# The ABC Transporter Encoded at the Pneumococcal Fructooligosaccharide Utilization Locus Determines the Ability To Utilize Long- and Short-Chain Fructooligosaccharides

Caroline M. Linke,[a] Shireen A. Woodiga,[a] Dustin J. Meyers,[a] Carolyn M. Buckwalter,[a,b] Hussam E. Salhi,[a,b] Samantha J. King[a,b]

Center for Microbial Pathogenesis, The Research Institute at Nationwide Children's Hospital, Columbus, Ohio, USA[a]; Department of Pediatrics, The Ohio State University College of Medicine, Columbus, Ohio, USA[b]

**Streptococcus pneumoniae is an important human pathogen that requires carbohydrates for growth. The significance of carbohydrate acquisition is highlighted by the genome encoding more than 27 predicted carbohydrate transporters. It has long been known that about 60% of pneumococci could utilize the fructooligosaccharide inulin as a carbohydrate source, but the mechanism of utilization was unknown. Here we demonstrate that a predicted sucrose utilization locus is actually a fructooligosaccharide utilization locus and imparts the ability of pneumococci to utilize inulin. Genes in strain TIGR4 predicted to encode an ABC transporter (SP_1796-8) and a β-fructosidase (SP_1795) are required for utilization of several fructooligosaccharides longer than kestose, which consists of two β(2-1)-linked fructose molecules with a terminal α(1-2)-linked glucose molecule. Similar to other characterized pneumococcal carbohydrate utilization transporter family 1 transporters, growth is dependent on the gene encoding the ATPase MsmK. While the majority of pneumococcal strains encode SP_1796-8 at this genomic location, 19% encode an alternative transporter. Although strains encoding either transporter can utilize short-chain fructooligosaccharides for growth, only strains encoding SP_1796-8 can utilize inulin. Exchange of genes encoding the SP_1796-8 transporter for those encoding the alternative transporter resulted in a TIGR4 strain that could utilize short-chain fructooligosaccharide but not inulin. These data demonstrate that the transporter encoded at this locus determines the ability of the bacteria to utilize long-chain fructooligosaccharides and explains the variation in inulin utilization between pneumococcal strains.**

**D**espite the development of effective vaccines, *Streptococcus pneumoniae* is still a major human pathogen (1). *S. pneumoniae* is the leading cause of community-acquired pneumonia and a major cause of otitis media, sinusitis, bacteremia, and meningitis. Colonization of the human oronasopharynx is an essential precursor for disease. Most colonization events are asymptomatic, but due to a poorly understood interplay with the host, pneumococci can spread beyond this site and cause disease. One critical requirement for pneumococci to successfully colonize and cause disease is efficient nutrient acquisition. Pneumococci require carbohydrates as a source of carbon for growth. This requirement likely drives the maintenance of more than 27 predicted carbohydrate transporters in the pneumococcal genome (2). The ability of *S. pneumoniae* to utilize different carbohydrate sources at different times is almost certainly an important factor in the ability of the organism to successfully colonize and cause disease. Even though carbohydrate utilization is central to the pneumococcal lifestyle, there remains limited information regarding metabolic pathways, and it is as yet unknown what carbon sources are used by the bacteria during stages of infection.

Pneumococci encode two major families of carbohydrate transporters, the phosphoenolpyruvate phosphotransferase system (PTS) and ATP-binding cassette (ABC) transporters (2). There are two predicted pneumococcal sucrose transporters encoded within the TIGR4 genome (3), the PTS encoded by *scrT* and the ABC transporter encoded by *susT1T2X* (Fig. 1A). These transporters were initially identified as sucrose transporters due to the presence of genes in these loci encoding glycosyl hydrolase 32 (GH32) family proteins that contain the signature β-fructosidase motif (NDPNG) and contribute to growth on sucrose (*scrH* and *susH*) (3–5). Subsequently, it was demonstrated that growth of an

*scrT* mutant on sucrose was greatly reduced. Although loss of *susX*, encoding the substrate binding protein of the predicted ABC transporter, did not reduce growth on sucrose, growth of a mutant deficient in both *scrT* and *susX* was further reduced (6). These data suggest that *scrT* encodes the major sucrose transporter. Furthermore, while SusT1T2X may contribute to sucrose utilization, it seems likely that the primary substrate is a different carbohydrate. This is supported by the finding that sucrose induces expression of the *scr* locus but not the *sus* locus (3). The previous report that the *sus* locus contributes to pneumonia and the *scr* locus contributes to colonization (3) also suggests that they may transport different carbohydrates.

The presence of a gene encoding a β-fructosidase (*susH*) within the *sus* locus supports the hypothesis that SusT1T2X transports a carbohydrate containing sucrose. It had previously been demonstrated that components of the Sus transporter clustered phylogenetically with components of the Msm transporter encoded by *Lactobacillus acidophilus*, which is required for efficient growth of the bacterium on the oligofructose [β(2-1)-linked fructose] family of fructooligosaccharides (FOSs) (7). *S. pneumoniae*
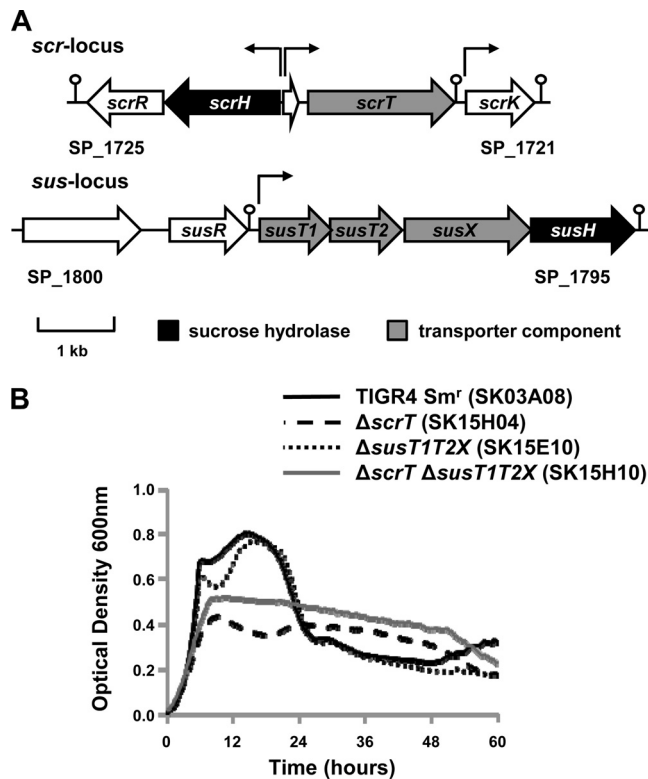
FIG 1 (A) Schematic depicting the two putative sucrose utilization loci in *S. pneumoniae* strain TIGR4. Open reading frames predicted within the TIGR4 sequence are represented by block arrows. Arrows above the schematic indicate predicted transcriptional start sites, and the predicted terminators are denoted by a stem-loop (2). In the *scr* locus, *scrT* encodes the sucrose PTS, *scrH* encodes a sucrose hydrolase, *scrK* encodes a predicted fructokinase, and *scrR* encodes a predicted LacI family repressor. In the *sus* locus, *susT1* and *susT2* encode predicted membrane-spanning permeases, *susX* encodes a predicted substrate binding protein, *susH* encodes a sucrose hydrolase, and *susR* encodes a predicted LacI family repressor. (B) *scrT* contributes to growth on sucrose. Strains were grown for 60 h on CDM supplemented with 12 mM sucrose as the sole carbon source. Growth was measured by determination of the optical density at 600 nm. A glucose control was included in each experiment to confirm that mutants and genetically reconstituted strains had no generalized growth defects (data not shown). Data for a CDM no-carbohydrate control were subtracted from each data set. Each experiment was performed at least three independent times in triplicate, and average data from one experiment are shown.

has previously been shown to utilize the FOS inulin as a carbon source (8). Inulin consists of a chain of fructose molecules with a terminal sucrose. Here, we demonstrate that the *sus* locus is required for utilization of FOS but not sucrose under the conditions tested. Furthermore, diversity in the *sus* locus determines the ability of pneumococcal strains to utilize different lengths of FOS chains.

## MATERIALS AND METHODS

**Bacterial strains, culture media, and chemicals.** The wild-type and genetically modified strains of *S. pneumoniae* utilized in this study are described in Table 1. Thirty-four clinical isolates representing 13 different serotypes were used in this study. Twenty-five of these strains were isolated by the Clinical Laboratory at Nationwide Children's Hospital during 2005 and 2006. Sinus, tracheal, and middle ear aspirates were cultured directly on chocolate agar plates and tryptic soy agar supplemented with 5% sheep blood. Agar plates were incubated at 37°C in 5% $CO_2$ for 4 days.

Blood culture was performed using a BacT/Alert blood culture system (bioMérieux, Durham, NC) with FAN aerobic (FA) and standard anaerobic (SN) blood culture bottles. Growth or a positive culture from the BacT/Alert culture system was confirmed to be *Streptococcus pneumoniae* using standard laboratory methods (13). The remaining nine isolates have been used in previous studies (12). Broth cultures were routinely grown at 37°C in Todd-Hewitt broth (Becton, Dickinson and Co., Sparks, MD) supplemented with 0.2% (wt/vol) yeast extract (Becton, Dickinson and Co.) (THY). C medium with 5% yeast extract (pH 8; C + Y) was used for transformations (14). *S. pneumoniae* was also grown at 37°C in 5% $CO_2$ overnight on tryptic soy (TS; Becton, Dickinson and Co.) plates with 1.5% agar that were spread with 5,000 U of catalase (Worthington Biochemical Corporation, Lakewood, NJ) and TS plates supplemented with 5% sheep blood (Becton, Dickinson and Co.). Bacteria were selected on TS plates containing streptomycin (200 μg ml$^{-1}$) or kanamycin (500 μg ml$^{-1}$). Unless otherwise specified, all chemicals, substrates, and enzymes were purchased from Sigma Chemicals (St. Louis, MO).

**Construction of mutants and genetically altered strains.** All mutants were constructed using a Janus cassette selection system (15). This method requires two rounds of transformation. The first introduced a Janus cassette containing genes encoding kanamycin resistance and streptomycin sensitivity ($rpsL^+$) into the genome of streptomycin-resistant (Sm$^r$) *S. pneumoniae* in place of the region to be deleted. The name of each primer used in construction of a mutant is a letter indicating the genetic location being mutated and a number that corresponds to the primer numbers described below (Table 2). For each mutant constructed, DNA fragments flanking the region to be deleted were generated using the corresponding primers 1 and 2 and primers 4 and 5 (Table 2). The flanking fragments are sequentially joined to the Janus cassette PCR product (primers J.F and J.R) using a variation of splicing by overlap extension (SOE) PCR (11), first described by Horton et al. (16). All genomic DNA was prepared as previously described (17). A high-fidelity proofreading polymerase (Phusion; New England BioLabs, Ipswich, MA) was used to minimize PCR-generated errors. The Janus construct was transformed into *S. pneumoniae*, and the transformants were selected for on kanamycin and confirmed by PCR using primers 7 and 8, which flank the mutant construct.

The second round of transformation replaced the Janus cassette with an engineered segment of DNA consisting of the fragments flanking the deleted region. The fragments were generated using primers 1 and 3 (upstream fragment) and primers 5 and 6 (downstream fragment) and spliced together via SOE PCR. The unmarked mutants were confirmed with primers flanking the construct (primers 7 and 8) and sequencing. The genetically reconstituted strains were generated by transforming the final mutants with the corresponding Janus construct and then subsequently with parental PCR products. Genetic reconstitution was confirmed by primers flanking the construct (primers 7 and 8).

The TIGR4 strain encoding the *sfuABC* locus (SK17C10 Δ*susT1T2X sfuABC*$^+$) was generated by transforming the Janus intermediate of *susT1T2X* with a PCR product spanning the *sfu* locus amplified from strain D39 (primers S.A and S.B). Successful generation of the strain was confirmed using primer pairs S.C-S.D and S.E-S.F.

**Growth assays.** Chemically defined medium (CDM) was prepared essentially as previously described (18), without the addition of carbohydrate. The CDM buffer was made at a 2.5× concentration to allow addition of sufficient carbohydrate to support bacterial growth. The medium was supplemented with no carbohydrate, glucose (12 mM), sucrose (12 mM), inulin (5 mg ml$^{-1}$), kestose (12 mM; glucose linked to two fructose molecules [GF2]; Wako Chemicals, Richmond, VA), nystose (GF3; 12 mM; Wako Chemicals), 1F-fructofuranosyl nystose (GF4; 12 mM; Wako Chemicals), or raftilose (2.7 mg ml$^{-1}$; Beneo-Orafti, Tienen, Belgium). *S. pneumoniae* strains were grown in THY to an optical density at 600 nm (OD$_{600}$) of 0.3 ± 0.005, and 1-ml aliquots were washed and resuspended in 65 μl of phosphate-buffered saline (PBS) and 65 μl of catalase (50,000 U ml$^{-1}$). Twenty-microliter aliquots of bacterial suspensions or PBS-

**TABLE 1** Strains used in this study

| Strain name | Serotype | Characteristic or genotype (site of isolation)[a] | Source or reference |
|---|---|---|---|
| SK03A08 | 4 | TIGR4 $rpsL(K56T)$, conferring Sm$^r$ | Bender et al. (9) |
| SK15J09 | 4 | TIGR4 $\Delta msmK\ rpsL(K56T)$ (Sm$^r$) | This study |
| SK16A03 | 4 | TIGR4 $\Delta msmK\ msmK^+\ rpsL(K56T)$ (Sm$^r$) | This study |
| SK15E10 | 4 | TIGR4 $\Delta susT1T2X\ rpsL(K56T)$ (Sm$^r$) | This study |
| SK15F04 | 4 | TIGR4 $\Delta susT1T2X\ susT1T2X^+\ rpsL(K56T)$ (Sm$^r$) | This study |
| SK15H04 | 4 | TIGR4 $\Delta scrT\ rpsL(K56T)$ (Sm$^r$) | This study |
| SK15H07 | 4 | TIGR4 $\Delta scrT\ scrT^+\ rpsL(K56T)$ (Sm$^r$) | This study |
| SK15H10 | 4 | TIGR4 $\Delta scrT\ \Delta susT1T2X\ rpsL(K56T)$ (Sm$^r$) | This study |
| SK16F10 | 4 | TIGR4 $\Delta scrT\ \Delta susT1T2X\ scrT^+\ susT1T2X^+\ rpsL(K56T)$ (Sm$^r$) | This study |
| SK16G02 | 4 | TIGR4 $\Delta susH\ rpsL(K56T)$ (Sm$^r$) | This study |
| SK16G05 | 4 | TIGR4 $\Delta susH\ susH^+\ rpsL(K56T)$ (Sm$^r$) | This study |
| SK17D03 | 4 | TIGR4 $\Delta scrT\ \Delta msmK\ rpsL(K56T)$ (Sm$^r$) | This study |
| SK17C10 | 4 | TIGR4 $\Delta susT1T2X\ sfuABC^+\ rpsL(K56T)$ (Sm$^r$) | This study |
| SK10G03 | 2 | D39 $rpsL(K56T)$ (Sm$^r$) | This study |
| C06_4 | 6A/B | Clinical isolate (blood) | This study |
| C06_5 | 15A | Clinical isolate (sinus) | Marion et al. (10) |
| C06_6 | 19A | Clinical isolate (ear) | This study |
| C06_8 | 19A | Clinical isolate (BAL fluid/tracheal aspirate) | This study |
| C06_9 | 19A | Clinical isolate (ear) | This study |
| C06_10 | 3 | Clinical isolate (BAL fluid/tracheal aspirate) | Marion et al. (10) |
| C06_12 | 19A | Clinical isolate (sinus) | This study |
| C06_14 | 19A | Clinical isolate (ear) | Marion et al. (10) |
| C06_18 | 22F | Clinical isolate (blood) | Burnaugh et al. (11) |
| C06_19 | 6A/B | Clinical isolate (BAL fluid/tracheal aspirate) | This study |
| C06_20 | 19A | Clinical isolate (ear) | This study |
| C06_21 | 19A | Clinical isolate (blood) | This study |
| C06_23 | 19A | Clinical isolate (blood) | This study |
| C06_25 | 3 | Clinical isolate (blood) | This study |
| C06_31 | 23F | Clinical isolate (BAL fluid/tracheal aspirate) | Burnaugh et al. (11) |
| C06_34 | 35F | Clinical isolate (BAL fluid/tracheal aspirate) | This study |
| C06_36 | 19A | Clinical isolate (BAL fluid/tracheal aspirate) | This study |
| C06_39 | 35F | Clinical isolate (sinus) | Burnaugh et al. (11) |
| C06_40 | 19A | Clinical isolate (blood) | This study |
| C06_44 | 19A | Clinical isolate (blood) | This study |
| C06_45 | 19A | Clinical isolate (blood) | This study |
| C06_48 | 19A | Clinical isolate (ear) | This study |
| C06_50 | 19A | Clinical isolate (sinus) | This study |
| C06_57 | 6A/B | Clinical isolate (BAL fluid/tracheal aspirate) | Burnaugh et al. (11) |
| C06_60 | 19F | Clinical isolate (ear) | This study |
| C141 | 19 | Clinical isolate (throat) | Müller-Graf et al. (12) |
| Cr3 | 4 | Clinical isolate (nasopharynx) | Müller-Graf et al. (12) |
| C17 | 6B | Clinical isolate (wound) | Müller-Graf et al. (12) |
| Cr31 | 11 | Clinical isolate (nasopharynx) | Müller-Graf et al. (12) |
| Cl50 | 23F | Clinical isolate (pleural fluid) | Müller-Graf et al. (12) |
| Cl32 | 15 | Clinical isolate (sputum) | Müller-Graf et al. (12) |
| Cl43 | 19 | Clinical isolate (sputum) | Müller-Graf et al. (12) |
| Cl15 | 9 | Clinical isolate (pernasal swab) | Müller-Graf et al. (12) |
| Cl25 | 12 | Clinical isolate (unknown) | Müller-Graf et al. (12) |

[a] $rpsL(K56T)$, a K-to-T change at position 56 encoded by $rpsL$; Sm$^r$, resistance to streptomycin; BAL, bronchoalveolar lavage.

catalase (no-bacteria control) were transferred to 180 µl of CDM supplemented with the appropriate carbon source. Medium supplemented with glucose served as a positive control, demonstrating in each experiment that bacteria were viable and that mutant strains showed no general growth defect relative to their parent strain. Medium with no added carbohydrate served as a negative control to show that in the absence of carbohydrate no growth was supported. Growth plates were incubated at 37°C for 24 or 60 h in a Bio-Tek Synergy HT plate reader, and the OD$_{600}$ was read every 20 min. All data were adjusted to a path length of 1 cm. As no increase in optical density was observed for any bacterial strain when incubated in medium lacking carbohydrate, at each time point the average of results from triplicate wells containing no carbohydrate was subtracted from results from all other wells containing the same bacterial strain. Results from triplicate wells were then averaged. Growth experiments were performed three times in triplicate, with the exception of the growth of clinical isolates on GF3 and inulin. Clinical strains that showed no growth on these carbohydrates were tested at least twice. Although the phenotypes of the mutants in the experiments were maintained, experiment-to-experiment variation was observed; therefore, representative data are graphed, and the results of additional growth experiments are reported as a range of maximal optical densities.

**RNA preparation and RT-PCR.** RNA was isolated by an acid-phenol extraction, with modifications for *S. pneumoniae* (19). Bacteria were grown as described above for growth assays. When samples reached an

**TABLE 2** Primers used in this study

| Group | Name | Primer sequence (5′ → 3′)[a] | Location (GenBank accession no.)[b] |
|---|---|---|---|
| Janus | J.F | CCGTTTGATTTTTAATGGATAATG | 7–30 (AY334019) |
| | J.R | GGGCCCCTTTCCTTATGCTT | 247511–247527 (AE005672) |
| aroE | A.F | GCCTTTGAGGCGACAGC | 1299642–1299626 (AE005672) |
| | A.R | TGCAGTTCARAAACATWTTCTAA | 1299164–1299186 (AE005672) |
| susT1T2X | U.1 | GCAGGTGGAAGAGAAGCAG | 1715812–1715830 (AE005672) |
| | U.2 | <u>CATTATCCATTAAAAATCAAACGGG</u>GACAGTCATCAATAACAAGAG (1) | 1715089–1715110 (AE005672) |
| | U.3 | <u>ATCATCCCACTCAGTATCAAG</u>GACAGTCATCAATAACAAGAG (3) | 1715089–1715110 (AE005672) |
| | U.4 | <u>AAGCATAAGGAAAGGGGCCC</u>TTGATACTGAGTGGGATGAT (2) | 1711791–1711810 (AE005672) |
| | U.5 | CCTAGTAGAACGATACAGCC | 1711791–1711810 (AE005672) |
| | U.6 | TTGATACTGAGTGGGATGAT | 1711111–1711130 (AE005672) |
| | U.7 | TGCGTCATAATAAAGTTGATGG | 1715949–1715970 (AE005672) |
| | U.8 | ATCTAACTCGAAGTAATCTGGG | 1711011–1711032 (AE005672) |
| | U.9 | CAGGCTCTGCCATTGTCA | 1711373–1711390 (AE005672) |
| | U.10 | TCCCTCACGCTGATAA | 1710963–1710978 (AE005672) |
| scrT | C.1 | CTCAACATGGTAATGAGTATGC | 1627457–1627478 (AE005672) |
| | C.2 | <u>CATTATCCATTAAAAATCAAACGG</u>TTCGCCCTTACGACTATTAT (1) | 1627038–1627057 (AE005672) |
| | C.3 | <u>CACATTAGAGAGGATTGATTAGA</u>TTCGCCCTTACGACTATTAT (4) | 1627038–1627057 (AE005672) |
| | C.4 | <u>AAGCATAAGGAAAGGGGCCC</u>TCTAATCAATCCTCTCTAATGTG (2) | 1625085–1625107 (AE005672) |
| | C.5 | GACCAGCTGCATAACCTTCTA | 1625085–1625107 (AE005672) |
| | C.6 | TCTAATCAATCCTCTCTAATGTG | 1624384–1624404 (AE005672) |
| | C.7 | GGGTCGTTGAGAAGTCCTGTT | 1627484–1627504 (AE005672) |
| | C.8 | ACCTGTACGAGCTTCCAAACT | 1624361–1624381 (AE005672) |
| | C.9 | CAATTTCCAACAACAACTCC | 1624845–1624864 (AE005672) |
| | C.10 | CATAACCTTCTAGACATCCC | 1624393–1624412 (AE005672) |
| susH | H.1 | CACAATACGCTCCACTCC | 1712186–1712203 (AE005672) |
| | H.2 | <u>CATTATCCATTAAAAATCAAACGG</u>CTTATCTCCCATAATCAACCTC (1) | 1711680–1711701 (AE005672) |
| | H.3 | <u>CTTGCTTCACCTTGATTGAC</u>CTTATCTCCCATAATCAACCTC (5) | 1711680–1711701 (AE005672) |
| | H.4 | <u>AAGCATAAGGAAAGGGGCCC</u>GTCAATCAAGGTGAAGCAAG (2) | 1710427–1710446 (AE005672) |
| | H.5 | GAAACTGGGATATCTGAACTG | 1710427–1710446 (AE005672) |
| | H.6 | GTCAATCAAGGTGAAGCAAG | 1709871–1709891 (AE005672) |
| | H.7 | AGCTCGTACAAACGGTATGG | 1712276–1712295 (AE005672) |
| | H.8 | GTTCTCAAGCTAGTGGCATC | 1709730–1709749 (AE005672) |
| | H.9 | GAACTGATGAAGAACAATGAAAG | 1709958–1709980 (AE005672) |
| | H.10 | TAAAATAGAAGCAAACAAAACTAAC | 1709603–1709627 (AE005672) |
| msmK | M.1 | TAGACAAAAGCAGAAACAAGGATC | 1486427–1486450 (AE005672) |
| | M.2 | <u>CATTATCCATTAAAAATCAAACGG</u>ATGTCTTCTTTGCTGTATTTACGC (1) | 1485931–1485954 (AE005672) |
| | M.3 | <u>CAAGTCAAATCCAAGCTCAAC</u>ATGTCTTCTTTGCTGTATTTACGC (6) | 1485931–1485954 (AE005672) |
| | M.4 | <u>GGAAAGGGGCCCAGGTGG</u>TTGAGCTTGGATTTGACTTG (2) | 1485183–1485203 (AE005672) |
| | M.5 | GTTGAGCTTGGATTTGACTTG | 1485183–1485203 (AE005672) |
| | M.6 | TGACCTGCTTCTGACATTTGA | 1484618–1484638 (AE005672) |
| | M.7 | CTACACAAAAATAAGCTCCATAAT | 1486508–1486531 (AE005672) |
| | M.8 | TTCCCCTATTACACGCAACCT | 1484490–1484510 (AE005672) |
| qRT-PCR | Q.A | TGAGAAACAACCATTTCGGTCA | 1711389–1711410 (AE005672) |
| | Q.B | GCACTGGGGACATGCTAAAAGT | 1711467–1711488 (AE005672) |
| | Q.C | TTATGACAGCCAGTTCACCGTTAT | 1713143–1713166 (AE005672) |
| | Q.D | GCAAAGTCGGATTGGTAGTTGG | 1713049–1713070 (AE005672) |
| | Q.E | GTTTCGGTGAGATGGAGGTTT | 1866738–1866758 (AE005672) |
| | Q.F | TTTGGAATTGGTTTGCCTTTTG | 1866617–1866638 (AE005672) |
| | Q.G | TTAGCCCGTTATAGTGAAGGTC | 762192–762213 (AE005672) |
| | Q.H | CTCTACTCGATCCCCCATCAA | 762360–762380 (AE005672) |
| Locus screen | L.A | TTCGTTGCAGAGACACT | 1604550–1604534 (CP000410) |
| | L.B | TCCACCACTACCACTAC | 1602594–1602611 (CP000410) |
| | L.C | AATCCAATCACTCCCTAAA | 1714982–1715000 (AE005672) |
| Locus exchange | S.A | ACATCTGTATCGTCAAGAG | 1604834–1604815 (CP000410) |
| | S.B | ATAATCTGTATCATTTCCTAA | 1598592–1598613 (CP000410) |
| | S.C | TGAAGAAGTTGTTTGCCA | 1605023–1605040 (CP000410) |
| | S.D | CCTTCTAACCAGTTCCTCT | 1603434–1603452 (CP000410) |
| | S.E | GTCTATGTAGTCGCTCAGAAAT | 1599725–1599746 (CP000410) |
| | S.F | ATCCGTGGGTCACAGGTT | 1598107–1598124 (CP000410) |

[a] Underlining indicates the reverse complement sequences of primers J.F (1), J.R (2), U.6 (3), C.6 (4), H.6 (5), and M.6 (6).
[b] Locations are given as nucleotide positions in the sequences with the indicated GenBank accession numbers.

$OD_{600}$ of $0.3 \pm 0.005$, they were combined with 10 ml acid phenol and 100 µl 10% SDS and incubated at 90°C until the phases merged. Samples were then cooled on ice and centrifuged for 30 min at $3,200 \times g$ and 4°C. Two additional extractions were performed, first with equal volumes of pH 4.5 acid phenol-chloroform (5:1; Ambion, Austin, TX) and then with chloroform. RNA was precipitated with 10 ml isopropanol and 1 ml of 3 M sodium acetate at −20°C overnight. RNA was concentrated and washed twice with 1 ml 70% ethanol and then vacuum dried for 20 min. The resulting RNA was resuspended in 100 µl $H_2O$, quantified, and adjusted to 100 µg in a 41-µl volume to provide a final reaction volume of 50 µl for DNase treatment. Nucleic acid was incubated with 5 µl of 10× DNase I buffer, 6 U DNase I (New England BioLabs), and 20 U Super RNaseIN (Promega, Madison, WI) at 37°C for 1 h. Purification of RNA was performed with a Qiagen RNeasy minikit (Valencia, CA) as per the manufacturer's instruction. DNase- and RNase-free reagents were used throughout. cDNA was generated with SmartScribe reverse transcriptase (RT) according to the manufacturer's recommendations (Clontech, Mountain View, CA). Parallel samples were processed without the addition of RT as a negative control.

The Janus system of mutant generation is designed to generate unmarked nonpolar mutants; however, to confirm that there were no polar effects, we demonstrated that the mutations had no effect on the transcription of the distal gene by RT-PCR. cDNA was amplified with primers designed within the gene distal to the mutation: primers C.9 and C.10 for the *scrT* mutant (*scrK*), primers U.9 and U.10 for the *susT1T2X* mutant (*susH*), and primers H.9 and H.10 for the *susH* (SP_1794) mutant. Primers designed within *aroE* (A.F and A.R) were used as a control.

Quantitative RT-PCR (qRT-PCR) was conducted to measure gene expression profiles under various growth conditions. The Brilliant SYBR green qRT-PCR master mix (Stratagene, La Jolla, CA) was used according to the manufacturer's instructions. Reactions were performed on a Mx3005 multiplex quantitative PCR system (Stratagene) using the SYBR green protocol. Cycling parameters were as follows: 95°C for 10 min and then 40 cycles of 95°C (30 s), 60°C (1 min), and 72°C (30 s), followed by a melting curve analysis. Housekeeping genes *rpoB* (primers Q.E and Q.F) and *gyrB* (primers Q.G and Q.H) were used for all experiments. Primers within the genes *susH* (primers Q.A and Q.B) and *susX* (primers Q.C and Q.D) were used for analysis. The median cycle threshold ($C_T$) value was used for analysis, and all $C_T$ values were normalized to expression of *rpoB* and *gyrB*. The delta-delta $C_T$ method was used to compare expression of test genes relative to expression of controls. Statistical analysis was performed using Student's *t* test, and a *P* value below 0.05 was considered significant. Primer sequences for housekeeping and test genes can be found in Table 2.

**PCR screening for the transporter encoded.** Analysis of strains whose genomes were sequenced revealed that they contained either a complete *fus* locus or a complete alternative transporter locus. On the basis of this information, primers were designed to distinguish between strains encoding each transporter. One primer (L.A) was designed to bind DNA encoding the predicted activator from all strains (TIGR4 SP_1800 and D39 SPD_1587). Two additional primers were designed, one of which was specific for each transporter (L.B binds within *fusC*, and L.C binds within SPD_1585). Two reactions (with primer pairs L.A-L.B and L.A-L.C) were performed for each strain. For each strain tested, a product was observed for only one primer pair, and the sizes of all products were consistent with the predicted sizes.

## RESULTS

***susT1T2X* is required for utilization of fructooligosaccharides.** To investigate the contribution of *scrT* and *susT1T2X* to growth on sucrose, we generated unmarked nonpolar single and double deletions in these loci in strain TIGR4 and tested their ability to utilize sucrose as a sole carbon source. As previously reported for a D39 derivative, the *scrT* mutant showed a significant reduction

in growth on sucrose (maximal $OD_{600}$ of ≤0.44 compared to one of ≥0.69 for the parental strain), while the *susT1T2X* mutant showed no significant alteration in growth profile (Fig. 1B) (6). In contrast to the previously published data, in this strain background, we observed no consistent reduction in the growth of the double mutant (Δ*scrT* Δ*susT1T2X*; maximal $OD_{600}$, ≥0.52) on sucrose over that observed for the *scrT* mutant (Fig. 1B) (6). In individual experiments, the maximal optical density reached by the double transporter mutant could be greater or less than that achieved by the *scrT* mutant, but this seemed to be due to variability in the assay and not due to any phenotypic difference between the strains. Genetically reconstituted strains showed no significant reduction in growth compared to the parental strain (data not shown). These data suggest that, under these experimental conditions, the contribution of the SusT1T2X transporter to growth on sucrose is minimal and that the primary substrate of the transporter is likely another carbohydrate. The residual growth of the *scrT* mutants suggests that another pneumococcal transporter(s) can contribute to growth on sucrose.

The presence of a gene encoding a β-fructosidase within the same transcriptional unit as *susT1T2X* suggests that the transporter may contribute to uptake of a carbohydrate related to sucrose [glucose α(1-2) fructose]. The inulin family of FOSs consists of a chain of β(2-1)-linked repeating fructose molecules, often with a terminal α(1-2)-linked glucose molecule. As this linkage of glucose is the same as that in sucrose, we hypothesized that SusT1T2X acts as a FOS transporter.

It has previously been reported that pneumococci could ferment inulin. In 1905 it was first proposed that the ability of pneumococci to grow on inulin could be used as a diagnostic tool, and yet the mechanism of transport remained unknown (8). To determine if growth on inulin is *susT1T2X* dependent, the ability of the parental strain, a *susT1T2X* mutant, and the genetically reconstituted strain to utilize inulin as a sole carbon source was assessed. Although the maximal optical density achieved on inulin was variable, likely due to the differences in fructose chain lengths in different preparations, the parental and genetically reconstituted strains consistently demonstrated significant growth ($OD_{600}$ range, 0.25 to 0.55) (Fig. 2A; data not shown). In contrast, the *susT1T2X* mutant showed no growth on inulin ($OD_{600}$, ≤0.016). These data support the suggestion that the ability of pneumococci to utilize inulin for growth is conferred by the presence of the *sus* locus.

To determine if *susT1T2X* contributes to growth on other FOSs, we assessed the ability of the mutant (Δ*susT1T2X*), the parental strain, and the genetically reconstituted strain to grow on defined fragments containing the same linkages as inulin and an oligofructose (raftilose) which consists of two to eight β(2-1)-linked fructose molecules. The parental strain and the genetically reconstituted strain showed significant growth on kestose (GF2), nystose (GF3), 1F-fructofuranosyl nystose (GF4), and oligofructose. The growth profiles of TIGR4 Sm$^r$ on these different FOSs varied (Fig. 2; Table 3). The reasons for these differences are currently unknown, although we predict that they reflect, at least in part, differences in the regulation of various utilization systems, different amounts of carbohydrates available for utilization by the bacterium, the affinity of the carbohydrates for the transporters, and the switch from simple to more complex carbohydrates. It does appear that carbohydrates transported by a PTS have a shorter lag phase than those utilizing the ABC transporter; how-
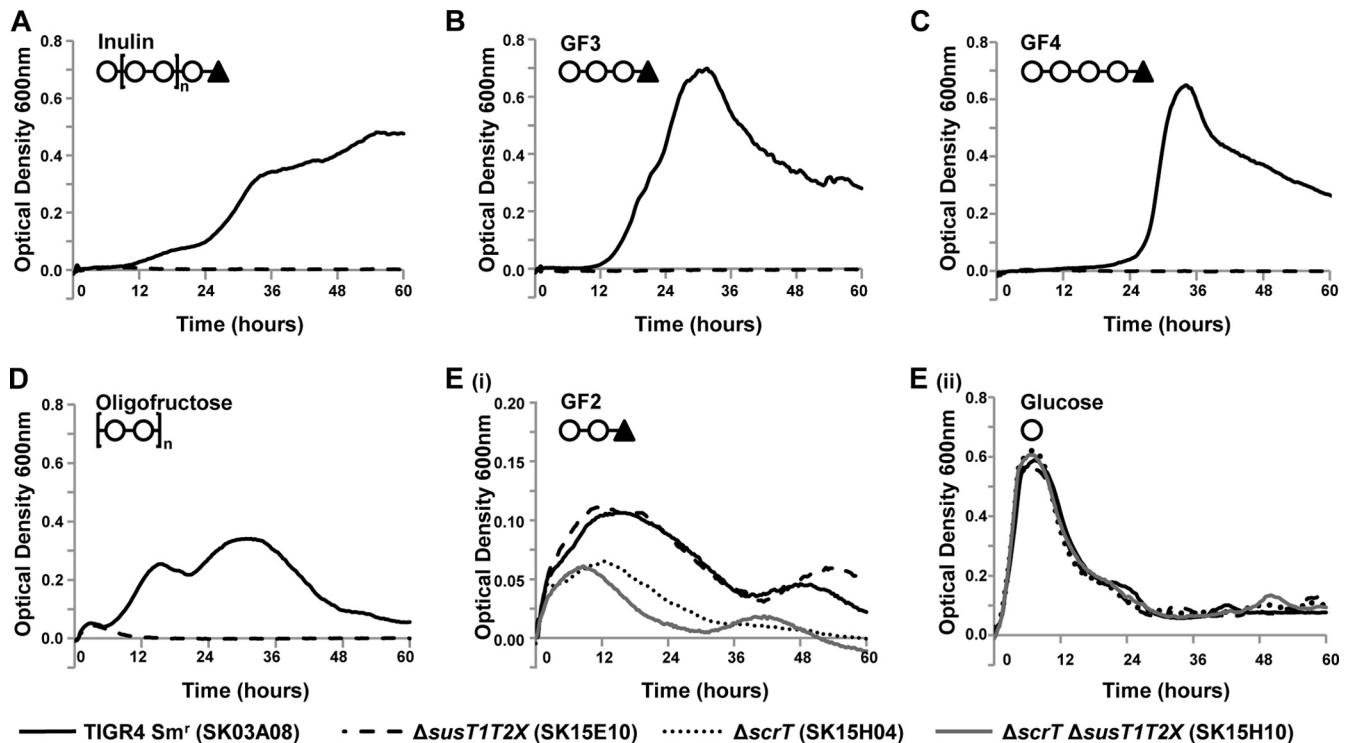
FIG 2 *susT1T2X* is required for growth on GF3 (nystose; glucose linked to three fructose molecules), GF4 (1F-fructofuranosyl nystose; glucose linked to four fructose molecules), inulin, and oligofructose. The parental strain (TIGR4 Sm$^r$ SK03A08; all panels), *susT1T2X* mutant (SK15E10; all panels), *scrT* mutant (SK15H04; panel E). and double mutant ($\Delta scrT \Delta susT1T2X$ SK15H10; panel E) were grown for 60 h on CDM supplemented with 5 mg ml$^{-1}$ inulin (A), 12 mM GF3 (B), 12 mM GF4 (C), 2.7 mg ml$^{-1}$ oligofructose (D), 12 mM GF2 (kestose; glucose linked to two fructose molecules) (Ei), or 12 mM glucose (Eii) as the sole carbon source. Symbols on the top left of the graphs indicate the structure of the carbohydrate source: ○, fructose; ▲, glucose. Growth was measured by determination of the optical density at 600 nm. A glucose control was included in each experiment to confirm that the mutants and genetically reconstituted strains had no generalized growth defects (Eii and data not shown). Data for a CDM no-carbohydrate control were subtracted from each data set. Each experiment was performed at least three independent times in triplicate, and the average data from one replicate are shown. Data graphed in panels Ei and Eii are from the same replicate.

ever, it remains unknown if this is responsible for the differences in growth profiles. Long lag phases have been observed during growth of pneumococci in chemically defined media when provided with some carbohydrates as a sole carbon source (10, 20). It is possible that the long lag could indicate selection for a mutant that enables growth on FOSs. While this is possible, we believe that it is unlikely, as we observed growth of the parental strain and the

genetically reconstituted strain in each independent experiment. Furthermore, despite the long lag phase previously observed during growth on sialic acid, transport assays have demonstrated that transport occurs rapidly (20).

Pneumococcal growth on GF3 and GF4 (OD$_{600}$ ≥ 0.6) reached approximately the same maximal OD$_{600}$ as growth on glucose, while growth on oligofructose and GF2 was less robust (oligofructose, OD$_{600}$ ≥ 0.28; GF2, OD$_{600}$ ≥ 0.121) (Fig. 2B to E). Growth on oligofructose, GF3, and GF4 was eliminated in a *susT1T2X* mutant (OD$_{600}$ ≤ 0.05). In contrast, mutation of *susT1T2X* did not affect the growth on GF2. As GF2 was the shortest FOS tested, consisting of a sucrose molecule plus an additional fructose, it seemed possible that this carbohydrate would be transported through ScrT. Studies demonstrated that growth of an *scrT* mutant on GF2 was reduced but not eliminated (Fig. 2E). A double mutant ($\Delta scrT \Delta susT1T2X$) showed no further reduction in growth over that of an *scrT* mutant, indicating that under these conditions *susT1T2X* does not contribute to growth on GF2. These data demonstrate that *susT1T2X* is critical for growth of TIGR4 on FOSs containing more than two fructose molecules. Given the variability in the maximal OD$_{600}$ observed for oligofructose and inulin, presumably due to the variation in chain lengths, we selected the defined carbohydrate GF3 for further growth studies.

TABLE 3 Maximum OD$_{600}$s and doubling time ranges of TIGR4 Sm$^r$ and transporter mutants on FOSs

| Carbohydrate source | Genotype (strain) | Range of maximum OD$_{600}$s[a] | Doubling time during maximal growth rate (h)[a] |
|---|---|---|---|
| Glucose | TIGR4 Sm$^r$ (SK03A08) | 0.55–0.68 | 1.00–2.67 |
| Inulin | TIGR4 Sm$^r$ (SK03A08) | 0.44–0.50 | 4.33–18.33 |
| GF3 | TIGR4 Sm$^r$ (SK03A08) | 0.64–0.69 | 2.50–5.17 |
| GF4 | TIGR4 Sm$^r$ (SK03A08) | 0.60–0.75 | 2.33–4.33 |
| Oligofructose | TIGR4 Sm$^r$ (SK03A08) | 0.34–0.56 | 2.00–4.67 |
| GF2 | TIGR4 Sm$^r$ (SK03A08) | 0.11–0.29 | 3.00–3.70 |
| GF2 | $\Delta susT1T2X$ (SK15E10) | 0.09–0.21 | 3.00–5.70 |
| GF2 | $\Delta scrT$ (SK15H04) | 0.07–0.09 | 3.70–4.00 |
| GF2 | $\Delta scrT \Delta susT1T2X$ (SK16F10) | 0.05–0.11 | 2.00–3.00 |

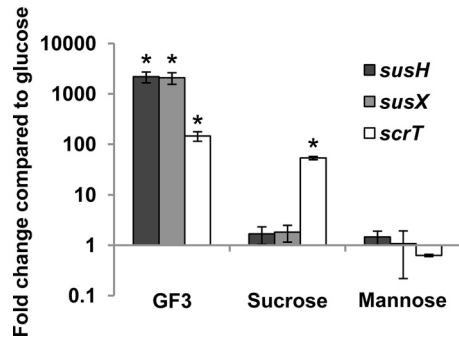[a] Ranges represent values from three independent experiments.

FIG 3 GF3 (nystose; glucose linked to three fructose molecules) induces the expression of both the *sus* locus and *scrT*, while sucrose induces the expression of only *scrT*. TIGR4 Sm$^r$ (SK03A08) was grown in CDM supplemented with 12 mM glucose, 12 mM GF3, 12 mM sucrose, or 12 mM mannose, and *susH*, *susX*, and *scrT* expression was measured by qRT-PCR. Fold change is represented as a comparison to bacteria grown in glucose. Statistical significance was tested by a two-tailed Student's *t* test; *, $P \leq 0.005$.
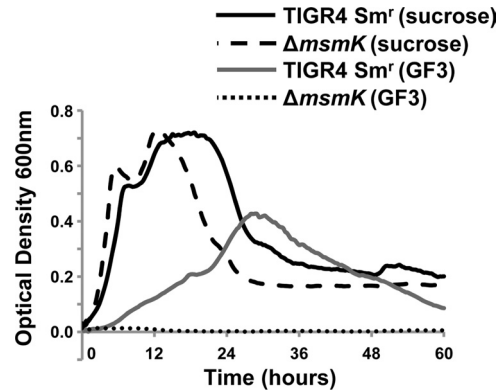


FIG 4 *msmK* is required for growth on GF3 (nystose; glucose linked to three fructose molecules). Growth of TIGR4 Sm$^r$ (SK03A08) and the *msmK* mutant (SK15J09) on CDM supplemented with 12 mM GF3 or 12 mM sucrose as the sole carbon source is shown. A glucose control was included in each experiment to confirm that the mutants and genetically reconstituted strains had no generalized growth defects (data not shown). The data for a CDM no-carbohydrate control were subtracted from each data set. The experiment was performed three independent times in triplicate, and average data from one experiment are shown.

The data generated using pneumococcal mutants strongly support the hypothesis that SusT1T2X transports FOSs. To further support the role for the *sus* locus in FOS utilization, the transcription of *susX* and *susH* during growth on sucrose, GF3, and the control carbohydrate mannose was compared in the parental strain TIGR4 Sm$^r$. While growth on neither sucrose nor mannose significantly increased expression of *susX* or *susH* compared to that with growth on glucose, GF3 induced expression more than 2,000-fold (Fig. 3). Together, these data strongly support the hypothesis that *susT1T2X* encodes components of a FOS transporter. As expected, expression of *scrT* was induced by growth on sucrose. Although *scrT* is not required for growth on GF3, expression of the gene was increased during growth on this carbohydrate. These data suggest that while the inducer of *scrT* is a by-product of both sucrose and GF3 utilization, the inducer of the *sus* locus is specific to GF3.

**FOS utilization requires *msmK*.** In addition to SusT1T2X, *S. pneumoniae* TIGR4 encodes five additional carbohydrate uptake transporter family 1 (CUT1) systems (2). Each of these loci lacks a gene(s) encoding the ATPase required to energize transport. We previously identified the ATPase MsmK, encoded by a gene at a distinct site in the genome, that is required for growth on the substrates of three characterized CUT1 ABC transporters (21). It was reported that MsmK did not contribute to pneumococcal growth on sucrose (6, 21). As a result it was suggested that the ATP-binding cassette protein energizing SP_1796-98 was likely not MsmK (6). In light of the identification of FOSs as the primary substrates of the transporter, we tested the growth of the parental strain, the *msmK* mutant, and the genetically reconstituted strain on GF3 and sucrose (Fig. 4). Consistent with previous reports, the *msmK* mutant and the genetically reconstituted strains were not significantly reduced in growth on sucrose. In contrast, the *msmK* mutant did not grow (maximal OD$_{600}$, $\leq 0.014$) on GF3, whereas the parental strain and genetically reconstituted grew to a maximal OD$_{600}$ of $\geq 0.37$. These data parallel those for the *susT1T2X* mutant and demonstrate that MsmK energizes SusT1T2X. Thus, MsmK has been shown to energize four of the six CUT1 family transporters found in TIGR4. To test the role of CUT1 family ABC transporters in the residual growth of the *scrT* mutant on GF2, a double mutant ($\Delta scrT$ $\Delta msmK$) was generated. This strain showed

no further reduction in growth compared to the single *scrT* mutant (data not shown). These data suggest that residual growth is not due to a CUT1 family transporter.

**FOS utilization is dependent on *susH*.** SusH and ScrH are both members of the GH32 family, which includes enzymes that can hydrolyze sucrose and FOSs (http://www.cazy.org) (4, 5). Previously published work demonstrated that the presence of sucrose upregulated expression of *scrH*, but not that of *susH*, and that *susH* could to some extent compensate for the loss of *scrH* (3). The contribution of these enzymes to growth on FOSs had not been addressed. A *susH* mutant was generated, and while the parental and genetically reconstituted strains could utilize the FOS GF3 as a sole carbon source (maximal OD$_{600}$, $\geq 0.359$), a *susH* mutant was unable to grow on this carbohydrate (maximal OD$_{600}$, $\leq 0.016$) (Fig. 5). In agreement with the previously published data, all strains showed equivalent growth on sucrose (Fig. 5) (3). These data demonstrate that *susH*, but not *scrH*, is required for growth on FOSs. Open reading frames in the FOS utilization locus were named *sus*, for sucrose utilization system, on the basis of the presence of a sucrose binding box and their previously proposed contribution to growth on sucrose (6). Given the finding that FOSs are the primary substrates of the *sus* locus, we propose renaming the genes *fus*, for fructooligosaccharide utilization system (Fig. 6A).

**The ability to utilize different FOSs varies between pneumococcal strains.** Although inulin fermentation has been used as a diagnostic for pneumococci, it has also been reported that there is strain-to-strain variation in the ability of pneumococci to ferment inulin (6, 22, 23). To investigate whether diverse pneumococcal strains can utilize FOS as a sole carbon source, 34 clinical isolates representing 13 different serotypes were grown on GF3. Although the growth rate and maximal OD$_{600}$ varied, all except two of the strains showed substantial growth on GF3, reaching maximal OD$_{600}$s ranging from 0.34 to 0.97 (Table 4). The percentage of strains that could utilize GF3 as a sole carbon source (94%) was much higher than the approximately 60% of strains previously reported to ferment inulin (22, 23). To determine if there were
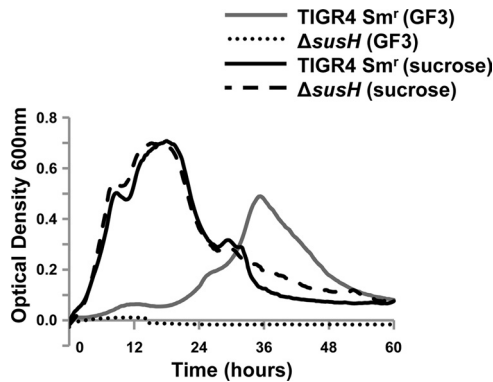
FIG 5 *susH* is required for growth on GF3 (nystose; glucose linked to three fructose molecules). Growth of TIGR4 Smr (SK03A08) and Δ*susH* (SK16G02) on CDM supplemented with 12 mM GF3 or 12 mM sucrose as the sole carbon source is shown. A glucose control was included in each experiment to confirm that the mutants and genetically reconstituted strains had no generalized growth defects (data not shown). The data for a CDM no-carbohydrate control were subtracted from each data set. The experiment was performed three independent times in triplicate, and average values from one experiment are shown.

differences in the ability of strains to utilize different FOSs, the clinical strains were tested for growth on inulin. Twenty-six of these 34 strains (76%) were able to use inulin as a sole carbon source. These data indicate that while GF3 can be utilized by the majority of pneumococci, a significant proportion of strains cannot utilize the longer FOS inulin.

**Differences in the ability of pneumococcal strains to utilize inulin are due to the transporter encoded at this genomic locus.** It has previously been reported that the *fus* locus is not part of the core genome (3, 24). Iyer and Camilli (3) reported that while the hydrolase and regulator were present in all sequenced genomes, the transporter genes were replaced by genes encoding an alternative carbohydrate-specific ABC transporter in sequenced strains of serotypes 1, 2, and 19F (Fig. 6A). Comparison of TIGR4, which encodes the Fus transporter, and D39, which encodes the alternative transporter, revealed that there were low levels of sequence similarity between the predicted amino acid sequences of the two transporters (2, 25). The substrate binding proteins essentially share no sequence similarity, while the two permease components share a maximum of 30% amino acid identity over the full length of the predicted amino acid sequence. The presence of the genes encoding the Fus transporter in 104 of the 123 pneumococcal
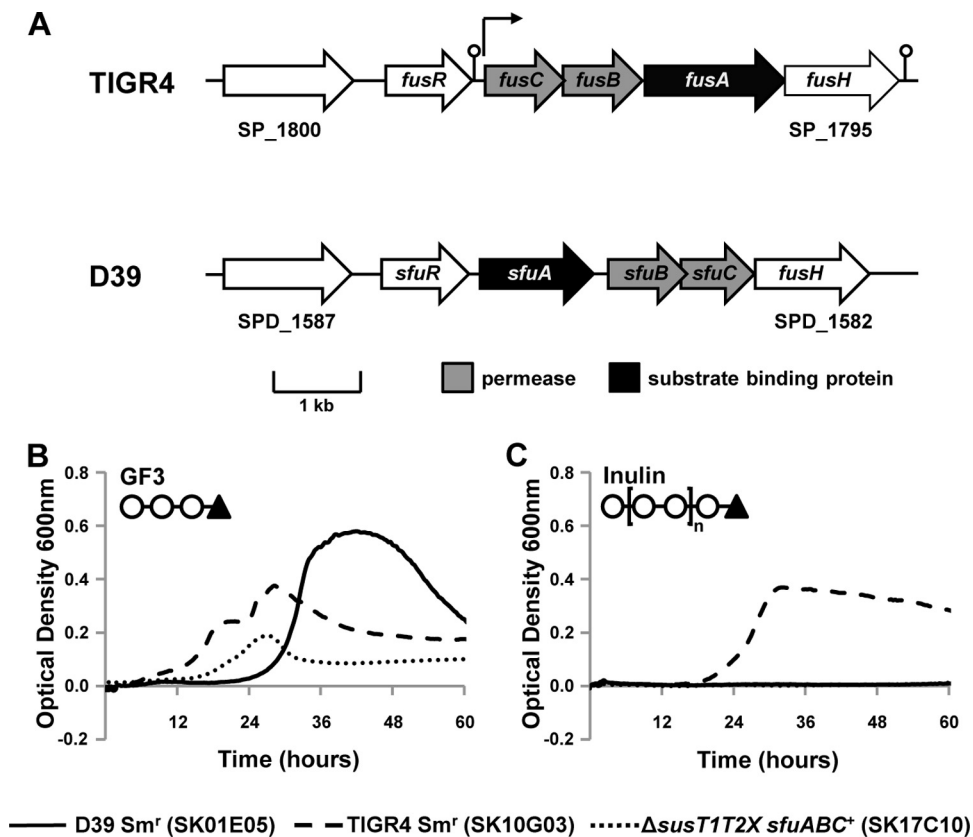


FIG 6 *fusABC* is required for growth on inulin. (A) Schematic of putative fructooligosaccharide ABC transporters in *S. pneumoniae* strains TIGR4 and D39. Open reading frames predicted within the TIGR4 and D39 sequences are represented by block arrows. Arrows above the schematic indicate predicted transcriptional start sites, and the predicted terminators are denoted by a stem-loop (2). A gene predicted to encode a LacI family repressor is present in both loci but shows high diversity (*fusR* and *sfuR*), while the sucrose hydrolase gene *fusH* is conserved across pneumococcal strains. Also conserved is the first gene of each genomic locus, predicted to encode a transcriptional activator. Pneumococcal strains were grown for 60 h on CDM supplemented with 12 mM GF3 (B), 5 mg ml$^{-1}$ inulin (C), 12 mM glucose (not shown), or no carbohydrate (not shown) as the sole carbon source. Symbols on the top left of the graphs indicate the structure of the carbohydrate source: ○, fructose; ▲, glucose. Growth was measured by determination of the optical density at 600 nm. Each experiment was performed at least three independent times in triplicate, and the average data from one experiment are shown.

TABLE 4 Clinical isolate growth on short- and long-chain FOSs

| Strain | OD$_{600}$[a] | | FOS transporter allele[c] |
|---|---|---|---|
| | GF3[b] | Inulin | |
| C06_4 | 0.82 | 0.25 | *fus* |
| C06_5 | 0.34 | − | *sfu* |
| C06_6 | 0.85 | 0.64 | *fus* |
| C06_8 | 0.93 | 0.59 | *fus* |
| C06_9 | − | − | *sfu* |
| C06_10 | 0.81 | 0.79 | *fus* |
| C06_12 | 0.81 | 0.34 | *fus* |
| C06_14 | 0.49 | 0.24 | *fus* |
| C06_18 | 0.97 | 0.6 | *fus* |
| C06_19 | 0.50 | − | *sfu* |
| C06_20 | 0.78 | 0.4 | *fus* |
| C06_21 | 0.62 | − | *sfu* |
| C06_23 | 0.81 | 0.35 | *fus* |
| C06_25 | 0.34 | 0.76 | *fus* |
| C06_31 | 0.89 | − | *sfu* |
| C06_34 | 0.67 | 0.33 | *fus* |
| C06_36 | 0.55 | 0.82 | *fus* |
| C06_39 | 0.88 | 0.39 | *fus* |
| C06_40 | 0.75 | 0.39 | *fus* |
| C06_44 | 0.35 | 0.79 | *fus* |
| C06_45 | 0.54 | 0.25 | *fus* |
| C06_48 | 0.46 | 0.62 | *fus* |
| C06_50 | 0.84 | 0.39 | *fus* |
| C06_57 | 0.85 | − | *sfu* |
| C06_60 | 0.80 | 0.45 | *fus* |
| Cl41 | 0.79 | 0.66 | *fus* |
| Cr3 | 0.59 | 0.72 | *fus* |
| Cl7 | 0.73 | 0.41 | *fus* |
| Cr31 | 0.66 | 0.33 | *fus* |
| Cl50 | 0.76 | 0.49 | *fus* |
| Cl32 | − | − | *fus* |
| Cl43 | 0.35 | − | *sfu* |
| Cl15 | 0.70 | 0.41 | *fus* |
| Cl25 | 0.65 | 0.54 | *fus* |

[a] The maximum OD$_{600}$ is reported as a measure of growth on defined medium containing the denoted carbohydrate; −, no growth, defined as a maximum OD$_{600}$ of less than 0.03.
[b] GF3 is nystose (glucose linked to three fructose molecules).
[c] As determined by PCR analysis.

genomes available at the time of the analysis suggests that these genes are the most commonly present at this genomic location (see Table S1 in the supplemental material). The alternative carbohydrate-specific ABC transporter is encoded by the remaining 19 pneumococcal genomes. All five serotype 1 strains encoded the alternative transporter, consistent with the report by Iyer and Camilli (3). In contrast, we found no correlation between the alternative transporter and serotype 19F. We are unable to determine a correlation between serotype 2 and the alternative transporter, as D39 and the derivative R6, which encode the alternative transporter, were the only two serotype 2 strains included. Although the number of strains included was relatively small for this kind of analysis, our data suggest that the presence of the *fus* locus correlates with serotypes 3, 6B, 9V, and 14.

In light of the observation that not all pneumococcal strains could utilize inulin for growth, we wanted to investigate whether the ABC transporter encoded by the strains correlated with the ability to utilize FOSs of different lengths. PCR analysis using primers specific for each transporter demonstrated that 27 of the

34 strains contained the genes encoding the Fus transporter and the remaining 7 contained the genes encoding the alternative transporter. All the strains that grew on inulin encoded the Fus transporter, while seven of the eight strains that showed no growth on inulin encoded the alternative transporter (Table 4). These data suggest that the presence of the genes encoding the Fus transporter confers the ability to utilize both short- and longer-chain FOSs and that the alternative transporter can contribute to growth only on the shorter-length FOSs. As a result of these data, we propose naming the alternative transporter short-chain fructooligosaccharide utilization system (Sfu). Two strains were unable to utilize GF3 under these assay conditions. One of these two strains was shown by PCR to encode the Fus transporter and grew well on glucose but could not utilize inulin or GF3 for growth. In addition, one strain encoding the alternative transporter grew poorly on all defined media tested, including glucose. It is possible that the strain that could utilize glucose but not FOS has a mutation in the FOS utilization pathway.

Sequence homology searches revealed that the genes encoding the Sfu transporter are closely related to a predicted ABC transporter encoded by *Streptococcus mitis*. Comparison of the sequence of D39 which encodes the alternative ABC transporter and the sequence of *S. mitis* strain NCTC 12261 revealed 98.5% identity between the nucleotide sequences and 99% identity between the predicted amino acid sequences of the substrate binding proteins. Furthermore, each permease component encoded by D39 shared greater than 98.8% identity at both the nucleotide and amino acid levels with the putative permeases encoded by NCTC 12261. These data suggest that the *sfu* locus may have been introduced into pneumococci through homologous recombination with *S. mitis*.

To confirm that the Sfu transporter confers the ability of strains encoding this transporter to utilize GF3, we constructed a TIGR4 strain encoding this transporter (Δ*susT1T2X sfuABC*⁺). As predicted, this strain was able to utilize GF3 (minimum OD$_{600}$, ≥0.1487) but not inulin (maximum OD$_{600}$, ≤0.0133) for growth (Fig. 6B). It is not clear why the growth on GF3 is less robust than that observed for either TIGR4 or D39; however, it is possible that there are other genetic or regulatory differences between strains encoding Fus and Sfu transporters that are required for efficient growth on GF3. These data demonstrate that while the ability to utilize both short- and long-chain FOSs is encoded at the *fus* locus, a transporter with the ability to utilize only short-chain FOSs is encoded at the *sfu* locus.

## DISCUSSION

Our data are consistent with previously published reports demonstrating that ScrT contributes to sucrose utilization, yet they expand beyond these observations to demonstrate that ScrT also contributes to growth on the FOS GF2 (3, 6). The residual growth of the *scrT* mutant on sucrose indicates that an additional transporter(s) contributes to growth on sucrose. Although Iyer and Camilli (3) have suggested a role for FusABC (SusT1T2X) in sucrose utilization and Bidossi et al. (6) demonstrated a role for the locus in utilization of sucrose by strain D39, under our assay conditions, we found no contribution of FusABC (SusT1T2X) to growth on sucrose in a Δ*scrT* background (3, 6). We also demonstrate that mutation of *msmK* does not further reduce the growth of the *scrT* mutant on GF2 (data not shown). As *msmK* has been demonstrated to energize all pneumococcal CUT1 family trans-

porters of known function (21), these data suggest that residual transport is due to the PTS.

It was previously reported that FusH (SusH) can compensate for the loss of ScrH during growth on sucrose (3). Our data demonstrate that *fusH* is absolutely required for growth on GF3, indicating that neither ScrH nor any other pneumococcal gene product expressed under these growth conditions can compensate for the absence of this enzyme.

It has long been known that some pneumococcal strains can ferment the FOS inulin, but the mechanism of utilization was unknown (8). Some other bacterial species are known to metabolize FOSs, including bifidobacteria and lactobacilli, present in the intestinal microflora, and oral streptococci (26, 27). Transporters for FOSs include ABC transporters, PTS, and major facilitator superfamily proteins (7, 28, 29). Our data suggest that utilization of FOSs longer than GF2 by pneumococci is dependent on one of two ABC transporters encoded at the same genomic location.

Some bacteria encode extracellular enzymes that degrade FOSs prior to transport into the cell (28, 30). We hypothesize that FOSs are transported by pneumococci intact because the β-fructosidase encoded at the *sus* locus is predicted to be intracellular and further examination of the pneumococcal genome provides no evidence of extracellular enzymes that could degrade FOSs (2). These data suggest that the transporter itself dictates the size of the FOS that can be utilized by *S. pneumoniae*. This hypothesis is also supported by data demonstrating that the ability to utilize long-chain FOSs depends on the transporter encoded at the locus. Previous reports suggest that approximately 60% of pneumococci can ferment inulin. Consistent with this, we show that 79% of clinical isolates can ferment inulin, corresponding almost exactly with the presence or absence of genes encoding the Fus transporter (22, 23). Furthermore, a TIGR4 strain encoding the Sfu transporter was unable to utilize inulin for growth.

The utilization of FOSs by the Fus transporter was found to be dependent on *msmK*. It had previously been suggested that import by this transporter was not energized by MsmK because an *msmK* mutant had no growth defect on sucrose (6, 21). We believe that growth on sucrose is primarily via the PTS ScrT, which is ATPase independent, and therefore, loss of the ATPase does not affect growth under these conditions. In contrast, growth with GF3 as the sole carbohydrate source relies on ABC transport, and therefore, loss of MsmK resulted in a significant growth defect. Fus is the fourth CUT1 family transporter which has been demonstrated to be energized by MsmK (21, 31). Although putative substrates have been identified for the remaining two CUT1 family transporters encoded by TIGR4, these carbohydrates can be substrates of multiple transporters, making determination of MsmK dependence complex (6). At present, data support the hypothesis that MsmK energizes all pneumococcal CUT1 family transporters.

The genes encoding the Sfu transporter, encoded by 19% of isolates, appear to have been introduced by recombination with *S. mitis*. It is known that *S. mitis* can utilize FOS, but no transporter has been identified (27). It is not particularly surprising that the FOS utilization locus has evolved through recombination with *S. mitis* DNA. *S. mitis* is closely related to *S. pneumoniae*, and recombination has been shown to lead to evolution of pneumococci, including increases in antibiotic resistance (32, 33). The maintenance of both the Fus and Sfu transporters in the population suggests that both transporters provide the same function *in vivo*, despite the different substrate specificities *in vitro*. Although the

possibility is less likely, it is also possible that each transporter provides a selective advantage at different stages of infection.

Conservation of the ability to utilize FOSs despite the recombination events that occur at this locus suggests that there is a selective pressure to maintain this function. These data suggest that FOSs are relevant carbohydrate sources *in vivo*; however, the question of where *S. pneumoniae* encounter FOSs remains. FOSs, including inulin, are not synthesized by humans and are generally considered prebiotics that are degraded by the bacterial community in the gut. It is known that pneumococci can colonize the oropharynx, where the bacterium may encounter dietary carbohydrates; however, the concentration of such carbohydrates in the oronasopharynx has yet to be determined. Previous studies demonstrated that the *fus* (*sus*) locus contributes to *in vivo* fitness in the lung but not the nasopharynx (3). It seems unlikely that FOSs are present in the lung, although it is possible that the transporter has additional related substrates.

There is a growing body of literature identifying complex dietary carbohydrates as potential substrates for pneumococcal transporters (6, 34–36). It is challenging to determine whether these carbohydrates are the physiologically relevant substrates of these transporters. It is known that some bacterial species, including *Streptococcus mutans*, express extracellular polysaccharides consisting of β(2-1)-linked fructose, which is essentially the same structure as oligofructan (37, 38). The relationship between these carbohydrate structures and inulin suggests that bacteria present in the same niche as *S. pneumoniae* may express fructans that can be utilized by pneumococci as a carbohydrate source for growth. Previous studies have suggested that pneumococci can utilize capsular carbohydrates produced by other bacteria for growth; for example, the *Streptococcus pyogenes* hyaluronic acid capsule can be used by pneumococci as a sole carbon source (10).

In summary, this work demonstrates the mechanism by which pneumococci utilize FOSs and determines the reason for the previously identified variation in the ability of pneumococci to utilize inulin.

## ACKNOWLEDGMENTS

## REFERENCES

1. **Weinberger DM, Malley R, Lipsitch M.** 2011. Serotype replacement in disease after pneumococcal vaccination. Lancet **378**:1962–1973.
2. **Tettelin H, Nelson KE, Paulsen IT, Eisen JA, Read TD, Peterson S, Heidelberg J, DeBoy RT, Haft DH, Dodson RJ, Durkin AS, Gwinn M, Kolonay JF, Nelson WC, Peterson JD, Umayam LA, White O, Salzberg SL, Lewis MR, Radune D, Holtzapple E, Khouri H, Wolf AM, Utterback TR, Hansen CL, McDonald LA, Feldblyum TV, Angiuoli S, Dickinson T, Hickey EK, Holt IE, Loftus BJ, Yang F, Smith HO, Venter JC, Dougherty BA, Morrison DA, Hollingshead SK, Fraser CM.** 2001. Complete genome sequence of a virulent isolate of *Streptococcus pneumoniae*. Science **293**:498–506.
3. **Iyer R, Camilli A.** 2007. Sucrose metabolism contributes to in vivo fitness of *Streptococcus pneumoniae*. Mol. Microbiol. **66**:1–13.
4. **Henrissat B.** 1991. A classification of glycosyl hydrolases based on amino acid sequence similarities. Biochem. J. **280**(Pt 2):309–316.
5. **Henrissat B, Bairoch A.** 1993. New families in the classification of glycosyl

hydrolases based on amino acid sequence similarities. Biochem. J. **293** (Pt 3):781–788.

6. **Bidossi A, Mulas L, Decorosi F, Colomba L, Ricci S, Pozzi G, Deutscher J, Viti C, Oggioni MR.** 2012. A functional genomics approach to establish the complement of carbohydrate transporters in *Streptococcus pneumoniae.* PLoS One 7:e33320. doi:10.1371/journal.pone.0033320.

7. **Barrangou R, Altermann E, Hutkins R, Cano R, Klaenhammer TR.** 2003. Functional and comparative genomic analyses of an operon involved in fructooligosaccharide utilization by *Lactobacillus acidophilus.* Proc. Natl. Acad. Sci. U. S. A. **100**:8957–8962.

8. **Hiss PH.** 1905. A contribution to the physiological differentiation of Pneumococcus and Streptococcus, and to methods of staining capsules. J. Exp. Med. **6**:317–345.

9. **Bender MH, Weiser JN.** 2006. The atypical amino-terminal LPNTG-containing domain of the pneumococcal human IgA1-specific protease is required for proper enzyme localization and function. Mol. Microbiol. **61**:526–543.

10. **Marion C, Stewart JM, Tazi MF, Burnaugh AM, Linke CM, Woodiga SA, King SJ.** 2012. *Streptococcus pneumoniae* can utilize multiple sources of hyaluronic acid for growth. Infect. Immun. **80**:1390–1398.

11. **Burnaugh AM, Frantz LJ, King SJ.** 2008. Growth of *Streptococcus pneumoniae* on human glycoconjugates is dependent upon the sequential activity of bacterial exoglycosidases. J. Bacteriol. **190**:221–230.

12. **Müller-Graf CD, Whatmore AM, King SJ, Trzcinski K, Pickerill AP, Doherty N, Paul J, Griffiths D, Crook D, Dowson CG.** 1999. Population biology of *Streptococcus pneumoniae* isolated from oropharyngeal carriage and invasive disease. Microbiology **145**(Pt 11):3283–3293.

13. **Versalovic J, Carroll KC, Funke G, Jorgensen JH, Landry ML, Warnock DW** (ed). 2011. Manual of clinical microbiology, 10th ed. ASM Press, Washington, DC.

14. **Lacks S, Hotchkiss RD.** 1960. A study of the genetic material determining an enzyme in *Pneumococcus.* Biochim. Biophys. Acta **39**:508–518.

15. **Sung CK, Li H, Claverys JP, Morrison DA.** 2001. An *rpsL* cassette, Janus, for gene replacement through negative selection in *Streptococcus pneumoniae.* Appl. Environ. Microbiol. **67**:5190–5196.

16. **Horton RM, Hunt HD, Ho SN, Pullen JK, Pease LR.** 1989. Engineering hybrid genes without the use of restriction enzymes: gene splicing by overlap extension. Gene **77**:61–68.

17. **Whatmore AM, Barcus VA, Dowson CG.** 1999. Genetic diversity of the streptococcal competence (*com*) gene locus. J. Bacteriol. **181**:3144–3154.

18. **Kloosterman TG, Bijlsma JJ, Kok J, Kuipers OP.** 2006. To have neighbour's fare: extending the molecular toolbox for *Streptococcus pneumoniae.* Microbiology **152**:351–359.

19. **Johnston JW.** 2009. Example of use of TaqMan real-time RT-PCR to analyze bacterial gene transcript levels: *Haemophilus influenzae.* Curr. Protoc. Microbiol. **Chapter 1**:Unit 1D.1.

20. **Marion C, Burnaugh AM, Woodiga SA, King SJ.** 2011. Sialic acid transport contributes to pneumococcal colonization. Infect. Immun. **79**:1262–1269.

21. **Marion C, Aten AE, Woodiga SA, King SJ.** 2011. Identification of an ATPase, MsmK, which energizes multiple carbohydrate ABC transporters in *Streptococcus pneumoniae.* Infect. Immun. **79**:4193–4200.

22. **Langvas-Nielsen A.** 1944. Fermentation power of pneumococci. Acta Pathol. Microbiol. Scand. **21**:370–373.

23. **Morch-Lund E.** 1949. The fermenting power of pneumococci. Acta Pathol. Microbiol. Scand. **26**:709–714.

24. **Obert C, Sublett J, Kaushal D, Hinojosa E, Barton T, Tuomanen EI, Orihuela CJ.** 2006. Identification of a candidate *Streptococcus pneumoniae* core genome and regions of diversity correlated with invasive pneumococcal disease. Infect. Immun. **74**:4766–4777.

25. **Lanie JA, Ng WL, Kazmierczak KM, Andrzejewski TM, Davidsen TM, Wayne KJ, Tettelin H, Glass JI, Winkler ME.** 2007. Genome sequence of Avery's virulent serotype 2 strain D39 of *Streptococcus pneumoniae* and comparison with that of unencapsulated laboratory strain R6. J. Bacteriol. **189**:38–51.

26. **Endo H, Tamura K, Fukasawa T, Kanegae M, Koga J.** 2012. Comparison of fructooligosaccharide utilization by *Lactobacillus* and *Bacteroides* species. Biosci. Biotechnol. Biochem. **76**:176–179.

27. **Hartemink R, Quataert MC, van Laere KM, Nout MJ, Rombouts FM.** 1995. Degradation and fermentation of fructo-oligosaccharides by oral streptococci. J. Appl. Bacteriol. **79**:551–557.

28. **Goh YJ, Zhang C, Benson AK, Schlegel V, Lee JH, Hutkins RW.** 2006. Identification of a putative operon involved in fructooligosaccharide utilization by Lactobacillus paracasei. Appl. Environ. Microbiol. **72**:7518–7530.

29. **Schouler C, Taki A, Chouikha I, Moulin-Schouleur M, Gilot P.** 2009. A genomic island of an extraintestinal pathogenic *Escherichia coli* strain enables the metabolism of fructooligosaccharides, which improves intestinal colonization. J. Bacteriol. **191**:388–393.

30. **Kuzuwa S, Yokoi KJ, Kondo M, Kimoto H, Yamakawa A, Taketo A, Kodaira K.** 2012. Properties of the inulinase gene levH1 of *Lactobacillus casei* IAM 1045; cloning, mutational and biochemical characterization. Gene **495**:154–162.

31. **Tyx RE, Roche-Hakansson H, Hakansson AP.** 2011. Role of dihydrolipoamide dehydrogenase in regulation of raffinose transport in *Streptococcus pneumoniae.* J. Bacteriol. **193**:3512–3524.

32. **Donati C, Hiller NL, Tettelin H, Muzzi A, Croucher NJ, Angiuoli SV, Oggioni M, Dunning Hotopp JC, Hu FZ, Riley DR, Covacci A, Mitchell TJ, Bentley SD, Kilian M, Ehrlich GD, Rappuoli R, Moxon ER, Masignani V.** 2010. Structure and dynamics of the pan-genome of *Streptococcus pneumoniae* and closely related species. Genome Biol. **11**:R107. doi:10.1186/gb-2010-11-10-r107.

33. **Dowson CG, Coffey TJ, Kell C, Whiley RA.** 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.

34. **McAllister LJ, Ogunniyi AD, Stroeher UH, Paton JC.** 2012. Contribution of a genomic accessory region encoding a putative cellobiose phosphotransferase system to virulence of *Streptococcus pneumoniae.* PLoS One 7:e32385. doi:10.1371/journal.pone.0032385.

35. **McKessar SJ, Hakenbeck R.** 2007. The two-component regulatory system TCS08 is involved in cellobiose metabolism of *Streptococcus pneumoniae* R6. J. Bacteriol. **189**:1342–1350.

36. **Shafeeq S, Kloosterman TG, Kuipers OP.** 2011. CelR-mediated activation of the cellobiose-utilization gene cluster in *Streptococcus pneumoniae.* Microbiology **157**:2854–2861.

37. **Birkhed D, Rosell KG, Granath K.** 1979. Structure of extracellular water-soluble polysaccharides synthesized from sucrose by oral strains of *Streptococcus mutans*, *Streptococcus salivarius*, *Streptococcus sanguis* and *Actinomyces viscosus.* Arch. Oral Biol. **24**:53–61.

38. **Ebisu S, Kato K, Kotani S, Misaki A.** 1975. Structural differences in fructans elaborated by *Streptococcus mutans* and *Strep. salivarius.* J. Biochem. **78**:879–887.