

## ARTICLE OPEN



# Pneumococcal capsule blocks protection by immunization with conserved surface proteins

Tonia Zangari<sup>1,4</sup>, M. Ammar Zafar<sup>1,2,4</sup>, John A. Lees<sup>1,3</sup>, Annie R. Abruzzo<sup>1</sup>, Gavyn Chern Wei Bee<sup>1</sup> and Jeffrey N. Weiser<sup>1</sup>

Vaccines targeting *Streptococcus pneumoniae* (Spn) are limited by dependence on capsular polysaccharide and its serotype diversity. More broadly-based approaches using common protein antigens have not resulted in a licensed vaccine. Herein, we used an unbiased, genome-wide approach to find novel vaccine antigens to disrupt carriage modeled in mice. A Tn-Seq screen identified 198 genes required for colonization of which 16 are known to express conserved, immunogenic surface proteins. After testing defined mutants for impaired colonization of infant and adult mice, 5 validated candidates (StkP, PenA/Pbp2a, PgdA, HtrA, and LytD/Pce/CbpE) were used as immunogens. Despite induction of antibody recognizing the Spn cell surface, there was no protection against Spn colonization. There was, however, protection against an unencapsulated Spn mutant. This result correlated with increased antibody binding to the bacterial surface in the absence of capsule. Our findings demonstrate how the pneumococcal capsule interferes with mucosal protection by antibody to common protein targets.

npj Vaccines (2021)6:155; <https://doi.org/10.1038/s41541-021-00413-5>

## INTRODUCTION

*Streptococcus pneumoniae* (Spn, the pneumococcus) is a common member of the human upper respiratory tract (URT) flora, found in 25–65% of children and <7% of adults<sup>1</sup>. Access by the organism to normally sterile sites can result in inflammatory diseases such as otitis media, community-acquired pneumonia, and bacterial meningitis. Because of the high burden of disease and increasing resistance to many common antimicrobial agents, the World Health Organization included Spn as one of its 12 priority pathogens.

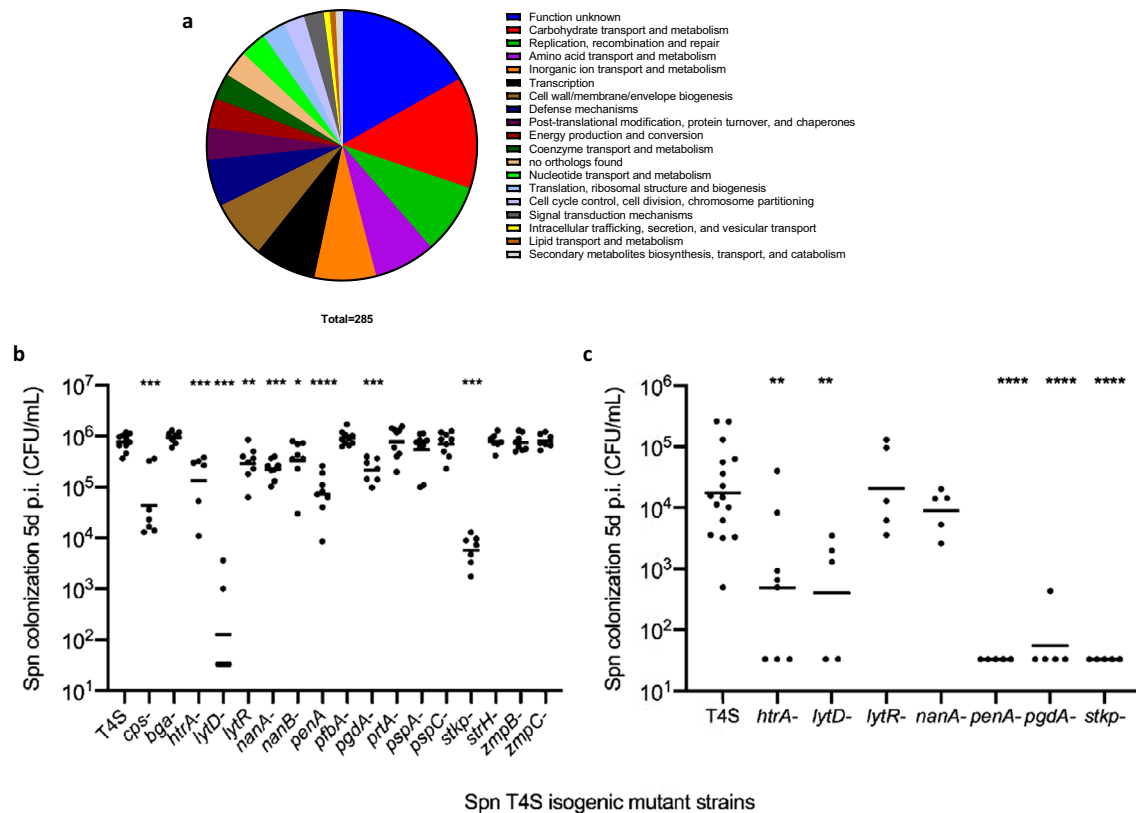
Capsular polysaccharide-based conjugate vaccines (PCV) have reduced the incidence of invasive Spn disease for included serotypes<sup>2</sup>. An unexpected benefit of PCV is decreased acquisition of carriage among immunized children that has lowered overall rates of spread within the community<sup>3</sup>. In the U.S., these “herd immune” effects of PCV have been particularly impactful for pneumonia in the elderly<sup>4</sup>. Protection by PCV, however, is incomplete, as currently licensed formulations contain only 10 to 20 of the >98 pneumococcal capsular polysaccharide (CPS) serotypes. The 2013 recommendation of PCV for adults has not further decreased the burden of Spn pneumonia, which is more likely to be due to non-PCV serotypes. It has proven immunologically challenging to further increase the number of included CPS serotypes in PCV. Other issues are the high cost of this complex vaccine and rising rates of non-vaccine serotypes (“serotype replacement”) in both carriage and, in studies in some locations, disease<sup>5,6</sup>. Several phase I and II clinical trials of protein-based or whole-cell vaccines have failed to progress, emphasizing the need for new approaches. Since it is estimated that >80% of the public health benefit of PCV is attributable to herd immunity<sup>7</sup>, for any novel strategy to be as effective as PCV its effects would also need to be amplified by disrupting transmission within the population. The only demonstrated means of interrupting pneumococcal community transmission is through a reduction in carriage. While protein-based products have been licensed for another

encapsulated pathogen, *Neisseria meningitidis*, it is still unclear how effectively this vaccine protects against carriage<sup>8</sup>. Additionally, because of the widespread use and effectiveness of PCV, a reduction in carriage is likely the only practical, tractable outcome for clinical efficacy studies leading to licensure of novel Spn vaccines<sup>9</sup>.

Spn carriage in the upper respiratory tract (URT), the first step in pneumococcal disease<sup>9</sup>, is characterized by sequential and overlapping episodes, each lasting for weeks to months. Spn carriage rates decline after early childhood for both common and uncommon serotypes<sup>10</sup>. This broad, serotype-independent decline in prevalence with host age is best explained by the gradual accumulation of immunity to conserved Spn antigens. Previous research, using murine models of URT colonization, provide evidence of the development of serotype-independent protection following pneumococcal colonization<sup>11–14</sup>. In this regard, Spn contains highly conserved structural features (such as teichoic acids) and multiple enzymes required to build and remodel its cell surface. Under routine growth conditions in vitro or during invasive infection, the thick layer of CPS effectively covers up many of these antigens. It has been suggested, however, that during carriage there is a down-regulation of CPS, which otherwise blocks adherence and interaction of these bacterial factors with the epithelial surface<sup>15</sup>.

Anti-capsular antibody (Ab) is most effective at impacting colonization when present the time of acquisition, perhaps to promote mucociliary clearance before the organism can access the epithelial surface or because of the small numbers of incoming bacteria at this time<sup>16</sup>. Targeting genes required during colonization would minimize the possibility that these genes could be lost or mutate in response to vaccine-induced immune pressure. A further consideration is that Spn, like other successful mucosal pathogens of the URT, expresses a protease that cleaves human IgA1, the most abundant immunoglobulin in this niche<sup>17</sup>. Thus, the generation of protease-insensitive IgG might be critical for

<sup>1</sup>Department of Microbiology, New York University Grossman School of Medicine, New York, NY, USA. <sup>2</sup>Present address: Department of Microbiology and Immunology, Wake Forest School of Medicine, Winston-Salem, NC, USA. <sup>3</sup>Present address: Department of Infectious Disease Epidemiology, Medical Research Council Centre for Global Infectious Disease Analysis, Imperial College London, London, UK. <sup>4</sup>These authors contributed equally: Tonia Zangari, M. Ammar Zafar. <sup>✉</sup>email: [Jeffrey.weiser@nyulangone.org](mailto:Jeffrey.weiser@nyulangone.org)



**Fig. 1** Colonization by TnSeq mutants and analysis of candidate genes. **a** Functional annotation of genes identified in the TnSeq screen with diminished infant mouse colonization ( $P < 0.05$ ). **b** Defined mutants in the 16 genes identified in the TnSeq screen expressing antibody-binding targets were constructed and tested for their ability to colonize infant mice in comparison to the parental (T4S) and unencapsulated (T4cps-) strain controls. **c** Mutants with diminished colonization of infant mice ( $P < 0.01$ ) were also tested in adult mice. Symbols represent colonization density (CFU/ml upper respiratory track lavage) of individual mice at 5 days post-infection with median values shown. Statistical comparisons using the Mann–Whitney test for each mutant were to T4S strain. \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ , \*\*\*\* $P < 0.0001$ .

blocking the establishment of carriage. PCV induces high levels of IgG that access the mucosal surface of the URT by transcytosis<sup>18,19</sup>. Studies in mice and experimental human colonization experiments demonstrate that mucosal IgG to CPS blocks carriage acquisition through its agglutinating function rather than through opsonization<sup>16,20</sup>. This observation established different mechanisms for Ab-mediated protection from carriage and invasive disease<sup>21,22</sup>. Another consideration is that antibody to Spn surface enzymes could interfere with their function and thereby inhibit bacterial growth and viability.

Given these competing immunological factors in both the host and pathogen, empirical studies of the effect of protein vaccine targets in disease models are needed to inform future iterations of the pneumococcal vaccine. We, therefore, tested the hypothesis that specific antibody raised to conserved, immunogenic surface proteins could block Spn colonization. We took an unbiased, whole genomic approach to screen for genes required for URT colonization, as a first step in identification of candidate antigens, which we then validated in our murine infection model.

## RESULTS AND DISCUSSION

### TnSeq screen to identify loci contributing to colonization

To assess genes impacting Spn colonization using a TnSeq screen, we employed an infant mouse model of pneumococcal colonization<sup>23</sup>, thereby reducing the problem of population bottlenecks (stochastic loss of clones) which confounds whole-genome mutagenesis studies<sup>24</sup>. In contrast to adults, infant mice become stably colonized following a much lower inoculum (50% infecting

dose  $< 5$  CFU)<sup>25</sup>. Camilli et al. described the use of TnSeq technology in Spn, and we previously described the use of this method in an analysis of genes affecting Spn shedding in our infant mouse model<sup>26–28</sup>. In the current report, we include the analysis of this dataset for genes affecting colonization. Briefly, *Mariner* inserts in TA sites in genomic DNA were generated in vitro and transformed into Spn T4S (a streptomycin-resistant derivative of the extensively-annotated type 4 clinical isolate TIGR4). Twenty-eight independently-generated input pools, containing ~500 unique insertions, were each inoculated into two (biological replicates) 4-day-old pups; URT lavage samples were collected 5 days post-inoculation. The 28 “input” pools were then compared with the corresponding “output” pool. Genomic DNA was isolated from each pool and used to amplify the transposon and insertion site for Illumina sequencing followed by bioinformatic analysis for statistically significant loss of clones in the output pool. In total, the input pool included  $> 16,800$  mutations in non-essential genes (2020 of 2248 TIGR4 open reading frames). We identified a significant loss of fitness for colonization ( $P$  value cutoff of  $< 0.02$ ) in 198 non-essential genes (TnSeq “hits”; Table S1). [Note: *nanA* ( $P = 0.028$ ) was included in the list because of its potentially overlapping function with *nanB*, making our list 199 candidates<sup>29</sup>]. These “hits” included genes across multiple functional classes (Fig. 1a and Table S1 shown for  $P < 0.05$ ,  $n = 285$ ).

### Prioritization and confirmation of candidate loci

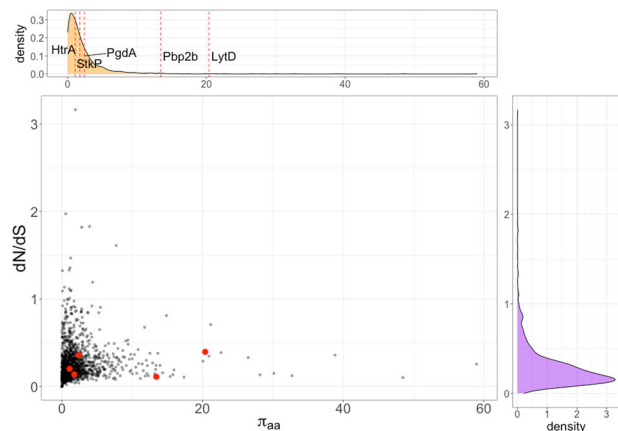
Next, we prioritized immunogenic targets by cross-referencing to a list of 208 binding targets identified from a proteome microarray<sup>30</sup>. These Ab-binding targets (ABTs), which were

**Table 1.** Antibody-binding targets among the TnSeq-identified genes affecting carriage.

Gene ID	Gene	P-value	Function
SP_0071	<i>zmpC</i>	4.66E-06	Zinc metalloprotease
SP_1732	<i>stkp</i>	1.18E-05	Serine/Threonine protein kinase
SP_0930	<i>cbpE</i>	0.0001	Choline binding protein E, Also LytD
SP_1833	<i>pfbA</i>	0.0001	Plasmin and fibronectin-binding protein A; cell wall surface anchor family protein
SP_0648	<i>bgaA</i>	0.0001	Beta-galactosidase
SP_1942	<i>lytR</i>	0.0002	Putative transcriptional regulator
SP_1673	<i>penA</i>	0.0046	Penicillin-binding protein; pbp2B
SP_0664	<i>zmpB</i>	0.0053	Zinc metalloprotease
SP_0057	<i>strH</i>	0.0065	Beta-N-acetylhexosaminidase, exoglycosidase
SP_0117	<i>pspA</i>	0.0067	Pneumococcal surface protein A
SP_2239	<i>htrA</i>	0.007	Serine protease
SP_2190	<i>pspC/cbpA</i>	0.0077	Choline binding protein A
SP_1479	<i>pgdA</i>	0.0086	Peptidoglycan N-acetylglucosamine deacetylase A
SP_1687	<i>nanB</i>	0.011	Sialidase; neuraminidase B
SP_0641	<i>prtA</i>	0.0122	Serine protease, subtilase family
SP_1693	<i>nanA</i>	0.028	Sialidase; neuraminidase A

identified using naturally-acquired serum IgG, were enriched for cell-surface association. No essential gene products were found among the ABTs. Sixteen of our 199 TnSeq ‘hits’ were in ABTs (Table 1). We then constructed defined, knock-out mutants of these candidate genes in T4S and tested these strains for their ability to colonize infant mice (Fig. 1b). Colonization density at 5 days post-inoculation was compared to the T4S parent strain and with a previously described capsule-deficient mutant (T4Δ*cps*) included as a positive control for colonization deficiency<sup>31</sup>. Seven of the 16 candidates showed a significant defect in colonization of infant mice compared to the parent strain ( $P < 0.01$ ). We then assessed colonization of the validated candidates in adult mice as, optimally, a novel Spn vaccine should protect both infants and adults. Colonization of adult mice occurs at a lower density and with increased animal-to-animal variation (Fig. 1c). Of the seven mutant strains attenuated for infant colonization ( $P < 0.01$ ), five were also impaired in colonization of adults ( $P < 0.01$ ). Interestingly, the remaining candidates did not include several well-recognized Spn virulence determinants (PspA, PspC/CpbA, NanA) identified in the TnSeq screen.

Each of the remaining five candidate genes are part of the Spn core genome (which consists of 1194 genes found throughout the species)<sup>32</sup>. For each core gene, we produced multiple sequence alignments and quantified selection by calculating the ratio of non-synonymous to synonymous changes (dN/dS) using a maximum-likelihood method (Fig. 2). All of the selected antigens were under negative selection, typical of other functionally conserved bacterial core genes. This suggests rapid evolution away from protein-targeted vaccines via point mutations would be unlikely in any target. We also quantified allelic diversity in the population by calculating  $\pi$  between amino acids<sup>33</sup>, which can be interpreted as the average number of coding changes between a pair of sequences. When a gene has multiple common alleles,  $\pi_{aa}$  will be larger due to their sequence differences (though dN/dS may still be low, as the function of either allele is still conserved). In this analysis, Pbp2b and LytD showed above-average population wide diversity. Pbp2b is a recombination hotspot due to its involvement in penicillin resistance, leading to multiple mosaic

**Fig. 2** Selection and diversity in core pneumococcal genes and the selected antigens. Top panel: marginal density distribution of  $\pi_{aa}$  (a measure of the average number of coding changes between a pair of sequences from different isolates) showing the position of the selected antigens with red dashed lines, as labeled. Right panel: marginal density distribution of dN/dS (ratio of non-synonymous to synonymous changes). Central panel: scatter plot of average pairwise sequence diversity  $\pi_{aa}$  vs selection strength for all core genes (black) and tested antigens (red).

alleles<sup>34</sup>. Thus, all candidate immunogenic proteins required for efficient colonization are highly conserved among Spn isolates with PenA and LytD displaying multiple functional conserved alleles.

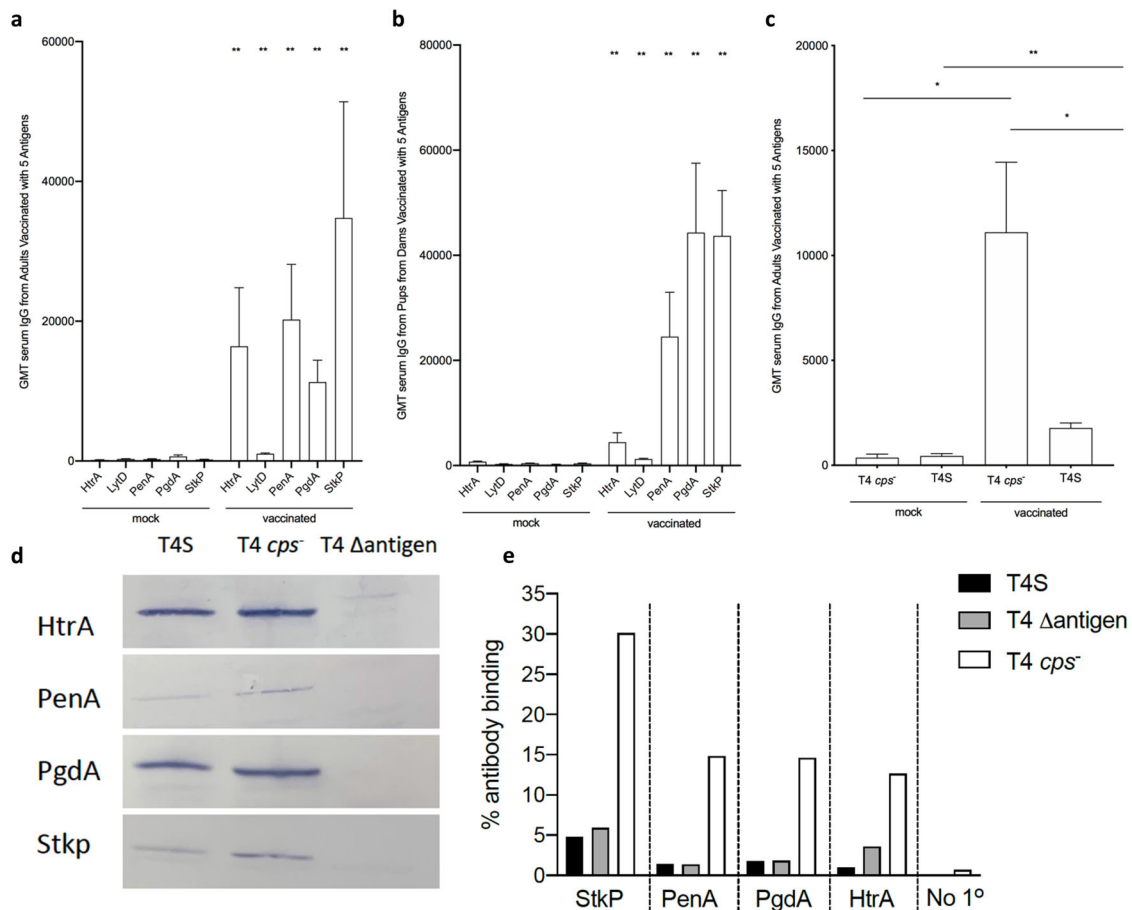
### Description of protein candidates

All five Spn protein candidates have well-characterized functions on the bacterial cell surface as follows:

StkP is a Ser/Thr protein kinase that functions in the cell division process as a global cell cycle regulator<sup>35</sup>. Four extracellular PASTA domains interact with peptidoglycan fragments at the division septum and activate penicillin-binding protein 2a through StkP-mediated phosphorylation of MacP<sup>36,37</sup>. StkP was also identified in a prior screen of the Spn proteome for increased reactivity with patient convalescent sera following invasive pneumococcal infection<sup>38</sup>. Specific Ab to recombinantly-expressed StkP showed surface staining of encapsulated TIGR4 by FACS analysis. Immunization of mice with the C-terminal portion of StkP (the 314 extracellular AA residues in alum adjuvant) was as effective as PCV in protecting mice against development of Spn pneumonia. However, following a phase I clinical study, because there was no increased killing in an opsonophagocytic assay with sera from StkP immunized individuals, further testing of this antigen was discontinued without consideration of other mechanisms of Ab-mediated protection. In a recent longitudinal study of 976 mother-infant pairs, higher levels of cord blood IgG to StkP were associated with reduced Spn acquisition among infants<sup>39</sup>.

PenA is a monofunctional Class B penicillin-binding protein (Pbp2b) with transpeptidase activity needed for normal cell elongation. AA substitutions in PenA extracellular catalytic domains are associated with low level  $\beta$ -lactam resistance due to decreased drug binding affinity<sup>40</sup>. The selection for  $\beta$ -lactam resistance could explain the higher level of sequence variation for this candidate. A recent report showed that immunization with recombinant PenA conferred significant protection against lethal challenge from diverse Spn strains following intranasal (IN) challenge<sup>41</sup>.

PgdA is an N-acetylglucosamine deacetylase that modifies the pneumococcal peptidoglycan backbone<sup>42</sup>. Deacetylation of >80% of GlcNAc residues by PgdA renders the organism resistant to the hydrolytic activity of host lysozyme, an abundant antimicrobial on



**Fig. 3 Effect of immunization on antibody responses.** Dams were vaccinated with the five-antigen combination (StkP, PenA, PgdA, HtrA, and LytD) with adjuvant and compared to adjuvant alone controls (mock). Serum from **a** adults ( $n = 6$ ) and **b** pups from vaccinated dams ( $n = 5$ ) was used to quantify serum IgG responses to individual antigens by ELISA. Statistical significance of ELISA values was determined using a Mann–Whitney test. **c** Adult sera were also used in a whole bacterial cell ELISA using the unencapsulated mutant (T4*cps*<sup>-</sup>) or its encapsulated parent strain (T4S). Values represent geometric mean titers  $\pm$  S.E.M. ELISA values were compared using an unpaired Student's *t* test. \* $P < 0.05$ , \*\* $P < 0.01$ . **d** Pooled sera ( $n = 5$ ) from adult mice immunized with the individual antigen shown was used in Western analysis on whole-cell lysates of strain T4S, the unencapsulated mutant T4*cps*<sup>-</sup> or the mutant construct corresponding to the sera tested. **e** Representative flow cytometric comparison of antibody binding to the surface of strain T4S, the unencapsulated mutant T4*cps*<sup>-</sup> or the mutant construct corresponding to the sera tested. Samples tested were pooled sera ( $n = 5$ ) from adult mice immunized with the individual antigen shown below or a no sera control.

the mucosal surface of the URT. We have shown that without *pgdA*, Spn grows normally *in vitro* but is unfit for colonization in a lysozyme-dependent manner<sup>43</sup>. The relatively greater contribution of *pgdA* to colonization in adult compared to infant mice (Fig. 1) correlates with higher expression of lysozyme in adult mice<sup>44</sup>.

HtrA is a serine protease that plays a crucial role in bacterial growth under stress conditions and protein quality control on the bacterial surface<sup>45,46</sup>. For instance, HtrA controls competence and bacteriocin activity by regulating levels of quorum-sensing peptide pheromones that control these functions<sup>47</sup>.

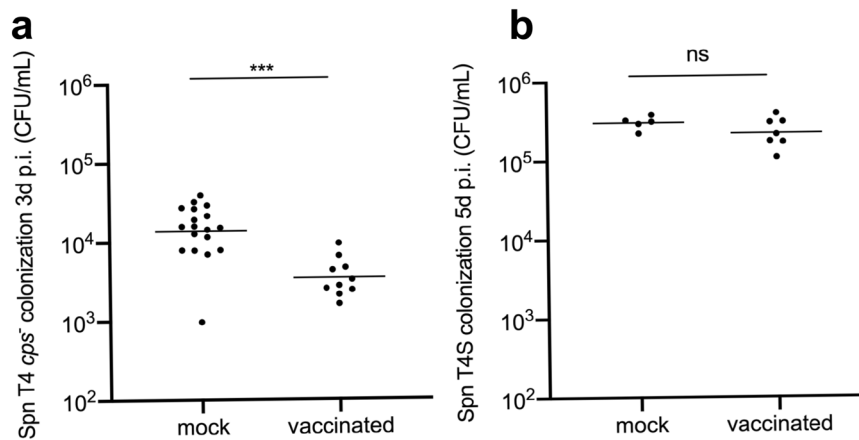
LytD, also known as CbpE or Pce, is a phosphorylcholine esterase that binds to and hydrolyses phosphorylcholine residues on wall- and lipo-teichoic acid protruding from the cell surface<sup>48</sup>. Our group has shown that this enzyme also cleaves and inactivates host platelet activating factor (PAF), a phosphorylcholine-containing substrate<sup>49</sup>. PAF is a potent activator of neutrophils and its depletion by LytD impairs their ability to clear Spn.

#### Immunization of adult mice with candidate proteins

The five protein candidates (StkP, PenA, PgdA, HtrA, and LytD) were then tested for their ability to induce immunity that protects

against respiratory tract infection. Briefly, the full-length gene (except as noted below) without the signal peptide (based on the TIGR4 reference genome sequence)<sup>50</sup> was cloned and expressed with N- or C-terminal His tags. StkP contains transmembrane domains and only the C-terminal extracellular domain (AA 345–659) was expressed. Once the purity of recombinant protein candidates was confirmed, these were used to subcutaneously immunize adult mice, with mice receiving one candidate protein plus TiterMax Gold<sup>®</sup> adjuvant. Our initial focus was on adult mice because we could not induce adaptive immunity directly while mice were still infants. In pilot experiments, we found high titers of serum IgG to the recombinant protein by ELISA, but did not see a significant difference in colonization density compared to adjuvant-only controls following IN challenge with strain T4S. However, colonization levels of both antigen+adjuvant and adjuvant-alone groups were lower than naïve age-matched control mice, suggesting a confounding antigen-independent effect of immunization (Supplementary Fig. 1A). Prolonged effects of adjuvants on respiratory tract infection by Spn has been reported previously<sup>51</sup>.

To further optimize protection, we immunized adult mice with a combination of all five candidate proteins + TiterMax Gold<sup>®</sup>



**Fig. 4 Effect of immunization on infant mouse colonization.** Dams were vaccinated with the five-antigen combination (StkP, PenA, PgdA, HtrA, and LytD) with adjuvant and compared to adjuvant alone controls (mock). After the immune response was confirmed, they were bred and pups were challenged with the **a** unencapsulated mutant (T4cps-) or **b** its encapsulated parent strain (T4S). Symbols represent colonization density (CFU/ml upper respiratory track lavage) for individual mice at 3 (T4cps-) or 5 (T4S) days post-infection with median values shown. Mann–Whitney test used for comparisons. \*\*\* $P < 0.001$ , n.s., non-significant.

adjuvant. This resulted in high titers of serum IgG to each of the candidate proteins as measured by ELISA, although the response to LytD was less robust compared to the other four antigens (Fig. 3a). Importantly, a whole bacterial cell ELISA to strain T4S was used to demonstrate that the levels of serum IgG following immunization increased antibody recognition of the Spn surface (Fig. 3c). Whole bacterial cell ELISA with serum from the adults immunized with individual proteins showed that antibody to no single protein accounted for this surface recognition. Western analyses of whole-cell lysates confirmed that immunization generated specific antibody for each of the four antigens against which antibody was generated at high titer (Fig. 3d). Total IgG obtained from immune serum (at a concentration as high as 800  $\mu\text{g}/\text{mL}$ ) did not show agglutinating activity of either strain T4S or T4 $\Delta\text{cps}$ . Moreover, immune serum showed no inhibitory effect on in vitro growth of strain T4S in nutrient medium (Supplementary Fig. 1B). The lack of an effect on growth was not unexpected as only the T4 $\Delta\text{stkp}$  mutant demonstrated impaired growth characteristics.

### Protection from pneumococcal colonization

The confounding effect of the adjuvant on colonization was also observed when adult mice immunized with the five-antigen combination were IN challenged (Fig. S1A). Due to the difficulties in assessing pneumococcal colonization in immunized and adjuvant-only adult mice, we examined protection from colonization in infant mice, who acquire maternal antibody in utero and post-natum<sup>52</sup>. Serum IgG levels to each of the five recombinant proteins were comparable between infants from vaccinated dams and adults immunized with the combination of antigens (Fig. 3b). Four-day-old pups from immunized dams and pups from adjuvant-only dams were challenged IN with a low dose ( $10^3$  CFU) of strain T4S. Expected levels of colonization at 5 days post-inoculation were achieved in adjuvant-only pups, but there was no protection in the vaccinated pups (Fig. 4) despite high levels of serum IgG. We also tested the effect of immunization on protection among pups using a model of intra-litter transmission in the setting of influenza A co-infection where on average <10 bacteria establish infection<sup>25</sup>. Transmission of strain T4S from inoculated index pups was observed in 54.5% of contact pups in litters from immunized dams compared to 78.6% in litters from dams given adjuvant alone ( $P = 0.175$ , Table S2). This result contrasted with the complete protection previously observed in pups from dams immunized with PCV<sup>52</sup>. The lack of robust

protection with the combination of five candidate antigens led us to question whether the pneumococcal capsule might be interfering with antibody binding to its targets in vivo. To examine this possibility, we repeated the whole bacterial cell ELISA with an unencapsulated mutant of strain T4S (T4 $\Delta\text{cps}$ ). Using the same pool of sera as in Fig. 3a, we found that surface binding of IgG from vaccinated mice was far greater to the unencapsulated strain compared to the isogenic encapsulated parent strain (Fig. 3c). Similarly, a comparison of antibody binding using flow cytometry on pneumococci showed that immune, high titer sera generated to the individual proteins recognized the bacterial surface but only when not expressing capsule (Fig. 3e). Unlike challenge with T4S, when pups received  $10^3$  CFU T4 $\Delta\text{cps}$ , immunity to the combination of five Spn proteins was associated with decreased colonization density compared to adjuvant-only pups (Fig. 4). [Note: colonization density was assessed at 3 rather than 5 days p.i. because of its shorter duration for unencapsulated strains<sup>53</sup>]. Western analysis confirmed that protection in immunized infant mice against strain T4 $\Delta\text{cps}$  colonization was not due to increased antigen expression in the absence of capsule (Fig. 3d).

### Summary

Our unbiased genome-wide approach using Tn-Seq identified multiple conserved, immunogenic, surface exposed proteins required for Spn colonization. After identifying and validating these targets, we tested whether antibody-mediated disruption of their function could mitigate pneumococcal carriage. We demonstrated that, in mice, antibodies generated by systemic immunization with a combination of five recombinant pneumococcal surface proteins was sufficient to reduce Spn colonization in the nasopharynx. However, this protective effect was observed only when the pneumococcal capsule was not present—a result that correlated with increased binding of serum antibody to the surface unencapsulated pneumococci<sup>54</sup>. Protection against the unencapsulated strain in infant mice, who develop immunity through passive transfer, makes a role for T cell rather than antibody-mediated protection unlikely. Of course, we cannot exclude the possibility that the antibody generated by immunization with these targets could be protective against encapsulated Spn if induced at higher titers. It is also possible that our approach, while not sufficient to block colonization, could be effective at later stages of pathogenesis or that antibody directed against selected other conserved, immunogenic surface targets alone or in combination could provide protection against encapsulated

Spn. In this regard, leading protein vaccine candidates, including PspA and PspC(CbpA), were identified in our TnSeq screen. As these have variable domains, it appears they are under immune pressure, perhaps because they extend beyond the capsule unlike the five candidates we tested. Mutants in *pspA* and *pspC*, however, showed wild-type levels colonization and, therefore, were not considered further. It also remains plausible that a combination of capsular polysaccharide and common protein antigens could be used to optimize and broaden protection<sup>55,56</sup>. That said, our study provides a demonstration of the challenge of serotype-independent approaches to prevent Spn infection: the effectiveness of the capsule in shielding conserved surface antigens that would otherwise be targets of protective humoral immunity.

## METHODS

### Bacterial strains and culture

*S. pneumoniae* strains were grown in tryptic soy (TS) broth (BD, Franklin Lakes, NJ) at 37 °C, without aeration, to an optical density at 620 nm (OD<sub>620</sub>) of 1.0; or incubated on TS agar plates supplemented with 100 µl of catalase (30,000 U/ml; Worthington Biochemical) and antibiotic (as indicated) at 37 °C in 5% CO<sub>2</sub>, overnight. Antibiotics used in this study are as follows: streptomycin (str; 200 µg/ml); kanamycin (kan; 125 µg /mL or 500 µg/mL); erythromycin (erm; 1 or 2 µg/ml); spectinomycin (spec; 200 µg/l); and chloramphenicol (cam; 2 µg/ml).

### Mutant construction and protein purification

The SpnΔ*lytR* SpnΔ*nanB*, SpnΔ*penA*, SpnΔ*pfbA*, and SpnΔ*stkP* knockout strains were constructed in a str-resistant *S. pneumoniae* type T4 strain (T4S, P2406) using the Janus cassette<sup>57</sup>. The Janus cassette was amplified from genomic DNA, obtained using a MasterPure DNA purification kit (Illumina), from Spn T4 strain P2408<sup>58</sup>, with upstream and downstream regions (~1 kb) flanking to the gene of interest added via isothermal assembly. Strain P2406 was then transformed with the PCR product, and the transformants were selected on 5% blood agar plates supplemented with kan (500 µg/mL). The gene replacements with the Janus cassette were confirmed by PCR. The SpnΔ*zmpB* and SpnΔ*zmpC* Janus-insertions strains were kindly provided by Dr. Pyong Woo Park<sup>59</sup>, and the deletions were moved to our T4S parent strain by transforming genomic DNA obtained from them into P2406, and selecting the transformants on TS-kan (500 µg /mL) plates and 5% sheep's blood agar plates. Strains SpnΔ*pgdA* (P1831; kan<sup>R</sup>), SpnΔ*pspA* (P2205; kan<sup>R</sup>), SpnΔ*bga* (P2167; erm<sup>R</sup>), SpnΔ*htrA* (P1539; kan<sup>R</sup>), SpnΔ*strH* (P2168; spec<sup>R</sup>), SpnΔ*nanA* (P2082; cam<sup>R</sup>) and SpnΔ*pce/lytD* (P2421; kan<sup>R</sup>), were constructed by transforming genomic DNA obtained from them into P2406, and the transformants selected on TS-agar plates with appropriate antibiotics. The *cbpA/pspC* deletion was constructed by transforming genomic DNA obtained from P1465 in to P2406 and selecting on TS-agar plates supplemented with erythromycin (2 µg/ml) and 5% blood. DNA obtained from this strain was back-transformed in to P2406 and transformants were selected on plates as described above, to generate strain P2513 (*cbpA::erm*<sup>R</sup>; Erm<sup>R</sup>, Strep<sup>R</sup>).

For purification of recombinant proteins, the pET-28a(+) vector (Novagen, Madison, WI), which contains a his<sub>6</sub> tag at the N-terminus, was used. The open reading frames for genes *htrA*, *penA*, *pgdA*, *pce/lytD*, and *stkP* were amplified by PCR using Q5 polymerase (NEB) with genomic DNA as template and primers that contained restriction enzyme sites for NdeI and XhoI (*htrA*, *pgdA*, *stkP*, and *penA*) and NheI and XhoI (*pce/lytD*), and were cloned in to the pET-28a(+) vector. The resulting plasmids were transformed into a derivative of *E. coli* strain BL21(DE3)pLysS that had been modified to have a deletion of *ompA* (making strain E548). DegP has homology to HtrA and has been shown to co-purify with rHtrA<sup>46</sup>; therefore, we deleted the *E. coli* gene coding for the protease DegP by P1vir phage transduction of DNA from strain (E936) with *degP::kan* from Keio collection<sup>60</sup> followed by excision of the kanamycin-resistance cassette by expression of the FLP recombinase from temperature sensitive plasmid pcp20 (E947). For rPce, the first 26 amino acids of *pce/lytD* were not included, as they likely contain the signal sequence. As described previously<sup>38</sup>, only the C-terminal immunogenic PASTA domain region of StkP (AA 345–659) was cloned in to the expression vector.

Recombinant proteins were purified with the HisTrap column (GE Healthcare) by FPLC using Äkta Start (GE Healthcare)<sup>35,38,46,61–63</sup>. Briefly, *E. coli* cell pellets were resuspended in binding buffer (20 mM NaH<sub>2</sub>PO<sub>4</sub>,

500 mM NaCl, 20 mM Imidazole, 5% glycerol, 0.5% Triton, 10 mM 2-mercaptoethanol pH 7.4 and 20 mg/ml lysozyme), loaded on the HisTrap column, washed to remove non-specific bound protein (10–15% of elution buffer mixed with binding buffer), and fractions were eluted in 1 ml (elution buffer is binding buffer with the addition of 500 mM imidazole). For rStkP C-terminal purification, the cell pellet was first resuspended in binding buffer with 20 µl of RNase A (0.5 mg/ml) and 20 µl of DNase I (150 units/ml), incubated on ice for 60 min, and vortexed every 10 min. Samples were then sonicated (10 times for 30 s at 10% output, with 30 s on ice in between each sonication step) and pelleted by centrifugation (10,000 × *g* for 30 min at 4 °C). The pellet, containing the insoluble fraction, was solubilized in 8 M urea, 50 mM Tris-Cl, pH 8, sonicated (as described above) and pelleted by centrifugation (15,000 *g* × 10 min, RT). The supernatant was loaded on to the HisTrap column, washed, and eluted as described above. Fractions containing the proteins were pooled together and dialyzed with the following dialysis buffer: rPenA (25 mM Tris-Cl pH 8.0, 5% glycerol, 100 mM NaCl); rPce/LytD (25 mM Tris-Cl pH 8.0, 5% glycerol, 100 mM NaCl); rStkP-C terminal (25 mM Tris-Cl pH 7.4, 10% glycerol, 100 mM NaCl); rHtrA (PBS); and rPgdA (50 mM Na<sub>2</sub>HPO<sub>4</sub>, 150 mM NaCl, and 0.5 mM CoCl<sub>2</sub> pH 7.5). Recombinant proteins were concentrated using Amicon Ultra centrifuge filters (Millipore; according to the manufacturer's instructions). Protein concentration was assessed by Bradford protein assay (Bio-Rad cat. no. 5000006; bovine serum albumin Protein Standard II cat. no. 5000007).

### Mouse studies

Wild-type C57BL/6J mice (Jackson Laboratories, Bar Harbor, ME) were used in all experiments. Mice were maintained and bred in a conventional animal facility. Pups were housed with a dam for the course of the studies. Following infection, all mice appeared healthy, showed normal activity, and gained weight similar to uninfected controls. These studies were conducted in accordance with the recommendations of the Guide for the Care and Use of Laboratory Animals. All mouse studies were approved by the Institutional Animal Care and Use Committee of the New York University Medical Center. All procedures were in compliance with Biosafety in Microbiological and Biomedical Laboratories.

### Transposon library preparation and in vivo selection

Library construction using the mariner transposon (Tn) was carried out as previously described; we generated 28 independent libraries<sup>28</sup>. Each library contained ~500 random transposon insertion mutants in nonessential genes. Pools were stored at –80 in 20% glycerol-TS until use. For mouse infections, a transposon pool was grown in TS broth in static culture at 37 °C. Two 4-day-old pups (representing biological replicates) were inoculated IN, without anesthesia, with 4000–8000 CFU Spn in 3 µL PBS, such that each pup received ~10-fold excess of the Tn mutant strains (*Input*).

On day 5 post-inoculation, pups were euthanized by CO<sub>2</sub> and cardiac puncture; retrotracheal lavages were collected from the nares (200 µL PBS). Lavages were diluted and plated to TS-str plates supplemented with catalase and incubated at 37 °C + 5% CO<sub>2</sub> to determine colonization density. The remainder of the lavage was plated on TS-str plates supplemented with catalase and incubated overnight at 37 °C + 5% CO<sub>2</sub>. After overnight incubation, the colonies were collected from the plate and transferred to 5 mL sterile PBS (Output). Genomic DNA was isolated and samples were prepared as described<sup>28</sup>. A clean up step was added to the ligation mixture to remove excess adaptor using AMPure XP beads (Beckman Coulter) before using PCR to amplify the transposon insertion site.

### Sequencing and TnSeq bioinformatics analysis

Transposon-junction DNA fragments were subjected to single-end 50 rapid run sequencing on an Illumina HiSeq2500 Instrument (Illumina, San Diego). The resulting reads were filtered, mapped, and normalized as described<sup>64</sup>. Reads were aligned using Bowtie<sup>65</sup> to the *S. pneumoniae* TIGR4 reference genome<sup>50</sup> and insertion sites were called for each sample based on the alignment counts. Insertion sites within 80% of gene length from the transcription start sites were treated as candidates that could affect gene functions. Finally, TnseqDiff<sup>66</sup> was used to compare the insertions between output (nasal lavage) and input (inoculum) samples to identify candidate genes required for URT colonization. Functional/biological annotation was carried out using EggNOG(v5)<sup>67</sup>.

## Sequence conservation of candidate antigens

To quantify the conservation and selection pressure on the chosen antigens, we used the publicly available sequences from the same study in which their IgG binding was measured<sup>30</sup>. This consisted of 616 *S. pneumoniae* samples isolated from the nasopharynx of asymptomatic child carriers, over seven years spanning the rollout of PCV-7<sup>68</sup>. We used a previously-generated core-genome alignment, with clusters of orthologous genes (COG) as defined within this<sup>69</sup>. We then calculated dN/dS values for each COG as follows: we first aligned each COG at the codon level using RevTrans<sup>70</sup> and at the amino acid level using MUSCLE<sup>71</sup>. We used the nucleotide alignment to generate a maximum likelihood tree using a GTR rate model with IQ-TREE<sup>72</sup>. These alignments and the tree were then input into HYPHY, where we used Single-Likelihood Ancestor Counting<sup>73</sup> to calculate dN/dS values for each COG. We calculated the pairwise-diversity of the translated sequences ( $p_{aa}$ ) using dendropy on the amino-acid alignments<sup>74</sup>.

## Vaccination and spn challenge of adult mice

Four-week-old male and female mice were vaccinated with 5 µg of each recombinant protein + an equal volume of TiterMax Gold<sup>®</sup> adjuvant or a combination of all 5 proteins (5 µg each; 25 µg total) + TiterMax Gold adjuvant; mock mice received an equal amount of adjuvant only. The injection was divided and mice were vaccinated by subcutaneous injection in 2–3 different body sites (back of the neck, base of tail, and flank), two times at an interval of 14 days. Mice were challenged with Spn 8-weeks post-vaccination (6-weeks post-boost). Briefly, mice were infected intranasally (IN) with 10<sup>5</sup> CFU Spn T4S in 10 µL of sterile phosphate-buffered saline (PBS) by instillation on the nares with a blunt pipette tip, without anesthesia. On day 5 post-inoculation, mice were euthanized by CO<sub>2</sub> and cardiac puncture; retrotracheal lavages were collected from the nares (in 200 µL PBS). Lavages were diluted and plated to TS-str plates supplemented with catalase and incubated at 37 °C + 5% CO<sub>2</sub> to determine colonization density. The presence of specific IgG in the serum of mice was confirmed by ELISA (as described below).

## Maternal immunization and Spn challenge of pups

Immunity to the recombinant pneumococcal proteins was induced in adult female mice as described above. The female mice were then paired with males, and their litters were used in experiments to assess the role of immunity in protecting pups from pneumococcal challenge. Four-day-old pups were infected with 10<sup>3</sup> CFU of Spn T4S or T4Δcps in 3 µL of PBS by IN instillation with a blunt pipette tip, without anesthesia, and returned to their mother for the duration of the experiment. Three- or five-days post-inoculation, pups were euthanized by CO<sub>2</sub> asphyxiation and cardiac puncture and the trachea was lavaged with 200 µL of sterile PBS, which was collected from the nose. Tenfold serial dilutions of the lavage sample were plated to TS-str plates and enumerated after overnight incubation. The presence of anti-pneumococcal IgG in serum of mothers and their pups was confirmed by an ELISA (as described below). Transmission of Spn T4S among pups from vaccinated (5 antigen+adjuvant) compared to mock dams (adjuvant alone) in the setting of influenza A (strain x31) co-infection was examined using a previously described protocol with a ratio of ~1 to 4 index:contact pups<sup>52</sup>.

## ELISA

An ELISA to assess the titer of anti-pneumococcal IgG in the serum of immunized mothers and their pups, using whole bacteria as the capture antigen, was done as previously described<sup>11</sup>. Briefly, Immulon 2HB flat-bottom microtiter plates were coated with 100 µL of PBS-washed whole bacteria (*S. pneumoniae* T4S or T4Δcps), diluted in coating buffer (0.015 M Na<sub>2</sub>CO<sub>3</sub>, 0.035 M NaHCO<sub>3</sub>) to a final OD<sub>620</sub> of 0.1, and incubated at 4 °C overnight. The plates were then washed three times with wash solution (0.05% Brij-35 Surfact-Amps [Thermo, Fisher Scientific]–PBS) and blocked for 1 h at room temperature with 1% bovine serum albumin (BSA)–PBS. Serial serum dilutions, in PBS, were applied to the plates and incubated at 4 °C overnight. The plates were then washed three times with wash solution. Mouse IgG was detected with alkaline phosphatase (AP)-conjugated goat anti-mouse IgG antibody (Sigma A5153) at a 1:4,000 dilution in PBS, incubated at 37 °C for 1.5 h. The plates were then washed three times with wash buffer. The plate was developed with pNPP

substrate tablets (Kirkegaard & Perry Laboratories Inc., Gaithersburg, MD) dissolved in diethanolamine (DEA [Thermo, Fisher Scientific]). Absorbance at 405 nm was read at 1 h. The endpoint titers were determined by calculating the dilution at which the absorbance is equal to 0.1.

An antigen ELISA assessed the presence of anti-pneumococcal protein IgG in the serum of mice vaccinated with recombinant proteins. Purified recombinant proteins (the same used to vaccinate mice), assessed individually, were diluted to a final concentration of 2 µg/ml in PBS and applied to Nunc-Maxisorp flat-bottom microtiter plates. Plates were incubated overnight at 4 °C and then washed three times with PBS and blocked for 2 h at RT with 1% BSA–PBS. Serially diluted serum samples, in 1% BSA–PBS, were applied to the plate and incubated at 37 °C for 1 h. The plate was washed with PBS, and 100 µL of AP-conjugated goat anti-mouse IgG (diluted 1:2,500 in PBS) was applied to each well. The plate was incubated at 37 °C for 1 h. The wells were washed with PBS, and the plate was developed as described above.

## Western blot analysis

Bacteria were grown to an OD<sub>620</sub> = 0.8, centrifuged, washed in PBS. After resuspension of the pellet in NuPAGE LDS sample buffer with NuPAGE reducing agent (Invitrogen), lysates were prepared by treatment at 100 °C for 5 min and storage at –20 °C. Lysates were separated using 10% Bolt Bis-Tris Plus gels (Thermo Fisher Scientific NW00100BOX) and transferred onto PVDF membranes via a dry blotting system (iBlot 2, Thermo Fisher Scientific). Membranes were blocked with 1% bovine serum albumin/PBS for one hr. Subsequent steps were carried out in TBS/0.1% Tween-20. Blots were incubated with pooled immune sera (1:500) overnight at 4 °C and then, following 3 washes, with an alkaline phosphatase-coupled goat anti-mouse IgG (1:5000, Sigma A3688) for 2 h at RT. Protein bands were visualized after 3 washes using BCIP/NBT as substrate. For detection of each antigen, all blots derive from the same experiment and were processed in parallel (Supplementary Fig. 2).

## Flow cytometric analysis

Bacterial cultures were grown to an O.D. of 0.2 at 620 nm, resuspended in PBS and incubated with 1% sera from adult mice immunized with each protein candidate at 37 °C for 30 min. Cells were washed, then incubated with an APC-conjugated anti-mouse secondary antibody (Invitrogen) at 1:100 in 2% BSA in PBS at 4 °C for 30 min. Cells were washed, then fixed in 2% paraformaldehyde at room temperature for 20 min. Samples were run on an LSRII flow cytometer (BD Biosciences) and FACS analysis was performed using FlowJo software (Tree Star). Antibody binding was determined by proportion of bacteria with APC fluorescence signal with the gating strategy shown in Supplementary Fig. 3.

## Statistical analysis

Statistical analysis was done using GraphPad Prism 8.0 (GraphPad Software, Inc., San Diego, CA).

## Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this article.

## DATA AVAILABILITY

All datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

## CODE AVAILABILITY

Code used to analyze selection and diversity in the gene sequences is available on github (MIT license) [https://github.com/johnlees/pneumo\\_core](https://github.com/johnlees/pneumo_core) and archived at <https://doi.org/10.5281/zenodo.5606975>. External software used (and version): RevTrans (v.1.10); MUSCLE (v.3.8.31); IQ-TREE (v.1.5.5); HYPHY (v.2.3); dendropy (v4.4.0).

Received: 22 June 2021; Accepted: 9 November 2021;

Published online: 20 December 2021

## REFERENCES

- Weiser, J. N., Ferreira, D. M. & Paton, J. C. Streptococcus pneumoniae: transmission, colonization and invasion. *Nat. Rev. Microbiol.* **16**, 355–367 (2018).
- Whitney, C. G. et al. Decline in invasive pneumococcal disease after the introduction of protein-polysaccharide conjugate vaccine. *N. Engl. J. Med.* **348**, 1737–1746 (2003).
- O'Brien, K. L. et al. Effect of pneumococcal conjugate vaccine on nasopharyngeal colonization among immunized and unimmunized children in a community-randomized trial. *J. Infect. Dis.* **196**, 1211–1220 (2007).
- Griffin, M. R., Zhu, Y., Moore, M. R., Whitney, C. G. & Grijalva, C. G. U.S. hospitalizations for pneumonia after a decade of pneumococcal vaccination. *N. Engl. J. Med.* **369**, 155–163 (2013).
- Klugman, K. The significance of serotype replacement for pneumococcal disease and antibiotic resistance. *Adv. Exp. Med Biol.* **634**, 121–128 (2009).
- Ladhani, S. et al. Rapid increase in non-vaccine serotypes causing invasive pneumococcal disease in England and Wales, 2000–17: a prospective national observational cohort study. *Lancet Infect. Dis.* **18**, 441–451 (2018).
- Simonsen, L. et al. Impact of pneumococcal conjugate vaccination of infants on pneumonia and influenza hospitalization and mortality in all age groups in the United States. *MBio* **2**, e00309–e00310 (2011).
- Pizza, M., Bekkat-Berkani, R. & Rappuoli, R. Vaccines against meningococcal diseases. *Microorganisms* **8**, 1521 (2020).
- Simell, B. et al. The fundamental link between pneumococcal carriage and disease. *Expert Rev. Vaccines* **11**, 841–855 (2012).
- Lipsitch, M. et al. Are anticapsular antibodies the primary mechanism of protection against invasive pneumococcal disease? *PLoS Med.* **2**, e15 (2005).
- Roche, A., King, S. & Weiser, J. Live attenuated Streptococcus pneumoniae strains induce serotype-independent mucosal and systemic protection in mice. *Infect. Immun.* **75**, 2469–2475 (2007).
- Wilson, R. et al. Naturally acquired human immunity to pneumococcus is dependent on antibody to protein antigens. *PLoS Pathog.* **13**, e1006137 (2017).
- Chan, W. et al. A novel, multiple-antigen pneumococcal vaccine protects against lethal Streptococcus pneumoniae challenge. *Infect. Immun.* **87**, e00846–00818 (2019).
- Moffitt, K. et al. T(H)17-mediated protection against pneumococcal carriage by a whole-cell vaccine is dependent on toll-like receptor 2 and surface lipoproteins. *Clin. Vaccin. Immunol.* **22**, 909–916 (2015).
- Hammerschmidt, S. et al. Illustration of pneumococcal polysaccharide capsule during adherence and invasion of epithelial cells. *Infect. Immun.* **73**, 4653–4667 (2005).
- Roche, A. M., Richard, A. L., Rahkola, J. T., Janoff, E. N. & Weiser, J. N. Antibody blocks acquisition of bacterial colonization through agglutination. *Mucosal Immunol.* **8**, 176–185 (2015).
- Janoff, E. N. et al. Pneumococcal IgA1 protease subverts specific protection by human IgA1. *Mucosal Immunol.* **7**, 249–256 (2014).
- Richards, L., Ferreira, D. M., Miyaji, E. N., Andrew, P. W. & Kadioglu, A. The immunising effect of pneumococcal nasopharyngeal colonisation; protection against future colonisation and fatal invasive disease. *Immunobiology* **215**, 251–263 (2010).
- Pennington, S. H. et al. Polysaccharide-specific memory b cells predict protection against experimental human pneumococcal carriage. *Am. J. Respir. Crit. Care Med.* **194**, 1523–1531 (2016).
- Mitsi, E. et al. Agglutination by anti-capsular polysaccharide antibody is associated with protection against experimental human pneumococcal carriage. *Mucosal Immunol.* **10**, 385–394 (2017).
- Jochems, S., Weiser, J., Malley, R. & Ferreira, D. The immunological mechanisms that control pneumococcal carriage. *PLoS Pathog.* **13**, e1006665 (2017).
- Riegler, A., Mann, B., Orihuela, C. & Tuomanen, E. Opening the OPK assay gatekeeper: harnessing multi-modal protection by pneumococcal vaccines. *Pathogens* **8**, 203 (2019).
- Zafar, M. A., Kono, M., Wang, Y., Zangari, T. & Weiser, J. N. Infant mouse model for the study of shedding and transmission during Streptococcus pneumoniae mono-infection. *Infect. Immun.* **84**, 2714–2722 (2016).
- Li, Y., Thompson, C. M., Trzcinski, K. & Lipsitch, M. Within-host selection is limited by an effective population of Streptococcus pneumoniae during nasopharyngeal colonization. *Infect. Immun.* **81**, 4534–4543 (2013).
- Kono, M. et al. Single cell bottlenecks in the pathogenesis of Streptococcus pneumoniae. *PLoS Pathog.* **12**, e1005887 (2016).
- van Opijnen, T., Bodi, K. L. & Camilli, A. Tn-seq: high-throughput parallel sequencing for fitness and genetic interaction studies in microorganisms. *Nat. Methods* **6**, 767–772 (2009).
- van Opijnen, T., Lazinski, D. W. & Camilli, A. Genome-wide fitness and genetic interactions determined by Tn-seq, a high-throughput massively parallel sequencing method for microorganisms. *Curr. Protoc. Mol. Biol.* **106**, 7.16.11–24 (2014).
- Zafar, M. et al. Identification of pneumococcal factors affecting its shedding shows the dlt locus promotes inflammation and transmission. *mbio* **10**, e01032–01019 (2019).
- Hammond, A. et al. Neuraminidase B controls neuraminidase A-dependent mucus production and evasion. *PLoS Pathog.* **17**, e1009158 (2021).
- Croucher, N. J. et al. Diverse evolutionary patterns of pneumococcal antigens identified by pangenome-wide immunological screening. *Proc. Natl Acad. Sci. USA* **114**, E357–e366 (2017).
- Zafar, M. A., Hamaguchi, S., Zangari, T., Cammer, M. & Weiser, J. N. Capsule type and amount affect shedding and transmission of Streptococcus pneumoniae. *MBio* **8**, e00989–00917 (2017).
- Corander, J. et al. Frequency-dependent selection in vaccine-associated pneumococcal population dynamics. *Nat. Ecol. evolution* **1**, 1950–1960 (2017).
- Nei, M. & Li, W. Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proc. Natl Acad. Sci. USA* **76**, 5269–5273 (1979).
- Feil, E. et al. Recombination within natural populations of pathogenic bacteria: short-term empirical estimates and long-term phylogenetic consequences. *Proc. Natl Acad. Sci. USA* **98**, 182–187 (2001).
- Beilharz, K. et al. Control of cell division in Streptococcus pneumoniae by the conserved Ser/Thr protein kinase StkP. *Proc. Natl Acad. Sci. USA* **109**, E905–E913 (2012).
- Fenton, A. K. et al. Phosphorylation-dependent activation of the cell wall synthase PBP2a in Streptococcus pneumoniae by MacP. *Proc. Natl Acad. Sci. USA* **115**, 2812–2817 (2018).
- Zucchini, L. et al. PASTA repeats of the protein kinase StkP interconnect cell constriction and separation of Streptococcus pneumoniae. *Nat. Microbiol.* **3**, 197–209 (2018).
- Giefing, C. et al. Discovery of a novel class of highly conserved vaccine antigens using genomic scale antigenic fingerprinting of pneumococcus with human antibodies. *J. Exp. Med.* **209**, 117–131 (2008).
- Ojal, J., Goldblatt, D., Tigoi, C. & Scott, J. A. G. Effect of maternally derived anti-protein and anticapsular IgG antibodies on the rate of acquisition of nasopharyngeal carriage of pneumococcus in newborns. *Clin. Infect. Dis.* **66**, 121–130 (2018).
- Kotevska, V. et al. Phenotypes and genes of resistance of pneumococci to penicillin isolated from children. *Prilozi* **30**, 143–154 (2009).
- Wang, Y. et al. Novel immunoprotective proteins of Streptococcus pneumoniae identified by opsonophagocytosis killing screen. *Infect Immun* **86**, <https://doi.org/10.1128/iai.00423-18> (2018).
- Blair, D. E., Schuttelkopf, A. W., MacRae, J. I. & van Aalten, D. M. Structure and metal-dependent mechanism of peptidoglycan deacetylase, a streptococcal virulence factor. *Proc. Natl Acad. Sci. USA* **102**, 15429–15434 (2005).
- Davis, K., Akinbi, H., Standish, A. & Weiser, J. Resistance to mucosal lysozyme compensates for the fitness deficit of peptidoglycan modifications by Streptococcus pneumoniae. *PLoS Pathog.* **4**, e1000241 (2008).
- Lokken-Toyli, K. et al. Decreased production of epithelial-derived antimicrobial molecules at mucosal barriers during early life. *Mucosal Immunol.* <https://doi.org/10.1038/s41385-41021-00438-y> (2021).
- Fan, K., Zhang, J., Zhang, X. & Tu, X. Solution structure of HtrA PDZ domain from Streptococcus pneumoniae and its interaction with YZF-COOH containing peptides. *J. Struct. Biol.* **176**, 16–23 (2011).
- Cassone, M., Gagne, A. L., Spruce, L. A., Seeholzer, S. H. & Sebert, M. E. The HtrA protease from Streptococcus pneumoniae digests both denatured proteins and the competence-stimulating peptide. *J. Biol. Chem.* **287**, 38449–38459 (2012).
- Kochan, T. J. & Dawid, S. The HtrA protease of Streptococcus pneumoniae controls density-dependent stimulation of the bacteriocin blp locus via disruption of pheromone secretion. *J. Bacteriol.* **195**, 1561–1572 (2013).
- Lagartera, L. et al. Crystallization and preliminary X-ray diffraction studies of the pneumococcal teichoic acid phosphorylcholine esterase Pce. *Acta Crystallogr. Sect. F., Struct. Biol. Crystallization Commun.* **61**, 221–224 (2005).
- Hergott, C. B. et al. Bacterial exploitation of phosphorylcholine mimicry suppresses inflammation to promote airway infection. *J. Clin. Invest.* **125**, 3878–3890 (2015).
- Tettelin, H. et al. Complete genome sequence of a virulent isolate of Streptococcus pneumoniae. *Science* **293**, 498–506 (2001).
- Kuipers, K. et al. Antigen-independent restriction of pneumococcal density by mucosal adjuvant cholera toxin subunit B. *J. Infect. Dis.* **214**, 1588–1596 (2016).
- Zangari, T., Wang, Y. & Weiser, J. N. Streptococcus pneumoniae transmission is blocked by type-specific immunity in an infant mouse model. *MBio* **8**, e00188–00117 (2017).
- Nelson, A. L. et al. Capsule enhances pneumococcal colonization by limiting mucus-mediated clearance. *Infect. Immun.* **75**, 83–90 (2007).
- Hyams, C., Camberlein, E., Cohen, J. M., Bax, K. & Brown, J. S. The Streptococcus pneumoniae capsule inhibits complement activity and neutrophil phagocytosis by multiple mechanisms. *Infect. Immun.* **78**, 704–715 (2010).



55. Colijn, C., Corander, J. & Croucher, N. Designing ecologically optimized pneumococcal vaccines using population genomics. *Nat. Microbiol.* **5**, 473–485 (2020).
56. Rowe, H. et al. A cross-reactive protein vaccine combined with PCV-13 prevents *Streptococcus pneumoniae*- and *Haemophilus influenzae*-mediated acute otitis media. *Infect. Immun.* **87**, e00253–00219 (2019).
57. Sung, C., Li, H., Claverys, J. & Morrison, D. An rpsL cassette, janus, for gene replacement through negative selection in *Streptococcus pneumoniae*. *Appl Environ. Microbiol.* **67**, 5190–5196 (2001).
58. Lemon, J. K. & Weiser, J. N. Degradation products of the extracellular pathogen *Streptococcus pneumoniae* access the cytosol via its pore-forming toxin. *MBio* **6**, e02110–e02114 (2015).
59. Chen, Y., Hayashida, A., Bennett, A., Hollingshead, S. & Park, P. *Streptococcus pneumoniae* sheds syndecan-1 ectodomains through ZmpC, a metalloproteinase virulence factor. *J. Biol. Chem.* **282**, 159–167 (2007).
60. Baba, T. et al. Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol. Syst. Biol.* **2**, 2006.0008 (2006).
61. Kuipers, K. et al. Age-related differences in IL-1 signaling and capsule serotype affect persistence of *Streptococcus pneumoniae* colonization. *PLoS Pathog.* **14**, e1007396 (2018).
62. Morlot, C. et al. Interaction of penicillin-binding protein 2x and Ser/Thr protein kinase StkP, two key players in *Streptococcus pneumoniae* R6 morphogenesis. *Mol. Microbiol.* **90**, 88–102 (2013).
63. Benachour, A. et al. The lysozyme-induced peptidoglycan N-acetylglucosamine deacetylase PgdA (EF1843) is required for *Enterococcus faecalis* virulence. *J. Bacteriol.* **194**, 6066–6073 (2012).
64. Goodman, A., Wu, M. & Gordon, J. Identifying microbial fitness determinants by insertion sequencing using genome-wide transposon mutant libraries. *Nat. Protoc.* **6**, 1969–1980 (2011).
65. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009).
66. Zhao, L., Anderson, M., Wu, W., Mobley, H. & Bachman, M. TnseqDiff: identification of conditionally essential genes in transposon sequencing studies. *BMC Bioinform.* **18**, 326 (2017).
67. Huerta-Cepas, J. et al. eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Res* **47**, <https://doi.org/10.1093/nar/gky1085> (2019).
68. Croucher, N. J. et al. Population genomics of post-vaccine changes in pneumococcal epidemiology. *Nat. Genet* **45**, 656–663 (2013).
69. Croucher, N. J. et al. Population genomic datasets describing the post-vaccine evolutionary epidemiology of *Streptococcus pneumoniae*. *Sci. Data* **2**, 150058 (2015).
70. Wernersson, R. & Pedersen, A. RevTrans: Multiple alignment of coding DNA from aligned amino acid sequences. *Nucleic Acids Res* **31**, 3537–3539 (2003).
71. Edgar, R. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792–1797 (2004).
72. Nguyen, L. T., Schmidt, H. A., Von Haeseler, A. & Minh, B. Q. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* **32**, 268–274 (2015).
73. Kosakovsky Pond, S. & Frost, S. Not so different after all: a comparison of methods for detecting amino acid sites under selection. *Mol. Biol. Evol.* **22**, 1208–1222 (2005).

74. Sukumaran, J. & Holder, M. DendroPy: a Python library for phylogenetic computing. *Bioinformatics* **26**, 1569–1571 (2010).

## ACKNOWLEDGEMENTS

The authors thank Weisheng Wu, BRCF Bioinformatics Core, University of Michigan. This work was supported by grants from the National Institutes of Health, R01 AI150893, R01 AI038446, and R21 AI50867 to J.N.W.

## AUTHOR CONTRIBUTIONS

T.Z., M.A.Z., J.A.L., and J.N.W. wrote the paper. T.Z., M.A.Z., G.C.W.B., and A.R.A. performed the experiments. T.Z., M.A.Z., J.A.L., G.C.W.B., and J.N.W. analyzed the data. T.Z., M.A.Z., J.A.L., and J.N.W. conceived and designed the experiments. J.N.W. supervised the research.

## COMPETING INTERESTS

Dr. Weiser receives royalties from GlaxoSmithKline for a pneumococcal vaccine. He also consults for or has consulted for Sanofi, Matrivax, and Merck & Co about pneumococcal vaccines.

## ADDITIONAL INFORMATION

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41541-021-00413-5>.

**Correspondence** and requests for materials should be addressed to Jeffrey N. Weiser.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021