

A Homologue of *aliB* Is Found in the Capsule Region of Nonencapsulated *Streptococcus pneumoniae*

Lucy J. Hathaway,[†] Patricia Stutzmann Meier,[†] Patrick Bättig,
Suzanne Aebi, and Kathrin Mühlemann*

*Institute for Infectious Diseases, University of Bern, and University
Hospital, Bern, Switzerland*

Received 7 July 2003/Accepted 12 March 2004

The epidemiology, phylogeny, and biology of nonencapsulated *Streptococcus pneumoniae* are largely unknown. Increased colonization capacity and transformability are, however, intriguing features of these pneumococci and play an important role. Twenty-seven nonencapsulated pneumococci were identified in a nationwide collection of 1,980 nasopharyngeal samples and 215 blood samples obtained between 1998 and 2002. On the basis of multilocus sequence typing and capsule region analysis we divided the nonencapsulated pneumococci into two groups. Group I was closely related to encapsulated strains. Group II had a clonal population structure, including two geographically widespread clones able to cause epidemic conjunctivitis and invasive diseases. Group II strains also carried a 1,959-bp homologue of *aliB* (*aliB*-like ORF 2) in the capsule region, which was highly homologous to a sequence in the capsule region of *Streptococcus mitis*. In addition, strains of the two major clones in group II had an additional sequence, *aliB*-like ORF 1 (1,968 to 2,004 bp), upstream of *aliB*-like ORF 2. Expression of *aliB*-like ORF 1 was detected by reverse transcription-PCR, and the corresponding RNA was visualized by Northern blotting. A gene fragment homologous to *capN* of serotypes 33 and 37 suggests that group II strains were derived from encapsulated pneumococci some time ago. Therefore, loss of capsule expression *in vivo* was found to be associated with the importation of one or two *aliB* homologues in some nonencapsulated pneumococci.

Streptococcus pneumoniae is an important human pathogen that causes meningitis, sepsis, pneumonia, and upper respiratory tract infections, such as acute otitis. This gram-positive bacterium is usually surrounded by a polysaccharide capsule, which acts as a virulence factor by protecting the bacterium from phagocytosis. The capsule is the main target for protective antibodies, and pneumococcal strains are grouped into serotypes based on the antigenic properties of their capsules. The capsule is also the main target of currently used pneumococcal vaccines.

Some *S. pneumoniae* strains do not appear to bind anticapsular antibodies and, when cultured on agar, form rough colonies rather than the smooth colonies formed by bacteria with capsules. Nonencapsulated strains are less virulent than encapsulated strains in a mouse model (12), but nonencapsulated strains can also be responsible for disease, as described by Martin et al. (17), who found that an outbreak of conjunctivitis on a college campus was due to a nonencapsulated clone.

Little is known about the epidemiology and population genetics of nonencapsulated pneumococci. Whatmore et al. (30) described two groups of *S. pneumoniae* which lack typical features of pneumococci: (i) strains which, based on the sequence of housekeeping genes, appear to be genetically distinct from typical encapsulated strains and lack one or more of the three characteristics bile solubility, optochin sensitivity, and capsule expression expected of *S. pneumoniae*, and (ii) strains which

lack capsule expression but genetically appear to be closely related to typical encapsulated strains.

Nonencapsulated strains adhere better to respiratory epithelial cells and have an advantage when the nasopharynx is colonized (1, 6, 27). Also, the lack of a capsule makes them more transformable (20, 28). Encapsulated pneumococci may therefore take advantage of down-regulation or transitory loss of capsule expression while they colonize the nasopharynx (25, 26). A mechanism that has been shown to prevent capsule expression *in vitro* is inactivation of capsule genes by duplications which can later be excised, causing reversion to the encapsulated phenotype (25, 26). Also, a single mutation can lead to the loss of capsule expression for serotype 3 (3).

In this paper we describe a novel finding. Loss of capsule expression *in vivo* was observed to be associated with the importation of an *aliB* homologue in some nonencapsulated pneumococci.

(Some of the data were presented at the 103rd General Meeting of the American Society for Microbiology, Washington, D.C., 18 to 22 May 2003.)

MATERIALS AND METHODS

Bacterial strains. The colonizing strains used for this study originated from a strain collection (1,980 isolates) obtained from a nationwide surveillance for nasopharyngeal isolates of *S. pneumoniae* in 1998 and 1999 and in 2001 and 2002 (19). The invasive strains of *S. pneumoniae* used (an invasive strain was defined as a culture from a sterile body site) originated from nationwide collection of all invasive isolates in 1998 and 1999 (23) and between March and December 2002 (24). *S. pneumoniae* was grown and identified and antibiotic susceptibility was determined as described previously (19). Bacteria were routinely grown on Columbia sheep blood agar (CSBA) plates at 37°C in a 5% CO₂-enriched atmosphere or in brain heart infusion (BHI) broth containing 5% fetal calf serum (FCS) at 37°C in ambient air and were stored at –80°C by using Protect bacterial preservers (TSC, Heywood, United Kingdom).

* Corresponding author. Mailing address: Institute for Infectious Diseases, University of Bern, Friedbühlstrasse 51, CH-3010 Bern, Switzerland. Phone: 41 632 32 59. Fax: 41 632 87 66. E-mail: kathrin.muehlemann@ifik.unibe.ch.

[†] L.J.H. and P.S.M. contributed equally to this work.

TABLE 1. Primers used for characterization of the cap region

Primer ^a	Sequence ^b	Use	Reference
FI3	TCTTAGTTCATGGGATGCTTCTGTGTG	Long-range PCR, Southern blot probe cap110.58, sequencing of cap region, and in vitro translation of <i>aliB</i> -like ORF 1	11
FI4	CGCTGAACTTTTGTAGTTGCTGTCTGGTCAAC	Long-range PCR and sequencing of cap region	11
aliA_b832	TGCAGGTTTGTGGTATCTTGAC	Long-range PCR, Southern blot probe cap110.58, and sequencing of cap region	This study
cpsA3	ATCCTTGTCAGCTCTGTGTC	PCR for Southern blot probe cap111.46	14
cpsB2	TCACTTGCAACTACATGAAC	PCR for Southern blot probe cap111.46	14
104_b832.12	GTATCTGAACCTGATTGTCCGC	In vitro translation of <i>aliB</i> -like ORF 1	This study
110_FI3.4	AACACTTGGAAACGGAGAATG	PCR detection of <i>aliB</i> -like ORF 1 and RT-PCR	This study
104_b832.14	GCCCTTTGTTATACCTAGATGTTTC	PCR detection of <i>aliB</i> -like ORF 1 and RT-PCR	This study
104_FI3.6	AGATGCCAAATGGTTCACGG	PCR detection of <i>aliB</i> -like ORF 2 and PCR for Southern blot <i>aliB</i> -like ORF 2 probe	This study
104_b832.10	GAAATCTTTTAACAAATAAGGTCCG	PCR detection of <i>aliB</i> -like ORF 2 and PCR for Southern blot <i>aliB</i> -like ORF 2 probe	This study
Rif-399f	GATGAYATCGAYCACCTCGGAAA	RT-PCR, control	18
Rif-525r	GATGTTAGGTCCTTCAGGTGTCTC	RT-PCR, control	18
dexB_f145_XbaI	GCTCTAGAGCTGGCTTCTCCGTTTATGACA	PCR for mutant construction	This study
dexB_b1504_BamHI	CGGGATCCCGAGACAGATTTGACGTTTCCTTC	PCR for mutant construction	This study
capN_F1_ClaI	CCATCGATGGTTCGTTGGTGTGTTTGTCA	PCR for mutant construction	This study
capN_B1_XhoI	CCGCTCGAGGCACTGCTTTGAGGTTGTAGATAA	PCR for mutant construction	This study
cat_Spn_F1_HindIII	CCCAAGCTTGGATTTTTCGCTACGCTCAA	PCR for mutant construction	This study
cat_Spn_B1_HindIII	CCCAAGCTTTCACCTTATTCAGGCGTAGACCAG	PCR for mutant construction	This study

^a Primers used for primer walking sequencing of the different cap regions are not shown.

^b Recognition sites for restriction enzymes within the primers are underlined.

Serotyping. Isolates were serotyped by using the Quellung reaction with specific antisera from the Statens Serum Institute (Copenhagen, Denmark). Strains which did not react with any of the pool sera were retyped after growth under anaerobic conditions to enhance capsule expression (29). All nontypeable strains were sent to the Statens Serum Institute for retyping.

DNA methods. Chromosomal DNA were obtained from *S. pneumoniae* isolates as described previously (18). Southern blotting was performed essentially as described previously (21), with the following exceptions. Probes were labeled nonradioactively by using the ECL direct nucleic acid labeling and detection system (Amersham Biosciences, Duerbendorf, Switzerland) according to the manufacturer's instructions. The same system was used to detect signals. Chromosomal DNA was digested with restriction enzyme PvuII. The primers and probes used for Southern blotting are described in Table 1 and Fig. 1. Probes were amplified by using Taq DNA polymerase (Roche Molecular Biochemicals, Rotkreuz, Switzerland). Restriction enzymes were purchased from New England Biolabs (Frankfurt am Main, Germany).

Multilocus sequence typing (MLST). Amplification, sequencing, and analysis of the seven housekeeping genes were carried out as described previously (8).

PCR amplification and sequencing of the cap region. The cap region of individual strains was amplified with forward primer FI3 and reverse primer FI4 or aliA_b832 (Table 1) by using the Expand Long Template PCR system (Roche) according to the manufacturer's instructions. Amplification was performed by using the following cycling conditions: primary denaturation for 2 min at 92°C, followed by 10 cycles consisting of 92°C for 10 s, 65°C for 30 s, and 68°C for 17 min and then 20 cycles in which each extension cycle was prolonged by 20 s. PCR products were purified with a QIAquick PCR purification kit (QIAGEN, Basel, Switzerland). The entire cap region of strains 104.72, 106.44, 110.58, and 208.56 was sequenced by primer walking performed by Microsynth GmbH (Balgach, Switzerland) or in house with a ABI PRISM 310 genetic analyzer (PE Biosystems, Rotkreuz, Switzerland) by using a Big Dye terminator cycle sequencing Ready Reaction kit (PE Biosystems). DNA sequences were analyzed and aligned by using the Lasergene software (DNASTAR Inc., Madison, Wis.). Open reading frames (ORF) within the cap regions were analyzed further by performing a translated BLAST search (blastx) at the National Center for Biotechnology

Information website (<http://www.ncbi.nlm.nih.gov/BLAST/>). PCRs for detection of *aliB*-like ORF 1 and *aliB*-like ORF 2 were carried out by using Taq DNA polymerase (Roche). The DNA molecular weight standard used was Marker X (Roche).

RNA isolation. Bacteria were subcultured in BHI broth containing 5% FCS and were incubated at 37°C until the optical density at 600 nm was 0.6. Twice the culture volume of RNAprotect (QIAGEN) was added to prevent further transcription and to prevent RNA degradation. After vortexing and 5 min of incubation at room temperature, the bacteria were pelleted by centrifugation for 10 min at 5,000 × *g*. The pellet was resuspended in 200 μl of TE buffer (10 mM Tris-Cl, 1 mM EDTA; pH 8) containing 600 μg of lysozyme. After incubation for 10 min at room temperature with repeated vortexing, 700 μl of RLT buffer (QIAGEN RNeasy mini kit) containing β-mercaptoethanol was added, and the tube was vortexed. The mixture was transferred to a 1.5-ml tube containing 0.05 g of 100-μm-diameter acid-washed glass beads (Sigma) and vortexed for 5 min. The mixture was then centrifuged, and the RNA was extracted from the supernatant by using a QIAGEN RNeasy mini kit according to the manufacturer's instructions. The RNA recovered was treated with DNase I (Stratagene Europe, Amsterdam, The Netherlands) used according to the manufacturer's instructions to remove any contaminating DNA.

RT-PCR. Reverse transcription (RT)-PCR was performed by using a QIAGEN One Step RT-PCR kit according to the manufacturer's instructions or in two steps to allow inclusion of a reverse transcriptase negative control to confirm the absence of contaminating DNA. RT was performed by using an Omniscript kit (QIAGEN) according to the manufacturer's instructions, and the PCR was performed under the following conditions by using Taq DNA polymerase (Roche): 94°C for 3 min and then 30 cycles of 94°C for 1 min, 55°C for 1 min, and 72°C for 1 min, followed by 10 min at 72°C. The primers used are shown in Table 1. The PCR products were visualized in an agarose gel by staining with ethidium bromide.

Northern blotting. Northern blotting was performed by standard methods (21), except that the nylon membrane was probed with a chemiluminescently labeled PCR product. Labeling of the probe and visualization of the blot were performed according to the manufacturer's instructions (ECL direct nucleic acid

TABLE 2. MLST of 27 nonencapsulated clinical pneumococcal isolates

Strain	Source ^a	Clone ^b	Penicillin MIC (mg/liter)	Alleles ^d							Related clone ^h
				<i>aro</i>	<i>ddl</i>	<i>gdh</i>	<i>gki</i>	<i>spi</i>	<i>rec</i>	<i>xpt</i>	
208.80	NP	A ^c	0.125	8	53	37	9	2	1	12	NT, Norway
207.63	NP	A	0.094	8	53	37	9	2	29	12	NT, Norway
101.43	NP	A	0.19	8	53	37	9	2	29	12	NT, Norway
104.58	NP	A	0.094	8	53	37	9	2	29	12	NT, Norway
104.72	NP	A	1.0	8	53	37	9	2	29	12	NT, Norway
106.44	NP	A	2.0	8	53	37	9	2	29	12	NT, Norway
110.58	NP	A	0.19	8	53	37	9	2	29	12	NT, Norway
111.81	NP	A	0.125	8	53	37	9	2	29	12	NT, Norway
202.66	NP	A	0.125	8	53	37	9	2	29	12	NT, Norway
203.15	NP	A	0.094	8	53	37	9	2	29	12	NT, Norway
205.09	NP	A	0.064	8	53	37	9	2	29	12	NT, Norway
305.33	NP	A ^c	0.094	8	53 ^e	37	9	64 ^f	29	12	NT, Norway
110.41	NP	s	0.012	11	14	3	9	2	29 ^g	19	NT, Norway
B104.72	BL	B	<0.06	8	71	5	2	2	27	11	NT, United Kingdom
208.15	NP	B	<0.06	8	71	5	2	2	27	11	NT, United Kingdom
208.37	NP	s	<0.06	NM	NM	NM	NM	NM	NM	NM	
301.77	NP	s	<0.06	NM	NM	NM	NM	NM	NM	NM	
208.56	NP	s	0.125	NM	NM	NM	NM	NM	NM	NM	
209.68	NP	s	<0.06	NM	NM	NM	NM	NM	ND	NM	
108.46	NP	C	0.25	2	14	5	9	17	1	21	15A
105.81	NP	C	0.25	2	14	5	9	17	1	21	15A
B105.04	BL	s	0.064	2	31	10	9	27	18	1	3
B105.27	BL	s	<0.06	7	8	25	8	25	6	6	6A
B102.34	BL	s	<0.06	10	14	10	9	13	43	1	6A
B105.28	BL	s	<0.06	7	14	5	1	14	8	11	14, Denmark
Bil86	BL	s	<0.06	NM	NM	NM	NM	ND	NM	NM	
Bil700	BL	s	<0.06	NM	NM	NM	NM	NM	NM	NM	

^a Clinical specimens were collected from the nasopharynx (NP) or blood (BL).

^b MLST allelic patterns allowed isolates to be grouped into three clusters, designated clones A, B, and C; some of the isolates produced individual patterns (s).

^c Strains 208.80 and 305.33 belong to clonal complex A.

^d *aroE* encodes shikimate dehydrogenase, *gdh* encodes glucose-6-phosphate dehydrogenase, *gki* encodes glucose kinase, *recP* encodes transketolase, *spi* encodes signal peptidase I, *xpt* encodes xanthine phosphoribosyltransferase, and *ddl* encodes D-alanine-D-alanine ligase. The numbers indicate allele types. NM, no match; ND, not determined.

^e 98% homology with this allele.

^f 96% homology with this allele.

^g 98% homology with this allele.

^h Serotype and/or country of isolation for the nearest match in the MLST database. NT, nontypeable.

labeling and detection system; Amersham Biosciences). The molecular weight standard used was RNA Ladder (New England Biolabs).

In vitro translation of ORF 1. PCR amplicons of ORF 1 were generated with primers FI3 and 104_b832.12 (Table 1). PCR products were purified with a QIAquick gel extraction kit (QIAGEN) and were used in an *Escherichia coli* S30 extract system for linear templates (Promega, Madison, Wis.). Proteins were expressed from 2 µg of PCR-amplified template DNA. Samples of each reaction mixture were resolved by sodium dodecyl sulfate—10% polyacrylamide gel electrophoresis and were blotted onto polyvinylidene difluoride membranes (Immobilon-P; Millipore Corp., Bedford, Mass.). Proteins were detected by using the Transcend chemiluminescent translation detection system (Promega) according to the manufacturer's instructions. The sizes of proteins were estimated by comparison with a prestained marker (Bio-Rad Laboratories, Inc.).

Construction of an *aliB*-like ORF-deficient mutant of strain 110.58. Transformation of *E. coli* and *S. pneumoniae* was performed as described previously (10, 18). Plasmid pBluescript KSII (Stratagene) was used for plasmid construction.

Within pBluescript KSII the following genes were cloned into the multiple cloning site. First, a 1,381-bp fragment of the *dexB* ORF of strain 110.58, flanked by the chloramphenicol acetyltransferase (*cat*) gene from plasmid pJ53copG7 (kindly provided by M. Espinosa, Centro de Investigaciones Biológicas, Madrid, Spain) (13), was inserted. Downstream of *cat* a 883-bp fragment of the *capN*-like ORF of strain 110.58 was inserted. The primers used to amplify the individual genes are described in Table 1. The whole *dexB-cat-capN* construct (3.1 kb) was amplified by PCR by using primers dexB_f145_XbaI and capN_B1_XhoI (Table 1). The *dexB-cat-capN* amplicon was then used to transform competent cells of *S. pneumoniae* strain 110.58. Recombinant clones were selected on CSBA plates containing 3 mg of chloramphenicol per liter. Positive recombinants were analyzed by PCR and sequencing to confirm correct deletion of *aliB*-like ORF 1 and *aliB*-like ORF 2.

Adherence of *S. pneumoniae* to human lung epithelial cell line A549. A549 epithelial cells were inoculated into 24-well tissue culture plates (10⁵ cells/well) in minimum essential medium (MEM) (Gibco, Paisley, United Kingdom) containing 10% FCS, 1% glutamine (Gibco), and 1% penicillin/streptomycin (Biochrom KG, Berlin, Germany) and were cultured until complete confluence was reached at 37°C in the presence of 5% CO₂. *S. pneumoniae* strain 110.58 and an *aliB*-like ORF-deficient mutant of this strain grown in BHI broth containing 5% FCS to the logarithmic phase were washed in MEM and then resuspended to obtain 10⁷ bacteria in 0.5 ml of MEM, which was added to each well containing washed A549 cells (the concentration of bacteria was confirmed by plating appropriate dilutions of the inoculum). The bacteria were centrifuged onto the A549 cells at 390 × g for 5 min at room temperature. After 1 h of incubation at 37°C in the presence of 5% CO₂, the monolayers were washed five times with phosphate-buffered saline, and the cells were detached by 5 min of incubation with 200 µl of trypsin-EDTA (Gibco) diluted 1:5 in phosphate-buffered saline. Appropriate dilutions were plated on CSBA plates for subsequent counting of colonies. A total of eight experiments were performed, and each strain was analyzed in triplicate in each experiment.

RESULTS

Nonencapsulated strains. A total of 1,980 noninvasive and 215 invasive *S. pneumoniae* isolates were evaluated for capsule expression. Forty-five isolates (2.1% of the 2,195 isolates examined) produced rough colonies. Of these 45 isolates, 27 (1.2% of all isolates) could not be typed by the Quellung reaction, indicating that they did not express capsule polysac-

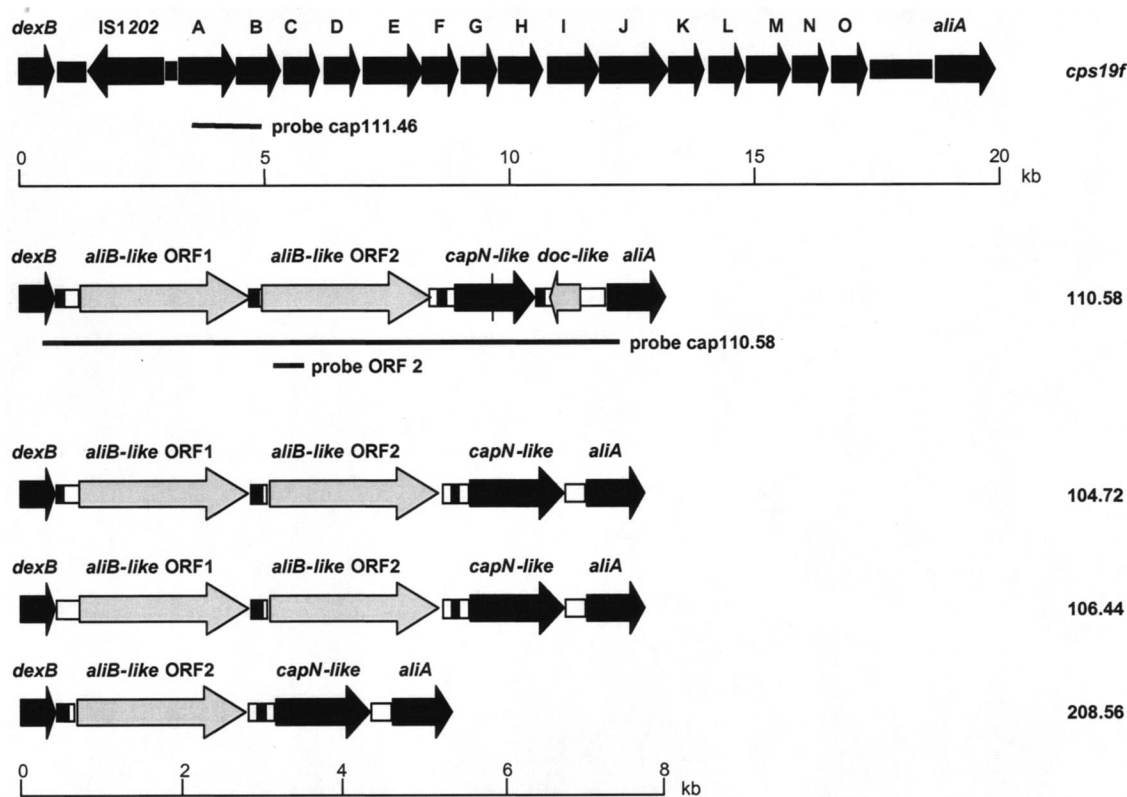


FIG. 1. Schematic representation of cap regions of nonencapsulated *S. pneumoniae* strains. Cap regions (i.e., the DNA sequence between the *dexB* and *aliA* genes) were sequenced and analyzed for nonencapsulated *S. pneumoniae* strains 110.58, 104.72, 106.44, and 208.56. For comparison, the cap region of a serotype 19F strain is shown at the top (9). Regions and ORFs found in nonencapsulated strains exhibiting high levels of homology to the cap region (including *dexB* and *aliA*) of encapsulated *S. pneumoniae* strains are indicated by black arrows (ORF) or boxes. The *capN*-like ORF of strain 110.58 contains a premature stop codon within the ORF represented by a vertical line. The *aliB*-like ORF 1 and *aliB*-like ORF 2, as well as the *doc*-like ORF found in the cap region of the nonencapsulated strains, are indicated by gray arrows. Regions with no homology to database entries are indicated by open boxes. The black boxes indicate complete and incomplete BOX elements found between *dexB* and *aliB*-like ORF 1 (strains 110.58 and 104.72), between *aliB*-like ORF 2 and a *capN*-like ORF (all nonencapsulated strains), and between a *capN*-like ORF and *aliA* (strain 110.58). Probes cap111.46 and cap110.58 and the *aliB*-like ORF 2 probe used for Southern blot analysis are indicated by lines. The *cps19f* gene (9) and the cap region of nonencapsulated strains are drawn to different scales; scale bars are indicated beneath the corresponding regions.

charide, but the remaining 18 (0.8%) isolates could be assigned to a serotype after repeated typing. Of the 27 truly nonencapsulated isolates, 7 were invasive and 20 colonized the nasopharynx.

MLST analysis. The 27 nonencapsulated strains were characterized by MLST. Three clusters were identified. Clone A comprised 10 isolates with identical MLST patterns. Three isolates differed from this pattern by one or two alleles (and so were included in clonal complex A) or by five alleles (Table 2). Clone A strains were isolated from patients of different ages and from different geographical locations around Switzerland (data not shown). None of the clone A isolates were susceptible to penicillin, and all of them were resistant to erythromycin and cotrimoxazole (Table 2). Clones B and C each contained two isolates with identical MLST patterns.

Of the 18 difficult-to-type isolates, all but three belonged to one of three clones (data not shown) which were eventually characterized as serotypes 1, 38, and 6A.

Characterization of the cap region. Southern blotting was performed by probing with PCR products representing a conserved region of the cap locus of a typical encapsulated sero-

type 19F strain, strain 111.46 (probe cap111.46; 1,795 bp) or the cap locus of nonencapsulated strain 110.58 (probe cap110.58; 7.4 kb) (Fig. 1 and 2A and B and Table 3). Probe cap111.46 bound the 18 encapsulated strains (data not shown) and 8 of the 27 nonencapsulated strains, including clone C strains with a single band (Table 3). This indicated that cap genes were present and corresponded to the observation made by Arrecubieta et al. about a serotype 3 strain not expressing the capsule due to a point mutation in the cap operon (3). Probe cap110.58 bound all strains that produced distinctive banding patterns, which correlated with the MLST results (Table 2). Because of the apparent differences in the cap locus, sequence analysis was performed with three representative strains belonging to clone A and strain 208.56. This analysis revealed high homology in this region for all four strains (Fig. 1). BLAST analysis revealed that capsule genes, which are usually found in encapsulated pneumococci, were absent, except for a region with homology to the *capN* gene of serotypes 33F and 37 (15). However, we found two apparent ORFs which were 1,968 to 2,004 and 1,959 bp long, exhibited 61 to 64.5% homology to the ATP-binding cassette (ABC) transporter gene

TABLE 3. Characterization of the cap region of nonencapsulated pneumococcal isolates by Southern blotting, PCR, and sequencing analysis

Strain	Clone	Southern blotting with the following probes:			PCR for <i>aliB</i> -like ORF 1 ^d	PCR for <i>aliB</i> -like ORF 2 ^e
		cap111.46 ^a	cap110.58 ^b	<i>aliB</i> -like ORF 2 ^c		
104.72	A	n	T1	ND	pos	pos
106.44	A	n	T1	ND	pos	pos
110.58	A	n	T1	1b	pos	pos
207.63	A	n	T1	1b	pos	pos
208.80	A ^f	n	T1	1b	pos	pos
305.33	A ^f	n	T1	1b	pos	pos
110.41	S ^g	n	T1	1b	pos	pos
B104.72	B	n	T2	1b	pos	pos
208.15	B	n	T2	1b	pos	pos
208.37	s	n	S	1b	w	pos
301.77	s	n	S	1b	w	pos
208.56	s	n	S	1b	w	pos
209.68	s	n	S	1b	w	pos
108.46	C	1b	T3	n	neg	neg
105.81	C	1b	T3	n	w	neg
B105.04	s	1b	T4	n	w	neg
B105.27	s	1b	S	n	w	neg
B102.34	s	1b	S	n	neg	neg
B105.28	s	1b	T4	n	w	neg
Bil86	s	1b	S	1b	neg	neg
Bil700	s	1b	S	1b	neg	pos

^a Southern blots were probed with a conserved region of the cap locus from encapsulated serotype 19F strain 111.46 (Fig. 1). 1b, 1 band; n, no bands.

^b Southern blots were probed with the cap locus of nonencapsulated strain 110.58 (Fig. 1). T1, T2, T3, and T4, four different banding patterns; S, sporadic bands.

^c Southern blots were probed with a region of *aliB*-like ORF 2. 1b, 1 band; n, no bands; ND, not determined.

^d PCR was performed with primers binding within *aliB*-like ORF 1. pos, positive for PCR product; neg, no PCR product; w, weak, probably nonspecific, band.

^e PCR was performed with primers binding within *aliB*-like ORF 2. pos, positive for PCR product; neg, no PCR product.

^f Strains 208.80 and 305.33 belong to clonal complex A.

^g s, individual pattern.

aliB (and 62% homology to each other), and are not normally found in the capsule-encoding region of encapsulated strains. These homologous regions were designated *aliB*-like ORF 1 and *aliB*-like ORF 2. Strain 208.56 lacked *aliB*-like ORF 1 but possessed *aliB*-like ORF 2. The sequence of the region between *aliB*-like ORF 1 and *aliB*-like ORF 2 also matched the sequence between the *dexB* and *capA* genes in serotype 33F (85 to 91% homology), and the sequence following *aliB*-like ORF 2 also matched the BOX element CBBA sequence (77 to 78% homology over 159 bp). In strains 110.58 and 104.72 there was also a 106-bp region with homology to a BOX AC element.

Having identified these novel ORFs in the cap locus, we used a probe specific for *aliB*-like ORF 2 (370 bp) for Southern blotting of three clone A strains and all nontypeable strains not belonging to clone A. This probe hybridized only to the isolates which failed to bind to probe cap111.46 (except for strains Bil86 and Bil700) (Table 3 and Fig. 2C). PCRs performed with primers binding within *aliB*-like ORF 2 (Table 1) confirmed these results. Amplification products were obtained for all strains that reacted with the *aliB*-like ORF 2 probe except strain Bil86 (Table 3). Therefore, nontypeable strains could be divided into two groups. Group I appeared to have the con-

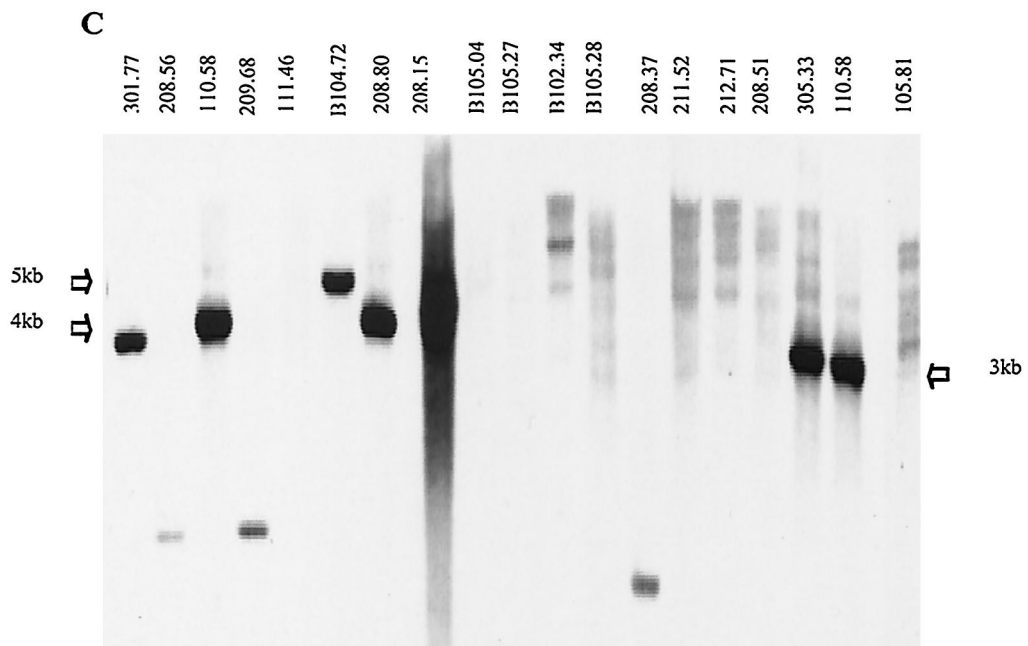
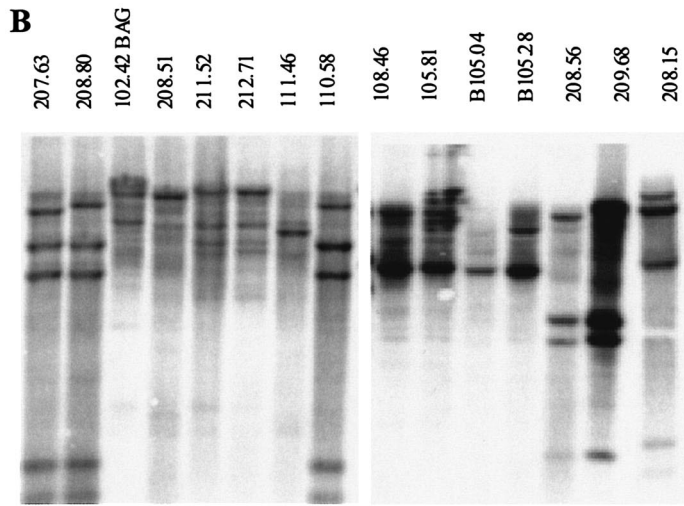
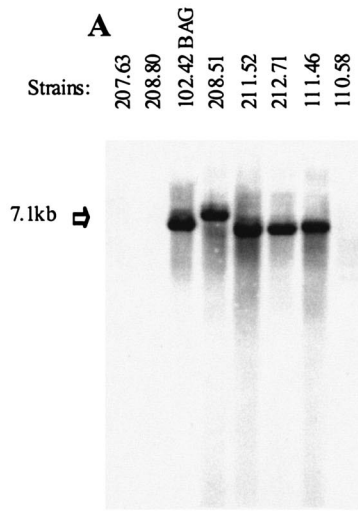
served capsule gene region (binding to probe cap111.46) and not *aliB*-like ORF 2. Group II seemed to lack typical capsule genes but possessed *aliB*-like ORF 2. *aliB*-like ORF 1, however, was restricted to strains in clonal complexes A and B plus strain 110.41; an intense band was produced by PCR performed with *aliB*-like ORF 1 primers (Table 1) for all six representatives of clonal complex A tested, as well as strain 110.41 (which differs from clone A at five alleles, as determined by MLST) and both members of clone B. All other nonencapsulated isolates either gave a nondetectable or weak amplification product. The weak band was interpreted as a nonspecific band, since it was also observed for strain 208.56, in which no *aliB*-like ORF 1 region was found by sequencing, and strain 111.46, a typical serotype 19F isolate.

BLAST analysis of *aliB*-like ORFs. The closest match (66% nucleotide homology over 1,977 bp) obtained for *aliB*-like ORF 1 was the match with a sequence fragment of the partially sequenced *Streptococcus gordonii* genome (<http://www.tigr.org/tdb/mdb/mdbinprogress.html>), and *aliB*-like ORF 1 also exhibited 80% homology over 242 bp to the *S. gordonii* oligopeptide-binding lipoprotein gene *hppH*. BLAST analysis revealed 93% nucleotide homology over 1,886 bp of *aliB*-like ORF 2 with a sequence fragment of the *Streptococcus mitis* genome between *dexB* and the first capsule gene (unfinished *S. mitis* strain NCTC 12261 genome [<http://www.tigr.org/tdb/mdb/mdbinprogress.html>]). In addition, the hypothetical protein gene *spr0311*, adjacent to *dexB* in the capsule region of the *S. pneumoniae* R6 genome, exhibited 92% homology over 163 bp to a region at the 5' end of *aliB*-like ORF 2. In a short region before the *cps2A* gene of R6 30 of 32 bp matched the residues in a region toward the 3' end of *aliB*-like ORF 2, which may indicate possible sites of insertion or deletion. The G+C content of *aliB*-like ORF 1 and *aliB*-like ORF 2 was around 40%, which is equivalent to the G+C content of the whole *S. pneumoniae* R6 genome. Genes of *S. mitis* have a similar average value, making it difficult to determine the sites of possible recombination from differences in the G+C contents. However, the homology between the capsule regions of *S. pneumoniae* strain 110.58 and *S. mitis* DNA stops at the 3' end of *aliB*-like ORF 2, suggesting that that this may be a site of recombination.

Gene expression analysis of *aliB*-like ORFs. RT-PCR showed that three representatives of clone A (strains 110.58, 104.72, and 106.44) transcribed *aliB*-like ORF 1 (Fig. 3A). No transcription product was found for clone C isolates 105.81 and 108.46 or the encapsulated serotype 19F strain 111.46 when they were grown under the same conditions. By using Northern blotting (Fig. 3B) an *aliB*-like ORF 1 probe was found to give multiple bands with nonencapsulated strain 110.58 (clone A) and a particularly strong band around 3,000 bases. There was also a weak band of a similar size for encapsulated isolate 111.46, most likely due to nonspecific binding to a transcript homologous to *aliB*-like ORF 1. No expression of *aliB*-like ORF 2 was detected in any of the strains tested by RT-PCR or Northern blotting (data not shown).

In vitro translation assays showed that translation products for *aliB*-like ORF 1 were obtained from isolates 110.58, 104.72, and 106.44 representing the major clone A cluster (Fig. 3C).

Adherence of *S. pneumoniae* to lung epithelial cells. The mean percentage of cells in the strain 110.58 inoculum which



adhered to A549 epithelial cells was 13.5% (standard deviation, 4.8%), compared to 10.6% (standard deviation, 5.9%) for the *aliB*-like ORF 1- and *aliB*-like ORF 2-deficient mutant, indicating that there was no significant difference in adherence between these two strains.

DISCUSSION

Just over 1% of the colonizing and invasive *S. pneumoniae* strains analyzed in this study were rough forms (i.e., nonencapsulated). These nontypeable strains otherwise had the characteristics of typical *S. pneumoniae* strains (bile soluble, susceptible to optochin, and binding to a species-specific 16S rRNA probe) rather than those of the atypical pneumococci described by Whatmore et al. (30). MLST and analysis of the cap locus was used to divide the nonencapsulated *S. pneumoniae* strains into two groups. Group I strains appear to be closely related to encapsulated strains, since they have typical capsule genes and produce the MLST patterns of typeable strains. Group II strains are less closely related to known encapsulated pneumococci as determined by MLST and contain, instead of the capsule genes, a sequence with homology to the *aliB* gene. *aliB* encodes a peptide-binding molecule associated with an ABC transporter (2, 4). All strains in group II of nonencapsulated pneumococci apparently possess this *aliB* homologue, which we designated *aliB*-like ORF 2. However, some strains had, in addition, a second ORF, *aliB*-like ORF 1 upstream of *aliB*-like ORF 2. All but one of the strains with *aliB*-like ORF 1 belonged to clonal complex A or clone B.

The origin of the *aliB*-like ORFs is not known. Based on BLAST analysis, *aliB*-like ORF 2 may have been acquired by horizontal gene transfer from *S. mitis*, and *aliB*-like ORF 1 may have originated from *S. gordonii*, similar to the acquisition of penicillin-binding proteins by *S. pneumoniae* from the oral commensal organisms *S. mitis* and *Streptococcus oralis* (7, 22). The simultaneous occurrence of *aliB*-like ORF 1 and *aliB*-like ORF 2 in some isolates may be explained by repeated transformation. Duplication within the pneumococcal cap region seems less likely, given the 38% sequence divergence between the two ORFs.

The association of *aliB*-like ORF 1 with two clones is intriguing. Clone A was more prevalent, and at least one isolate has been found in Norway (<http://spneumoniae.mlst.net/>). In this study, clone A strains did not cluster by time, age, or geographical region, suggesting that clone A is widely distributed. Strains of this clone were also characterized by reduced susceptibility to penicillin; in fact, this clone contained 10% of

the nonsusceptible nasopharyngeal isolates in Switzerland (data not shown). Clone B matched a clone which caused an outbreak of conjunctivitis on an American college campus (17). Therefore, the clones possessing *aliB*-like ORF 1 may have a selective advantage. Whether such an advantage is related to the presence of *aliB*-like ORF 1 is a matter for speculation at this time. The function of the putative AliB-like protein is not known, but true AliB has been described as a membrane-bound lipoprotein which binds oligopeptides to deliver them to the Ami permease ABC transporter for uptake into the cell (2, 4). It has been proposed that, in addition to taking up peptide for nutrition, this may be a way of sensing the environment and that this ABC transporter may affect a gene regulator which controls pathways involved in diverse processes, such as induction of competence or adherence (2, 4). Mutation of the *ami* locus results in a reduced ability of *S. pneumoniae* to bind to resting human lung pneumocytes, which represent the alveolar site of infection, and to human endothelial cells (5).

aliB-like ORF 1, but not *aliB*-like ORF 2, was expressed at the RNA level and could be translated in vitro, and so the protein may be expressed and have a function in vivo. No difference in adherence to human lung epithelial cells between a wild-type strain and a mutant of this strain lacking the *aliB*-like ORFs has been found so far. Also, no difference in transformability has been detected (data not shown). Further experiments to determine possible functions of AliB-like molecules are being performed. The presence of sequences with homology to BOX sequences (106 bp with homology to BOX AC and 159 bp with homology to BOX CBBA) in the capsule region of nonencapsulated strains possessing *aliB*-like ORF 1 and *aliB*-like ORF 2 may indicate the presence of functional genes, since it has been proposed that BOX elements are involved in regulation of virulence genes (16). Nontypeable *S. pneumoniae* strains are often considered to be avirulent. However, the fact that seven of the invasive strains in this study were nonencapsulated and the fact that nontypeable strains can cause outbreaks of disease (17) show that this is not necessarily always the case.

The phylogenetic origin of group II of the nonencapsulated strains carrying *aliB*-like ORFs in the cap locus is not known. A small region with homology to the *capN* gene of serotypes 33F and 37 downstream of *aliB*-like ORF 2 indicates that this group may have been derived from one of these serotypes or a close relative some time ago. Also, one strain which shared two alleles with clone A and had one allele that was closely related to a clone A allele contained *aliB*-like ORF 1 and *aliB*-like

FIG. 2. (A) Southern blot probed with cap111.46 bound with a single band to DNA of encapsulated pneumococcal isolates (strains 102.42 BAG, 208.51, 211.52, 212.71, and 111.46) but not to DNA of nonencapsulated isolates (strains 207.63, 208.80, and 110.58). (B) Probe cap110.58, representing the capsule region of a nonencapsulated clone A strain, bound to DNA from all strains with several bands. The banding pattern for the nonencapsulated strains belonging to the clone A complex (strains 207.63, 208.80, and 110.58) is identical or very similar. The banding patterns for the five encapsulated strains (strains 102.42 BAG, 208.51, 211.52, 212.71, and 111.46) represented four different hybridization patterns, in accordance with the expected polymorphism of the *cap* operon in these strains. The banding patterns for nonencapsulated strains representing clone B (strain 208.15) and clone C (strains 108.46 and 105.81) and for strains not belonging to a clone (strains B105.04, B105.28, 208.56, and 209.68) are also shown. (C) Probe specific for *aliB*-like ORF 2 bound to some of the nonencapsulated strains (strains 301.77, 208.56, 110.58, 209.68, B104.72, 208.80, 208.15, 208.37, 305.33, and 110.58) but not to nonencapsulated strains B105.04, B105.27, B102.34, B105.28, and 105.81 or to encapsulated strains (strains 111.46, 211.52, 212.71, and 208.51). Molecular weights were determined by using a standard run on the gel, which was visualized before blotting.

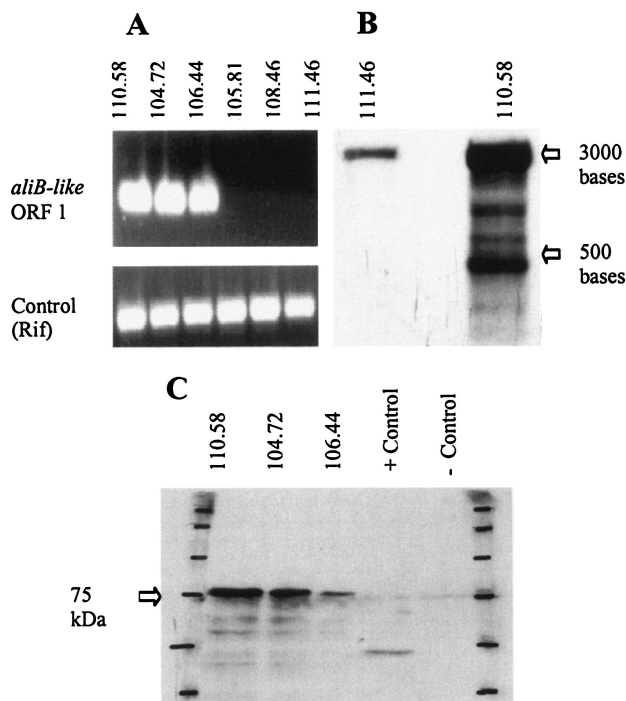


FIG. 3. (A) RT-PCR detected *aliB*-like ORF 1 expression in non-encapsulated clone A isolates (strains 110.58, 104.72, and 106.44) but not in the nonencapsulated clone C isolates (strains 105.81 and 108.46) or the encapsulated serotype 19F control strain (strain 111.46). The lower panel shows the results for a positive control for RT-PCR performed with equal amounts of PCR product produced for all strains with Rif primers binding to cDNA of *rpoB*. (B) Northern blot showing expression of RNA recognized by the *aliB*-like ORF 1 probe in non-encapsulated clone A strain 110.58. There is a weak band for encapsulated serotype 19F control strain 111.46, which may have been due to a cross-reaction. Molecular weights were determined by using a standard run on the gel, which was visualized before blotting. (C) In vitro translation produced a protein from the *aliB*-like ORF 1 region for all three nonencapsulated clone A strains tested (strains 110.58, 104.72, and 106.44). The position of the protein standard was marked on the blot (first and last lanes); positive and negative controls were produced according to the instructions of the kit manufacturer.

ORF 2. This indicates that clone A and related strains may have persisted for quite a long time, allowing extensive recombination. Group I of nonencapsulated strains described in this study exhibits a closer phylogenetic relationship to encapsulated pneumococci that have the cap region. This suggests that they may have arisen from encapsulated pneumococci by mutation in the cap region. Different mechanisms for loss of capsule expression have been described. Capsule gene inactivation by duplication has been demonstrated in vitro for serotypes 3, 8, and 37 (25, 26). Arrecubieta et al. (3) described a single mutation in the *cap3A* gene of serotype 3 that led to a loss of capsule expression. *S. pneumoniae* may take advantage of such (partially) reversible mechanisms for capsule loss to enhance survival at the mucosal level (25, 27).

In conclusion, invasive and noninvasive nonencapsulated *S. pneumoniae* strains were identified and could be assigned to two groups based on the structure of the cap region and MLST. Group I was closely related to encapsulated pneumococci. Group II had a clonal population structure and was

found to carry genes with homology to the *aliB* gene in the cap region, a novel feature. Therefore, loss of capsule expression in vivo was found to be associated with the importation of one or two *aliB* homologues in some nonencapsulated pneumococci.

ACKNOWLEDGMENTS

This study was supported by grant 3200-067998 from the Swiss National Science Foundation to K. Mühlemann.

We thank Jacques Billes for providing two of the invasive strains used in this study and Thomas Geiser for providing the A549 cell line. We are indebted to Martin Täuber for support and encouragement.

REFERENCES

- Adamou, J. E., T. M. Wizemann, P. Barren, and S. Langermann. 1998. Adherence of *Streptococcus pneumoniae* to human bronchial epithelial cells (BEAS-2B). *Infect. Immun.* **66**:820–822.
- Alloing, G., P. de Philip, and J. P. Claverys. 1994. Three highly homologous membrane-bound lipoproteins participate in oligopeptide transport by the Ami system of the gram-positive *Streptococcus pneumoniae*. *J. Mol. Biol.* **241**:44–58.
- Arrecubieta, C., R. Lopez, and E. Garcia. 1994. Molecular characterization of *cap3A*, a gene from the operon required for the synthesis of the capsule of *Streptococcus pneumoniae* type 3: sequencing of mutations responsible for the unencapsulated phenotype and localization of the capsular cluster on the pneumococcal chromosome. *J. Bacteriol.* **176**:6375–6383.
- Claverys, J.-P., B. Grossiord, and G. Alloing. 2000. Is the Ami-AliA/B oligopeptide permease of *Streptococcus pneumoniae* involved in sensing environmental conditions? *Res. Microbiol.* **151**:457–463.
- Cundell, D. R., B. J. Pearce, J. Sandros, A. M. Naughton, and H. R. Masure. 1995. Peptide permeases from *Streptococcus pneumoniae* affect adherence to eucaryotic cells. *Infect. Immun.* **63**:2493–2498.
- Cundell, D. R., J. N. Weiser, J. Shen, A. Young, and E. I. Tuomanen. 1995. Relationship between colonial morphology and adherence of *Streptococcus pneumoniae*. *Infect. Immun.* **63**:757–761.
- Dowson, C. G., T. J. Coffey, C. Kell, and R. A. Whitley. 1993. Evolution of penicillin resistance in *Streptococcus pneumoniae*; the role of *Streptococcus mitis* in the formation of a low affinity PBP2B in *S. pneumoniae*. *Mol. Microbiol.* **9**:635–643.
- Enright, M., and B. Spratt. 1998. A multilocus sequence typing scheme for *Streptococcus pneumoniae*: identification of clones associated with serious invasive disease. *Microbiology* **144**:3049–3060.
- Garcia, E., D. Llull, R. Munoz, M. Mollerach, and R. Lopez. 2000. Current trends in capsular polysaccharide biosynthesis of *Streptococcus pneumoniae*. *Res. Microbiol.* **151**:429–435.
- Hanahan, D. 1983. Studies on transformation of *Escherichia coli* with plasmids. *J. Mol. Biol.* **166**:557–580.
- Iannelli, F., B. J. Pearce, and G. Pozzi. 1999. The type 2 capsule locus of *Streptococcus pneumoniae*. *J. Bacteriol.* **181**:2652–2654.
- Kim, J. O., and J. N. Weiser. 1998. Association of intrastrain phase variation in quantity of capsular polysaccharide and teichoic acid with the virulence of *Streptococcus pneumoniae*. *J. Infect. Dis.* **177**:368–377.
- Lacks, S. A., P. Lopez, B. Greenberg, and M. Espinosa. 1986. Identification and analysis of genes for tetracycline resistance and replication functions in the broad-host-range plasmid pLS1. *J. Mol. Biol.* **192**:753–765.
- Lawrence, E., C. Arias, B. Duke, D. Beste, K. Broughton, A. Efstratiou, R. George, and L. Hall. 2000. Evaluation of serotype prediction by *cpsA-cpsB* gene polymorphism in *Streptococcus pneumoniae*. *J. Clin. Microbiol.* **38**:1319–1323.
- Llull, D., R. Lopez, E. Garcia, and R. Munoz. 1998. Molecular structure of the gene cluster responsible for the synthesis of the polysaccharide capsule of *Streptococcus pneumoniae* type 33F. *Biochim. Biophys. Acta* **1443**:217–224.
- Martin, B., O. Humbert, M. Camara, E. Guenzi, J. Walker, T. Mitchell, P. Andrew, M. Prudhomme, G. Alloing, and R. Hakenbeck. 1992. A highly conserved repeated DNA element located in the chromosome of *Streptococcus pneumoniae*. *Nucleic Acids Res.* **20**:3479–3483.
- Martin, M., J. H. Turco, M. E. Zegans, R. R. Facklam, S. Sodha, J. A. Elliott, J. H. Pryor, B. Beall, D. D. Erdmann, Y. Y. Baumgartner, P. A. Sanchez, J. D. Schwartzman, J. Montero, A. Schuchat, and C. G. Whitney. 2003. An outbreak of conjunctivitis due to atypical *Streptococcus pneumoniae*. *N. Engl. J. Med.* **348**:1112–1121.
- Meier, P. S., S. Utz, S. Aebi, and K. Mühlemann. 2003. Low-level resistance to rifampicin in *Streptococcus pneumoniae*. *Antimicrob. Agents Chemother.* **47**:863–868.
- Mühlemann, K., H. Matter, M. Tauber, and T. Bodmer. 2002. Nationwide surveillance of nasopharyngeal *Streptococcus pneumoniae* isolates from children with respiratory infection, Switzerland, 1998–1999. *J. Infect. Dis.* **187**:589–596.
- Pearce, B. J., F. Iannelli, and G. Pozzi. 2002. Construction of new unencap-

- sulated (rough) strains of *Streptococcus pneumoniae*. Res. Microbiol. **153**:243–247.
21. **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.
 22. **Sibold, C., J. Henrichsen, A. Konig, C. Martin, L. Chalkley, and R. Hakenbeck.** 1994. Mosaic *pbpX* genes of major clones of penicillin-resistant *Streptococcus pneumoniae* have evolved from *pbpX* genes of a penicillin-sensitive *Streptococcus oralis*. Mol. Microbiol. **12**:1013–1023.
 23. **Swiss Federal Office of Health.** 2002. Invasive Pneumokokken-Infektionen in der Schweiz 1998–1999: Empfindlichkeit auf Antibiotika und Abdeckung durch 7- und 9-valente Konjugat-Impfstoffe. Bundesamt Gesundheit Bull. **31**:544–549.
 24. **Swiss Federal Office of Health.** 2003. Überwachung der Infektionen durch invasive und nicht invasive Pneumokokken im Jahr 2002. Bundesamt Gesundheit Bull. **22**: 368–372.
 25. **Waite, R. D., J. K. Struthers, and C. G. Dowson.** 2001. Spontaneous sequence duplication within an open reading frame of the pneumococcal type 3 capsule locus causes high-frequency phase variation. Mol. Microbiol. **42**: 1223–1232.
 26. **Waite, R. D., D. W. Penfold, J. K. Struthers, and C. G. Dowson.** 2003. Spontaneous sequence duplications within capsule genes *cap8E* and *its* control phase variation in *Streptococcus pneumoniae* serotypes 8 and 37. Microbiology **149**:497–504.
 27. **Weiser, J. N., R. Austrian, P. K. Sreenivasan, and H. R. Masure.** 1994. Phase variation in pneumococcal opacity: relationship between colonial morphology and nasopharyngeal colonization. Infect. Immun. **62**:2582–2589.
 28. **Weiser, J. N., and M. Kapoor.** 1999. Effect of intrastrain variation in the amount of capsular polysaccharide on genetic transformation of *Streptococcus pneumoniae*: implications for virulence studies of encapsulated strains. Infect. Immun. **67**:3690–3692.
 29. **Weiser, J. N., D. Bae, H. Epino, S. B. Gordon, M. Kapoor, L. A. Zenewicz, and M. Shchepetov.** 2001. Changes in availability of oxygen accentuate differences in capsular polysaccharide expression by phenotypic variants and clinical isolates of *Streptococcus pneumoniae*. Infect. Immun. **69**:5430–5439.
 30. **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.