



# The Challenges of Using Oropharyngeal Samples To Measure Pneumococcal Carriage in Adults

 Laura K. Boelsen,<sup>a,b</sup>  Eileen M. Dunne,<sup>a,b,\*</sup>  Katherine A. Gould,<sup>c,d</sup> F. Tupou Ratu,<sup>e</sup>  Jorge E. Vidal,<sup>f</sup>  Fiona M. Russell,<sup>a,b</sup>  
 E. Kim Mulholland,<sup>a,b,g</sup>  Jason Hinds,<sup>c,d</sup>  Catherine Satzke<sup>a,b,h</sup>

<sup>a</sup>Infection and Immunity, Murdoch Children's Research Institute, Royal Children's Hospital, Parkville, Victoria, Australia

<sup>b</sup>Department of Paediatrics, The University of Melbourne, Parkville, Victoria, Australia

<sup>c</sup>Institute for Infection and Immunity, St. George's University of London, London, United Kingdom

<sup>d</sup>BUGS Bioscience, London Bioscience Innovation Centre, London, United Kingdom

<sup>e</sup>Ministry of Health and Medical Services, Suva, Fiji

<sup>f</sup>Department of Microbiology and Immunology, University of Mississippi Medical Center, Jackson, Mississippi, USA

<sup>g</sup>London School of Hygiene & Tropical Medicine, London, United Kingdom

<sup>h</sup>Department of Microbiology and Immunology, The University of Melbourne at the Peter Doherty Institute for Infection and Immunity, Parkville, Victoria, Australia

**ABSTRACT** *Streptococcus pneumoniae* (the pneumococcus) carriage is commonly used to measure effects of pneumococcal vaccines. Based on findings from culture-based studies, the World Health Organization recommends both nasopharyngeal (NP) and oropharyngeal (OP) sampling for detecting adult carriage. Given evidence of potential confounding by other streptococci, we evaluated molecular methods for pneumococcal identification and serotyping from 250 OP samples collected from adults in Fiji, using paired NP samples for comparison. Samples were screened using *lytA* quantitative PCR (qPCR), as well as pneumococcal identification and serotyping conducted by DNA microarray. A subset of OP samples were characterized by latex sweep agglutination and multiplex PCR. Alternate qPCR assays (*piaB* and *bguR*) for pneumococcal identification were evaluated. The *lytA* qPCR was less specific and had poor positive predictive value (PPV) in OP samples (88% and 26%, respectively) compared with NP samples (95% and 64%, respectively). Using additional targets *piaB* and/or *bguR* improved qPCR specificity in OP, although the PPV (42 to 53%) was still poor. Using microarray, we found that 102/107 (95%) of OP samples contained nonpneumococcal streptococci with partial or divergent complements of pneumococcal capsule genes. We explored 91 colonies isolated from 11 OP samples using various techniques, including multiplex PCR, latex agglutination, and microarray. We found that nonpneumococcal streptococci contribute to false positives in pneumococcal serotyping and may also contribute to spurious identification by qPCR. Our results highlight that molecular approaches should include multiple loci to minimize false-positive results when testing OP samples. Regardless of method, pneumococcal identification and serotyping results from OP samples should be interpreted with caution.

**IMPORTANCE** *Streptococcus pneumoniae* (the pneumococcus) is a significant global pathogen. Accurate identification and serotyping are vital. In contrast with World Health Organization recommendations based on culture methods, we demonstrate that pneumococcal identification and serotyping with molecular methods are affected by sample type. Results from oropharyngeal samples from adults were often inaccurate. This is particularly important for assessment of vaccine impact using carriage studies, particularly in low- and middle-income countries where there are significant barriers for disease surveillance.

**KEYWORDS** PCR, *Streptococcus pneumoniae*, carriage, genotypic, identification, nasopharyngeal, oropharyngeal, serotyping

**Citation** Boelsen LK, Dunne EM, Gould KA, Ratu FT, Vidal JE, Russell FM, Mulholland EK, Hinds J, Satzke C. 2020. The challenges of using oropharyngeal samples to measure pneumococcal carriage in adults. *mSphere* 5:e00478-20. <https://doi.org/10.1128/mSphere.00478-20>.

**Editor** Ana Cristina Gales, Escola Paulista de Medicina/Universidade Federal de São Paulo

**Copyright** © 2020 Boelsen et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Catherine Satzke, [catherine.satzke@mcri.edu.au](mailto:catherine.satzke@mcri.edu.au).

\* Present address: Eileen M. Dunne, Centers for Disease Control and Prevention, Atlanta, USA.

**Received** 20 May 2020

**Accepted** 1 July 2020

**Published** 29 July 2020

The Gram-positive bacterium *Streptococcus pneumoniae* (the pneumococcus) causes a range of diseases, including otitis media, pneumonia, and meningitis. Pneumococci are a significant cause of morbidity and mortality worldwide, particularly in young children and older adults (1, 2). Colonization, and subsequent carriage, of pneumococci in the upper airways (particularly the nasopharynx and oropharynx) is considered a prerequisite for pneumococcal disease and transmission (3). Carriage of pneumococci in healthy individuals is generally asymptomatic (3, 4). Pneumococcal conjugate vaccines (PCVs) targeting the pneumococcal polysaccharide capsule (current pediatric formulations contain 10 and 13 serotypes) have substantially reduced pneumococcal disease caused by vaccine-type pneumococci (5). In some settings, a subsequent increase in disease caused by nonvaccine serotypes has partially offset the benefits of PCVs (6, 7). PCVs also reduce carriage of vaccine-type pneumococci (8, 9). This, in turn, reduces transmission of pneumococci to unvaccinated individuals and therefore protects unvaccinated age groups within a population through indirect (or herd) effects.

Carriage studies are used to measure direct and indirect effects of pneumococcal vaccination in a population, particularly in low- or middle-income settings that lack robust disease surveillance systems (10). In children, pneumococci are primarily found in the nasopharynx, and so carriage is normally determined by testing nasopharyngeal (NP) swab specimens (11, 12). In contrast, pneumococcal carriage in adults is more evenly distributed between the nasopharynx and the oropharynx (13–16). Therefore, the World Health Organization (WHO) recommends collecting both oropharyngeal (OP) and NP samples for the detection of pneumococcal carriage in adults (11, 12). However, the 2013 recommendations were based on carriage studies using culture-based methodology, where the sensitivity of sampling the nasopharynx alone ranged between 58% and 81% compared with sampling both sites. The WHO highlighted that further research was needed to ascertain whether this recommendation is appropriate when using molecular techniques, which are becoming more common globally (11). More recently, two studies have compared NP and OP sampling in adults and adolescents for the detection of pneumococci using molecular methods (17, 18). Both studies concluded that sampling the oropharynx was superior to sampling the nasopharynx. A recent review examining studies of upper respiratory tract carriage of *S. pneumoniae* in adults, including those that utilized molecular methods, also encourages the sampling of other respiratory sites, such as OP and/or saliva, in addition to NP sampling (19).

However, there is some evidence that the use of molecular methods, particularly for serotyping of *S. pneumoniae* in OP samples, may yield false-positive results (20, 21). This may be particularly problematic in serotyping of *S. pneumoniae* in OP samples due to the presence of capsule biosynthesis genes in nonpneumococcal species similar to pneumococcal capsule loci. In this study, we examined whether molecular methods can be used to accurately identify and serotype pneumococci in adult OP samples. To do this, we used molecular methods for identification and serotyping on 250 paired adult NP and OP swabs collected as part of a vaccine evaluation study in Fiji (22). We assessed whether the use of multiple quantitative PCR (qPCR) targets could improve molecular detection of pneumococci from OP samples, and investigated how the presence of nonpneumococcal streptococci may confound molecular detection and serotyping of pneumococci.

## RESULTS AND DISCUSSION

The aim of our study was to examine the performance of molecular methods for identification and serotyping of pneumococci in adult OP samples. To do this, we evaluated the detection of pneumococci applying molecular methods to 250 paired NP and OP samples, collected from the same participant but stored and processed individually. A subset of OP samples and isolates was further explored to investigate potential false-positive results caused by nonpneumococcal streptococci.

***S. pneumoniae* identification in NP and OP samples.** Samples were initially screened for pneumococci using DNA extracted directly from skim milk-tryptone-glucose-glycerol (STGG) medium using *lytA* qPCR, with samples classified as positive

**TABLE 1** Pneumococcal carriage results obtained from paired nasopharyngeal and oropharyngeal swabs<sup>a</sup>

<i>lytA</i> qPCR result	Microarray result	NP <sup>b</sup> samples (n = 250)	OP <sup>c</sup> samples (n = 250)
Positive ( $C_T < 35$ )	Positive	13	7
	Negative	0	2
	No growth <sup>d</sup>	1	0
Equivocal ( $C_T$ 35–40)	Positive	7	3
	Negative	0	30
	No growth <sup>e</sup>	12	1
Negative (no $C_T$ )		217	207

<sup>a</sup>Samples were screened by *lytA* quantitative real-time PCR, and positive and equivocal samples were further tested by culture on selective agar and DNA microarray.

<sup>b</sup>NP, nasopharyngeal.

<sup>c</sup>OP, oropharyngeal.

<sup>d</sup>One NP sample was nonculturable and therefore not tested by microarray.

<sup>e</sup>Equivocal samples with no growth were considered negative for pneumococci (54).

(cycle threshold value [ $C_T$ ], <35), equivocal ( $C_T$  35 to 40), or negative (no  $C_T$ ) (Table 1). Forty-three (17%) of the 250 OP samples screened were either *lytA* positive (9/250 [4%]) or equivocal (34/250 [14%]), while 33 (13%) of the 250 NP samples tested were either *lytA* positive (14/250 [6%]) or equivocal (19/250 [8%]). All *lytA*-positive and equivocal samples were tested for the presence of pneumococci using DNA microarray following culture enrichment on selective agar (23).

For NP samples, microarray detected pneumococci in 13/14 (93%) *lytA*-positive samples and 7/19 (37%) *lytA*-equivocal samples. In contrast, microarray detected pneumococci in 7/9 (78%) *lytA*-positive OP samples, compared with only 3/34 (9%) *lytA*-equivocal OP samples. The proportion of *lytA*-equivocal samples that were found to be true positives for pneumococci was higher for NP samples than for OP samples ( $P = 0.025$ ), suggesting that the *lytA* gene is a poor target for the detection of pneumococci in adult OP samples. Our findings are consistent with a study examining carriage in Kenyan adults using 40 combined NP/OP samples, in which the real-time *lytA* PCR positivity rate (60%) was much higher than what could be confirmed using culture (12.5%) (20).

When cultured on selective media, OP samples typically produced a greater diversity of colony morphologies (and therefore may be more likely to contain a number of different gentamicin-resistant species [see Fig. S1 in the supplemental material]) than NP samples. Consistent with this, pneumococci were only detected by microarray in 10/42 (24%) of OP samples that had bacterial growth on gHBA (horse blood agar plates containing 5  $\mu$ g/ml of gentamicin) plates, compared with 20/20 (100%) NP samples (two additional NP samples were not tested by microarray, as they lacked  $\alpha$ -hemolytic growth). OP samples also had a greater density of bacterial growth on gHBA plates than NP samples: 41/42 (98%) of OP samples had growth on a 1:100 dilution plate, compared with 13/22 (59%) of NP samples ( $P < 0.001$ ), consistent with a study with Dutch adults (17).

**OP sample characterization.** We hypothesized that the presence of nonpneumococcal streptococci in the oropharynx may contribute to the *lytA* signal and may lead to false identification of pneumococci. We therefore sought to confirm the presence of nonpneumococcal streptococci using microarray which, though not designed specifically for identification of individual streptococci, can confirm the presence of streptococcal species. We tested 30 OP samples using microarray (5 of which were *lytA* positive and 25 *lytA* equivocal) and found that while only 5 (17%) contained pneumococci, all contained complex mixtures of nonpneumococcal species, including *Streptococcus salivarius*, *Streptococcus mitis*, *Streptococcus oralis*, *Streptococcus infantis*, and, less commonly, *Streptococcus parasanguinis*, *Streptococcus anginosus*, and *Streptococcus sanguinis*. The detection of streptococci by microarray also suggests that the reasons for

**TABLE 2** Evaluation of *lytA*, *bguR*, and *piaB* qPCR (alone and in combination) for detection of pneumococci in 250 OP samples<sup>a</sup>

Target gene(s)	No. of:				PPV (%) <sup>b</sup>	Specificity (%)	Sensitivity (%)
	True positives	False positives	False negatives	True negatives			
<i>lytA</i>	10	29	1	210	26	88	91
<i>bguR</i>	10	68	1	171	13	72	91
<i>piaB</i>	10	20	1	219	33	92	91
<i>lytA</i> + <i>bguR</i>	10	14	1	225	42	94	91
<i>lytA</i> + <i>piaB</i>	9	11	2	228	45	95	82
<i>bguR</i> + <i>piaB</i>	9	10	2	229	47	96	82
<i>lytA</i> + <i>bguR</i> + <i>piaB</i>	9	8	2	231	53	97	82

<sup>a</sup>A positive is defined as any  $C_T$  of  $<40$ .

<sup>b</sup>Positive predictive value (PPV), specificity, and sensitivity were compared to DNA microarray, the study gold standard for pneumococcus-positive samples.

discrepancies between *lytA* qPCR and microarray in pneumococcal detection are not due to poor sensitivity of microarray. While *S. pneumoniae* has distinct *lytA* alleles compared with the mitis group, there is evidence that *lytA* qPCR can both misidentify commensal streptococci as pneumococci and fail to identify pneumococci (24, 25). Genetic exchange with streptococci has played a major role in evolution of *S. pneumoniae*, and many *S. pneumoniae* virulence genes are also present in *S. mitis*, *S. oralis*, and *S. infantis* (26). As bacterial diversity in the oropharynx is higher than in the nasopharynx (27), there is greater opportunity for genetic exchange between streptococci. This genetic exchange, particularly within members of the mitis group streptococci, means that the reliance on one single loci for identification may result in erroneous species identification (28).

**Alternative qPCR-based *S. pneumoniae* identification in OP samples.** Next, we aimed to improve qPCR-based identification of pneumococci in OP samples by evaluating alternative target genes, which could be potentially used alone or in combination with *lytA*. We assessed the seven potential targets (Tables S2 and S3) by conventional and real-time PCR against a set of 30 reference isolates (6 *S. pneumoniae* and 24 nonpneumococcal streptococcal isolates) (Tables S4 and S5). Based upon the results, the genes *bguR* and *piaB* were further tested.

The *bguR* and *piaB* qPCR assays were applied to DNA extracted from 250 OP samples, which had been previously tested with *lytA* qPCR. One hundred forty-two samples were negative for all three targets. One hundred eight samples were positive or equivocal for either *bguR*, *piaB*, and/or *lytA* and were subsequently cultured on gHBA plates and analyzed by microarray. Based on microarray results, the gold standard for pneumococcal identification in our study by detection of multiple species-specific genes, 11 (10%) out of 108 samples contained *S. pneumoniae*. Note that the matched NP sample was also positive for pneumococci for two of these 11 OP samples; therefore, by using OP samples, nine additional pneumococcal carriers were detected.

The three qPCR targets were evaluated for sensitivity, positive predictive value (PPV), and specificity individually and in combination using a  $C_T$  of  $<40$  as a cutoff for positivity (Table 2).

Individually, none of the targets performed well: *piaB* had 33% PPV and 92% specificity, *lytA* had 26% PPV and 88% specificity, and *bguR* had 13% PPV and 72% specificity. All three targets had one false negative (to give a sensitivity of 91%). When two targets were combined, the PPV and specificity improved (ranging from 42 to 47% and 94 to 96%, respectively), although when *piaB* was used with *lytA* or *bguR* there were two false negatives, reducing the sensitivity of these combinations to 82%. Combining all three targets further improved both PPV (53%) and specificity (97%) but yielded lower sensitivity (82%) than for single targets. It should be noted that the PPV of all approaches and the specificity of most approaches investigated were inferior to the use of *lytA* in NP samples (64% PPV and 95% specificity). Our findings were in contrast to those of a recent study exploring the use of *lytA*, *piaB*, and SP2020 (*bguR*) assays for the detection of pneumococci (29). When tested on 80 OP samples, there were few inconsistencies between assays and no apparent issues with the *lytA* assay, and

combining assays eliminated false positives. The reasons for the discrepancies between studies is unclear.

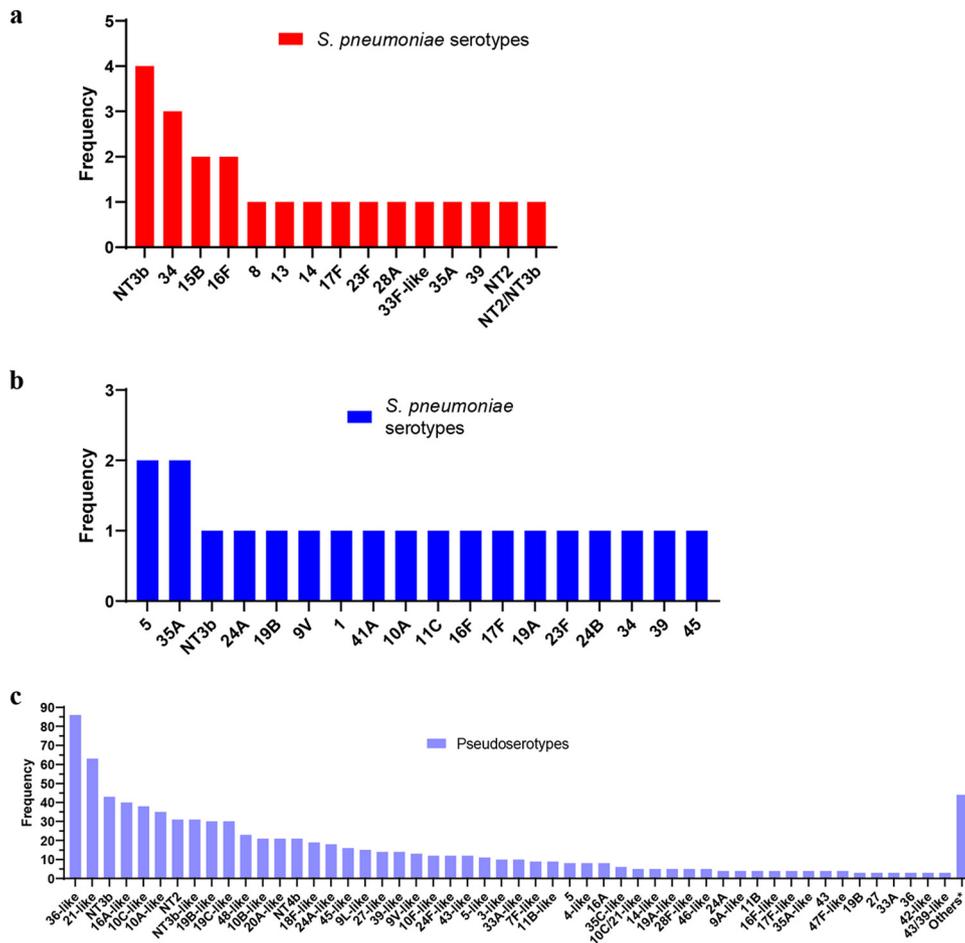
The primary aim of this study was to evaluate targets for use in a screening assay, and therefore, we focused primarily on specificity and PPV. Any sample that was negative by all three targets was deemed a true negative. For each individual assay, there was a subset of samples that were negative for that assay and were tested further (69, 30, and 78 for *lytA*, *bguR*, and *piaB*, respectively). However, one limitation of our study is that we lacked a complete measure of false negatives and therefore did not fully evaluate sensitivity and negative predictive value.

If these assays were used for identification rather than detection of pneumococcus-positive adult OP samples, a more stringent cutoff would be more appropriate. When we applied a cutoff  $C_T$  of  $<35$  to our OP samples, *lytA* was the best target (99% specificity, 78% PPV, and 64% sensitivity) and the use of multiple targets was not appropriate (all other target/s had  $\leq 36\%$  sensitivity) (Table S6). We also considered a sequential screening approach whereby only those samples with a  $C_T$  of  $<40$  for one target would be further tested by other targets, which gave 37 to 41% PPV, 93 to 95% specificity, and 82 to 91% sensitivity (Table S7). We also explored the possibility that false-positive samples could be identified through differences in signal intensity between two targets; however, no clear differences were observed in positive samples versus negative samples for any target combination. Although we had only a small number of pneumococcus-positive OP samples, our results indicate that qPCR may be unreliable for accurate detection of pneumococci in OP samples even when multiple targets are used.

**S. pneumoniae serotyping in NP and OP samples.** Next, we conducted molecular serotyping of pneumococci from NP and OP samples using DNA microarray. The microarray detects multiple targets for each of the pneumococcal capsule genes in order to assign a serotype, and it is the best method for detecting multiple serotypes in NP samples from children owing to its high sensitivity and PPV (95.8% and 93.9%, respectively) (23). Serotyping of NP samples from adults using microarray was relatively straightforward (Fig. 1a), with only one (18/20 [90%]) or two (2/20 [10%]) serotypes detected in each sample. Additionally, all samples were clearly positive for pneumococci and no other streptococci were detected.

In contrast, OP samples were highly complex and challenging to serotype, with the presence of partial or divergent sets of capsule genes that were likely to be present in nonpneumococcal species, here referred to as pseudoserotypes. Serotyping was attempted on 108 OP samples, with a total of 876 serotypes ( $n = 20$ ) and pseudoserotypes ( $n = 856$ ) detected (Fig. 1b and c). Only one OP sample had no growth on blood agar plates and was therefore unable to be further tested. Of the 107 OP samples, 55 (51%) had 10 serotypes and/or pseudoserotypes detected, only 5 samples had no serotypes or pseudoserotypes (streptococci were detected in all 5, so this was not due to an error in the microarray) and no OP samples had only pneumococci and pneumococcal serotypes. The most common pseudoserotype was similar to serotype 36 ("36-like," indicating that a partial or divergent set of serotype 36 pneumococcal capsule genes was detected) and was present in 83% (89/107) of OP samples.

**Additional characterization of OP samples.** We next investigated whether the presence of pseudoserotypes in nonpneumococcal streptococci in the OP samples could lead to false-positive serotyping results using other methods, and/or potentially produce capsule similar to true pneumococci. We selected 11 OP samples for this detailed analysis, representing a variety of their prior *lytA* qPCR and microarray serotyping results. For example, two samples were *lytA* positive but had no pneumococci detected by microarray, while another had a complete (or near complete) set of capsule genes for serotype 9V but had no pneumococci detected. Only 2 of the 11 samples were found to contain pneumococci using microarray. We tested the 11 OP samples with antibody-based and molecular-based serotyping methods (latex sweep agglutination and multiplex PCR [mPCR], respectively) (Table 3). When latex sweep aggluti-



**FIG 1** (a) Frequency of pneumococcal serotypes detected in NP samples. (b and c) Pneumococcal capsular genes detected in OP samples from *S. pneumoniae* (serotypes) (b) or other species (pseudoserotypes) (c). \*, others include 9V, 9L, 18A-like, 18C-like, 20-like, 33F-like, 37, 41F-like, 47A, NT3a, and NT4a (all  $n = 2$ ) and 1, 41A, 7A-like, 7F, 9A, 9N-like, 10F, 13-like, 17A-like, 17-like, 20A, 22A, 24B-like, 28A-like, 32F-like, 33B-like, 35B, 35F-like, 40, 41A-like, 46, and 48 (all  $n = 1$ ).

nation was used, 6 of 11 samples had a serotype detected. For mPCR, all 11 samples had a serotype detected; however, for 2 samples, the *cpsA* pneumococcal control gene was absent. While all three methods (microarray, latex sweep agglutination and mPCR) detected pseudoserotypes, only microarray designated these as coming from nonpneumococcal species.

**OP isolate characterization.** We recultured these 11 OP samples on gHBA plates and isolated a representative of each colony morphology identified in the sample ( $n \geq 5$  for each sample). The resultant 91 isolates from these 11 OP samples were further characterized by the identification tests *lytA*, *bguR* and *piaB* real-time PCR, optochin susceptibility, bile solubility, and matrix-assisted laser desorption ionization–time of flight mass spectrometry (MALDI-TOF MS), as well as latex agglutination for serotyping (Table S8 and S9). An isolate was deemed *S. pneumoniae* if at least three out of four identification tests were indicative of *S. pneumoniae*; if only two tests were indicative of *S. pneumoniae*, DNA microarray was used to resolve identification. Only three of these isolates were identified as *S. pneumoniae*. Nearly half (41/88 [47%]) of the nonpneumococcal isolates were identified as *S. mitis/S. oralis* (not distinguished by MALDI-TOF MS), with *S. parasanguinis* (12/88 [14%]) and *S. salivarius* (11/88 [13%]) also common (Fig. 2a). For 11 of the 88 isolates, no identification was obtained by MALDI-TOF MS.

Serotyping by latex agglutination was attempted on all 91 isolates; however, 33 isolates did not emulsify in saline and so were not able to be tested. We did not identify

**TABLE 3** Detailed microbiological analysis of 11 OP samples using *lytA* qPCR, microarray, latex sweep agglutination, and multiplex PCR

Sample ID	<i>lytA</i> qPCR result (C <sub>T</sub> ) <sup>a</sup>	Microarray PathID <sup>b</sup>	Microarray serotyping <sup>c</sup>	Latex sweep agglutination	mPCR <i>cpsA</i> result <sup>d</sup>	mPCR serotyping <sup>e</sup>
FVEP-002-004	Equivocal (39.52)	SP-3/5	9A-like* (34%) + 7F-like* (26%) + 36-like* (17%) + NT4b* (17%) + 27-like* (6%)	4 + 9L + 36 + 15	Positive	7F/7A + 24F/24A/24B + 10F/10C/33C + 9N/9L
FVEP-002-418	Equivocal (38.48)	SP-3/5	NT4a* (53%) + 24A-like* (21%) + NT4b* (11%) + 43/39-like* (9%) + 16A-like* (3%) + 28A-like* (2%) + 4-like* (1%)	24A	Negative	33F/33A/37 + 35B + 24F/24A/24B + 4 + 10F/10C/33C + 35A/35C/42 + 9N/9L
FVEP-002-496	Equivocal (37.93)	SP-3/5	10C/21-like* (41%) + NT4a* (20%) + 36-like* (12%) + 39-like* (8%) + 16A-like* (8%) + 48-like* (4%) + 24A-like* (4%) + 45-like* (2%) + 19B-like* (1%)	NSD <sup>f</sup>	Positive	22F/22A + 24F/24A/24B + 10A + 10F/10C/33C + 20 + 13
FVEP-002-460	Equivocal (35.82)	SP-5/5	NT4b* (35%) + 10C/21-like* (29%) + 48-like* (11%) + 41A (7%) + 20-like* (5%) + 23F (4%) + 35A-like* (4%) + 5-like* (3%) + 16A-like* (2%)	NSD	Positive	6A/6B/6C/6D + 6C/6D + 22F/22A + 33F/33A/37 + 35B + 4 + 23F + 10A + 10F/10C/33C + 5 + 35A/35C/42 + 34 + 21 + 20
FVEP-002-002	Equivocal (37.79)	SP-3/5	NT4b* (46%) + 33A-like* (40%) + 9V (5%) + 18F-like (4%) + 35A-like (4%) + 19B-like (1%)	12 + 33 + 35B	Positive	6A/6B/6C/6D + 19A + 33F/33A/37 + 38/25F/25A + 35B + 10A + 10F/10C/33C + 35A/35C/42
FVEP-002-084	Equivocal (37.57)	SP-3/5	NT4b* (62%) + 36-like* (25%) + 5-like* (7%) + 19B-like* (5%) + 9N-like* (1%)	19B	Positive	4 + 5 + 9N/9L
FVEP-002-488	Equivocal (39.82)	SP-3/5	NT4b* (45%) + 48-like* (36%) + 19B-like* (11%) + 43/39-like* (8%)	NSD	Negative	38/25F/25A + 4 + 23F + 39 + 10F/10C/33C + 34 + 21 + 20 + 6A/6B/6C/6D
FVEP-002-078	Positive (28.44)	SP-3/5	NT4b* (26%) + 10C/21-like* (26%) + 7F-like* (26%) + 16A-like* (11%) + 45-like* (6%) + 9L-like* (5%)	35B	Positive	15A/15F + 7F/7A + 24F/24A/24 + 4 + 10F/10C/33C + 9N/9L + 21 + 20
FVEP-002-080	Positive (34.41)	SP-5/5	NT4b* (32%) + 48-like* (27%) + 24B (20%) + 19B-like* (13%) + 1 (8%)	19B	Positive	24F/24A/24B + 1 + 9N/9L + 13
FVEP-002-422	Positive (34.70)	SP-3/5	NT4b* (42%) + 10C/21-like* (42%) + 19C-like* (12%) + 16A-like* (4%)	NSD	Positive	15B/15C + 10F/10C/33C + 21 + 17F
FVEP-002-486	Equivocal (35.29)	SP-3/5	NT4b* (71%) + 10C/21-like* (17%) + 11B-like* (5%) + 16A-like* (5%) + 19C-like* (2%)	NSD	Positive	15A/15F + 10F/10C/33C + 35A/35C/42 + 9N/9L + 31 + 20

<sup>a</sup>C<sub>T</sub>, cycle threshold value for *lytA* qPCR (C<sub>T</sub> of < 35, positive; C<sub>T</sub> of 35 to 40, equivocal; no C<sub>T</sub>, negative).

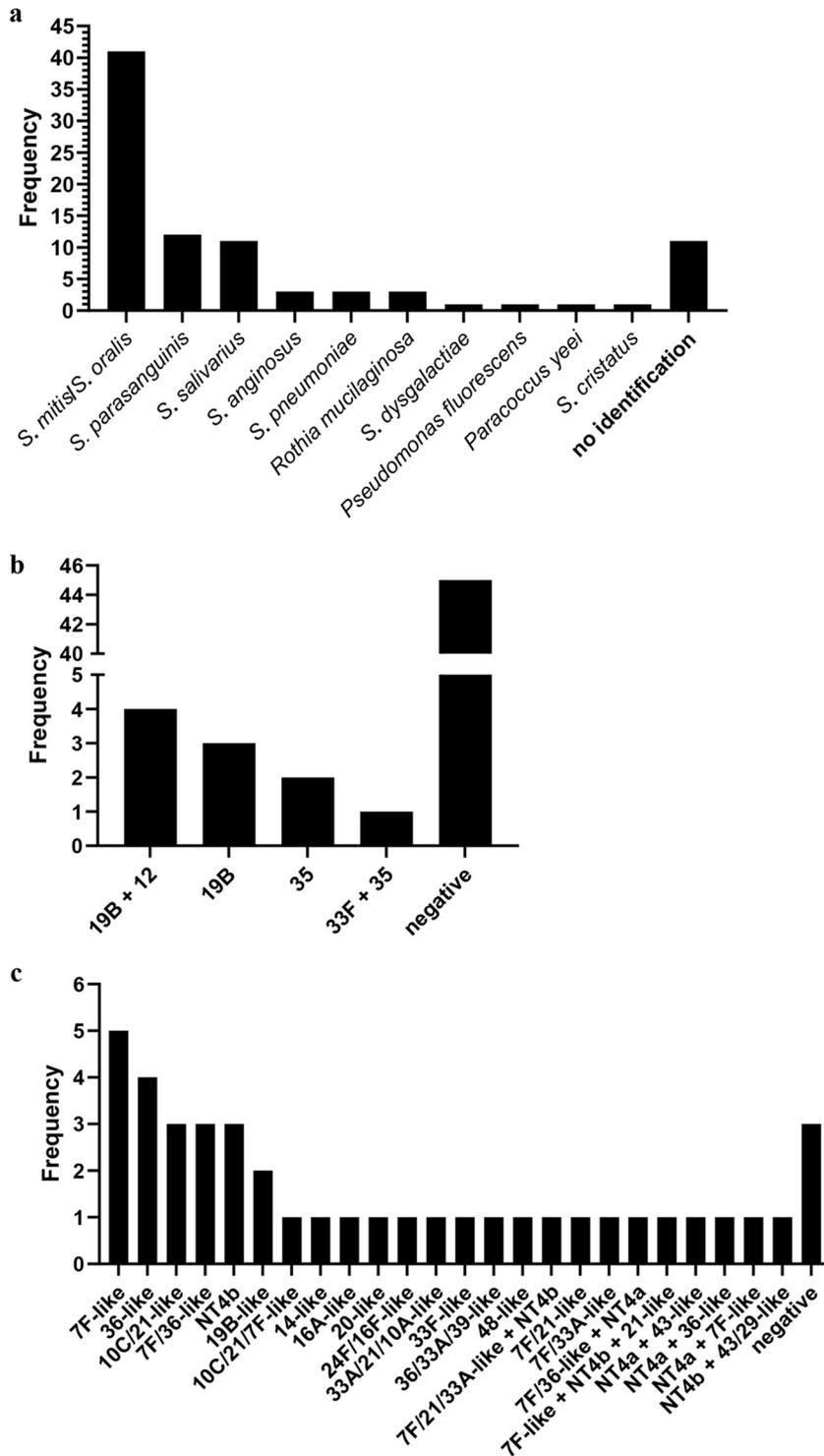
<sup>b</sup>SP-5/5 result for the PathID component on microarray indicates a sample positive for *S. pneumoniae*; SP-4/5 or less indicates that *S. pneumoniae* is not present.

<sup>c</sup>Pseudoserotypes are marked by "\*" indicating that the result is likely from a nonpneumococcal bacterial species and "-like\*" indicating a partial or divergent set of capsule genes.

<sup>d</sup>*cpsA* in mPCR is used as a positive control for pneumococci.

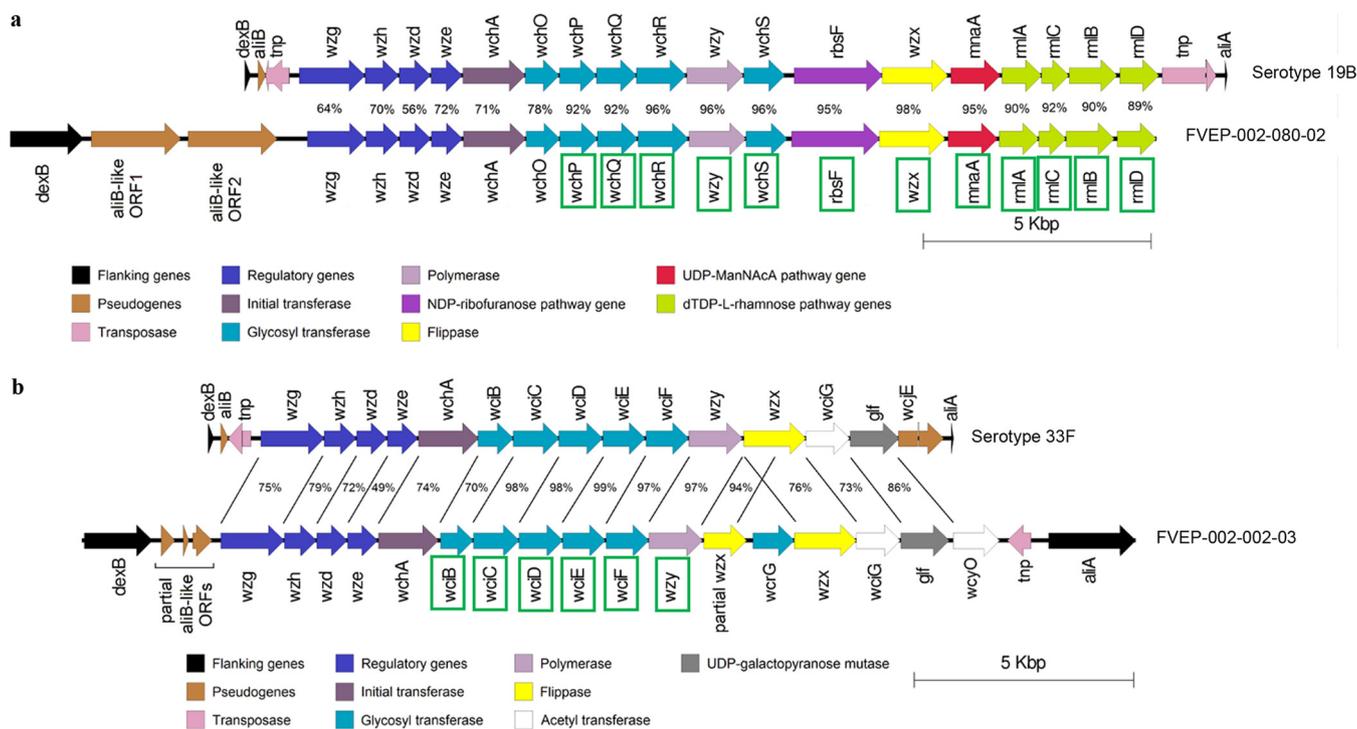
<sup>e</sup>Each sample also had additional mPCR products that did not correlate with an expected mPCR serotype product size that are not listed here.

<sup>f</sup>NSD, no serotype detected.



**FIG 2** MALDI-TOF MS identification (a), latex agglutination (b), and microarray serotyping (c) results for 88 nonpneumococcal isolates subcultured from 11 OP samples. These data do not include results from 33 isolates that were not able to be tested by latex agglutination or from 47 isolates that were not tested by microarray for practical reasons.

a serotype for 45 isolates. A serotyping result was obtained for 10 of 55 nonpneumococcal isolates, with results shown in Fig. 2b. When the 10 isolates were serotyped by Quellung, none were positive, although some slight cell enlargement was observed for some isolates. The reasons for this were not explored, but it is plausible that the capsule



**FIG 3** Comparison of the capsule gene locus for isolate FVEP-002-080-02 (a) with serotype 19B (GenBank accession no. [CR931676.1](#)) and the capsule gene locus for isolate FVEP-002-002-03 (b) with serotype 33F (GenBank accession no. [AJ006986.1](#)). Sequence identity (percent) between the two sets of genes is shown. Genes are categorized according to characteristics used by Bentley et al. (53). Green boxes around genes indicate that a gene was categorized as present by microarray serotyping.

polysaccharide produced may be thinner in these closely related species than pneumococcal capsule polysaccharide, or not covalently linked to the cell wall peptidoglycan, therefore making it difficult to observe the typical “swelling” in a Quellung reaction.

To further elucidate the source of the spurious identification and serotyping results, a subset of isolates ( $n = 41$ ) were serotyped by microarray. These included all isolates that were equivocal for *lytA* and positive for a serotype/serogroup in latex agglutination and at least one isolate for each unique MALDI-TOF MS result. Pseudoserotypes were detected for 38 of the 41 (93%) nonpneumococcal isolates tested, with results shown in Fig. 2c. We conducted whole-genome sequencing on two of the nonpneumococcal OP isolates (which had typed as 19B [ $n = 1$ ] or 33F [ $n = 1$ ] by latex agglutination) and compared the *cps* locus of these isolates with the *cps* locus of a corresponding true pneumococcal serotype. Isolate FVEP-002-080-02 (which serotyped as 19B by latex agglutination) was identified as *S. infantis* and had a full set of 19B capsule genes, although the first six genes shared <78% identity with the 19B reference isolate (Fig. 3a). Isolate FVEP-002-002-03, which serotyped as 33F using latex agglutination, was identified as *Streptococcus oralis* subsp. *tigurinus* (Fig. 3b). The *cps* region of this isolate was similar to that of 33F; however, two copies of *wzx* were present (one of which was truncated). A *wcrG* gene, which is normally found in serotype 10A, was also present. Excluding the truncated *wzx* gene and the *wcrG* gene, this *cps* locus was highly similar to the *S. oralis* subsp. *tigurinus* strain Az\_3a *cps* locus (30), including the presence of the *wcyO* gene (which is most similar to that of serotype 21 and is not typically present in serotype 33F pneumococci).

**General discussion.** Our findings from testing the isolates confirm that the non-pneumococcal streptococci in OP samples can contribute to false positives for pneumococcal serotyping. Pseudoserotypes were found in all of the 38 nonpneumococcal streptococci isolates tested. Importantly, our results suggest that nonpneumococcal streptococci present in the oropharynx may be falsely identified as pneumococci and

can also produce capsule similar to pneumococcal capsule. Consistent with this, several recent studies have reported closely related streptococci expressing capsular polysaccharide antigenically similar to pneumococcal capsule (30–32). Kilian et al. have suggested that the diversity of the *cps* locus emerged from unidirectional interspecies gene transfer from commensal streptococci to *S. pneumoniae* (33, 34). Recently, *Streptococcus mitis* strains that produce capsule immunologically consistent with pneumococcal serotype 1 have recently been reported (35), raising questions as to whether carriage of such strains would induce an immune response to pneumococci.

A key strength of our study was applying some of the best available methods for serotyping our OP samples. The use of multiple different methods also allowed us to demonstrate that specificity issues affect multiple types of serotyping methods and are not limited to one method alone. Samples were processed in batches that included both NP and OP samples and multiple extraction controls, thereby eliminating the possibility that underlying differences have emerged due to contamination or batch variability. A limitation of our study is that we used a culture enrichment step prior to pneumococcal identification using microarray, and it is possible that we may have missed some samples containing pneumococci. We chose microarray as a gold standard for pneumococcal identification in this study because it is not based on one single data point or gene sequence but includes targets over the entire genome, which provides greater confidence in results. It is plausible the presence of high-density nonpneumococcal species on culture plates for OP samples may mask the detection of pneumococci in downstream testing, or some pneumococci may be nonviable. However, Trzcinski et al. found that culture enrichment improved detection of pneumococci (17), and our overall findings are consistent with other studies which have shown the presence of pneumococcal gene homologs in other closely related streptococci (20, 21, 36). Despite our detailed investigation, we were also unable to definitively identify the source of spurious pneumococcal identification results. While we failed to find any nonpneumococcal isolate that was positive for the *lytA* gene, we did find 10 isolates which had some amplification. It is plausible that we may have missed those colonies which directly contributed to the *lytA* result. It is also possible that, cumulatively, the weak amplification in individual nonpneumococcal colonies contributes to equivocal results at a sample level.

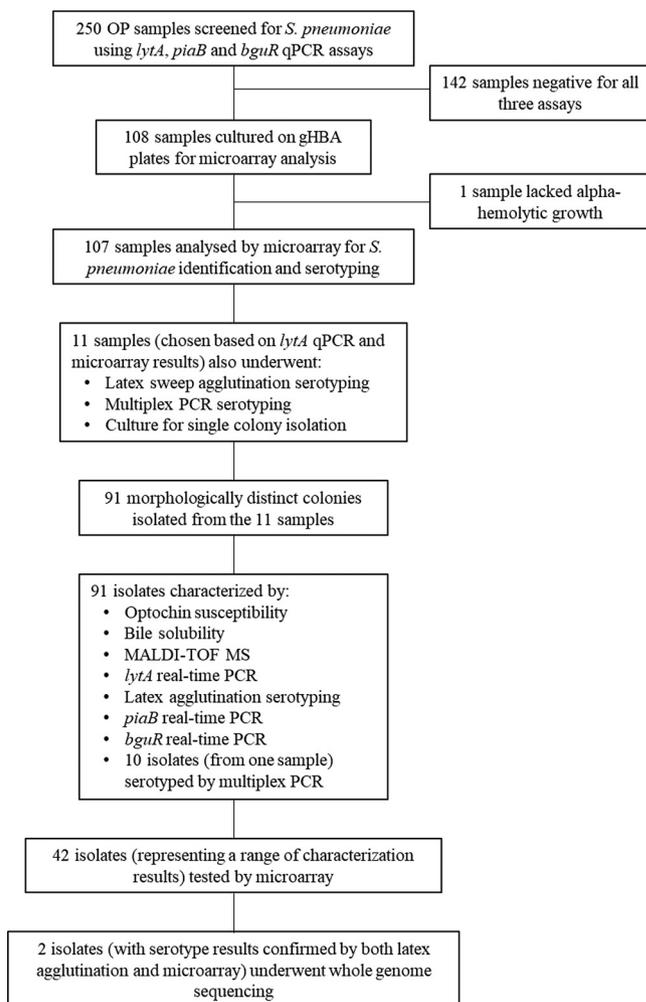
We should also note that our adult participants were mostly female (93%) and under the age of 65 (98%) and therefore may differ from adult participants described in studies aimed at understanding disease in older adults.

Taken together, our results highlight the challenging nature of conducting a pneumococcal carriage study in adults using OP samples. The presence of other commensal streptococci in the oropharynx interferes with identification and serotyping of true pneumococci within an OP sample, regardless of whether culture-based or molecular serotyping methods are used. For adult pneumococcal carriage studies, approaches which utilize multiple targets or loci, or involve the isolation and careful identification of pneumococci, are necessary if OP samples are to be included, although the latter may be impractical for a large study.

Given the challenges of identifying and serotyping pneumococci using OP samples from adults, we caution against their routine use in carriage studies.

## MATERIALS AND METHODS

**Swab collection, transport, and storage.** An NP and OP swab were collected from 250 adult caregivers of children (adult caregivers) in 2012 as part of the Fiji New Vaccine Evaluation Project, described previously (22); see Table S1 for study participant characteristics. In brief, NP and OP specimens were collected concurrently by the same personnel using flocked nylon swabs (11, 12) and were otherwise handled in the same manner. Each swab was placed in 1 ml of skim milk-tryptone-glucose-glycerol (STGG) medium in an individual cryovial and transported in a cool box to the Fiji Centre for Communicable Disease Control in Suva, Fiji. On arrival at the laboratory, swabs were vortexed for 20 s prior to aliquoting and storage at  $-80^{\circ}\text{C}$  within 8 h of collection, in line with WHO guidelines (11). Samples were shipped on dry ice to the Murdoch Children's Research Institute in Melbourne, Australia, and stored at  $-80^{\circ}\text{C}$ . For clarity, a flowchart showing the full set of experimental procedures conducted on OP samples and the derived isolates is presented in Fig. 4.



**FIG 4** Experimental procedures for OP samples and isolates. gHBA, gentamicin horse blood agar.

**DNA extraction from swab samples.** DNA was extracted from 100  $\mu$ l of STGG sample as previously described (22). Each extraction run contained approximately equal numbers of NP and OP samples and included STGG extraction controls that were checked for pneumococcal contamination using *lytA* qPCR.

**qPCR on swab samples.** The detection of *S. pneumoniae* by qPCR assays was performed on NP or OP swabs using DNA extracted as described above. Assays targeting the *lytA* and *piaB* genes were conducted as previously described (17, 37). An assay for the GntR-type transcriptional regulator gene *bguR* (38) (TIGR4 NCBI reference sequence, [NC\\_003028](#); locus tag, SP\_RS10220; old locus tag, SP\_2020), also referred to as SP2020 (19, 29), used forward primer (5'-AGTTTGCCTGTAGTCGAATGA-3'), reverse primer (5'-TTTGAGCTGCCACGAGAG-3') and the probe (5'-6-carboxyfluorescein [FAM]-AAACGTGGGCAG GGAACCTTTGTT-BHQ1-3') concentrations of 300 nM, 100 nM, and 200 nM, respectively. No-template controls (blanks) were included in each qPCR run, and runs were balanced with approximately equal numbers of NP and OP samples. Samples were considered positive when the cycle threshold value ( $C_T$ ) was <35, equivocal when the  $C_T$  was between 35 and 40, and negative when no  $C_T$  value was obtained in 40 cycles.

**Pneumococcal identification and molecular serotyping by DNA microarray.** Samples that were either positive or equivocal for the *lytA*, *piaB*, or *bguR* gene were cultured for microarray analysis. Samples were analyzed by a pneumococcal molecular serotyping microarray (Senti-SP1.4/1.5; BUGS Bioscience) as previously described (23). Microarray analysis was performed following culture amplification on horse blood agar plates containing 5  $\mu$ g/ml of gentamicin (gHBA; Oxoid brand, Thermo Fisher Scientific, Australia), and DNA extraction (for gHBA plates with  $\alpha$ -hemolytic growth only) using the QIAamp 96 DNA QIAcube HT kit (Qiagen) or the QIAamp DNA minikit (Qiagen).

**Alternative *S. pneumoniae* target genes.** We considered alternative target genes to *lytA* for the detection of *S. pneumoniae*. Fifteen targets identified by a literature search and/or used in the *S. pneumoniae* StrepID component of the microarray were initially considered. We excluded 4 targets (*rpoA*, *sodA*, *tuf*, and *recA*) because they required sequencing of the PCR product and 4 targets (*ply*, *psaA*, *spn9802*, and 16S rRNA genes) because they have been reported to lack the specificity to distinguish between closely related species or the sensitivity to detect all pneumococci. Genes *piaA*, *piaB*, and *ulaA*

were selected for further analysis as well as four additional genes used in the StrepID component of microarray (see Table S2 and S3). Targets were tested against a set of 30 reference isolates (6 *S. pneumoniae* and 24 nonpneumococcal streptococcal isolates) initially by PCR assays (see Table S4). Real-time qPCR assays for *lytA*, *piaB*, and *bguR* (described above) were also assessed against the reference isolates (see Table S5).

**Isolation of bacteria from OP samples.** Samples were cultured directly on gHBA, and a representative of each colony morphology (including nonhemolytic and  $\beta$ -hemolytic colonies) for each sample was then subcultured onto HBA and incubated overnight (37°C, 5% CO<sub>2</sub>) with an optochin disc. All OP isolates ( $n = 91$ ) were then tested for bile solubility (39). A laboratory strain of *S. pneumoniae* (PMP6, ATCC 6305) and *S. mitis* (PMP16) were used as controls for optochin sensitivity and bile solubility testing.

**DNA extraction on isolates.** A sweep of pure culture was inoculated into 100  $\mu$ l of phosphate-buffered saline using a 10- $\mu$ l sterile loop. Bacterial suspensions were pelleted at  $7,380 \times g$  for 10 min. Initially, DNA was extracted from all OP isolates using the QIAcube HT instrument (Qiagen). Prior to loading the instrument, pellets were resuspended in lysis buffer (20 mM Tris-HCl, 2 mM Na-EDTA, 1% [vol/vol] Triton X-100, 2 mg/ml of RNase A, and 20 mg/ml lysozyme) and incubated at 37°C for 60 min. Twenty microliters of 20-mg/ml proteinase K (Qiagen) and 200  $\mu$ l of buffer AL (Qiagen) were added before incubation at 56°C for a further 30 min.

However, for many of the isolates ( $n = 52$ ) there were issues during the extraction process resulting in poor-quality DNA. These isolates were extracted using the QIAamp DNA minikit (Qiagen) as described previously (40), with the following modifications: prior to the addition of the enzymatic lysis buffer, samples underwent a freeze-thaw step at  $-80^{\circ}\text{C}$  for 15 min, and DNA was eluted in 100  $\mu$ l of nuclease-free H<sub>2</sub>O (Ambion), with two elution steps of 50  $\mu$ l each.

**Real-time PCR on isolates.** Real-time PCR assays were performed on isolate DNA normalized to a concentration of 1 ng/ $\mu$ l. Real-time PCR assays targeting the *lytA*, *piaB*, and *bguR* genes were performed using the same conditions and primer and probe concentrations as the qPCR assays. Isolates were considered positive when the  $C_T$  was  $<30$ .

**MALDI-TOF MS.** All OP isolates were identified by MALDI-TOF MS using the Vitek MS (bioMérieux) system according to the manufacturer's instructions. Following an overnight culture of the isolates on HBA, approximately four colonies (i.e., a visible amount of bacterial growth) were applied to the Vitek MS-DS target slide (bioMérieux) using a sterile toothpick. Immediately after application to the slide, 1  $\mu$ l of Vitek MS-CHCA matrix solution (bioMérieux) was added to the bacteria on the slide. After allowing the matrix solution to air dry, the slide was loaded into the Vitek MS machine. Following run completion, results were analyzed using MYLA software (version 4; bioMérieux).

***S. pneumoniae* isolate identification.** An isolate was identified as *S. pneumoniae* if at least three of the following results were obtained: optochin sensitivity, bile solubility, a MALDI-TOF MS result of *S. pneumoniae*, and *lytA* real-time PCR positivity. Where only two of the four tests were indicative of *S. pneumoniae*, the StrepID component of microarray, based on detection of species-specific genes, was used to discriminate between *S. pneumoniae* and non-*S. pneumoniae*. Where none or one of the tests was indicative of *S. pneumoniae*, an isolate was considered non-*S. pneumoniae*.

**Serotyping by latex agglutination and Quellung reaction.** Latex agglutination was performed using reagents produced in-house (41) with Statens Serum Institute (SSI) antisera on a sweep of growth from the sample culture plate (latex sweep) or an isolate (latex agglutination) (39, 55). Quellung testing was performed as previously described (42) using SSI antisera on isolates for which pneumococcal serotypes were detected using latex agglutination. Reactions were visualized using a microscope at a magnification of  $\times 400$ .

**Serotyping by mPCR.** Serotyping multiplex PCR (mPCR) was performed using the method described by Carvalho et al. (43) and on the Centers for Disease Control and Prevention (CDC) *Streptococcus* Laboratory PCR Deduction of Pneumococcal Serotypes website (<https://www.cdc.gov/streplab/pcr.html>; accessed 14 July 2015) using 5 ng/ $\mu$ l of template DNA. PCR products were examined for size using an Agilent 2100 Bioanalyzer instrument (Agilent Technologies) and also run on an E-Gel 48 2% agarose gel (Invitrogen) according to the manufacturer's instructions. Results from both the E-Gel and Bioanalyzer methods were compared, and only products detected by both methods were considered true products. Products within 30 bp of an expected serotype product size were considered positive for the serotype.

**Whole-genome sequencing.** Two isolates which were not pneumococcus but yielded a pneumococcal serotype (using both latex agglutination and microarray) underwent whole-genome sequencing to identify any pneumococcal-like *cps* regions. Isolates FVEP-002-002-03 and FVEP-002-080-02 were sequenced using Illumina MiSeq V2 reagent kits ( $2 \times 150$  bp) on the MiSeq platform (Illumina). Sequence reads were down-sampled so that the assemblies would have approximately  $150\times$  coverage. Sequence assembly was performed using SPAdes version 3.5.0 (44). Assembled contigs were then submitted to the RAST server for annotation (45). Species identification was done using a combination of MetaPhlan v2.0 (46), Kraken (47), and eMLSA (48), plus using 16S sequences from the isolates with 16S SILVA Incremental Aligner (SINA) version 1.2.11 (49), Greengenes BLAST (50), and NCBI BLAST (51). Capsule genes were identified following RAST annotation and were aligned against 33F and 19B reference sequences (GenBank accession numbers [AJ006986.1](#) and [CR931676.1](#), respectively); sequences that aligned poorly were checked by NCBI BLAST. Figures for aligned sequences were generated using easyfig version 2.2.2 (52).

**Data analysis.** Data were analyzed and graphs were generated using GraphPad Prism version 7.02 and R version 3.2.4. A  $P$  value of  $<0.05$  was considered statistically significant. For differences in categorical data, Fisher's exact test was used. When evaluating qPCR assays (alone or in combination), samples found to contain pneumococci following culture amplification and DNA

microarray were considered true positives. The sensitivity, specificity, and positive predictive value (PPV) of assays were determined.

**Data availability.** The bacterial genomes for FVEP-002-002-03 and FVEP-002-080-02 analyzed during the current study are available in the NCBI Sequence Read Archive under accession number [PRJNA590329](https://www.ncbi.nlm.nih.gov/bioproject/590329) (<https://www.ncbi.nlm.nih.gov/bioproject/590329>). Other data and materials used are available upon reasonable request but may require ethical approval from the Fiji National Health Research and Ethics Review Committee and the University of Melbourne Health Sciences Human Ethics Subcommittee.

## SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

**FIG S1**, DOCX file, 0.7 MB.

**TABLE S1**, DOCX file, 0.01 MB.

**TABLE S2**, DOCX file, 0.02 MB.

**TABLE S3**, DOCX file, 0.01 MB.

**TABLE S4**, DOCX file, 0.02 MB.

**TABLE S5**, DOCX file, 0.02 MB.

**TABLE S6**, DOCX file, 0.01 MB.

**TABLE S7**, DOCX file, 0.01 MB.

**TABLE S8**, DOCX file, 0.01 MB.

**TABLE S9**, DOCX file, 0.04 MB.

## ACKNOWLEDGMENTS

We thank the study participants and their families and the staff involved in the New Vaccine Evaluation Project in Fiji. We thank scientists from the Microbiological Diagnostic Unit—Public Health Laboratory, Department of Microbiology and Immunology, The University of Melbourne at the Peter Doherty Institute for Infection and Immunity, Melbourne, for the MALDI-TOF MS used for species identification. We thank Ahmed Alamrousi for assistance with laboratory testing and Stephanie Eggers from the Translational Genomics Unit at Murdoch Children's Research Institute for the whole-genome sequencing. We also thank Fuminori Sakai from Emory University for kindly providing streptococcal isolates for the reference set.

F.M.R. and E.K.M. oversaw and F.T.R. led the field work for the New Vaccine Evaluation Project in Fiji. C.S. oversaw the current study, with J.H. and E.M.D. L.K.B., E.M.D., K.A.G., J.H., J.E.V., and C.S. conceived of and contributed to the design. L.K.B. conducted laboratory analyses with K.A.G. L.K.B. wrote the manuscript with E.M.D. and C.S. All authors read and approved the final manuscript.

Collection of the samples used in this research was funded through the New Vaccine Evaluation Project in Fiji (funded through Bill & Melinda Gates Foundation [OPP126272 and OPP1084341] and the Department of Foreign Affairs and Trade of the Australian Government and Fiji Health Sector Support Program [FHSSP]). L.K.B. was supported by a Fay Marles Scholarship, The University of Melbourne. C.S. was supported by an Australian NHMRC Career Development Fellowship (1087957) and a Veski Inspiring Women Fellowship. F.M.R. is supported by an NHMRC early career and TRIP fellowship. J.E.V. was in part supported by grants from the National Institutes of Health (NIH; R21AI112768-01A1 and 1R21AI144571-01). The Murdoch Children's Research institute was supported by the Victorian Government's Operational Infrastructure Support Program.

C.S., E.M.D., and E.K.M. are investigators on an unrelated collaborative research project funded by Pfizer. J.H. is an investigator on studies done on behalf of St George's, University of London, or BUGS Bioscience that are sponsored or funded by vaccine manufacturers, including Pfizer, GlaxoSmithKline, and Sanofi Pasteur. All other authors declare no competing interests.

## REFERENCES

- O'Brien KL, Wolfson LJ, Watt JP, Henkle E, Deloria-Knoll M, McCall N, Lee E, Mulholland K, Levine OS, Cherian T, Hib and Pneumococcal Global Burden of Disease Study Team. 2009. Burden of disease caused by *Streptococcus pneumoniae* in children younger than 5 years: global estimates. *Lancet* 374:893–902. [https://doi.org/10.1016/S0140-6736\(09\)61204-6](https://doi.org/10.1016/S0140-6736(09)61204-6).
- Drijckoningen JJC, Rohde G. 2014. Pneumococcal infection in adults:

- burden of disease. *Clin Microbiol Infect* 20(Suppl 5):45–51. <https://doi.org/10.1111/1469-0691.12461>.
3. Bogaert D, De Groot K, Hermans P. 2004. *Streptococcus pneumoniae* colonisation: the key to pneumococcal disease. *Lancet Infect Dis* 4:144–154. [https://doi.org/10.1016/S1473-3099\(04\)00938-7](https://doi.org/10.1016/S1473-3099(04)00938-7).
  4. Simell B, Auranen K, Käyhty H, Goldblatt D, Dagan R, O'Brien KL, Pneumococcal Carriage Group. 2012. The fundamental link between pneumococcal carriage and disease. *Expert Rev Vaccines* 11:841–855. <https://doi.org/10.1586/erv.12.53>.
  5. Cohen R, Cohen JF, Chalumeau M, Levy C. 2017. Impact of pneumococcal conjugate vaccines for children in high- and non-high-income countries. *Expert Rev Vaccines* 16:625–640. <https://doi.org/10.1080/14760584.2017.1320221>.
  6. Feikin DR, Kagucia EW, Loo JD, Link-Gelles R, Puhon MA, Cherian T, Levine OS, Whitney CG, O'Brien KL, Moore MR, Serotype Replacement Study Group. 2013. Serotype-specific changes in invasive pneumococcal disease after pneumococcal conjugate vaccine introduction: a pooled analysis of multiple surveillance sites. *PLoS Med* 10:e1001517. <https://doi.org/10.1371/journal.pmed.1001517>.
  7. Gladstone RA, Jefferies JM, Tocheva AS, Beard KR, Garley D, Chong WW, Bentley SD, Faust SN, Clarke SC. 2015. Five winters of pneumococcal serotype replacement in UK carriage following PCV introduction. *Vaccine* 33:2015–2021. <https://doi.org/10.1016/j.vaccine.2015.03.012>.
  8. Tsaban G, Ben-Shimol S. 2017. Indirect (herd) protection, following pneumococcal conjugated vaccines introduction: a systematic review of the literature. *Vaccine* 35:2882–2891. <https://doi.org/10.1016/j.vaccine.2017.04.032>.
  9. Davis SM, Deloria-Knoll M, Kassa HT, O'Brien KL. 2013. Impact of pneumococcal conjugate vaccines on nasopharyngeal carriage and invasive disease among unvaccinated people: review of evidence on indirect effects. *Vaccine* 32:133–145. <https://doi.org/10.1016/j.vaccine.2013.05.005>.
  10. Chan J, Nguyen CD, Dunne EM, Kim Mulholland E, Mungun T, Pomat WS, Rafai E, Satzke C, Weinberger DM, Russell FM. 2019. Using pneumococcal carriage studies to monitor vaccine impact in low- and middle-income countries. *Vaccine* 37:6299–6309. <https://doi.org/10.1016/j.vaccine.2019.08.073>.
  11. Satzke C, Turner P, Virolainen-Julkunen A, Adrian PV, Antonio M, Hare KM, Henao-Restrepo AM, Leach AJ, Klugman KP, Porter BD, Sá-Leão R, Scott JA, Nohynek H, O'Brien KL, WHO Pneumococcal Carriage Working Group. 2013. Standard method for detecting upper respiratory carriage of *Streptococcus pneumoniae*: updated recommendations from the World Health Organization Pneumococcal Carriage Working Group. *Vaccine* 32:165–179. <https://doi.org/10.1016/j.vaccine.2013.08.062>.
  12. O'Brien KL, Nohynek H, WHO Pneumococcal Vaccine Trials Carriage Working Group. 2003. Report from a WHO working group: standard method for detecting upper respiratory carriage of *Streptococcus pneumoniae*. *Pediatr Infect Dis J* 22:e1–e11. <https://doi.org/10.1097/01.inf.0000049347.42983.77>.
  13. Greenberg D, Broides A, Blancovich I, Peled N, Givon-Lavi N, Dagan R. 2004. Relative importance of nasopharyngeal versus oropharyngeal sampling for isolation of *Streptococcus pneumoniae* and *Haemophilus influenzae* from healthy and sick individuals varies with age. *J Clin Microbiol* 42:4604–4609. <https://doi.org/10.1128/JCM.42.10.4604-4609.2004>.
  14. Watt JP, O'Brien KL, Katz S, Bronsdon MA, Elliott J, Dallas J, Perilla MJ, Reid R, Murrow L, Facklam R, Santosham M, Whitney CG. 2004. Nasopharyngeal versus oropharyngeal sampling for detection of pneumococcal carriage in adults. *J Clin Microbiol* 42:4974–4976. <https://doi.org/10.1128/JCM.42.11.4974-4976.2004>.
  15. Lieberman D, Shleyfer E, Castel H, Terry A, Harman-Boehm I, Delgado J, Peled N, Lieberman D. 2006. Nasopharyngeal versus oropharyngeal sampling for isolation of potential respiratory pathogens in adults. *J Clin Microbiol* 44:525–528. <https://doi.org/10.1128/JCM.44.2.525-528.2006>.
  16. Levine H, Zarka S, Dagan R, Sela T, Rozhavski V, Cohen DI, Balicer RD. 2012. Transmission of *Streptococcus pneumoniae* in adults may occur through saliva. *Epidemiol Infect* 140:561–565. <https://doi.org/10.1017/S0950268811000884>.
  17. Trzciński K, Bogaert D, Wyllie A, Chu M, van der Ende A, Bruin JP, van den Dobbelen G, Veenhoven RH, Sanders E. 2013. Superiority of trans-oral over trans-nasal sampling in detecting *Streptococcus pneumoniae* colonization in adults. *PLoS One* 8:e60520. <https://doi.org/10.1371/journal.pone.0060520>.
  18. Principi N, Terranova L, Zampiero A, Manzoni F, Senatore L, Rios WP, Esposito S. 2014. Oropharyngeal and nasopharyngeal sampling for the detection of adolescent *Streptococcus pneumoniae* carriers. *J Med Microbiol* 63:393–398. <https://doi.org/10.1099/jmm.0.068726-0>.
  19. Argedas A, Trzciński K, O'Brien KL, Ferreira DM, Wyllie AL, Weinberger D, Danon L, Pelton SI, Azzari C, Hammit LL, Sá-Leão R, Brandileone M-C, Saha S, Suaya J, Isturiz R, Jodar L, Gessner BD. 2020. Upper respiratory tract colonization with *Streptococcus pneumoniae* in adults. *Expert Rev Vaccines* 19:353–366. <https://doi.org/10.1080/14760584.2020.1750378>.
  20. Carvalho MDG, Pimenta FC, Moura I, Roundtree A, Gertz RE, Li Z, Jagero G, Bigogo G, Junghae M, Conklin L, Feikin DR, Breiman RF, Whitney CG, Beall BW. 2013. Non-pneumococcal mitis-group streptococci confound detection of pneumococcal capsular serotype-specific loci in upper respiratory tract. *PeerJ* 1:e97. <https://doi.org/10.7717/peerj.97>.
  21. Carvalho MDG, Bigogo GM, Junghae M, Pimenta FC, Moura I, Roundtree A, Li Z, Conklin L, Feikin DR, Breiman RF, Whitney CG, Beall B. 2012. Potential nonpneumococcal confounding of PCR-based determination of serotype in carriage. *J Clin Microbiol* 50:3146–3147. <https://doi.org/10.1128/JCM.01505-12>.
  22. Dunne EM, Satzke C, Ratu FT, Neal EFG, Boelsen LK, Matanitobua S, Pell CL, Nation ML, Ortika BD, Reyburn R, Jenkins K, Nguyen C, Gould K, Hinds J, Tikoduadua L, Kado J, Rafai E, Kama M, Mulholland EK, Russell FM. 2018. Effect of ten-valent pneumococcal conjugate vaccine introduction on pneumococcal carriage in Fiji: results from four annual cross-sectional carriage surveys. *Lancet Glob Health* 6:e1375–e1385. [https://doi.org/10.1016/S2214-109X\(18\)30383-8](https://doi.org/10.1016/S2214-109X(18)30383-8).
  23. Satzke C, Dunne EM, Porter BD, Klugman KP, Mulholland EK, PneuCarriage project group. 2015. The PneuCarriage project: a multi-centre comparative study to identify the best serotyping methods for examining pneumococcal carriage in vaccine evaluation studies. *PLoS Med* 12:e1001903. <https://doi.org/10.1371/journal.pmed.1001903>.
  24. Lull D, López R, García E. 2006. Characteristic signatures of the *lytA* gene provide a basis for rapid and reliable diagnosis of *Streptococcus pneumoniae* infections. *J Clin Microbiol* 44:1250–1256. <https://doi.org/10.1128/JCM.44.4.1250-1256.2006>.
  25. Simões AS, Tavares DA, Rolo D, Ardanuy C, Goossens H, Henriques-Normark B, Linares J, de Lencastre H, Sá-Leão R. 2016. *lytA*-based identification methods can misidentify *Streptococcus pneumoniae*. *Diagn Microbiol Infect Dis* 85:141–148. <https://doi.org/10.1016/j.diagmicrobio.2016.03.018>.
  26. 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, Massignani V. 2010. Structure and dynamics of the pan-genome of *Streptococcus pneumoniae* and closely related species. *Genome Biol* 11:R107. <https://doi.org/10.1186/gb-2010-11-10-r107>.
  27. Stearns JC, Davidson CJ, McKeon S, Whelan FJ, Fontes ME, Schryvers AB, Bowdish DME, Kellner JD, Surette MG. 2015. Culture and molecular-based profiles show shifts in bacterial communities of the upper respiratory tract that occur with age. *ISME J* 9:1246–1259. <https://doi.org/10.1038/ismej.2014.250>.
  28. Jensen A, Scholz CFP, Kilian M. 2016. Re-evaluation of the taxonomy of the Mitis group of the genus *Streptococcus* based on whole genome phylogenetic analyses, and proposed reclassification of *Streptococcus dentisani* as *Streptococcus oralis* subsp. *dentisani* comb. nov., *Streptococcus tigurinus* as *Streptococcus oralis* subsp. *tigurinus* comb. nov., and *Streptococcus oligofermentans* as a later synonym of *Streptococcus cristatus*. *Int J Syst Evol Microbiol* 66:4803–4820. <https://doi.org/10.1099/ijsem.0.001433>.
  29. Tavares DA, Handem S, Carvalho RJ, Paulo AC, de Lencastre H, Hinds J, Sá-Leão R. 2019. Identification of *Streptococcus pneumoniae* by a real-time PCR assay targeting SP2020. *Sci Rep* 9:3285. <https://doi.org/10.1038/s41598-019-39791-1>.
  30. Skov Sørensen UB, Yao K, Yang Y, Tettelin H, Kilian M. 2016. Capsular polysaccharide expression in commensal *Streptococcus* species: genetic and antigenic similarities to *Streptococcus pneumoniae*. *mBio* 7:e01844-16. <https://doi.org/10.1128/mBio.01844-16>.
  31. Sheppard CL, Kapatai G, Broughton K, Schaefer U, Hannah M, Litt DJ, Fry NK. 2017. Clinical streptococcal isolates, distinct from *Streptococcus pneumoniae*, but containing the  $\beta$ -glucosyltransferase *tts* gene and expressing serotype 37 capsular polysaccharide. *PeerJ* 5:e3571. <https://doi.org/10.7717/peerj.3571>.
  32. Pimenta F, Gertz RE, Park SH, Kim E, Moura I, Milucky J, Roupael N, Farley MM, Harrison LH, Bennett NM, Bigogo G, Feikin DR, Breiman R, Lessa FC, Whitney CG, Rajam G, Schiffer J, da Gloria Carvalho M, Beall B. 2018. *Streptococcus infantis*, *Streptococcus mitis*, and *Streptococcus oralis*

- strains with highly similar *cps5* loci and antigenic relatedness to serotype 5 pneumococci. *Front Microbiol* 9:3199. <https://doi.org/10.3389/fmicb.2018.03199>.
33. Kilian M, Poulsen K, Blomqvist T, Håvarstein LS, Bek-Thomsen M, Tettelin H, Sørensen U. 2008. Evolution of *Streptococcus pneumoniae* and its close commensal relatives. *PLoS One* 3:e2683. <https://doi.org/10.1371/journal.pone.0002683>.
  34. Kilian M, Riley DR, Jensen A, Brüggemann H, Tettelin H. 2014. Parallel evolution of *Streptococcus pneumoniae* and *Streptococcus mitis* to pathogenic and mutualistic lifestyles. *mBio* 5:e01490-14. <https://doi.org/10.1128/mBio.01490-14>.
  35. Lessa FC, Milucky J, Roupael NG, Bennett NM, Talbot HK, Harrison LH, Farley MM, Walston J, Pimenta F, Gertz RE, Rajam G, Carvalho MDG, Beall B, Whitney CG. 2018. *Streptococcus mitis* expressing pneumococcal serotype 1 capsule. *Sci Rep* 8:17959. <https://doi.org/10.1038/s41598-018-35921-3>.
  36. Johnston C, Hinds J, Smith A, Van Der Linden M, Van Eldere J, Mitchell TJ. 2010. Detection of large numbers of pneumococcal virulence genes in streptococci of the mitis group. *J Clin Microbiol* 48:2762–2769. <https://doi.org/10.1128/JCM.01746-09>.
  37. Carvalho MDG, Tondella ML, McCaustland K, Weidlich L, McGee L, Mayer LW, Steigerwalt A, Whaley M, Facklam RR, Fields B, Carlone G, Ades EW, Dagan R, Sampson JS. 2007. Evaluation and improvement of real-time PCR assays targeting *lytA*, *ply*, and *psaA* genes for detection of pneumococcal DNA. *J Clin Microbiol* 45:2460–2466. <https://doi.org/10.1128/JCM.02498-06>.
  38. Shafeeq S, Kuipers OP, Kloosterman TG. 2013. Cellobiose-mediated gene expression in *Streptococcus pneumoniae*: a repressor function of the novel GntR-type regulator BguR. *PLoS One* 8:e57586. <https://doi.org/10.1371/journal.pone.0057586>.
  39. Castillo D, Harcourt B, Hatcher C, Jackson M, Katz L, Mair R, Mayer L, Novak R, Rahalison L, Schmink S, Theodore MJ, Thomas J, Vuong J, Wang X, McGee L. 2011. Laboratory methods for the diagnosis of meningitis caused by *Neisseria meningitidis*, *Streptococcus pneumoniae*, and *Haemophilus influenzae*, 2nd ed. World Health Organization, Geneva, Switzerland.
  40. Dunne EM, Manning J, Russell FM, Robins-Browne RM, Mulholland EK, Satzke C. 2012. Effect of pneumococcal vaccination on nasopharyngeal carriage of *Streptococcus pneumoniae*, *Haemophilus influenzae*, *Moraxella catarrhalis*, and *Staphylococcus aureus* in Fijian children. *J Clin Microbiol* 50:1034–1038. <https://doi.org/10.1128/JCM.06589-11>.
  41. Ortika BD, Habib M, Dunne EM, Porter BD, Satzke C. 2013. Production of latex agglutination reagents for pneumococcal serotyping. *BMC Res Notes* 6:49. <https://doi.org/10.1186/1756-0500-6-49>.
  42. Habib M, Porter BD, Satzke C. 2014. Capsular serotyping of *Streptococcus pneumoniae* using the Quellung reaction. *J Vis Exp* 84:e51208.
  43. Carvalho MDG, Pimenta FC, Jackson D, Roundtree A, Ahmad Y, Millar EV, O'Brien KL, Whitney CG, Cohen AL, Beall BW. 2010. Revisiting pneumococcal carriage by use of broth enrichment and PCR techniques for enhanced detection of carriage and serotypes. *J Clin Microbiol* 48:1611–1618. <https://doi.org/10.1128/JCM.02243-09>.
  44. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Pribelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 19:455–477. <https://doi.org/10.1089/cmb.2012.0021>.
  45. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formisano K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O. 2008. The RAST server: rapid annotations using subsystems technology. *BMC Genomics* 9:75. <https://doi.org/10.1186/1471-2164-9-75>.
  46. Segata N, Waldron L, Ballarini A, Narasimhan V, Jousson O, Huttenhower C. 2012. Metagenomic microbial community profiling using unique clade-specific marker genes. *Nat Methods* 9:811–814. <https://doi.org/10.1038/nmeth.2066>.
  47. Wood DE, Salzberg SL. 2014. Kraken: ultrafast metagenomic sequence classification using exact alignments. *Genome Biol* 15:R46. <https://doi.org/10.1186/gb-2014-15-3-r46>.
  48. Bishop CJ, Aanensen DM, Jordan GE, Kilian M, Hanage WP, Spratt BG. 2009. Assigning strains to bacterial species via the internet. *BMC Biol* 7:3. <https://doi.org/10.1186/1741-7007-7-3>.
  49. Pruesse E, Peplies J, Glöckner FO. 2012. SINA: accurate high-throughput multiple sequence alignment of ribosomal RNA genes. *Bioinformatics* 28:1823–1829. <https://doi.org/10.1093/bioinformatics/bts252>.
  50. DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K, Huber T, Dalevi D, Hu P, Andersen GL. 2006. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl Environ Microbiol* 72:5069–5072. <https://doi.org/10.1128/AEM.03006-05>.
  51. Johnson M, Zaretskaya I, Raytselis Y, Merezhuk Y, McGinnis S, Madden TL. 2008. NCBI BLAST: a better web interface. *Nucleic Acids Res* 36:W5–W9. <https://doi.org/10.1093/nar/gkn201>.
  52. Sullivan MJ, Petty NK, Beatson SA. 2011. Easyfig: a genome comparison visualizer. *Bioinformatics* 27:1009–1010. <https://doi.org/10.1093/bioinformatics/btr039>.
  53. Bentley SD, Aanensen DM, Mavroidi A, Saunders D, Rabinowitz E, Collins M, Donohoe K, Harris D, Murphy L, Quail MA, Samuel G, Skovsted IC, Kalltoft MS, Barrell B, Reeves PR, Parkhill J, Spratt BG. 2006. Genetic analysis of the capsular biosynthetic locus from all 90 pneumococcal serotypes. *PLoS Genet* 2:e31. <https://doi.org/10.1371/journal.pgen.0020031>.
  54. Satzke C, Dunne EM, Choummanivong M, Ortika BD, Neal EFG, Pell CL, Nation ML, Fox KK, Nguyen CD, Gould KA, Hinds J, Chanthongthip A, Xeuatvongsa A, Mulholland EK, Sychareun V, Russell FM. 2019. Pneumococcal carriage in vaccine-eligible children and unvaccinated infants in Lao PDR two years following the introduction of the 13-valent pneumococcal conjugate vaccine. *Vaccine* 37:296–305. <https://doi.org/10.1016/j.vaccine.2018.10.077>.
  55. Porter BD, Ortika BD, Satzke C. 2014. Capsular serotyping of *Streptococcus pneumoniae* by latex agglutination. *J Vis Exp* 91:e51747.