection (54.5 %). According to the MLST data, two clonal complex groups were revealed as the important clones related to *cnm*-positive *S. mutans* from various origins of isolation. Moreover, the collagen-binding properties of *S. mutans* strains with the *cnm* gene were significantly greater than those of strains without the gene, although four *cnm*-negative strains classified into two sequence types (STs), ST110 and ST136, showed extremely high collagen-binding rates suggesting the presence of additional genes involved with collagen binding in these STs. Taken together, these results provided information on both epidemiological as well as evolutionary aspects of *S. mutans* possessing the *cnm* gene.

Received 14 April 2011

Accepted 14 June 2011

INTRODUCTION

Streptococcus mutans is a well-known pathogen of dental caries (Hamada & Slade, 1980), which consists of four serotypes, c, e, f and k (Linzer *et al.*, 1986; Nakano & Ooshima, 2009). The detection rate of serotype c is highest among the strains isolated from oral cavities, followed by

serotype e, while serotypes f and k are regarded as being present in minor proportions (Lapirattanakul *et al.*, 2009; Nakano & Ooshima, 2009). Various pieces of evidence show the association of *S. mutans* with some life-threatening diseases such as infective endocarditis (IE), although the major species related to IE belong to the *mitis* group of oral streptococci (Banas, 2004; Mylonakis & Calderwood, 2001). Moreover, continuous detections of this bacterium in specimens from patients suffering from such disease emphasize the importance of *S. mutans* in IE (Gauduchon *et al.*, 2001; Nomura *et al.*, 2006; Vose *et al.*, 1987; Ullman *et al.*, 1988). Indeed, studies to clarify specific components

Abbreviations: ECM, extracellular matrix; IE, infective endocarditis; MLST, multilocus sequence typing; ST, sequence type.

The GenBank/EMBL/DDBJ accession numbers for the 8 housekeeping gene sequences of the 52 *Streptococcus mutans* isolates from Thai people are JF800232–JF800647.

involved in IE in *S. mutans* have also been performed worldwide (Beg *et al.*, 2002; Jung *et al.*, 2009; Matsumoto-Nakano *et al.*, 2009; Miller-Torbert *et al.*, 2008).

IE is initiated by the binding of bacterial surface molecules to host extracellular matrix (ECM) proteins exposed on damaged heart tissue (Moreillon & Que, 2004). In *S. mutans*, collagen-binding adhesin or Cnm protein, encoded by the *cnm* gene, was reported as a strain-specific collagen-binding molecule (Sato *et al.*, 2004). The Cnm protein possesses a collagen-binding domain (CBD) in its structure, which can bind to both collagen and laminin. Our previous studies have indicated the predominant frequency of this *cnm* gene in minor serotypes f and k, which usually showed unique characteristics concerning the pathogenesis of IE (Nakano *et al.*, 2004a, 2007, 2008). An unambiguous typing method based on the nucleotide sequences of eight housekeeping genes, multilocus sequence typing (MLST), was developed for *S. mutans*, and showed a close evolutionary relationship among *cnm*-positive strains (Nakano *et al.*, 2007). MLST analysis of *cnm*-positive strains from various origins will provide more insight into *S. mutans* clones with collagen-binding ability due to the possession of the *cnm* gene.

Thus, in this study, we investigated the distribution of the *cnm* gene in a large number of clinical isolates of *S. mutans* from the Thai population. Then, the *cnm*-positive strains obtained were analysed by MLST to evaluate the clonal relationship between these strains and the *cnm*-positive *S. mutans* strains previously analysed using MLST.

METHODS

Samples and bacterial isolation. Stimulated saliva specimens were provided by 150 Thai subjects (59 male, 91 female; aged 5–70 years) who came to the Faculty of Dentistry, Mahidol University, during 2008 to 2010. This project was approved by the Ethics Committee of Mahidol University. Firstly, serial dilutions of the collected saliva were cultured on mitis salivarius agar (Difco) plates containing bacitracin (0.2 U ml⁻¹; Sigma), 0.001% (v/v) tellurite solution (Becton, Dickinson) and 15% (w/v) sucrose (MSB agar plates) (Gold *et al.*, 1973). After a 2 day incubation at 37 °C, five presumptive rough

colonies of *S. mutans* were randomly picked from each subject's culture plate, and confirmed to be *S. mutans* by biochemical tests including fermentation of 1% mannitol, sorbitol, raffinose or melibiose in phenol red broth base (Difco), and negative testing for dextran agglutination (Nakano *et al.*, 2004a). Collectively, a total of 750 strains of *S. mutans* were obtained and analysed in this study.

Genomic DNA extraction. Genomic DNA of all strains was extracted as described previously (Nakano *et al.*, 2004a). Briefly, the tested bacteria were grown in brain heart infusion (BHI; Difco) broth at 37 °C for 18 h; then the cells were collected and washed by centrifugation. Bacterial cell walls were digested with lysozyme (Bio Basic; 20 mg ml⁻¹) at 37 °C for 1 h. Thereafter, cell lysis solution, protein precipitation solution and RNase (Qiagen) were used to purify the DNA as recommended by the manufacturer. Finally, genomic DNA was precipitated by ethanol, and resuspended in DNA hydration solution (Qiagen).

Serotype determination. The serotype of each *S. mutans* strain was determined by both an immunodiffusion method using rabbit antisera specific for serotypes c, e, f and k, as well as PCR using serotype-specific sets of primers as described elsewhere (Nakano *et al.*, 2004b; Shibata *et al.*, 2003). The amplified fragments were visualized using electrophoresis in 1.5% agarose gel in the presence of 1 µg ethidium bromide ml⁻¹.

Distribution of *cnm* genes among *S. mutans* strains isolated from Thai subjects. The distribution of the *cnm* gene was evaluated by the PCR method with specific primers for the entire length of the *cnm* gene (*cnm*-1F, 5'-GAC AAA GAA ATG AAA GAT GT-3'; *cnm*-1R, 5'-GCA AAG ACT CTT GTC CCT GC-3'), as described previously (Nomura *et al.*, 2009). The amplified fragments were visualized in 0.7% agarose gel in the presence of 1 µg ethidium bromide ml⁻¹. For all detections, strains TW871 and MT8148 (Fujiwara *et al.*, 2001; Ooshima *et al.*, 1983) were used as positive and negative controls, respectively.

Analysis of clonal relationship using MLST. One *S. mutans* strain possessing the *cnm* gene from each subject was randomly selected and analysed by the MLST scheme, as described previously (Nakano *et al.*, 2007). Equal numbers of *S. mutans* strains without the *cnm* gene in each serotype were also subjected to MLST as the control strains for the Thai subjects. Briefly, internal fragments of eight housekeeping gene loci were amplified (Nakano *et al.*, 2007), and their nucleotide sequences determined (Table 1). The sequences for each housekeeping gene fragment were compared with the previously described data (Lapirattanakul *et al.*, 2008; Nakano *et al.*, 2007) in order to assign

Table 1. Eight housekeeping genes for MLST of *S. mutans*

Locus	Definition	Allele no. previously found*	New allele no. detected in this study (GenBank accession no.)
<i>tkt</i>	Transketolase	1–14	15–17 (JF800280–JF800283)
<i>glnA</i>	Glutamine synthetase type I	1–18	19–21 (JF800333–JF800335)
<i>gltA</i>	Glutamate synthase	1–26	27–30 (JF800381–JF800387)
<i>glk</i>	Glucose kinase	1–19	20–21 (JF800436–JF800439)
<i>aroE</i>	Shikimate 5-dehydrogenase	1–22	23–25 (JF800487–JF800491)
<i>murI</i>	Glutamate racemase	1–19	20 (JF800542–JF800543)
<i>lepC</i>	Signal peptidase I	1–25	26–30 (JF800588–JF800595)
<i>gyrA</i>	DNA gyrase subunit A	1–17	18–20 (JF800645–JF800647)

*Allele numbers detected in our previous studies (Nakano *et al.*, 2007; Lapirattanakul *et al.*, 2008).

allele numbers using a non-redundant database (<http://linux.mlst.net/nrdb/nrdb.htm>). Eight allele numbers for each strain defined the allelic profile and, consequently, sequence type (ST) of that strain. STs 1–108 were assigned in our previous studies (Lapirattanakul *et al.*, 2008; Nakano *et al.*, 2007). All new alleles defined in the present study are shown in bold in Table 2.

All allelic profiles determined in this project, together with our previous MLST data of 108 STs (Lapirattanakul *et al.*, 2008; Nakano *et al.*, 2007), were analysed with computational programs as follows. The START (sequence type analysis and recombinational tests) tools (<http://pubmlst.org/software/analysis/start2>) (Jolley *et al.*, 2001) were used to produce a matrix of pairwise differences in the allelic profiles and a dendrogram was constructed from the matrix using UPGMA. Then, related STs were grouped into lineages or clonal complexes using the eBURST (enhanced based upon related sequence types) program, version 3 (<http://eburst.mlst.net>) (Feil *et al.*, 2004). Two or more independent isolates sharing identical alleles at six or more loci were defined as members of a lineage. Each lineage was a 'group' or 'clonal complex', which designates STs that have diversified very recently from a common ancestor.

Collagen-binding assay. The collagen-binding properties of the strains isolated from Thai subjects, which were analysed by the MLST method, were evaluated by the method described previously (Nomura *et al.*, 2009). Briefly, type I collagen, 2 mg in 0.25 M acetic acid (Sigma), was coated onto the wells of 96-well tissue culture plates (Corning) and incubated overnight at 4 °C; then the plates were washed three times with PBS and 5 % BSA (Sigma) in PBS was used as a blocking agent at 37 °C for 1.5 h. Next, the wells were washed again with PBS containing 0.01 % Tween 20. Cells from overnight cultures of *S. mutans* grown in BHI broth were collected by centrifugation, resuspended and an aliquot of 2×10^9 cells was added to each well. After 3 h incubation at 37 °C, adherent cells were washed three times with PBS, and 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 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 A_{595} value. The results for each strain were expressed as a percentage as compared to the binding property of TW871, which was defined as 100 %. Assays for each tested strain were performed in triplicate and repeated twice.

Statistical analysis. Statistical analyses were performed using the SPSS statistical software package (SPSS). Collagen-binding rates between the *cnm*-positive and *cnm*-negative strains were compared using the Mann–Whitney-U test. Pearson's correlation, as well as Fisher's exact probability test, was used to compare *cnm*-positive proportion between each serotype. For all analyses, *P* values less than 0.05 were considered to be statistically significant.

RESULTS

Distribution of *cnm* genes among *S. mutans* isolated from Thai subjects

A total of 750 strains of *S. mutans* from 150 Thai people were classified mainly as serotype c, followed by serotype e, whereas serotypes f and k were detected only in minor proportions (Table 3). PCR results indicated that the *cnm* gene was detected in 26 of 150 subjects (17.3 %), in which all five strains picked up were positive for the *cnm* gene in half of these subjects. The other half of these subjects

possessed both *cnm*-positive and *cnm*-negative isolates. When the number of strains was considered, the *cnm* gene was found in 93 of all 750 strains (12.4 %) (Table 3). Most of the *cnm*-positive strains belonged to serotype f (54.5 %), and positive strains could also be identified in serotype c (8.9 %), e (11.3 %) and k (20.8 %). The prevalence of the *cnm* gene in serotype f was significantly greater than in the other serotypes ($P < 0.001$ between serotype f and c, as well as between serotype f and e; and $P < 0.01$ between serotype f and k).

Clonal relationship of *S. mutans* strains by MLST

We performed MLST analysis for 26 *cnm*-positive and 26 *cnm*-negative *S. mutans* strains isolated from Thai subjects, which were compared with the strains with the previously determined 108 STs. All 52 *S. mutans* strains from Thai subjects (20 strains of serotype c, 13 strains of serotype e, 12 strains of serotype f and 7 strains of serotype k) were classified into 39 STs, in which 34 STs were considered new (designated ST109 through ST142) (Table 2). The most common STs among these 52 Thai strains were ST129 and ST140, which contained 3 members for each ST. The association of one ST with different serotypes was also identified. ST129 was composed of serotypes c and e, while ST111 showed both serotypes c and f. Furthermore, the presence of the *cnm* gene was detected only in strain TLJ42-1, although strains TLJ42-1 and TLJ6-1 shared the same ST (ST110).

Fig. 1 shows a dendrogram constructed based on the allelic profiles of 142 STs. The assignment of STs into lineages with eBURST revealed 56 STs as singletons that were unrelated to any other STs, whereas the remaining 86 STs were grouped into 16 lineages. The largest group (group 2) contained 26 isolates representing 20 STs. The remaining 15 groups contained from 2 to 15 member STs. The distribution of *cnm*-positive strains among these clonal complex groups was observed in six groups (groups 2, 5, 7, 11, 12 and 16). While groups 7, 12 and 16 were small groups containing only 2 member STs, groups 5 and 11 contained 15 and 9 member STs, respectively, and approximately 70–80 % of STs in these groups possessed the *cnm* gene. Even though the isolates in this dendrogram originated from three main countries, Finland, Japan and Thailand, no obvious bias according to the geographical region was seen. Mixed origins of isolation were seen in most of the groups, and only five groups (groups 4, 6, 7, 13 and 16) with limited number of members showed all STs isolated from a single country.

Collagen-binding properties of *S. mutans* strains from Thai subjects

The collagen-binding properties for 26 *cnm*-positive and 26 *cnm*-negative *S. mutans* strains from Thai subjects were evaluated. The collagen-binding properties of strains with the *cnm* gene were significantly higher than those of strains

Table 2. Allelic profiles and collagen-binding rates of 26 *cnm*-positive and 26 *cnm*-negative *S. mutans* strains analysed by MLST in this study

All new alleles defined in this study are indicated in bold.

Strain	ST	Serotype	Allelic profile								Collagen-binding property*
			<i>tkt</i>	<i>glnA</i>	<i>gltA</i>	<i>glk</i>	<i>aroE</i>	<i>murI</i>	<i>lepC</i>	<i>gyrA</i>	
<i>cnm</i> positive											
TLJ13-1	109	c	1	5	27	3	4	5	19	1	M
TLJ42-1	110	c	3	3	8	20	4	2	1	10	EL
TLJ43-1	109	c	1	5	27	3	4	5	19	1	M
TLJ57-2	111	c	1	2	15	3	4	3	26	1	EH
TLJ82-1	112	c	3	5	5	20	23	2	26	1	EH
TLJ89-1	113	c	1	3	8	4	24	4	1	1	H
TLJ92-1	114	c	2	5	8	1	4	3	3	1	H
TLJ93-1	52	c	3	2	15	3	1	11	1	3	EL
TLJ114-1	52	c	3	2	15	3	1	11	1	3	EL
TLJ123-1	114	c	2	5	8	1	4	3	3	1	H
TLJ124-1	115	c	3	2	28	1	14	3	3	1	M
TLJ126-1	115	c	3	2	28	1	14	3	3	1	H
TLJ18-1	116	e	3	9	5	1	2	3	3	10	M
TLJ19-1	117	e	3	9	5	1	14	3	7	1	L
TLJ24-4	118	e	3	2	28	1	14	4	3	1	H
TLJ59-1	119	e	2	2	5	1	2	5	27	1	M
TLJ81-3	119	e	2	2	5	1	2	5	27	1	EL
TLJ115-2	120	e	3	2	5	1	2	3	3	10	EH
TLJ34-1	121	f	1	5	5	1	23	2	1	1	M
TLJ53-1	122	f	15	2	5	1	4	4	3	11	L
TLJ72-3	123	f	1	3	8	1	25	3	1	1	H
TLJ79-5	123	f	1	3	8	1	25	3	1	1	H
TLJ109-1	122	f	15	2	5	1	4	4	3	11	H
TLJ129-3	124	f	2	3	8	1	13	3	3	1	EH
TLJ151-1	111	f	1	2	15	3	4	3	26	1	L
TLJ60-1	125	k	3	3	8	1	4	3	1	1	H
<i>cnm</i> negative											
TLJ1-1	126	c	3	1	1	1	18	3	28	1	EL
TLJ2-1	127	c	3	2	29	3	8	3	5	1	EL
TLJ3-1	128	c	14	2	30	1	2	3	1	1	EL
TLJ6-1	110	c	3	3	8	20	4	2	1	10	EL
TLJ15-1	129	c	2	5	5	6	2	3	11	1	EL
TLJ55-1	130	c	6	3	20	4	4	5	5	4	EL
TLJ81-1	131	c	3	1	1	4	2	19	1	15	EL
TLJ152-1	5	c	1	1	1	3	2	4	1	3	EL
TLJ8-2	129	e	2	5	5	6	2	3	11	1	EL
TLJ17-3	132	e	1	16	23	13	2	18	22	15	EL
TLJ28-1	76	e	6	1	6	4	5	5	1	4	EL
TLJ56-1	129	e	2	5	5	6	2	3	11	1	EL
TLJ95-1	133	e	3	5	5	17	17	20	29	18	EL
TLJ128-1	134	e	1	19	26	3	2	18	1	15	EL
TLJ145-1	135	e	16	20	9	1	2	7	30	1	EL
TLJ9-1	136	f	3	21	1	21	2	20	11	19	EH
TLJ20-3	137	f	3	5	5	1	2	3	3	1	EL
TLJ97-1	138	f	2	9	5	6	4	4	3	1	EL
TLJ112-2	39	f	2	5	5	6	2	3	3	1	EL
TLJ135-1	139	f	1	2	15	3	1	11	11	20	EL
TLJ11-2	140	k	1	3	1	8	11	2	9	3	EH
TLJ26-1	141	k	1	2	1	3	14	3	11	1	EL
TLJ80-2	141	k	1	2	1	3	14	3	11	1	EL

Table 2. cont.

Strain	ST	Serotype	Allelic profile								Collagen-binding property*
			<i>tkt</i>	<i>glnA</i>	<i>gltA</i>	<i>glk</i>	<i>aroE</i>	<i>murI</i>	<i>lepC</i>	<i>gyrA</i>	
TLJ85-4	140	k	1	3	1	8	11	2	9	3	EH
TLJ106-1	142	k	17	2	1	1	4	5	1	1	EL
TLJ137-4	140	k	1	3	1	8	11	2	9	3	EH

*Collagen-binding property was classified based on the percentages of collagen-binding ability of each strain obtained as described in Methods. EH, Extremely high (>130 %); H, high (90–130 %); M, moderate (50–90 %); L, low (10–50 %); EL, extremely low (<10 %).

without the gene ($P<0.001$) (Fig. 2). Nevertheless, four *cnm*-negative strains with strong collagen-binding ability were also seen (Table 2). All of these strains except for TLJ9-1 belonged to the same clonal complex group based on MLST analysis. Furthermore, strains TLJ6-1 and TLJ42-1 possessed quite similar collagen-binding rates, although they showed the same ST (ST110), but a different *cnm* gene profile. However, the relatedness of *cnm*-positive strains without collagen-binding ability could not be clarified based on the MLST data in this study.

DISCUSSION

The detection of *S. mutans* in specimens taken from patients with IE has been reported by either conventional or molecular techniques (Gauduchon *et al.*, 2001; Nakano *et al.*, 2006; Nomura *et al.*, 2006; Vose *et al.*, 1987; Ullman *et al.*, 1988). Since IE is initiated by the colonization of the organism on the ECM of damaged endothelium in heart tissue (Moreillon & Que, 2004), bacterial molecules associated with adhesion to ECM are considered important. The Cnm protein, encoded by the *cnm* gene, was reported as one of the possible virulence factors involved in IE in *S. mutans* (Sato *et al.*, 2004). This protein was shown to be highly homologous to the collagen-binding protein of *Staphylococcus aureus*, Cna, which plays an important role in the animal model of IE (Hienz *et al.*, 1996). This aspect

was also emphasized by the finding that the collagen-binding abilities of strains with the *cnm* gene were significantly greater than those of strains without the gene (Nomura *et al.*, 2009). In addition, the Cnm protein was proved to be necessary for invasion of human coronary artery endothelial cells (Abranches *et al.*, 2011). The distribution of the strains with the *cnm* gene has previously been reported in two populations, Japanese and Finnish people, in which *S. mutans* with minor serotypes, f and k, showed high proportions of the gene (Nakano *et al.*, 2007; Nomura *et al.*, 2009). Subsequently, the presence of strains with the *cnm* gene was reported in 6 of 50 strains from 50 Thai subjects (Nakano *et al.*, 2010). In the present study, we intended to expand the analysis using large numbers of strains isolated from Thai subjects. We found that the distribution of the *cnm* gene in *S. mutans* from Thai subjects was highest in serotype f (54.5 %). Such distribution was also observed in *S. mutans* strains from Japanese and Finnish people (Nakano *et al.*, 2007; Nomura *et al.*, 2009). These results thus confirmed the association between serotype f and the presence of the *cnm* gene in *S. mutans*.

Recently, we constructed a MLST method for *S. mutans*, and used this technique to evaluate the clonal relationship of 102 isolates including 22 *cnm*-positive strains. Thus, 92 STs (ST1–ST92) were identified (Nakano *et al.*, 2007). In the following study in which the MLST method was used to demonstrate mother-to-child transmission, an additional 16 STs (ST93–ST108) were detected (Lapirattanakul *et al.*, 2008). In the present study, we further analysed 52 *S. mutans*, 26 *cnm*-positive and 26 *cnm*-negative strains, from our Thai subjects, and showed 34 new STs (ST109–ST142). When we considered the clonal complexes related to the *cnm* gene, six clonal complex groups (groups 2, 5, 7, 11, 12 and 16) contained *cnm*-positive strains as members. Although the association of groups 2, 5, 7 and 11 with *cnm*-positive strains has been found since our previous work (Nakano *et al.*, 2007), the increasing numbers of both isolates and STs possessing the *cnm* gene were clearly found among groups 5 and 11 in this study (Fig. 1). Notably, the serotypes of members in these two groups seemed to be different. While group 5 contained strains in serotypes c, f and k, the members in group 11 were serotypes c, e and f. Thus, some isolates in these two different lineages have acquired the *cnm* gene during their evolution, and

Table 3. Serotype distribution of 750 *S. mutans* isolates obtained from 150 subjects

Serotype	No. of isolates (total $n=750$)*	No. of <i>cnm</i> -positive strains†
c	531 (70.8 %)	47 (8.9 %)
e	151 (20.1 %)	17 (11.3 %)
f	44 (5.9 %)	24 (54.5 %)
k	24 (3.2 %)	5 (20.8 %)

*The numbers in parentheses indicate the percentages of each serotype from the total number of isolates.

†The numbers in parentheses indicate the percentages of *cnm*-positive strains in each serotype.

Fig. 1. Dendrogram created from 142 STs of *S. mutans* based on the MLST method constructed by Nakano *et al.* (2007). Clonal complex groups, serotypes, as well as the origins of isolates (A, USA; F, Finland; J, Japan; T, Thailand) are indicated.

successfully propagated the gene to initiate such clones. Given the binding ability to ECM of strains harbouring the *cnm* gene, groups 5 and 11 might be considered as hypervirulent clones of *S. mutans* in terms of IE. Therefore, more studies on the common characteristics of strains in these clones might provide an efficient way to track IE-related *S. mutans* strains. As for the sources of the isolates, *S. mutans* strains in the present study were mainly isolated from Finland, Japan and Thailand, with only one complete genome sequenced strain, UA159 (ST1), from the United States. However, most of the clonal complex groups were composed of isolates from more than one country of origin. The origins of strains in groups 5 and 11, which related to many *cnm*-positive strains, were also mixed.

In this study, we also determined the collagen-binding properties of the 52 *S. mutans* strains (26 *cnm*-positive and 26 *cnm*-negative strains), which were analysed by the MLST method. The collagen-binding properties of *cnm*-positive strains were significantly greater than those of *cnm*-negative strains ($P < 0.001$) (Fig. 2). Thus, the result of isolates from Thai subjects confirmed our previous data based on isolates of Japanese strains (Nomura *et al.*, 2009). However, we also found four *cnm*-negative strains posses-

Fig. 2. Comparison of collagen-binding properties of *cnm*-positive ($n=26$) and *cnm*-negative ($n=26$) strains. The collagen-binding activity of each strain was standardized to that of strain TW871, taken as 100 %. The binding rates between these two groups were found to be statistically significantly different (Mann–Whitney–U test; $P < 0.001$).

sing high collagen-binding rates (TLJ11-2, TLJ85-4, TLJ137-4 and TLJ9-1) (Table 2). Interestingly, three of these four strains were the same ST (ST140), defined by similar allelic profile by MLST. In fact, this ST140 belonged to clonal complex group 14, in which we suppose that at least one additional gene encoding collagen-binding protein in *S. mutans* could be present. Further studies should be focused on the unknown genes involved in collagen-binding activity that are possibly present in these strains. Even though most of the *S. mutans* strains sharing the same STs expressed comparable rates for collagen-binding ability, the exception was still seen in STs 111, 115, 119 and 122 (Table 2). Two isolates in these STs showed totally different collagen-binding rates. Moreover, our information based on MLST analysis could not clarify the relatedness of *cnm*-positive strains without collagen-binding ability. Since MLST analysis is a strain classification based on the nucleotide sequences of genes, the result gained might not provide enough information relating to the protein expression level. Thus, factors involved in the different expression of the *cnm* gene in strains with the same STs are still of interest for further investigation.

A lack of correlation between serotypes and STs has been demonstrated in many bacterial species, such as *Streptococcus pneumoniae*, *Neisseria meningitidis* and *Streptococcus agalactiae*, and the horizontal transfer of capsular biosynthesis genes related to serotyping is speculated to be the reason (Coffey *et al.*, 1991; Luan *et al.*, 2005; Pettigrew *et al.*, 2006; Swartley *et al.*, 1997). In fact, we also reported the association of one ST with two serotypes in our previous work (Lapirattanakul *et al.*, 2008; Nakano *et al.*, 2007). However, the MLST analysis in this study revealed more evidence of this aspect (Fig. 1). Furthermore, the results in the present study also indicated isolates of identical ST with a differing presence of the *cnm* gene (ST52 and ST110). The *cnm* gene, similar to genes encoding proteins related to serotype-specific polysaccharide, was described as a dispensable gene in *S. mutans*, as it is present in some – but not all – strains of *S. mutans* (Medini *et al.*, 2005; Waterhouse & Russell, 2006). Thus, it could be speculated that such a gene is easily transferred among unrelated strains through horizontal transfer, while the core genome, such as housekeeping genes used in MLST, is more faithfully transferred according to evolutionary manner. Since the data gained from MLST analysis are also an invaluable resource for studying bacterial population structure, the expansion of MLST analysis into strains from various geographical regions, and from both disease and non-disease related sources, is needed to precisely clarify the important effects of horizontal gene transfer involving recombination to *S. mutans* population structure (Didelot & Maiden, 2010).

In summary, in this work both prevalence and MLST analyses were performed to study *S. mutans* containing the *cnm* gene in a large number of clinical isolates. The data, therefore, provided overview information for both the epidemiological and evolutionary aspects of such strains. Moreover, the important clones related to *cnm*-positive

strains of *S. mutans* have also been indicated. Further studies focused on additional *S. mutans* clones involved in other aspects of IE pathogenesis will provide more candidate clones of *S. mutans* related to IE, and might help with the narrowing down of suitable clones for the practical monitoring of IE risks caused by *S. mutans* in patients.

ACKNOWLEDGEMENTS

This study was supported by a grant for new researchers at Mahidol University, a grant from the Faculty of Dentistry, Mahidol University, Thailand, and a fellowship from the international scientific exchange fund of the Japan Dental Association. This study was also supported by a Grant-in-Aid for Scientific Research from the Japan Society for the Promotion of Science (A) 19209063, as well as Grants-in-Aid for Young Scientists from the Ministry of Education, Culture, Sports, Science and Technology of Japan (A) 21689052 and (B) 21792067.

REFERENCES

- Abranches, J., Miller, J. H., Martinez, A. R., Simpson-Haidaris, P. J., Burne, R. A. & Lemos, J. A. (2011). The collagen-binding protein Cnm is required for *Streptococcus mutans* adherence to and intracellular invasion of human coronary artery endothelial cells. *Infect Immun* **79**, 2277–2284.
- Banas, J. A. (2004). Virulence properties of *Streptococcus mutans*. *Front Biosci* **9**, 1267–1277.
- Beg, A. M., Jones, M. N., Miller-Torbert, T. & Holt, R. G. (2002). Binding of *Streptococcus mutans* to extracellular matrix molecules and fibrinogen. *Biochem Biophys Res Commun* **298**, 75–79.
- Coffey, T. J., Dowson, C. G., Daniels, M., Zhou, J., Martin, C., Spratt, B. G. & Musser, J. M. (1991). Horizontal transfer of multiple penicillin-binding protein genes, and capsular biosynthetic genes, in natural populations of *Streptococcus pneumoniae*. *Mol Microbiol* **5**, 2255–2260.
- Didelot, X. & Maiden, M. C. (2010). Impact of recombination on bacterial evolution. *Trends Microbiol* **18**, 315–322.
- Feil, E. J., Li, B. C., Aanensen, D. M., Hanage, W. P. & Spratt, B. G. (2004). eBURST: inferring patterns of evolutionary descent among clusters of related bacterial genotypes from multilocus sequence typing data. *J Bacteriol* **186**, 1518–1530.
- Fujiwara, T., Nakano, K., Kawaguchi, M., Ooshima, T., Sobue, S., Kawabata, S., Nakagawa, I. & Hamada, S. (2001). Biochemical and genetic characterization of serologically untypable *Streptococcus mutans* strains isolated from patients with bacteremia. *Eur J Oral Sci* **109**, 330–334.
- Gauduchon, V., Benito, Y., Célar, M., Mouren, C., Delorme, V., Philippe-Bert, J., Etienne, J. & Vandenesch, F. (2001). Molecular diagnosis of recurrent *Streptococcus mutans* endocarditis by PCR amplification and sequencing. *Clin Microbiol Infect* **7**, 36–37.
- Gold, O. G., Jordan, H. V. & Van Houte, J. (1973). A selective medium for *Streptococcus mutans*. *Arch Oral Biol* **18**, 1357–1364.
- Hamada, S. & Slade, H. D. (1980). Biology, immunology, and cariogenicity of *Streptococcus mutans*. *Microbiol Rev* **44**, 331–384.
- Hien, S. A., Schennings, T., Heimdahl, A. & Flock, J. I. (1996). Collagen binding of *Staphylococcus aureus* is a virulence factor in experimental endocarditis. *J Infect Dis* **174**, 83–88.
- Jolley, K. A., Feil, E. J., Chan, M. S. & Maiden, M. C. (2001). Sequence type analysis and recombinational tests (START). *Bioinformatics* **17**, 1230–1231.

- Jung, C. J., Zheng, Q. H., Shieh, Y. H., Lin, C. S. & Chia, J. S. (2009). *Streptococcus mutans* autolysin AtlA is a fibronectin-binding protein and contributes to bacterial survival in the bloodstream and virulence for infective endocarditis. *Mol Microbiol* **74**, 888–902.
- Lapirattanakul, J., Nakano, K., Nomura, R., Hamada, S., Nakagawa, I. & Ooshima, T. (2008). Demonstration of mother-to-child transmission of *Streptococcus mutans* using multilocus sequence typing. *Caries Res* **42**, 466–474.
- Lapirattanakul, J., Nakano, K., Nomura, R., Nemoto, H., Kojima, A., Senawongse, P., Srisatjaluk, R. & Ooshima, T. (2009). Detection of serotype *k* *Streptococcus mutans* in Thai subjects. *Oral Microbiol Immunol* **24**, 431–433.
- Linzer, R., Reddy, M. S. & Levine, M. J. (1986). Immunochemical aspects of serotype carbohydrate antigens of *Streptococcus mutans*. In *Molecular Microbiology and Immunology of Streptococcus mutans*, pp. 29–38. Edited by S. Hamada, S. M. Michalek, H. Kiyono, L. Manaker & J. R. McGhee. Amsterdam: Elsevier.
- Luan, S. L., Granlund, M., Sellin, M., Lagergård, T., Spratt, B. G. & Norgren, M. (2005). Multilocus sequence typing of Swedish invasive group B streptococcus isolates indicates a neonatally associated genetic lineage and capsule switching. *J Clin Microbiol* **43**, 3727–3733.
- Matsumoto-Nakano, M., Tsuji, M., Inagaki, S., Fujita, K., Nagayama, K., Nomura, R. & Ooshima, T. (2009). Contribution of cell surface protein antigen c of *Streptococcus mutans* to platelet aggregation. *Oral Microbiol Immunol* **24**, 427–430.
- Medini, D., Donati, C., Tettelin, H., Massignani, V. & Rappuoli, R. (2005). The microbial pan-genome. *Curr Opin Genet Dev* **15**, 589–594.
- Miller-Torbert, T. A., Sharma, S. & Holt, R. G. (2008). Inactivation of a gene for a fibronectin-binding protein of the oral bacterium *Streptococcus mutans* partially impairs its adherence to fibronectin. *Microb Pathog* **45**, 53–59.
- Moreillon, P. & Que, Y. A. (2004). Infective endocarditis. *Lancet* **363**, 139–149.
- Mylonakis, E. & Calderwood, S. B. (2001). Infective endocarditis in adults. *N Engl J Med* **345**, 1318–1330.
- Nakano, K. & Ooshima, T. (2009). Serotype classification of *Streptococcus mutans* and its detection outside the oral cavity. *Future Microbiol* **4**, 891–902.
- Nakano, K., Nomura, R., Nakagawa, I., Hamada, S. & Ooshima, T. (2004a). Demonstration of *Streptococcus mutans* with a cell wall polysaccharide specific to a new serotype, *k*, in the human oral cavity. *J Clin Microbiol* **42**, 198–202.
- Nakano, K., Nomura, R., Shimizu, N., Nakagawa, I., Hamada, S. & Ooshima, T. (2004b). Development of a PCR method for rapid identification of new *Streptococcus mutans* serotype *k* strains. *J Clin Microbiol* **42**, 4925–4930.
- Nakano, K., Inaba, H., Nomura, R., Nemoto, H., Takeda, M., Yoshioka, H., Matsue, H., Takahashi, T., Taniguchi, K. & other authors (2006). Detection of cariogenic *Streptococcus mutans* in extirpated heart valve and atheromatous plaque specimens. *J Clin Microbiol* **44**, 3313–3317.
- Nakano, K., Lapirattanakul, J., Nomura, R., Nemoto, H., Alaluusua, S., Grönroos, L., Vaara, M., Hamada, S., Ooshima, T. & Nakagawa, I. (2007). *Streptococcus mutans* clonal variation revealed by multilocus sequence typing. *J Clin Microbiol* **45**, 2616–2625.
- Nakano, K., Nomura, R., Nemoto, H., Lapirattanakul, J., Taniguchi, N., Grönroos, L., Alaluusua, S. & Ooshima, T. (2008). Protein antigen in serotype *k* *Streptococcus mutans* clinical isolates. *J Dent Res* **87**, 964–968.
- Nakano, K., Nomura, R., Taniguchi, N., Lapirattanakul, J., Kojima, A., Naka, S., Senawongse, P., Srisatjaluk, R., Grönroos, L. & Alaluusua, S. (2010). Molecular characterization of *Streptococcus mutans* strains containing the *cnm* gene encoding a collagen-binding adhesin. *Arch Oral Biol* **55**, 34–39.
- Nomura, R., Nakano, K., Nemoto, H., Fujita, K., Inagaki, S., Takahashi, T., Taniguchi, K., Takeda, M., Yoshioka, H. & other authors (2006). Isolation and characterization of *Streptococcus mutans* in heart valve and dental plaque specimens from a patient with infective endocarditis. *J Med Microbiol* **55**, 1135–1140.
- Nomura, R., Nakano, K., Taniguchi, N., Lapirattanakul, J., Nemoto, H., Grönroos, L., Alaluusua, S. & Ooshima, T. (2009). Molecular and clinical analyses of the gene encoding the collagen-binding adhesin of *Streptococcus mutans*. *J Med Microbiol* **58**, 469–475.
- Ooshima, T., Izumitani, A., Sobue, S. & Hamada, S. (1983). Cariostatic effect of palatinose on experimental dental caries in rats. *Jpn J Med Sci Biol* **36**, 219–223.
- Pettigrew, M. M., Fennie, K. P., York, M. P., Daniels, J. & Ghaffar, F. (2006). Variation in the presence of neuraminidase genes among *Streptococcus pneumoniae* isolates with identical sequence types. *Infect Immun* **74**, 3360–3365.
- Sato, Y., Okamoto, K., Kagami, A., Yamamoto, Y., Igarashi, T. & Kizaki, H. (2004). *Streptococcus mutans* strains harboring collagen-binding adhesin. *J Dent Res* **83**, 534–539.
- Shibata, Y., Ozaki, K., Seki, M., Kawato, T., Tanaka, H., Nakano, Y. & Yamashita, Y. (2003). Analysis of loci required for determination of serotype antigenicity in *Streptococcus mutans* and its clinical utilization. *J Clin Microbiol* **41**, 4107–4112.
- Swartley, J. S., Marfin, A. A., Edupuganti, S., Liu, L. J., Cieslak, P., Perkins, B., Wenger, J. D. & Stephens, D. S. (1997). Capsule switching of *Neisseria meningitidis*. *Proc Natl Acad Sci U S A* **94**, 271–276.
- Ullman, R. F., Miller, S. J., Strampfer, M. J. & Cunha, B. A. (1988). *Streptococcus mutans* endocarditis: report of three cases and review of the literature. *Heart Lung* **17**, 209–212.
- Vose, J. M., Smith, P. W., Henry, M. & Colan, D. (1987). Recurrent *Streptococcus mutans* endocarditis. *Am J Med* **82**, 630–632.
- Waterhouse, J. C. & Russell, R. R. (2006). Dispensable genes and foreign DNA in *Streptococcus mutans*. *Microbiology* **152**, 1777–1788.