

## Discrimination of *Streptococcus pneumoniae* from Viridans Group Streptococci by Genomic Subtractive Hybridization

Nao Suzuki,<sup>1</sup> Mitsuko Seki,<sup>2,3</sup> Yoshio Nakano,<sup>4</sup> Yusuke Kiyoura,<sup>1</sup> Masao Maeno,<sup>2,3</sup> and Yoshihisa Yamashita<sup>4\*</sup>

Department of Oral Medical Science, Ohu University School of Dentistry, 31-1 Misumido, Tomitamachi, Koriyama 963-8611, Japan<sup>1</sup>; Department of Oral Health Sciences, Nihon University, School of Dentistry, 1-8-13 Kanda-Surugadai, Chiyoda-ku, Tokyo 101-8310, Japan<sup>2</sup>; Division of Functional Morphology, Dental Research Center, Nihon University, School of Dentistry, 1-8-13 Kanda-Surugadai, Chiyoda-ku, Tokyo 101-8310, Japan<sup>3</sup>; and Department of Preventive Dentistry, Kyushu University Faculty of Dental Science, 3-1-1 Maidashi, Higashi-ku, Fukuoka 812-8582, Japan<sup>4</sup>

Received 30 March 2005/Returned for modification 16 May 2005/Accepted 3 June 2005

**Two oligonucleotide primer sets for the discrimination of *Streptococcus pneumoniae* from “pneumococcus-like” oral streptococcal isolates by PCR were developed. Genomic subtractive hybridization was performed to search for differences between *Streptococcus pneumoniae* strain WU2 and the most closely related oral streptococcus, *Streptococcus mitis* strain 903. We identified 19 clones that contained *S. pneumoniae*-specific nucleotide fragments that were absent from the chromosomal DNA of typical laboratory strains of *S. mitis* and other oral bacteria. Subsequently, oligonucleotide PCR primers for the detection of *S. pneumoniae* were designed from the sequences of the subtracted DNA fragments, and the specificities of the 19 primer sets were evaluated by PCR using chromosomal DNAs extracted from four *S. pneumoniae* clinical isolates and from 20 atypical organisms classified as *S. mitis* or *S. oralis*, which harbored genes encoding the pneumococcal virulence factors autolysin (*lytA*) or pneumolysin (*ply*), as templates. Of the 19 primer sets, two (Spn9802 and Spn9828) did not amplify PCR products from any of the pneumococcus-like streptococcal strains that we examined. The genes containing the Spn9802 and Spn9828 sequences encoded proteins of unknown function that did not correspond to any previously described proteins in other bacteria. These new oligonucleotide primers may be very useful for early and correct diagnosis of *S. pneumoniae* infections.**

*Streptococcus pneumoniae* is a major cause of bacterial disease in humans, including pneumonia, otitis media, septicemia, and meningitis. Two naturally transformable viridans group streptococci, *Streptococcus mitis* and *Streptococcus oralis*, are closely related to *S. pneumoniae*, and classification of these organisms has long been considered difficult (32). Indeed, the nucleotide sequences of the 16S rRNA genes from *S. mitis* and *S. oralis* are over 99% identical to that of *S. pneumoniae* (17). As the potentially pathogenic *S. pneumoniae* is frequently detected in the oral cavity (31), and the commensals *S. mitis* and *S. oralis* are almost always detected (18), it is important to accurately discriminate among them in order to provide accurate diagnoses and treatments.

Various molecular assays have been employed to identify pneumococcal strains and to detect pneumococci directly from clinical samples. PCR-based assays for identifying *S. pneumoniae* have frequently targeted genes encoding pneumococcal virulence factors, including pneumolysin (*ply*) (7), autolysin (*lytA*) (20), pneumococcal surface antigen A (21), manganese-dependent superoxide dismutase (*sodA*) (16), and penicillin binding protein (24). Ideally, amplification of these target genes should be specific for *S. pneumoniae* isolates only. However, the occasional occurrence of viridans group identified as

*S. mitis* or *S. oralis* harbor genes encoding *S. pneumoniae* virulence factors has been reported (31).

In this study, genomic subtractive hybridization was performed to identify genomic differences between the two type strains, *S. pneumoniae* WU2 and *S. mitis* 903. In addition, we designed primers from the sequences of the DNA fragments obtained by subtractive hybridization and examined their diagnostic utility in PCR assays using chromosomal DNA from clinical *S. mitis* and *S. oralis* isolates harboring the *lytA* and/or the *ply* gene. The technique of subtractive hybridization was initially developed to isolate differences in cDNA pools, but it has also been successfully used to identify genomic differences between closely related strains of *Helicobacter pylori* (1), *Mycobacterium tuberculosis* (19), and *Neisseria meningitidis* (25). We describe here the application of this technique to *S. pneumoniae* and the identification of specific nucleotide sequences for discrimination of this pathogenic organism from pneumococcus-like isolates.

### MATERIALS AND METHODS

**Bacterial strains and culture conditions.** *S. pneumoniae* WU2 and *S. mitis* 903 were grown anaerobically (85% N<sub>2</sub>, 10% H<sub>2</sub>, 5% CO<sub>2</sub>) at 37°C in Todd-Hewitt broth (Difco Laboratories, Detroit, Mich.) or on 2% Todd-Hewitt broth agar plates. Chromosomal DNA was purified using the Puregene DNA isolation kit (Gentra Systems, Minneapolis, Minn.) according to the manufacturer's instructions. In addition to *S. pneumoniae* WU2 and *S. mitis* 903, chromosomal DNAs extracted from the following oral bacterial strains were used in PCR assays to evaluate the specificity of the DNA sequences amplified by two newly designed primers: *S. oralis* ATCC 10557, *Streptococcus gordonii* DL1, *Streptococcus sanguinis* ATCC 10556, *Streptococcus mutans* Xc, *Streptococcus sobrinus* 6715, *Streptococcus*

\* Corresponding author. Mailing address: Department of Preventive Dentistry, Kyushu University Faculty of Dental Science, Fukuoka 812-8582, Japan. Phone: 81-92-642-6450. Fax: 81-92-642-6354. E-mail: yoshi@dent.kyushu-u.ac.jp.

TABLE 1. Sequences of oligonucleotide primers

Primer	Sequence	Amplicon size (bp)
<i>lytA</i> sense	5'-CAA CCG TAC AGA ATG AAG CGG-3'	308
<i>lytA</i> antisense	5'-TTA TTC GTG CAA TAC TCG TGC G-3'	
<i>ply</i> sense	5'-ATT TCT GTA ACA GCT ACC AAC GA-3'	329
<i>ply</i> antisense	5'-GAA TTC CCT GTC TTT TCA AAG TC-3'	
Spn9802-143F	5'-CAA GTC GTT CCA AGG TAA CAA GTC T-3'	162
Spn9802-304R	5'-CTA AAC CAA CTC GAC CAC CTC TTT-3'	
Spn14.1-27F	5'-CAC CAT CAT AAA TGC TAG CCT TGA-3'	301
Spn14.1-327R	5'-TTG GAA GAT AGA ACC ATA GTT GCT TG-3'	
Spn9822-17F	5'-CAG GGA TTC GCT TAA TAG CTG GT-3'	202
Spn9822-218R	5'-CCT GCC AAG GTC AAT CCA TAA A-3'	
Spn24-4F	5'-TGT CCC AAT CTC TAA AGG ATG GAG-3'	225
Spn24-228R	5'-CAA CTT GCC CGA TTC ACT TTT C-3'	
Spn25.1-12F	5'-TAA CCA AAA CCT TTA GTT TAG GTT GAT TTT-3'	135
Spn25.1-146R	5'-CTC ACT ATT TAG AGA CTC TTG AGG TTT GTT-3'	
Spn25.2-6F	5'-GAG ACT CCG TGG AAC CTC TTT G-3'	130
Spn25.2-135R	5'-GAC GAA CTG CCA TAT AAG AAC CAT T-3'	
Spn27-11F	5'-CAA GAT TGA ACG TAT CCG CGA-3'	285
Spn27-295R	5'-TGG CCA GTT GAT TCG AGT TTG-3'	
Spn9827-70F	5'-TGG CCC TAA TAC AAC GAT GGA-3'	208
Spn9827-277R	5'-AGC AAG TGA GGG AAA GAC CGT-3'	
Spn9828-19F	5'-GGC ATT GTG AAT GGA TTG ATT G-3'	227
Spn9828-245R	5'-TCA TGT GCA TCC CAA GCT ACA-3'	
Spn34-5F	5'-GGG CCT GTA TCG TTC TAC TAG GG-3'	231
Spn34-235R	5'-TTT CTC TCC AAT CTA CCT TAC CCG-3'	
Spn38-5F	5'-AAT GGT CGT ATT GTT GCC GG-3'	266
Spn38-270R	5'-AAG TAA TCA ATC GCA TCT TGT AGG G-3'	
Spn53-415F	5'-GGG ATT CAA AAT ACA GTA GAG TCC TTG-3'	346
Spn53-760R	5'-TGA GCT TTG CAG CAT AAC GG-3'	
Spn63-191F	5'-GGA CTT TCT TTT TCC CAT CGG-3'	339
Spn63-529R	5'-CCC AGA TTT GTT TGA GCC AAC-3'	
Spn66-10F	5'-AAC AAT CGA ATC CCC TGG G-3'	312
Spn66-321R	5'-AAC TCC ACC AGT CAG GGA TGA-3'	
Spn70-49F	5'-CCT GCG TAG CCC ATA TCT CG-3'	253
Spn70-301R	5'-GGC AGG ATG CCA TTC GAC T-3'	
Spn76-33F	5'-CAT AGA TTC CCC AGA TCT CGG T-3'	351
Spn76-383R	5'-CAG GTC AGA GTG AGC TTT TGG TG-3'	
Spn79-68F	5'-CGC TAG TGA TTG GTG GTG AGC-3'	261
Spn79-328R	5'-CAA GTT CCT GCA AGG TTT CAA GA-3'	
Spn84-93F	5'-CTT GTC AGC CTT TGC TTG GAA-3'	332
Spn84-424R	5'-GCG CTA GAC AGA TCA GCA GGA-3'	
Spn86-24F	5'-CGA AAC ACC AGC AGA AGA GCT-3'	184
Spn86-207R	5'-AAT TTC GCC ACT ATT TAG CCA GAT-3'	

*tococcus salivarius* HT9R, *Streptococcus anginosus* FW73, *Fusobacterium nucleatum* ATCC 10953, *Haemophilus aphrophilus* NCTC 5908, *Eikenella corrodens* 1085, *Actinomyces naeslundii* ATCC 51655, *Actinomyces viscosus* ATCC 43146, *Porphyromonas gingivalis* ATCC 33277, *Tannerella forsythensis* ATCC 43037, *Prevotella intermedia* ATCC 25611, *Actinobacillus actinomycetemcomitans* Y4, and *Treponema denticola* ATCC 35404. *Escherichia coli* DH5 $\alpha$  [F<sup>-</sup>  $\phi$  80 lacZ $\Delta$ M15  $\Delta$ (lacZYA-argF)U169 deoR recA1 endA1 hsdR17 (r<sub>k</sub><sup>-</sup> m<sub>k</sub><sup>+</sup>) phoA supE44  $\lambda$ <sup>-</sup> thi-1, gyrA96, relA1] (Takara Bio Co., Shiga, Japan) was used for the DNA manipulations.

**Subtractive hybridization.** Genomic subtractive hybridization was performed as described by Tinsley and Nassif (30). Briefly, 200 ng of chromosomal DNA from *S. pneumoniae* WU2 was precipitated digested with Sau3AI, precipitated with ethanol-sodium acetate, and ligated with 0.2 nmol of an oligonucleotide R-adaptor pair (RBam12, 5'-GATCCTCGGTGA-3', RBam24, 5'-AGCACTC TCCAGCCTCTCACCGAG-3') for 16 h at 11°C. The ligated DNA fragments were purified using the QIAquick PCR purification kit (QIAGEN, Tokyo, Japan) and stored at -20°C. To prepare the subtractive driver DNA, chromosomal DNA from *S. mitis* 903 was digested with restriction enzymes, including HindIII, EcoRI, and MspI, to yield fragments ranging from approximately 3 to 10 kb. After digestion, the DNA was precipitated with ethanol-sodium acetate and dissolved in TE buffer (10 mM Tris-HCl, 1 mM EDTA; pH 8.0).

The first subtractive hybridization was performed in 10  $\mu$ l of a reaction mixture containing 2  $\mu$ g of the driver DNA from *S. mitis*, 20 ng of the R-adaptor-linked Sau3AI fragments from *S. pneumoniae*, 10 mM EPPS (3-[4-(2-hydroxyethyl)-1-

piperazinyl] propanesulfonic acid; Wako, Tokyo, Japan), 1 mM EDTA, and 1 M NaCl. The DNA mixture was denatured at 100°C for 2 min and then hybridized at 55°C for 18 h. A 2- $\mu$ l aliquot of the reaction mixture was placed on ice and diluted to 40  $\mu$ l with the PCR mixture containing 100 pmol of RBam24, 0.25 mM of each deoxynucleoside triphosphate, 5 U of TaKaRa *ExTaq* and 1 $\times$  *ExTaq* buffer (Takara Bio Co.) to fill in the ends corresponding to the RBam24 adaptor.

After denaturation at 94°C for 5 min, amplification by PCR was performed in a T3 thermocycler (Biometra, Göttingen, Germany) as follows: a total of 30 cycles, each of which comprised 94°C for 1 min, 70°C for 1 min, and 72°C for 3 min, followed by a final extension step at 72°C for 10 min. The RBam24 adaptor was removed from the PCR products by digestion with Sau3AI and the DNA fragments were purified using the QIAquick PCR purification kit. In a 40- $\mu$ l volume, 1 nmol of the second adaptors (JBam12, 5'-GATCCGTTTCATG-3'; JBam24, 5'-ACCGACGTCGACTATCCATGAACG-3') were ligated for 16 h at 11°C. The second-round subtractive hybridization was performed with 2 ng of DNA from the first round PCR products and 2  $\mu$ g of the driver DNA from *S. mitis*, ranging in size from 3 to 10 kb, prepared as described above. The amplification products of the second-round PCR, which appeared as a smear of products smaller than 800 bp upon electrophoresis with ethidium bromide, were digested with Sau3AI, cloned into BamHI-digested pBluescript II SK<sup>+</sup> vector (Stratagene, La Jolla, Calif.), and used to transform *E. coli* DH5 $\alpha$  by heat shock. The nucleotide sequences of the inserts were determined using an ABI PRISM 310 automatic sequencer (Applied Biosystems, Foster City, Calif.) and the manufacturer's services (Genenet, Saga, Japan).

TABLE 2. Characteristics of differential clones of *S. pneumoniae* genomic DNA

Clone	Gene no. <sup>a</sup>	ORF in strain R6 containing differential clone			Other homologues	
		ORF size (aa)	% GC	Gene product and function in strain R6 (accession no.)	Gene product and function (homologues)	E value <sup>b</sup> (% identity)
spn9802 <sup>c</sup>		107	29.6	ND <sup>d</sup>	Unknown ( <i>S. pneumoniae</i> TIGR4)	3e-35 (93%)
spn9822	spr1802	680	43.4	Unknown (NP_359394)	Unknown ( <i>Lactococcus lactis</i> )	4e-59 (25%)
spn9827	spr1781	427	46.9	UDP-N-acetylglucosamine 1-carboxyvinyltransferase (NP_359373)		
spn9828 <sup>c</sup>	spr1523	114	39.1	Unknown (NP_359116)	Unknown ( <i>S. pneumoniae</i> TIGR4)	8e-41 (90%)
spn14.1	spr0087	320	40.6	Unknown (NP_357681)	Unknown ( <i>Streptococcus mutans</i> )	5e-7 (23%)
spn24	spr1092	292	44.9	tRNA pseudouridine 5S synthase (NP_358685)		
spn25.1	spr1622	512	30.0	Unknown (NP_359214)	Transcriptional activator, putative ( <i>S. pneumoniae</i> TIGR4)	0.0 (99%)
spn25.2	spr1179	334	52.9	Unknown (NP_358772)	Predicted iron-dependent peroxidase ( <i>Actinobacillus pleuropneumoniae</i> )	1e-101 (53%)
spn27	spr0058	252	37.3	Unknown (NP_357652)	Transcriptional regulator, GntR family ( <i>S. pneumoniae</i> TIGR4)	1e-135 (100%)
spn34	spr1617	439	40.0	Sucrose-6-phosphate hydrolase (NP_359209)		
spn38	spr1782	345	41.8	Unknown (NP_3593740)	PDZ domain protein ( <i>Enterococcus faecalis</i> )	1e-85 (48%)
spn53	spr0897	589	37.1	Unknown (NP_358491)	Glycerophosphoryl diester phosphodiesterase family protein ( <i>E. faecalis</i> )	3e-75 (40%)
spn63	spr0367	182	42.6	Unknown (NP_357961)	Unknown ( <i>Streptococcus agalactiae</i> )	3e-37 (39%)
spn66	spr2025	279	40.5	ABC transporter ATP-binding protein—unknown substrate (NP_359616)	ABC transporter ( <i>S. mutans</i> )	1e-107 (70%)
spn70	spr1496	924	43.8	Unknown (NP_359089)	ABC transporter permease protein ( <i>L. lactis</i> )	0.0 (39%)
spn76	spr1933	283	37.2	Positive transcriptional regulator of glucosyltransferase and Spp phenotype (NP_359524)	RggD ( <i>Streptococcus gordonii</i> )	1e-62 (45%)
spn79	spr0072	179	44.8	Unknown (NP_357666)	Glycosyltransferase involved in cell wall biogenesis ( <i>Clostridium acetobutylicum</i> )	5e-42 (44%)
spn84	spr1685	318	43.2	ABC transporter membrane-spanning permease—ferric iron transport (NP_359277)	ABC-type enterochelin transport system, permease component ( <i>Enterococcus faecium</i> )	2e-56 (43%)
spn86	spr1715	311	36.2	Biotin-(acetyl-coenzyme A-carboxylase) ligase (NP_359307)		

<sup>a</sup> The gene number named by Hoskins et al. (15).

<sup>b</sup> The significance of the protein sequence homologies in GenBank with the BlastX algorithm.

<sup>c</sup> Insertion sequences absent from DNA of any tested pneumococcus-like oral streptococci.

<sup>d</sup> ND, none detected.

**Analysis of clones and design of PCR primers.** The homologies of the sequences of the genomic subtracted fragments with sequences in the GenBank database were determined using BLAST (2) by means of the server at the National Center for Biotechnology Information, Bethesda, Md. (<http://www.ncbi.nlm.nih.gov/BLAST>). The predicted complete open reading frames (ORFs) of *S. pneumoniae* R6 (GenBank accession number AE 007317) and *S. pneumoniae* TIGR4 (GenBank accession number AE 005672) corresponding to the subtracted genomic fragments from *S. pneumoniae* WU2 were used to search for protein homologies using BlastP (2). Oligonucleotide primer pairs were designed for 46 sequences of interest from the subtracted genomic *S. pneumoniae* WU2 DNA using Primer Express 1.5 software (Applied Biosystems).

The oligonucleotide primers for the *lytA* and *ply* genes were designed as previously described by Nagai et al. (22) and Salo et al. (26), respectively. The primers for detection *lytA* were 5'-CAACCGTACAGAATGAAGCGG-3' (sense) and 5'-TTATTCTGTCAATACTCGTGCG-3' (antisense). The primers for detection *ply* were 5'-ATTCTGTAAACAGCTACCAACGA-3' (sense) and 5'-GAATTCCTGTCTTTTCAAAGTC-3' (antisense).

**PCR assays with candidate primers.** The amplification reactions were conducted using a T3 thermocycler (Biometra) with the following temperature profile: initial denaturation at 94°C for 2 min, then 25 cycles consisting of 94°C for 10 seconds, 58°C for 15 seconds, and 72°C for 1 min, followed by a final extension step at 72°C for 5 min. The amplification products were loaded onto

1.5% (wt/vol) agarose gels, separated by electrophoresis, stained with ethidium bromide (0.5 µl/ml), and photographed under UV light.

**Southern hybridization.** Southern hybridization was performed 6 h under stringent conditions (hybridization fluid with 50% formamide at 50°C). Posthybridization washes were performed twice with 2× SCC (1× SCC is 0.15 M NaCl plus 0.015 M sodium citrate)—0.1% (wt/vol) sodium dodecyl sulfate (SDS) at room temperature for 15 min per wash and twice with 0.1× SCC—0.1% (wt/vol) SDS at room temperature for 15 min per wash. All other procedures that involved Southern hybridization were performed by the methods of Sambrook et al. (27).

**Clinical specimens.** Oral mucosa swab samples from 432 preschool children (5 to 6 years old) were screened by PCR using *lytA* and *ply* as primers. Prior to this study the parents of preschool children were explained the procedure of taking oral mucosa swabs and informed consents were obtained from the parents of all children involved in this study. From 14 oral mucosa swab samples of systematically healthy preschool children, a total of 24 streptococcal isolates suspected of being pneumococci were identified on the basis of colony morphology of alpha-hemolytic streptococci on blood agar plates.

Although PCR amplification using the *lytA*- and/or the *ply*-specific primer gave positive results for all of these isolates, most could not be classified as *S. pneumoniae* based on optochin resistance and bile insolubility in addition to the results of the biotyping tests from the API 20 STREP system (BioMérieux,

TABLE 3. Specificity of the primers designed from subtractive DNA

Primer set	Result for type strain		Result for atypical oral streptococcal strains																		Result for clinical pneumococcal isolate:					
	WU2	903	1 <sup>a</sup>	2	3	4	5	6 <sup>a</sup>	7	8	9	10	11	12	13	14 <sup>a</sup>	15	16	17	18	19	20	21	22	23	24
<i>lytA</i>	+	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	+	+	+	+	+
<i>ply</i>	+	-	+	-	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	-	+	+	+	+
Spn9802 <sup>b</sup>	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	+
Spn9822	+	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	+	-	+	+	-	+	+	+	+	+
Spn9827	+	-	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	+
Spn9828 <sup>b</sup>	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	+
Spn14.1	+	-	-	+	-	-	-	-	-	-	-	-	-	-	+	+	-	+	-	-	+	-	+	+	+	+
Spn24	+	-	+	-	+	+	+	-	+	+	+	+	+	+	-	+	+	+	-	-	-	+	+	+	+	+
Spn25.1	+	-	-	-	+	+	+	-	+	+	-	+	-	-	-	-	-	-	-	-	-	+	+	+	+	+
Spn25.2	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	-	-	-	-	-	-	-	-	-
Spn27	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	+	-	-	-	-	+	+	+	+
Spn34	+	-	+	-	-	-	-	+	-	-	-	-	+	-	-	-	-	-	-	-	-	-	+	+	+	+
Spn38	+	-	+	+	+	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Spn53	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	+	+	+	+
Spn63	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Spn66	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Spn70	+	-	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Spn76	+	-	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	+	+	+	+	+	+	+
Spn79	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Spn84	+	-	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	+	+	-	+	+	+	+
Spn86	+	-	+	-	+	+	+	+	+	+	-	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+

<sup>a</sup> Strains 1, 6, and 14 are classified as *S. oralis* on the basis of phenotypic tests and API biotyping and additional biochemical tests. The other strains are classified as *S. mitis*.

<sup>b</sup> No amplifying DNA from any atypical streptococci tested but amplifying DNA from all typical and clinically isolated pneumococci tested.

Tokyo, Japan). Optochin sensitivity and bile solubility tests were performed as previously described (31). API 20 STREP was used according to the manufacturer's instructions. Of the 24 strains tested, four were classified as *S. pneumoniae* on the basis of the total consideration through these conventional identification methods and 20 were classified as *S. mitis* or *S. oralis*.

**RESULTS**

**Subtractive hybridization.** Approximately 200 transformed *E. coli* colonies were obtained from cloning the second-round PCR products by direct genomic subtraction of *S. mitis* 903 from *S. pneumoniae* WU2. The nucleotide sequences of 92 randomly selected clones were determined. The inserted fragments of genomic subtracted DNA varied from 150 to 600 bp and contained diverse sequences; only three clones were identical and additional four clones contained two distinct Sau3AI fragments in a single clone. A nucleotide database search using BlastN (2) revealed that 68 of the Sau3AI fragments had high homology with the genome sequence of other *S. pneumoniae* strains, and 46 of these had no significant sequence homology with any DNA sequence in the GenBank database other than *S. pneumoniae*.

Initially, 46 pairs of *S. pneumoniae*-specific primers were designed from the sequences of the subtracted genomic fragments. The nucleotide sequences amplified by each primer pair shared no significant identity with any previously reported nucleotide sequences in a nucleotide database search (BlastN) (2). The specificities of the primers were examined by PCR using chromosomal DNA extracted from 18 oral bacterial species, including *S. mitis* 903 in addition to *E. coli* DH5 $\alpha$ , as listed in Materials and Methods. The PCR analysis showed that 19 primer pairs amplified only the chromosomal DNA from *S. pneumoniae* WU2 (data not shown). The oligonucleotide se-

quences of these primer sets are listed in Table 1. The predicted sizes of the PCR amplicons ranged from 130 to 351 bp.

**Sequence homologies.** The characteristics of the 19 *S. pneumoniae*-specific sequences that were shown to be absent in the other tested oral bacteria are summarized in Table 2. All 19 subtracted genomic fragments showed high sequence identity with the genomic sequences of *S. pneumoniae* R6 (15) in the GenBank database. A search of the protein database using BlastP (2) was performed using the predicted complete ORFs of *S. pneumoniae* R6 corresponding to each of the 19 subtracted genomic fragments from *S. pneumoniae* WU2. Of the five ORFs corresponding to the sequences of clones spn9802, spn9822, spn9828, spn14.1, and spn63 that encoded proteins of unknown function, two ORFs corresponding to spn9802 and spn9828 showed no significant homologies with any organisms other than *S. pneumoniae*. A DNA sequence identical to that of clone spn9802 was found in *S. pneumoniae* R6; however, its 107-amino-acid ORF constituted has not been deposited in the GenBank database. In *S. pneumoniae* TIGR4, a similar ORF that shared 93% identity with that derived from spn9802 was identified. The ORF was six amino acids longer than that of the strain R6 at the N terminus.

The spn25.1 and spn25.2 fragments were isolated in the same clone but represented independent Sau3AI fragments corresponding to 1596789 to 1598270 and 1177462 to 1178466, respectively, in the *S. pneumoniae* R6 genome sequence. The ORF containing spn25.1 showed high amino acid sequence identity with a putative transcriptional activator of *S. pneumoniae* TIGR4 and had no significant homology with known proteins from any other species. The ORF containing spn25.2 had 53% identity with a predicted iron-dependent peroxidase

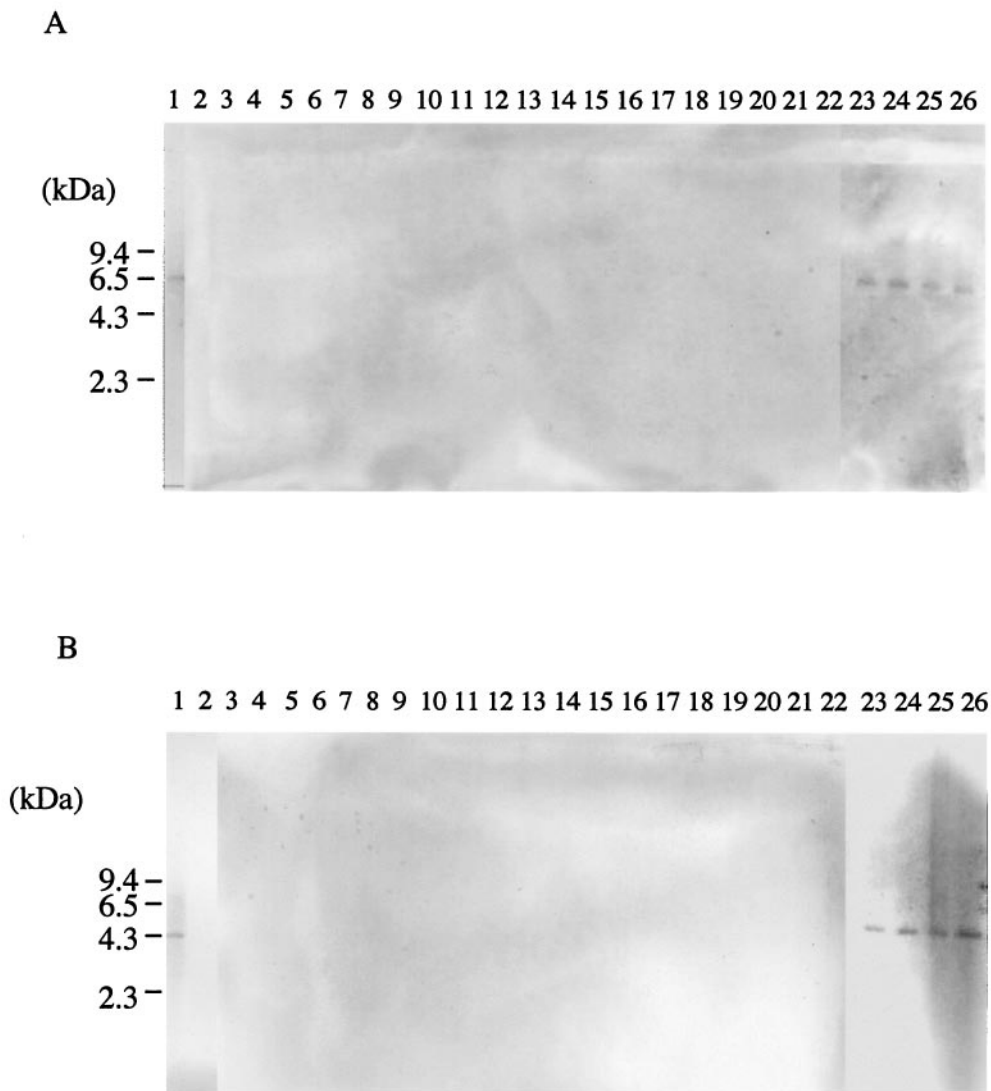


FIG. 1. Southern hybridization with a digoxigenin-labeled probe to the ORF sequence containing spn9802 (A) and spn9828 (B) of *S. pneumoniae* WU2. Chromosomal DNAs from *S. pneumoniae* WU2 (lane 1), *S. mitis* 903 (lane 2), atypical clinical streptococci (lane 3 to 22), and clinical isolated pneumococci (lane 23 to 26) were digested by either PstI (A) or NcoI (B). The lack of clear results in lanes 3 to 20 is due to negative reactions to the digoxigenin-labeled probes.

of *Actinobacillus pleuropneumoniae*, the etiologic agent of swine pleuropneumonia (10).

The two ORFs containing clones spn9827 and spn79 showed high identity with predicted proteins involved in cell wall biosynthesis. The gene containing clone spn9827 encoded UDP-*N*-acetylglucosamine 1-carboxyvinyltransferase, which catalyzes the first committed step of biosynthesis of cell wall peptidoglycan (6). The derived amino acid sequence from the ORF containing clone spn79 shared 44% identity with a glycosyltransferase involved in cell wall biogenesis in *Clostridium acetobutylicum* (23). Three clones (spn66, spn70, and spn84) encoding ABC-type transporters and two clones (spn27 and spn76) encoding transcriptional regulators were also found. The ORF containing spn53, which was a protein of unknown function from *S. pneumoniae*, shared 40% identity with the periplasmic enzyme GlpQ of *Enterococcus faecalis*. The *glpQ* gene encodes a glycerophosphoryl diester phosphodiesterase,

which is involved in the hydrolysis of deacetylated phospholipids and is expressed strongly after phosphate starvation (3).

The GC contents of the 19 ORFs examined ranged from 29.6 to 46.5% and most were similar to the predicted GC percentage (40%) of the *S. pneumoniae* R6 genome (15). However, the finding that the GC contents of spn9802 and spn25.1 were below 30% suggests that interspecific transfer of these genes from other species with lower GC content than the *S. pneumoniae* strain may have occurred.

The DNA fragments involved in the *lytA* and *ply* genes have not been isolated in 92 random selected colonies.

**Clinical specimens.** Nineteen primer sets designed from the subtracted DNA sequences were examined by PCR using as templates DNAs extracted from a total of 20 atypical *S. mitis* or *S. oralis* isolates harboring the *lytA* or *ply* genes classically associated with *S. pneumoniae* and four clinical isolates classified as *S. pneumoniae* (Table 3). Although chromosomal DNA

extracted from *S. mitis* 903 was not amplified with these 19 primer sets, several primer sets resulted in amplification of DNA from the pneumococcus-like clinical isolates. The primers for clones spn63 and spn66 yielded PCR products of the predicted size with all of the clinical specimens. The primer set designed from clone spn25.2 yielded no amplification with any *S. pneumoniae* clinical isolates, but the chromosomal DNAs from two pneumococcus-like isolates were amplified.

In contrast, none of the DNAs from pneumococcus-like isolates were amplified with the primers for clones spn9802 and spn9828, and the expected bands of 162 bp for spn9802 and 227 bp for spn9828 were observed from all strains of *S. pneumoniae* (data not shown). Furthermore, Southern hybridization analysis with the digoxigenin-labeled ORF fragments including spn9802 and spn9828 as probes showed that the same 6.5-kb PstI fragments (spn9802) and 4.3-kb NcoI fragments (spn9828) hybridized from chromosomal DNA from WU2 and all *S. pneumoniae* clinical isolates (Fig. 1). Consequently, our results revealed that two primer sets, Spn9802 and Spn9828, were able to discriminate *S. pneumoniae* from all of the pneumococcus-like streptococci tested.

## DISCUSSION

Genomic subtractive hybridization allows DNA sequences that are unique to the tester organism relative to the driver organism to be easily isolated and analyzed. Using this technique, we identified 19 DNA fragments that are present in the genomic DNA of *S. pneumoniae* WU2 and absent from *S. mitis* 903. All 19 primer sets designed from the sequences of these subtracted DNA fragments were specific for *S. pneumoniae* WU2, as shown by PCR using chromosomal DNAs from typical laboratory strains of oral bacteria as templates. However, when PCR was conducted using DNA from atypical clinical isolates, only two primer sets, Spn9802 and Spn9828, discriminated *S. pneumoniae* from pneumococcus-like strains. The ORFs containing clones spn9802 and spn9828 encoded proteins of unknown function and had no significant homologies with any organisms in the database other than *S. pneumoniae*. We are currently undertaking experiments to characterize these genes, which may play important roles in *S. pneumoniae*.

The clinical isolates of *S. mitis* and *S. oralis* showed various patterns by PCR amplification with the 19 primer sets tested (Table 3). Strains 1 to 9 were isolated from one individual and the other 11 strains of *S. mitis* and *S. oralis* were isolated from different individuals. Strains 3, 4, 5, 7, 8, and 10, and strains 17 and 18 showed the same amplification patterns by PCR and were classified as *S. mitis*. Strains 1 and 11 also showed the same pattern by PCR but were classified as *S. oralis* and *S. mitis*, respectively, indicating that our primer sets could not detect the genetic differences between *S. mitis* and *S. oralis* clinical isolates.

The remaining isolates were found to display distinct patterns. The extensive diversity seen within the *S. mitis* and *S. oralis* groups is consistent with a number of previous reports (11, 14). In addition, interspecies recombination events between these organisms and *S. pneumoniae* have been previously reported, such as the transfer of genes encoding penicillin-binding proteins (9) and the *lytA* or *ply* genes (31).

Pneumococci can be divided into at least 90 serotypes ac-

ording to the immunochemistry of their capsular polysaccharide (13) and approximately 90% of the invasive diseases worldwide are caused by 16 different serotypes (12). Therefore, more investigations using our primer sets in PCR assays with chromosomal DNA from many strains of *S. pneumoniae* and from various pneumococcus-like isolates are required. However, two primer sets designed in this study, Spn9802 and Spn9828, may contribute to the detection of *S. pneumoniae* by PCR as a means of providing definitive diagnoses, especially in combination with the classical primer sets, such as those for *lytA* and *ply*.

The viridans group streptococci, including *S. mitis* and *S. oralis*, constitute a major population of oral soft tissue, such as the ventral tongue, the floor of the mouth, buccal surfaces, and attached gingiva (18). These species are commensal organisms of healthy individuals but are also a frequent cause of bacterial endocarditis in patients with prosthetic valves (8) and bacteremia in neutropenic cancer patients (4). Several cell surface proteins of viridans group streptococci have been reported to contribute to platelet binding and endothelial cell invasion (5, 28). However, the mechanisms by which viridans group streptococci cause infections are not well understood.

The pathogenicity of clinical isolates that are most closely related to *S. mitis* and *S. oralis* but which harbor pneumococcal genes is unclear. Genomic subtractive hybridization has the potential to pinpoint regions of the chromosome that are likely to be involved in the differential virulence of bacterial pathogens (25). Our study indicates that this technique is also useful for designing species-specific primers for PCR detection, especially for organisms for which the whole genome sequence is not yet available.

Recently, to quantify early colonizer microorganisms in dental biofilms, we designed species-specific primers and probes for real-time TaqMan PCR using this technique in combination with Southern blots (29). The subtractive technique may be applicable to the rational design of new vaccines and to the study of microbial ecosystems consisting of related species in the oral cavity and the development of new diagnostic tests to distinguish pathogenic bacteria from closely related and atypical strains of lesser pathogenicity.

## ACKNOWLEDGMENTS

This investigation was supported primarily by the Sato Fund, Nihon University School of Dentistry (M.S.), by a Nihon University Research Grant for Assistants and Young Researchers (M.S.), by a Grant-in-Aid for Scientific Research (C) (no. 16592097) (M.S.), by a grant to promote multidisciplinary research projects from the Ministry of Education, Culture, Sports, Science and Technology of Japan, by research fellowships from the Japan Society for the Promotion of Science for Young Scientists (N.S.), and by a research grant from the Fukushima Society for the Promotion of Medicine (N.S.).

## REFERENCES

1. Akopyants, N. S., A. Fradkov, L. Diatchenko, J. E. Hill, P. D. Siebert, S. A. Lukyanov, E. D. Sverdlov, and D. E. Berg. 1998. PCR-based subtractive hybridization and differences in gene content among strains of *Helicobacter pylori*. Proc. Natl. Acad. Sci. USA 95:13108-13113.
2. Altschul, S. F., T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25:3389-3402.
3. Antelmann, H., C. Scharf, and M. Hecker. 2000. Phosphate starvation-inducible proteins of *Bacillus subtilis*: proteomics and transcriptional analysis. J. Bacteriol. 182:4478-4490.
4. Beighton, D., A. D. Carr, and B. A. Oppenheim. 1994. Identification of

- viridans streptococci associated with bacteraemia in neutropenic cancer patients. *J. Med. Microbiol.* **40**:202–204.
5. Bensing, B. A., I. R. Siboo, and P. M. Sullam. 2001. Proteins PblA and PblB of *Streptococcus mitis*, which promote binding to human platelets, are encoded within a lysogenic bacteriophage. *Infect. Immun.* **69**:6186–6192.
  6. Brown, E. D., E. I. Vivas, C. T. Walsh, and R. Kolter. 1995. MurA (MurZ), the enzyme that catalyzes the first committed step in peptidoglycan biosynthesis, is essential in *Escherichia coli*. *J. Bacteriol.* **177**:4194–4197.
  7. Corless, C. E., M. Guiver, R. Borrow, V. Edwards-Jones, A. J. Fox, and E. B. Kaczmarek. 2001. Simultaneous detection of *Neisseria meningitidis*, Haemophilus influenzae, and *Streptococcus pneumoniae* in suspected cases of meningitis and septicemia using real-time PCR. *J. Clin. Microbiol.* **39**:1553–1558.
  8. Douglas, C. W., J. Heath, K. K. Hampton, and F. E. Preston. 1993. Identity of viridans streptococci isolated from cases of infective endocarditis. *J. Med. Microbiol.* **39**:179–182.
  9. Dowson, C. G., V. Barcus, S. King, P. Pickerill, A. Whatmore, and M. Yeo. 1997. Horizontal gene transfer and the evolution of resistance and virulence determinants in *Streptococcus*. *Soc. Appl. Bacteriol. Symp. Ser.* **26**:42S–51S.
  10. Fittipaldi, N., A. Broes, J. Harel, M. Kobisch, and M. Gottschalk. 2003. Evaluation and field validation of PCR tests for detection of *Actinobacillus pleuropneumoniae* in subclinically infected pigs. *J. Clin. Microbiol.* **41**:5085–5093.
  11. Fitzsimmons, S., M. Evans, C. Pearce, M. J. Sheridan, R. Wientzen, G. Bowden, and M. F. Cole. 1996. Clonal diversity of *Streptococcus mitis* biovar 1 isolates from the oral cavity of human neonates. *Clin. Diagn. Lab. Immunol.* **3**:517–522.
  12. Hakenbeck, R., N. Balmelle, B. Weber, C. Gardès, W. Keck, and A. de Saizieu. 2001. Mosaic genes and mosaic chromosomes: intra- and interspecies genomic variation of *Streptococcus pneumoniae*. *Infect. Immun.* **69**:2477–2486.
  13. Henrichsen, J. 1995. Six newly recognized types of *Streptococcus pneumoniae*. *J. Clin. Microbiol.* **33**:2759–2762.
  14. Hohwy, J., and M. Kilian. 1995. Clonal diversity of the *Streptococcus mitis* biovar 1 population in the human oral cavity and pharynx. *Oral Microbiol. Immunol.* **10**:19–25.
  15. Hoskins, J., W. E. Alborn, Jr., J. Arnold, L. C. Blaszczyk, S. Burgett, B. S. DeHoff, S. T. Estrem, L. Fritz, D. J. Fu, W. Fuller, C. Geringer, R. Gilmour, J. S. Glass, H. Khoja, A. R. Kraft, R. E. Lagace, D. J. LeBlanc, L. N. Lee, E. J. Lefkowitz, J. Lu, P. Matsushima, S. M. McAhren, M. McHenry, K. McLeaster, C. W. Mundy, T. I. Nicas, F. H. Norris, M. O'Gara, R. B. Peery, G. T. Robertson, P. Rockey, P. M. Sun, M. E. Winkler, Y. Yang, M. Young-Bellido, G. Zhao, C. A. Zook, R. H. Baltz, S. R. Jaskunas, P. R. Rosteck, Jr., P. L. Skatrud, and J. I. Glass. 2001. Genome of the bacterium *Streptococcus pneumoniae* strain R6. *J. Bacteriol.* **183**:5709–5717.
  16. Kawamura, Y., R. A. Whiley, S.-E. Shu, T. Ezaki, and J. M. Hardie. 1999. Genetic approaches to the identification of the mitis group within the genus *Streptococcus*. *Microbiology* **145**:2605–2613.
  17. Kawamura, Y., X.-G. Hou, F. Sultana, H. Miura, and T. Ezaki. 1995. Determination of 16S rRNA sequences of *Streptococcus mitis* and *Streptococcus gordonii* and phylogenetic relationships among members of the genus *Streptococcus*. *Int. J. Syst. Bacteriol.* **45**:406–408.
  18. Mager, D. L., L. A. Ximenez-Fyvie, A. D. Haffajee, and S. S. Socransky. 2003. Distribution of selected bacterial species on intraoral surfaces. *J. Clin. Periodontol.* **30**:644–654.
  19. Mahairas, G. G., P. J. Sabo, M. J. Hickey, D. C. Singh, and C. K. Stover. 1996. Molecular analysis of genetic differences between *Mycobacterium bovis* BCG and virulent *M. bovis*. *J. Bacteriol.* **178**:1274–1282.
  20. McAvin, J. C., P. A. Reilly, R. M. Roudabush, W. J. Barnes, A. Salmen, G. W. Jackson, K. K. Beninga, A. Astorga, F. K. McCleskey, W. B. Huff, D. Niemeyer, and K. L. Lohman. 2001. Sensitive and specific method for rapid identification of *Streptococcus pneumoniae* using real-time fluorescence PCR. *J. Clin. Microbiol.* **39**:3446–3451.
  21. Morrison, K. E., D. Lake, J. Crook, G. M. Carlone, E. Ades, R. Facklam, and J. S. Sampson. 2000. Confirmation of *psaA* in all 90 serotypes of *Streptococcus pneumoniae* by PCR and potential of this assay for identification and diagnosis. *J. Clin. Microbiol.* **38**:434–437.
  22. Nagai, K., Y. Shibasaki, K. Hasegawa, T. A. Davies, M. R. Jacobs, K. Ubukata, and P. C. Appelbaum. 2001. Evaluation of PCR primers to screen for *Streptococcus pneumoniae* isolates and beta-lactam resistance, and to detect common macrolide resistance determinants. *J. Antimicrob. Chemother.* **48**:915–918.
  23. Nöbling, J., G. Breton, M. V. Omelchenko, K. S. Makarova, Q. Zeng, R. Gibson, H. M. Lee, J. Dubois, D. Qiu, J. Hitti, Y. I. Wolf, R. L. Tatusov, F. Sabathe, L. Doucette-Stamm, P. Soucaille, M. J. Daly, G. N. Bennett, E. V. Koonin, and D. R. Smith. 2001. Genome sequence and comparative analysis of the solvent-producing bacterium *Clostridium acetobutylicum*. *J. Bacteriol.* **183**:4823–4838.
  24. O'Neill, A. M., S. H. Gillespie, and G. C. Whiting. 1999. Detection of penicillin susceptibility in *Streptococcus pneumoniae* by *pbp2b* PCR-restriction fragment length polymorphism analysis. *J. Clin. Microbiol.* **37**:157–160.
  25. Perrin, A., X. Nassif, and C. Tinsley. 1999. Identification of regions of the chromosome of *Neisseria meningitidis* and *Neisseria gonorrhoeae* which are specific to the pathogenic *Neisseria* species. *Infect. Immun.* **67**:6119–6129.
  26. Salo, P., A. Ortvist, and M. Leinonen. 1995. Diagnosis of bacteremic pneumococcal pneumonia by amplification of pneumolysin gene fragment in serum. *J. Infect. Dis.* **171**:479–482.
  27. Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. *Molecular cloning: a laboratory manual*, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
  28. Stinson, M. W., S. Alder, and S. Kumar. 2003. Invasion and killing of human endothelial cells by viridans group streptococci. *Infect. Immun.* **71**:2365–2372.
  29. Suzuki, N., Y. Nakano, A. Yoshida, Y. Yamashita, and Y. Kiyoura. 2004. Real-time TaqMan PCR for quantifying oral bacteria during biofilm formation. *J. Clin. Microbiol.* **42**:3827–3830.
  30. Tinsley, C. R., and X. Nassif. 1996. Analysis of the genetic differences between *Neisseria meningitidis* and *Neisseria gonorrhoeae*: two closely related bacteria expressing two different pathogenicities. *Proc. Natl. Acad. Sci. USA* **93**:11109–11114.
  31. Whatmore, A. M., A. Efstratiou, A. P. Pickerill, K. Broughton, G. Woodard, D. Sturgeon, R. George, and C. G. Dowson. 2000. Genetic relationships between clinical isolates of *Streptococcus pneumoniae*, *Streptococcus oralis*, and *Streptococcus mitis*: characterization of “atypical” pneumococci and organisms allied to *S. mitis* harboring *S. pneumoniae* virulence factor-encoding genes. *Infect. Immun.* **68**:1374–1382.
  32. Whiley, R. A., and D. Beighton. 1998. Current classification of the oral streptococci. *Oral Microbiol. Immunol.* **13**:195–216.