AMERICAN SOCIETY FOR MICROBIOLOGY | Infection and Immunity

Host-Microbial Interactions | Full-Length Text

Bacterial surface lipoproteins mediate epithelial microinvasion by *Streptococcus pneumoniae*

Jia Mun Chan,[1] Elisa Ramos-Sevillano,[2] Modupeh Betts,[1] Holly U. Wilson,[1] Caroline M. Weight,[1] Ambrine Houhou-Ousalah,[1] Gabriele Pollara,[1] Jeremy S. Brown,[2] Robert S. Heyderman[1]

AUTHOR AFFILIATIONS See affiliation list on p. 18.

ABSTRACT *Streptococcus pneumoniae*, a common colonizer of the upper respiratory tract, invades nasopharyngeal epithelial cells without causing disease in healthy participants of controlled human infection studies. We hypothesized that surface expression of pneumococcal lipoproteins, recognized by the innate immune receptor TLR2, mediates epithelial microinvasion. Mutation of *lgt* in serotype 4 (TIGR4) and serotype 6B (BHN418) pneumococcal strains abolishes the ability of the mutants to activate TLR2 signaling. Loss of *lgt* also led to the concomitant decrease in interferon signaling triggered by the bacterium. However, only BHN418 *lgt::cm* but not TIGR4 *lgt::cm* was significantly attenuated in epithelial adherence and microinvasion compared to their respective wild-type strains. To test the hypothesis that differential lipoprotein repertoires in TIGR4 and BHN418 lead to the intraspecies variation in epithelial microinvasion, we employed a motif-based genome analysis and identified an additional 525 a.a. lipoprotein (pneumococcal accessory lipoprotein A; *palA*) encoded by BHN418 that is absent in TIGR4. The gene encoding *palA* sits within a putative genetic island present in ~10% of global pneumococcal isolates. While *palA* was enriched in the carriage and otitis media pneumococcal strains, neither mutation nor overexpression of the gene encoding this lipoprotein significantly changed microinvasion patterns. In conclusion, mutation of *lgt* attenuates epithelial inflammatory responses during pneumococcal-epithelial interactions, with intraspecies variation in the effect on microinvasion. Differential lipoprotein repertoires encoded by the different strains do not explain these differences in microinvasion. Rather, we postulate that post-translational modifications of lipoproteins may account for the differences in microinvasion.

IMPORTANCE *Streptococcus pneumoniae* (pneumococcus) is an important mucosal pathogen, estimated to cause over 500,000 deaths annually. Nasopharyngeal colonization is considered a necessary prerequisite for disease, yet many people are transiently and asymptomatically colonized by pneumococci without becoming unwell. It is therefore important to better understand how the colonization process is controlled at the epithelial surface. Controlled human infection studies revealed the presence of pneumococci within the epithelium of healthy volunteers (microinvasion). In this study, we focused on the regulation of epithelial microinvasion by pneumococcal lipoproteins. We found that pneumococcal lipoproteins induce epithelial inflammation but that differing lipoprotein repertoires do not significantly impact the magnitude of microinvasion. Targeting mucosal innate immunity and epithelial microinvasion alongside the induction of an adaptive immune response may be effective in preventing pneumococcal colonization and disease.

KEYWORDS *Streptococcus pneumoniae*, host-microbe interactions, microinvasion, bacterial lipoproteins, epithelium

Editor Nancy E. Freitag, University of Illinois Chicago, Chicago, Illinois, USA

Address correspondence to Robert S. Heyderman, r.heyderman@ucl.ac.uk, or Jia Mun Chan, elizabeth.jm.chan@ucl.ac.uk.

The authors declare no conflict of interest.

See the funding table on p. 18.

Received 2 November 2023
Accepted 27 March 2024
Published 17 April 2024

Streptococcus pneumoniae (pneumococcus) is a versatile pathobiont capable of asymptomatically colonizing the nasopharynx, causing localized infections of the middle ear, respiratory tract, and lungs, and causing disseminated invasive disease (e.g., bacteremic pneumonia and meningitis) with high mortality rates (1). *S. pneumoniae* has long been considered an extracellular pathogen despite the demonstration of bacterial invasion *in vitro* using epithelial and endothelial cell lines (1). However, controlled human infection with a serotype 6B strain revealed that the pneumococcus invades the nasopharyngeal epithelium of healthy carriers, stimulating epithelial inflammation without causing overt symptoms or disease (2–4). We have termed this phenomenon microinvasion, which is distinct from the invasion of deeper tissues or dissemination systemically which characterizes disease (2). Inflammation triggered by the epithelium-associated and intracellular bacteria, which peaks 9 days post inoculation, may be important for clearance and onward transmission (2).

In this study, we explored the hypothesis that surface expression of pneumococcal lipoproteins mediates epithelial microinvasion. Pneumococcal lipoproteins are post-translationally lipidated surface proteins, many of which function as metabolite transporters (5, 6). *S. pneumoniae* lipoproteins have also been shown to be major TLR2 ligands in macrophages, required for a Th17 response and for many of the dominant macrophage gene transcriptional responses, such as induction of IRAK-4-dependent protective cytokines (7–9). *S. pneumoniae* encodes over 30 lipoproteins, including the bifunctional adhesin/manganese transporter PsaA and the peptidoglycan hydrolase DacB (5, 10–12). Blocking lipidation by mutating the prolipoprotein diacylglyceryl transferase encoding gene *lgt* de-anchors lipoproteins from the cell surface, resulting in the release of immature preprolipoproteins into the extracellular milieu and abolishing the ability of the bacteria to activate TLR2 signaling (8, 9, 13). Mutating *lgt* also attenuates pneumococcal virulence and shortens colonization duration in murine models (8, 14).

To explore whether heterogeneity in surface-expression of pneumococcal lipoproteins also explains the differences in microinvasion seen between strains, we blocked lipoprotein lipidation by inactivation of *lgt* in two well-characterized strains: a highly invasive strain (TIGR4, serotype 4) and a less invasive strain (BHN418, serotype 6B) which was used in the controlled human challenge experiments (15, 16). It is important to note that pneumococcal strains from both serotypes can asymptomatically colonize as well as cause invasive disease in susceptible hosts, albeit to different extents (17). While attenuation of inflammatory responses was seen with both serotype 6B and serotype 4 *lgt* mutants, we observed intraspecies differences in the contribution of lipoproteins to microinvasion, with greater effects of lipoproteins with the less invasive 6B strain. Genomic analysis revealed the presence of a previously uncharacterized lipoprotein encoded within a genetic island found in BHN418 and approximately 10% of pneumococcal strains, but not in TIGR4. We designate this protein pneumocccal accessory lipoprotein A, or PalA, and investigated its role in mediating intraspecies differences in microinvasion.

RESULTS

Pneumococcal *lgt* mutants induce lower levels of TLR2 and interferon signaling compared to wild-type strains

In line with previous reports, mutation of *lgt* in both TIGR4 and BHN418 completely abolished the ability of these strains to trigger TLR2 signaling in HEK-Blue hTLR2 reporter cells, while genetic complementation of *lgt* at a chromosomal ectopic site restored wild-type (WT)-like ability to stimulate the TLR2 pathway (Fig. 1A) (9). Although macrophages respond to pneumococcal infections by activating TLR2 signaling pathways, it is unknown if nasopharyngeal epithelial cells respond in the same way (8, 9, 18). Using a transcriptional module reflective of TLR2 signaling and previously published transcriptomic data sets (2, 19), we found evidence of elevated TLR2-mediated

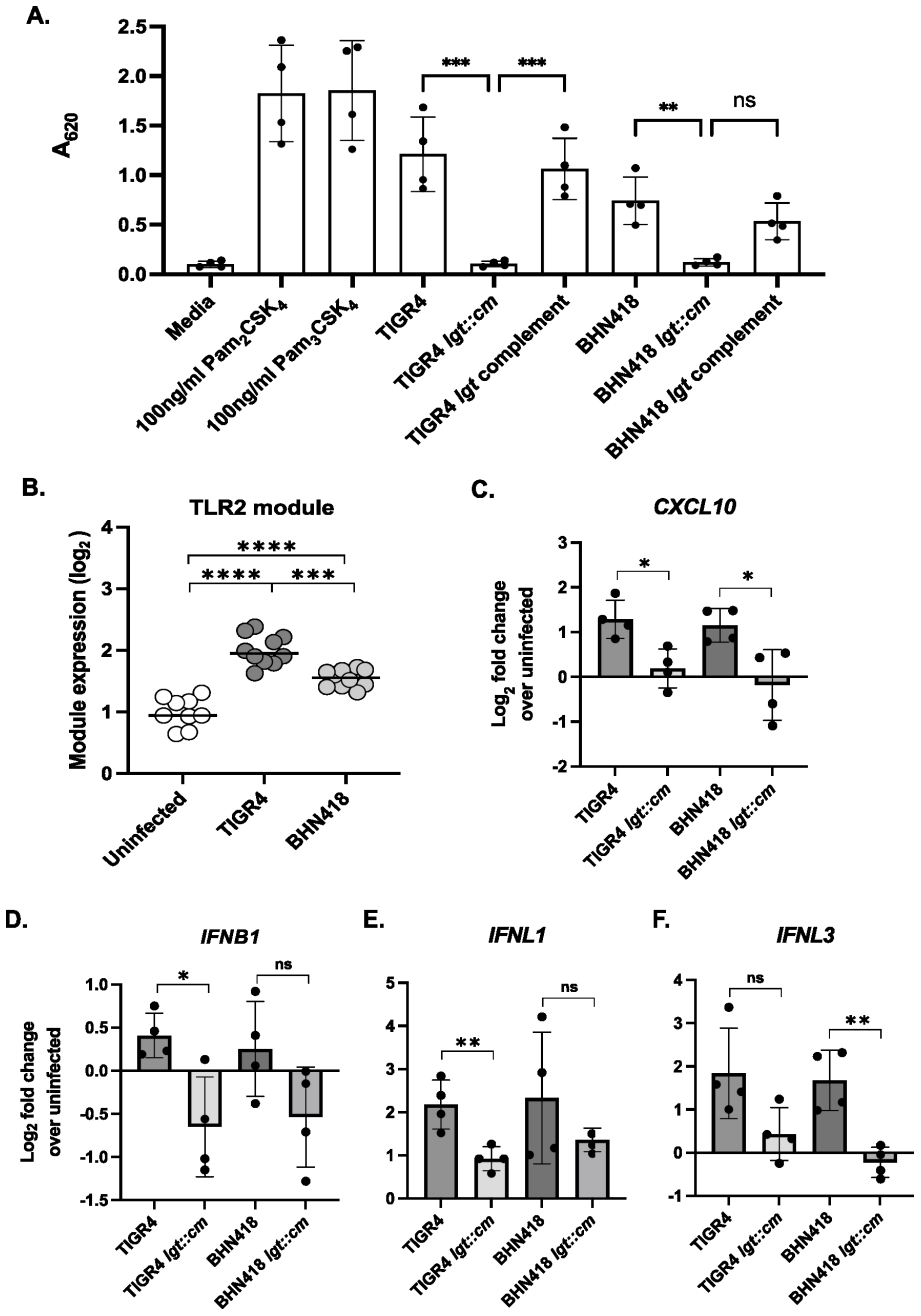

FIG 1 Pneumococcal *lgt* mutants were less inflammatory compared to WT strains. (A) SEAP reporter readout from HEK-Blue hTLR2 reporter cells treated with pneumococcal strains at MOI 10 for 16 hours. (B) Expression of a transcriptional module reflective of TLR2-mediated activity in Detroit 562 cells infected with TIGR4 and BHN418 for 3 hours. (C–F) Transcript levels of (C) *CXCL10*, (D) *IFNB1*, (E) *IFNL1*, and (F) *IFNL3*, quantified via qPCR using total RNA extracted from Detroit 562 cells after 6 hours of infection with pneumococcal strains. Statistical significance was determined using multiple comparison test with Bonferroni's correction (A), Mann-Whitney test (B), or Student's *t* test assuming equal variance (C–F). * indicates $P < 0.05$, ** indicates $P < 0.01$, *** indicates $P < 0.001$, and *** indicates $P < 0.0001$.

transcriptional activity in Detroit 562 nasopharyngeal epithelial cells infected with TIGR4 and BHN418 (Fig. 1B; Fig. S1).

TLR2 activation is necessary for the full induction of TLR4 by the *S. pneumoniae* virulence factor pneumolysin (20, 21). Transcriptomic analyses of human nasal biopsy samples from controlled pneumococcal challenge experiments and nasopharyngeal cell

lines infected with *S. pneumoniae* also showed upregulation of interferon signaling (6, 7). We, therefore, hypothesize that TLR2 activation potentiates interferon signaling in epithelial cells triggered by *S. pneumoniae* infection. Using qPCR, we observed that Detroit 562 cells infected with TIGR4 *lgt::cm* have reduced expression of *CXCL10*, *IFNB1*, and *IFNL1* compared to cells infected with WT TIGR4 (Fig. 1C through E), while cells infected with BHN418 *lgt::cm* have reduced expression of *CXCL10* and *IFNL3* compared to those infected with WT BHN418 (Fig. 1C and F). Our results suggest that lipoprotein-mediated TLR2 activation augments the epithelial interferon response during pneumococcal microinvasion.

Mutation of *lgt* attenuates epithelial microinvasion by *S. pneumoniae* serotype 6B but not by serotype 4

To determine if mutation of *lgt* and loss of TLR2 signaling impact on pneumococcal microinvasion, we infected confluent Detroit 562 nasopharyngeal cells (NPE) with serotype 6B (BHN418) and serotype 4 strains (TIGR4) for 3 hours (3 hpi), measuring the number of cell-associated, intracellular, and planktonic bacteria in the cell culture supernatant. Mutation of *lgt* significantly attenuated the ability of BHN418 but not TIGR4 to associate with and be internalized into Detroit 562 cells (Fig. 2A, B, D, and E). In concordance with prior reports, serotype 4 strains were more invasive compared to serotype 6B strains, with ~5 times more intracellular WT TIGR4 recovered compared to WT BHN418 (Fig. 2B and D) (15, 17). The *lgt* mutation also significantly reduced the number of planktonic BHN418 but not TIGR4 (Fig. 2F).

Genetic complementation of *lgt* in the BHN418 *lgt::cm* mutant did not fully restore microinvasion of NPE cells to WT-like levels, despite complementation in the HEK-Blue hTLR2 reporter assay (Fig. 1A and 2A through F). We double-checked the strain genotype using Illumina sequencing and confirmed *lgt* transcription (or lack thereof) in the TIGR4 and BHN418 strains using semi-quantitative PCR (Table S1; Fig. S2A and B). Leaky expression of the P_{IPTG} promoter in the absence of an inducer is sufficient to result in *lgt* expression (Fig. S2A and B). In agreement with prior studies, immunoblotting revealed substantially reduced retention but not complete loss of lipoproteins such as PiuA in whole-cell lysates in the *lgt::cm* mutant compared to wild-type and complementation strains (Fig. S3A and B) (13, 14). We conclude that the lack of complementation for the microinvasion phenotype is not due to failure in genetic complementation.

Mutation of *lgt* has been associated with growth defects in cation-limiting conditions, human blood, and mouse bronchoalveolar lavage fluid (14). Fewer planktonic BHN418 *lgt* mutant bacteria were also recovered from our NPE infection experiments (Fig. 2C). Time course sampling of planktonic pneumococci grown with Detroit 562 cells revealed a minor growth defect for the BHN418 *lgt* mutant starting at 3 hpi but not for the TIGR4 *lgt* mutant (Fig. 3A and B). To determine if the growth defect was dependent on the presence of NPE cells, time course sampling of planktonic BHN418 and its *lgt* mutant grown in infection medium and rich THY medium were performed. The growth defect was replicated in a cell-free medium and is therefore not dependent on the presence of NPE cells (Fig. 3C and D).

Our results indicate that inactivation of Lgt and, therefore, the lipoprotein processing pathway had greater consequences for BHN418 compared to TIGR4, except in their ability to trigger epithelial inflammation. These observations suggest that activation of the TLR2 pathway during pneumococcal-epithelial interactions is not dependent on the number of cell-associated or intracellular pneumococci. Additionally, within the timeframe of our assays, TLR2 signaling neither promotes nor inhibits epithelial microinvasion by *S. pneumoniae*.

BHN418 encodes a novel lipoprotein absent in TIGR4

One potential explanation for the intraspecies differences in microinvasion upon *lgt* mutation is the presence of one or more lipoproteins in BHN418 which are absent in TIGR4. This lipoprotein may play a role as an adhesin and/or be important for

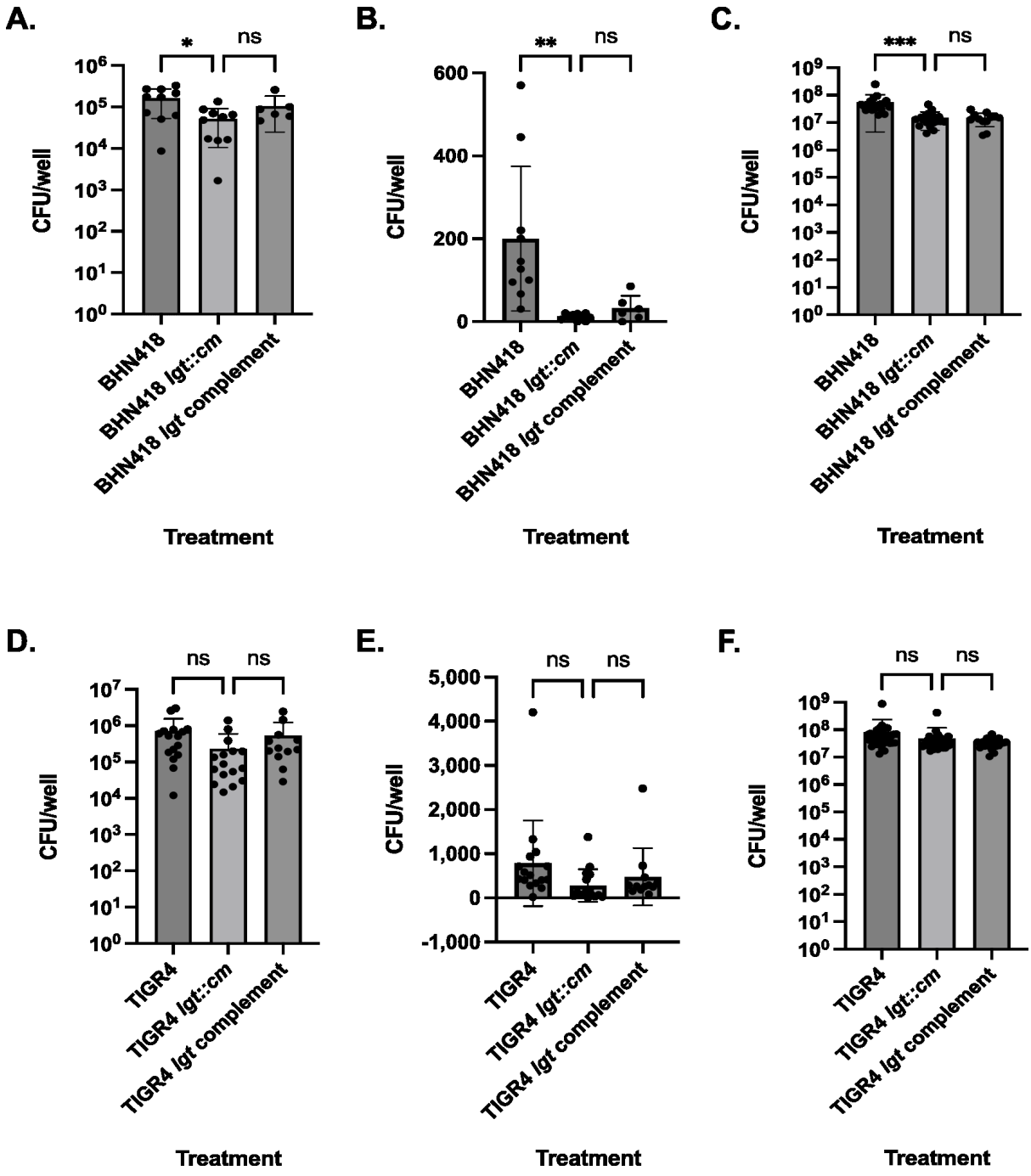

FIG 2 Mutation of *lgt* impaired epithelial microinvasion by *S. pneumoniae*, with greater defects for BHN418. (A–F) NPE microinvasion by WT and *lgt* mutants. Graphs show CFU numbers for BHN418-derived (A–C) and TIGR4-derived strains (D–F) associated with (A and D), internalized into (B and E), or growing in proximity with Detroit 562 NPE cells (C and F) 3 hours post infection. * indicates $P < 0.05$, ** indicates $P < 0.01$, and *** indicates $P < 0.001$.

nutrient transport and growth during infections. To address this hypothesis, we used a motif-based sequence toolkit, the MEME Suite, to compare the lipoprotein repertoires of BHN418 and TIGR4 (22). We identified 43 open-reading frames (ORFs) in TIGR4 and 44 ORFs in BHN418 with gene products that fit the criteria for a lipoprotein (detailed in Materials and methods; Table 1). Of these putative lipoprotein ORFs, only one is present in BHN418 but not in TIGR4. There are no lipoprotein ORFs present in TIGR4 that are not also present in BHN418.

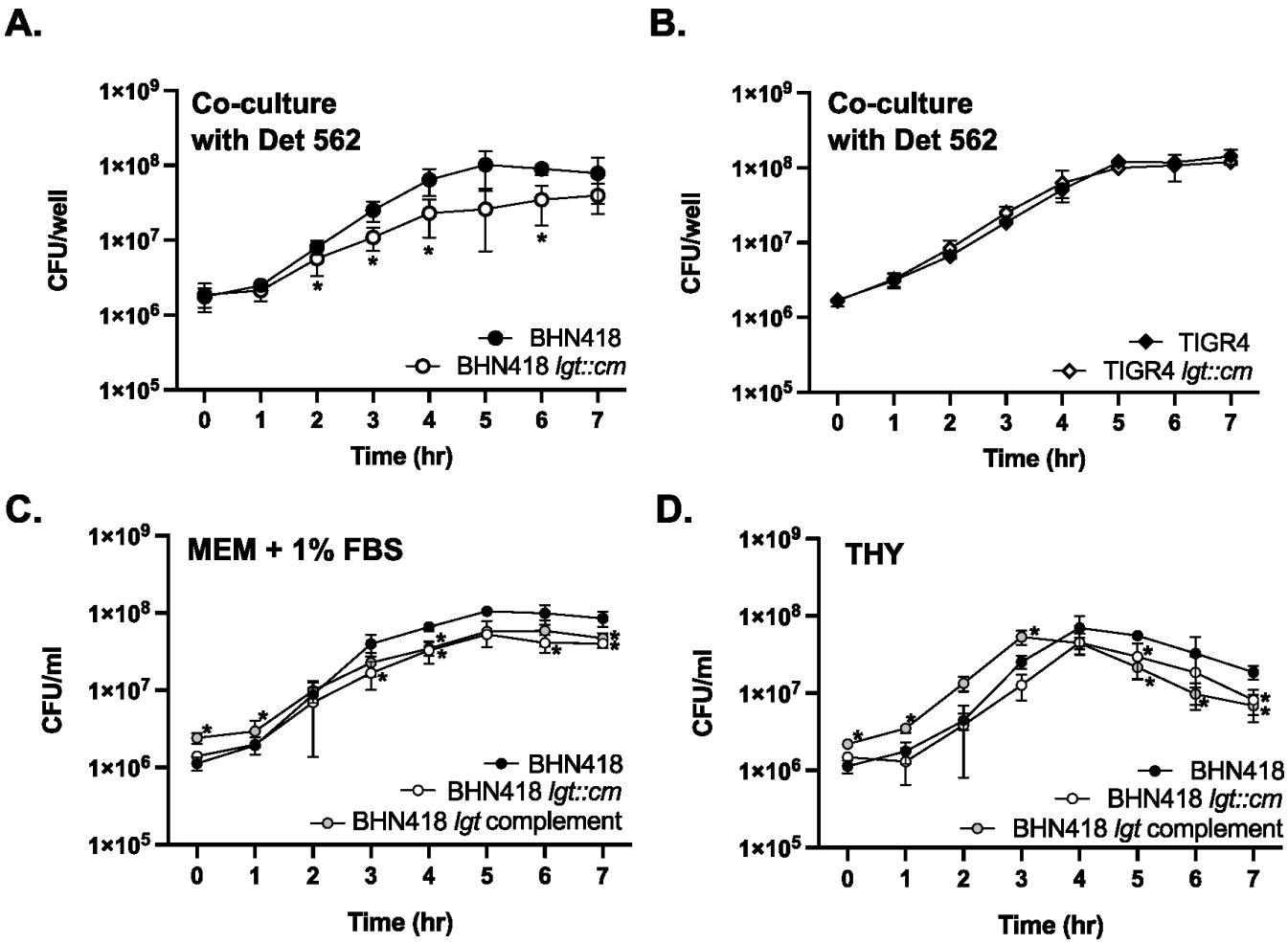

FIG 3 Growth of BHN418 *lgt::cm* compared to WT. (A and B) Growth of BHN418-derived (A) and TIGR4-derived (B) strains in infection medium (MEM + 1% FBS) in the presence of Detroit 562 NPE cells. (C and D) Growth of BHN418-derived strains in infection medium (C) and in the rich growth medium THY (D) in the absence of NPE cells. * indicates $P < 0.05$ by Student's *t* test.

The lipoprotein encoded by BHN418 but not TIGR4, encoded by the gene with the locus tag RSS80_03595 and which we named pneumococcal accesssory lipoprotein A (PalA), comprises of 525 amino acids with sequence and structural homology to extracellular solute binding domain proteins that deliver substrates to ABC family transporters (Fig. 4A and B). We modeled PalA's tertiary structure using AlphaFold2 and PHYRE, revealing a Type II periplasmic binding protein fold characterized by two subdomains connected with a hinge region (Fig. 4B; Fig. S4A) (44–46). The two predicted structures aligned well with each other, barring small conformational differences in the accessibility of the potential substrate binding pocket (Fig. S4B and C). The N-terminal extension (stalk-like structure) seen in Fig. 4B is likely cleaved post-lipidation (5).

Periplasmic (extracellular) binding proteins deliver substrates to ABC transporters, which are multi subunit proteins comprising of two transmembrane permease domains and two cytoplasmic ATPase domains (47). Two genes encoding ABC transporter permease domain proteins, annotated as *yteP* and *araQ*, were found ~3.3 kb and ~2.4 kb upstream of *palA*. We were unable to locate ORF(s) encoding for the ATPase domain proteins in the 10 kb region upstream or downstream of *palA*. Taken together, PalA likely binds to and delivers substrate(s) to YteP and/or AraQ. It is uncertain if YteP and/or AraQ co-opt the ATPase domains of ABC family transporters encoded elsewhere on the

TABLE 1 Lipoproteins encoded by TIGR4 and BHN418, identified bioinformatically using MEME suite

BHN418 locus tag	TIGR4 locus tag	Gene name	Description of gene product	Reference
RSS80_07140	SP_1500	*aatB*	Amino acid transporter	(23)
RSS80_10715	SP_2169	*adcA*	Adhesin competence protein A; zinc transporter	(24)
RSS80_04890	SP_1002	*adcAII*	Adhesin competence protein AII; zinc transporter	(25)
RSS80_01875	SP_0366	*aliA*	AmiA-like protein A; oligopeptide transporter	(26)
RSS80_07270	SP_1527	*aliB*	AmiA-like protein B; oligopeptide transporter	(26)
RSS80_09070	SP_1891	*amiA*	Aminopterin resistance locus protein A; oligopeptide transporter	(27)
RSS80_03115	SP_0629	*dacB*	L,D-carboxypeptidase	(10)
RSS80_03240	SP_0659	*etrx1*	Extracellular thioredoxin-like protein 1; thiol-disulfide oxidoreductase	(28)
RSS80_04875	SP_1000	*etrx2*	Extracellular thioredoxin-like protein 2; thiol-disulfide oxidoreductase	(29)
RSS80_06715	SP_1394	*glnH*	GlnH glutamine/polar amino acid ABC transporter substrate-binding protein	(30)
RSS80_00785	SP_0148	*gshT*	Glutathione transporter	(31)
RSS80_03685	SP_0749	*livJ*	Branched chain amino-acid transporter	(32)
RSS80_10385	SP_2108	*malX*	Maltosaccharide transporter	(33)
RSS80_00790	SP_0149	*metQ*	Methionine-binding lipoprotein Q	(34)
RSS80_05810	SP_1175	*phtA*	Pneumococcal histidine triad protein A	(35)
RSS80_05040	SP_1032	*piaA*	Pneumococcal iron acquisition protein A	(36)
RSS80_01305	SP_0243	*pitA*	Pneumococcal iron transporter protein A	(37)
RSS80_08955	SP_1872	*piuA*	Pneumococcal iron uptake protein A	(36)
RSS80_04120	SP_0845	*pnrA*	Nucleoside transporter	(38, 39)
RSS80_04795	SP_0981	*ppmA*	Putative proteinase maturation protein A; peptidyl-prolyl *cis–trans* isomerase	(40)
RSS80_07850	SP_1650	*psaA*	Pneumococcal surface adhesin A; manganese and zinc transporter	(11, 12)
RSS80_10265	SP_2084	*pstS*	Phosphate transport substrate-binding protein	(41)
RSS80_09095	SP_1897	*rafE*	Raffinose transporter	(42)
RSS80_03790	SP_0771	*slrA*	Streptococcal lipoprotein rotamase A;cyclophilin-type peptidyl-prolyl *cis–trans* isomerase	(43)
RSS80_04895[c]	SP_1003	*phtB*	Pneumococcal histidine triad protein B	(35)
RSS80_04895[c]	SP_1174	*phtD*	Pneumococcal histidine triad protein D	(35)
RSS80_06745	SP_1400	*pstS2*	phosphate binding protein	–
RSS80_10885	SP_2197	–[d]	ABC transporter binding protein	–
RSS80_00530	SP_0112	–	Amino acid-binding protein	–
RSS80_09960	SP_2041	–	Membrane protein insertase	–
RSS80_04185	SP_0857[b]	–	ABC transporter substrate-binding protein	–
RSS80_03070	SP_0620	–	Amino acid ABC transporter binding protein	–
RSS80_09570	SP_1975	–	Membrane protein insertase	–
RSS80_03435[b]	SP_0708[b]	–	ABC transporter substrate-binding protein (truncated)	–
RSS80_03595	**Not present**	**–**	**Extracellular solute binding protein**	**–**
RSS80_04430	SP_0899	–	Hypothetical protein	–
RSS80_05800[b]	Not assigned	–	ABC transporter substrate-binding protein (truncated)	–
RSS80_08015	SP_1683	–	ABC transporter sugar-binding protein	–
RSS80_08055	SP_1690	–	ABC transporter sugar-binding protein	–
RSS80_08595	SP_1796	–	Extracellular solute binding protein	–
RSS80_08765	SP_1826	–	ABC transporter substrate-binding protein	–
RSS80_00445	SP_0092	–	ABC transporter substrate-binding protein	–
RSS80_01055	SP_0191	–	Hypothetical protein	–
RSS80_01080	SP_0198	–	ABC transporter substrate-binding protein	–

[a]Bold: ORF present in BHN418 but not TIGR4. Not assigned: Homologous sequence present in genome but not annotated as ORF/CDS. Not present: Homologous sequence absent in genome. TIGR locus tag and gene name are based on TIGR4 genome annotation (Genbank accession number AE005672.3).
[b]Annotated as pseudogene, contained premature stop codon, or interrupted by insertion sequence.
[c]RSS80_04895 was the best match BLAST result for more than one TIGR4 CDS.
[d]–, no known common gene name for the designated gene/locus.

genome, in a similar strategy as the raffinose utilization system, or no longer function as transporters (42, 48).

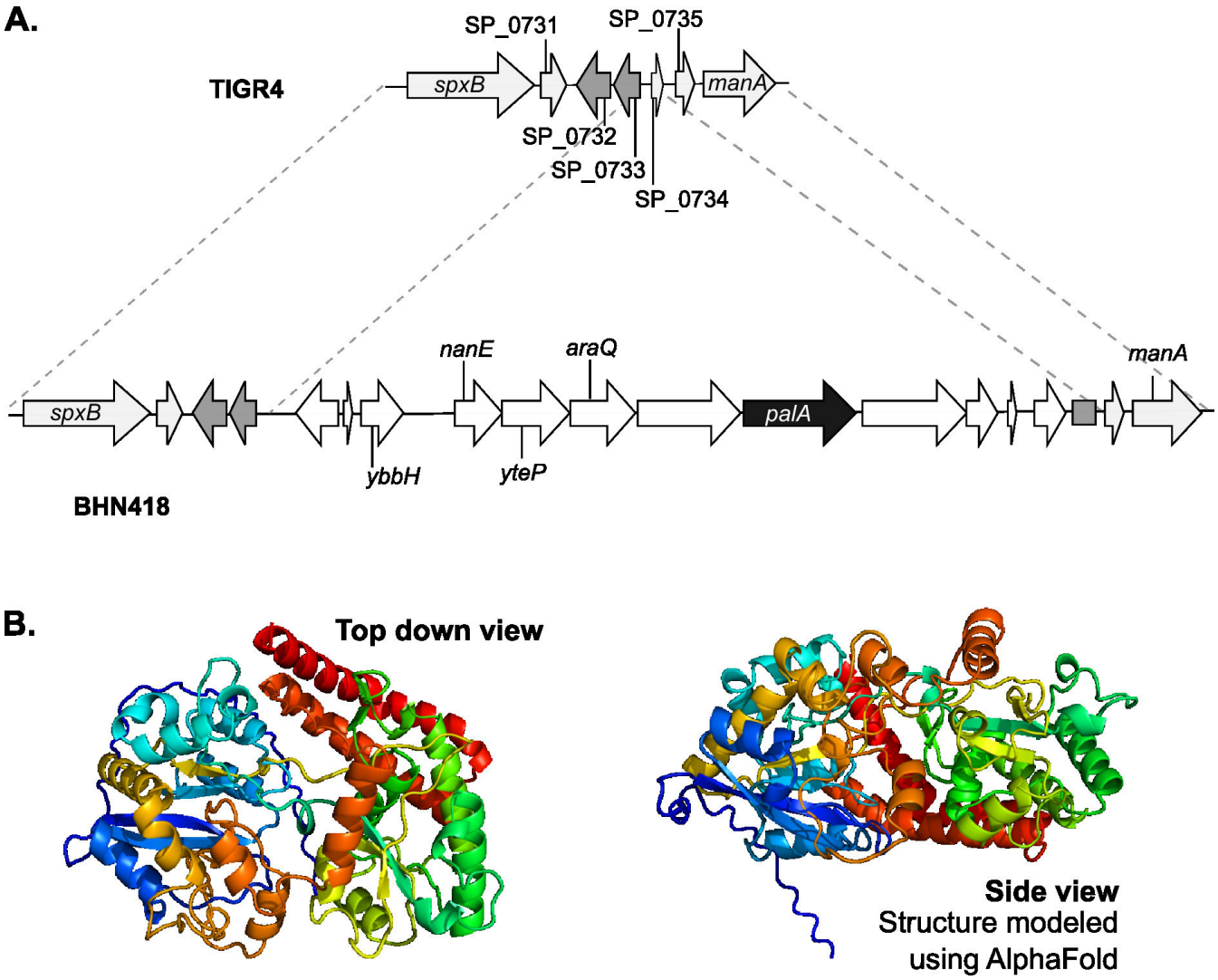

FIG 4 PalA is a lipoprotein present in BHN418 but not in TIGR4. (A) Genetic context of the *palA* gene (black arrow), which is within a multi-gene operon that is part of a putative genetic island. ORFs with homology to previously described genes are labeled with the gene name. (B) Predicted structure of PalA, modeled using AlphaFold2, which shows two subdomains connected by a hinge region with a possible central ligand-binding pocket. The extended stalk-like structure contains the signal peptide and lipoprotein processing sequence and is likely cleaved/absent in the mature lipoprotein.

This genetic context suggests that *palA* is the fifth gene in an operon encoding for carbohydrate import and utilization genes, which includes *yteP* and *araQ* (Fig. 4A). Upstream of the operon is a gene encoding a putative transcriptional regulator, *ybbH*, which may play a role in regulating the expression of the operon. The operon and *ybbH* sit within an 11.5 kb region with ~30% GC content, flanked by repetitive insertion sequences with homology to *IS630* elements. This region was likely acquired via horizontal gene transfer as the overall mean GC content of pneumococcal strains is around 40% (49). This putative genetic island is directly downstream of *spxB*, which encodes an important pneumococcal virulence factor involved in the production of H_2O_2 (50). Alignment of TIGR4 whole-genome sequencing reads to the BHN418 genome revealed the absence of the entire putative island in the TIGR4 genome (Fig. S5; 49).

To determine if *palA* is predominantly present in more carriage-type serotypes, such as serotype 6B, we examined the presence of *palA* in a well-curated data set of 2,806 carriage isolates from Malawi (51). Five hundred sixty-seven of these carriage isolates (20.5%) carry *palA* in their chromosome. Mapping the analysis results onto a hierarchal

clustering (Newick) tree showed that *palA* is present in specific lineages, with no clear association to capsular serotypes or sequence types (genetic relatedness, visualized as neighboring branches on a Newick tree; Fig. S6) (51). However, the presence of *palA* is enriched in certain serotypes, particularly serotype 6A (39/93, 41.9%), 6B (12/31, 38.7%), 10A (29/34, 85%), 15B (52.76, 68.4%), 16F (60/94, 63.8%), 23B (45/103, 43.7%), 35A (28/28, 100%), and 35B (60/114, 52.6%; Table S2). Additionally, the branching patterns of the phylogenetic tree for the Malawi carriage isolates support the inference that *palA* and its associated genetic island were acquired via horizontal gene transfer and expanded in specific lineages (Fig. S6).

PalA presence is enriched in carriage and ear isolates

Maintenance of this 11.5 kb genetic island is potentially costly and suggests that the island confers some form of advantage to isolates that carry it. *S. pneumoniae* is capable of colonizing and infecting multiple body sites including the nasopharynx, lungs, blood, cerebrospinal fluid, meninges, and middle ear. We, therefore, examined 51,379 genomes in the BIGSdb database to determine if there is an association between the presence of *palA* and the isolation site of the strain ("source") (52).

The *palA* gene presence is enriched in carriage isolates and in strains isolated from ear infections compared to strains isolated from invasive pneumococcal disease (IPD) or lower respiratory tract disease (Table 2). More than half of serotype 22F and 6A strains isolated from the ear carried *palA* and approximately 28% of all serotype 22F and 6A genomes in the database carry *palA*, in contrast to the overall *palA* prevalence rate of 9.97% (Table 2). By contrast, *palA* was detected in only 6.7% of hypervirulent serotype 1 genomes hosted on the BIGSdb database and in only 5 of 895 genomes belonging to the multidrug-resistant lineage GPSC10 on the Global Pneumococcal Sequencing database (~0.55%). In the BIGSdb database, the only serotype 4 strain isolated from the ear carried *palA* in its genome. These observations suggest that *palA* and/or its putative genetic island may facilitate spread to and cause infection of the ear, although *palA*'s presence is not necessary for colonization of the ear.

Mutation of *palA* does not alter pneumococcal colonization or microinvasion of the epithelium

To determine if PalA plays a role in epithelial microinvasion, we generated *palA* deletion and complementation mutants for testing in our NPE model. The deletion and complementation mutants were verified using Sanger sequencing, Illumina sequencing, and semi-quantitative PCR of *palA* transcript (Table S1; Fig. S2). Although there is a small reduction in the number of planktonic bacteria, the numbers of epithelial-associated and intracellular BHN418 *palA::kan* were not significantly different from that of WT BHN418 (Fig. 5A through C). Additionally, we did not observe a growth defect when BHN418 *palA::kan* was grown in THY or infection medium (Fig. 5D and E).

Mutation of *lgt* attenuates nasopharyngeal colonization density and duration in mice (14). We next asked if the presence of *palA* confers a survival advantage in a more complex and immune-replete environment such as the murine nasopharynx. Outbred CD-1 female mice were intranasally inoculated with wild-type BHN418 and the *palA*

TABLE 2 Presence of *palA* in whole-genome sequences of pneumococcal isolates on the BIGSdb database, stratified by site of isolation ("source")

Category	Source (BIGSdb label)	Proportion (%)
Carriage	"Nasopharynx," "pharynx," and "sputum"	14.57 (3,114/21,369)
Otitis	"Ear swab" and "middle ear fluid"	13.72 (129/940)
Pneumonia	"Lung aspirate," "sinus aspirate," "bronchoalveolar lavage," and "bronchi"	8.96 (25/279)
Invasive	"Blood," "cerebrospinal fluid," "joint fluid," and "pleural fluid"	6.54 (1,837/28,075)
Eye/pus/others	"eye swab," "pus," and "other"	2.65 (19/716)
Overall		9.97 (5,124/51,379)

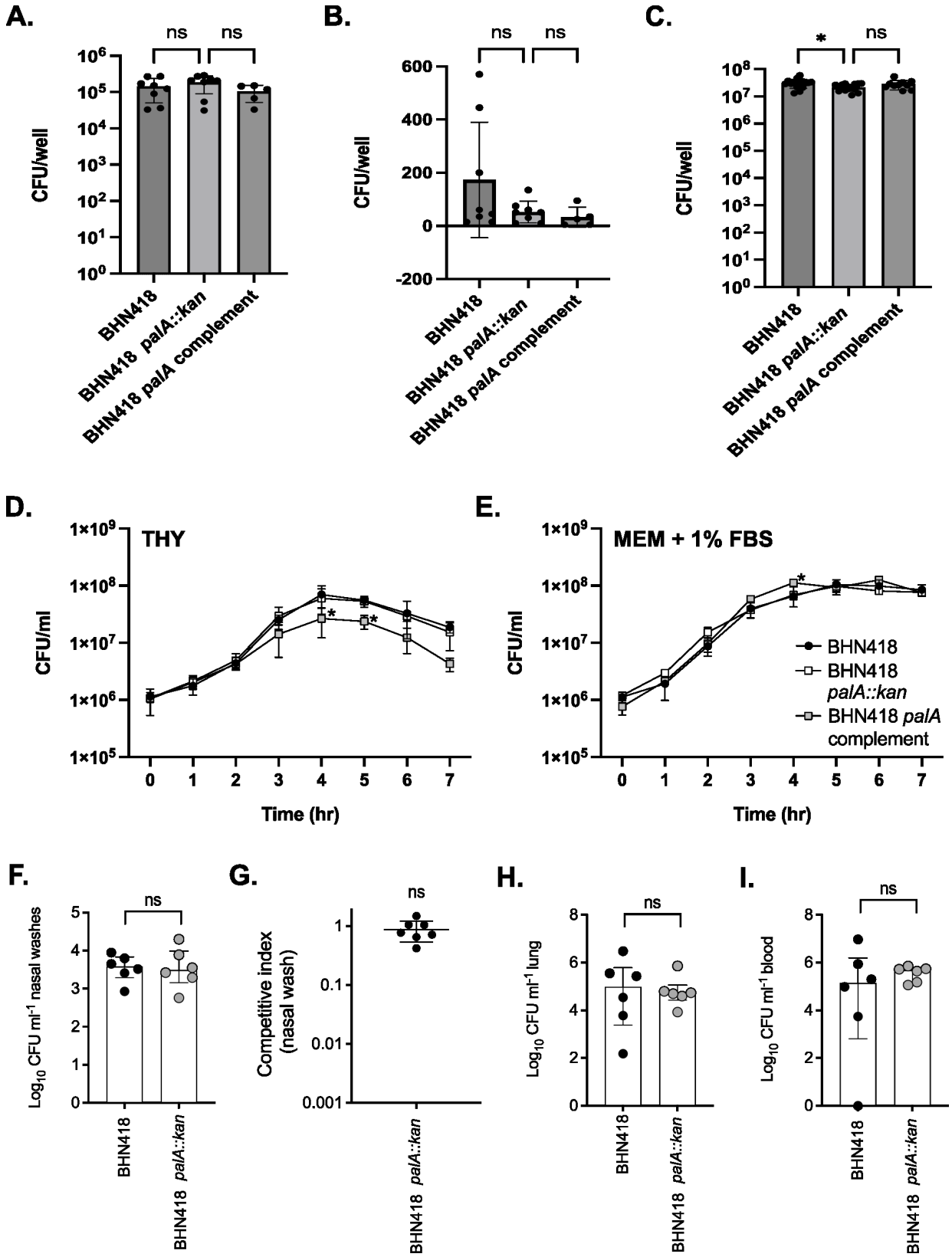

FIG 5 PalA was not essential for NPE microinvasion, murine colonization, and progression to disease. (A–C) NPE microinvasion by WT BHN418 and *palA* mutants, measured as NPE-associated bacteria (A), internalized bacteria (B), and planktonic bacteria growing in proximity with Detroit 562 NPE cells 3 hours post infection. (D and E) Growth of WT BHN418, the *palA* knock out, and complementation mutants in THY (D) and infection medium (E). (F–I) Recovery of pneumococci from mice intranasally inoculated with WT BHN418 and *palA::kan* mutant, recovered from nasal washes when inoculated singly (F) or competitively in a 1:1 ratio (G), as well as from the lungs (H) and bloodstream (I) when tested on a pneumonia model. * indicates $P < 0.05$.

mutant either singly or in a 1:1 competitive mix. After 7 days of colonization, similar CFU numbers for WT BHN418 and the *palA::kan* were recovered from nasal washes (Fig. 5F and G). Similar CFU numbers for BHN418 and *palA::kan* were also recovered in homogenized lungs and blood 24 hours post inoculation in a murine pneumonia model (Fig. 5H and I). We conclude that the presence of *palA* does not confer a colonization advantage in the murine nasopharynx or in the progression to bacteremic pneumonia.

To further probe the function of PalA, we heterologously expressed *palA* in a serotype 23F strain naturally lacking the island (P1121). Semi-quantitative PCR of *palA* transcripts demonstrated substantial *palA* expression driven by the inducible promoter (Fig. S2C). Expression of *palA* in P1121 did not increase the microinvasion potential of the resulting strains and reduced the number of planktonic bacteria in the cell culture supernatant (Fig. 6A through C). Moreover, the BHN418 *palA* knockout strains and the P1121 *palA* knock-in strains activated TLR2 signaling to similar levels as their respective wild-type strains (Fig. 6D). We, therefore, conclude that the presence of *palA* is not solely responsible for the observed strain-specific differences in Lgt-mediated epithelial microinvasion. Moreover, PalA does not contribute significantly to pneumococci's ability to activate TLR2.

DISCUSSION

In this study, we demonstrated that pneumococcal lipoproteins trigger inflammation during epithelial colonization at least partially via the TLR2-dependent pathway. We have previously shown that epithelial microinvasion can occur in the absence of disease, and there is heightened epithelial inflammation around the time of pneumococcal clearance in controlled human infection (2). In murine models, mutation of *lgt* reduces carriage duration and attenuates disease, associated with a concomitant reduction in inflammatory and immune responses (8, 9, 53). Although we have shown that BHN418 *lgt::cm* but not TIGR4 *lgt::cm* was significantly attenuated in epithelial adherence and microinvasion compared to their respective wild-type strains, this does not appear to be TLR2-dependent or due to differential lipoprotein repertoires encoded by the strains.

We additionally observed that the presence of *lgt* and therefore TLR2 activation heightens the epithelial interferon response elicited by pneumococcal microinvasion in the absence of immune cells. Induction of the interferon pathway upon pneumococcal challenge is thought to be dependent on sensing of intracellular pneumococci, pneumococcal DNA, or cellular DNA damage by the infected cells (54–59). In infant mice co-infected with pneumococci and influenzae, interferon signaling increases bacterial shedding while protecting against invasive disease, while TLR2 signaling limits bacterial shedding and transmission (18, 60, 61). Since TLR2 signaling augments the interferon response by human nasopharyngeal epithelial cells during mono-pneumococcal infection, it is unclear how these two pathways mediate the outcomes of pneumococcal microinvasion, colonization, or progression to disease in people. Nonetheless, our results suggest that mucosal innate immunity could be targeted alongside the induction of an adaptive immune response to prevent pneumococcal colonization, transmission, and invasive disease.

We have implicated lipoprotein expression in intraspecies differences in NPE cell microinvasion, but our genetic and mutational analysis lead us to suggest that this is not due to differences in lipoprotein repertoire but potentially due to differential post-translational lipidation kinetics. The lipoprotein PsaA functions as a bacterial adhesin that binds to host cell E-cadherin (62). In theory, we should have observed a significant reduction in the number of epithelial-associated bacteria for both the TIGR4 and BHN418 *lgt* mutants compared to wild-type due to the loss of PsaA surface presentation; however, our data did not reflect this hypothesis. Similar subtle effects have been seen, for example, when mutating the lipoprotein encoding gene *dacB* but not *lgt* alters murein sacculus composition, and when mutating *lgt* results in strain-dependent variable effects on growth in rich medium (5, 9, 14, 53, 63). The observation that genetic complementation of *lgt* in BHN418 restores the pneumococcal ability to trigger epithelial

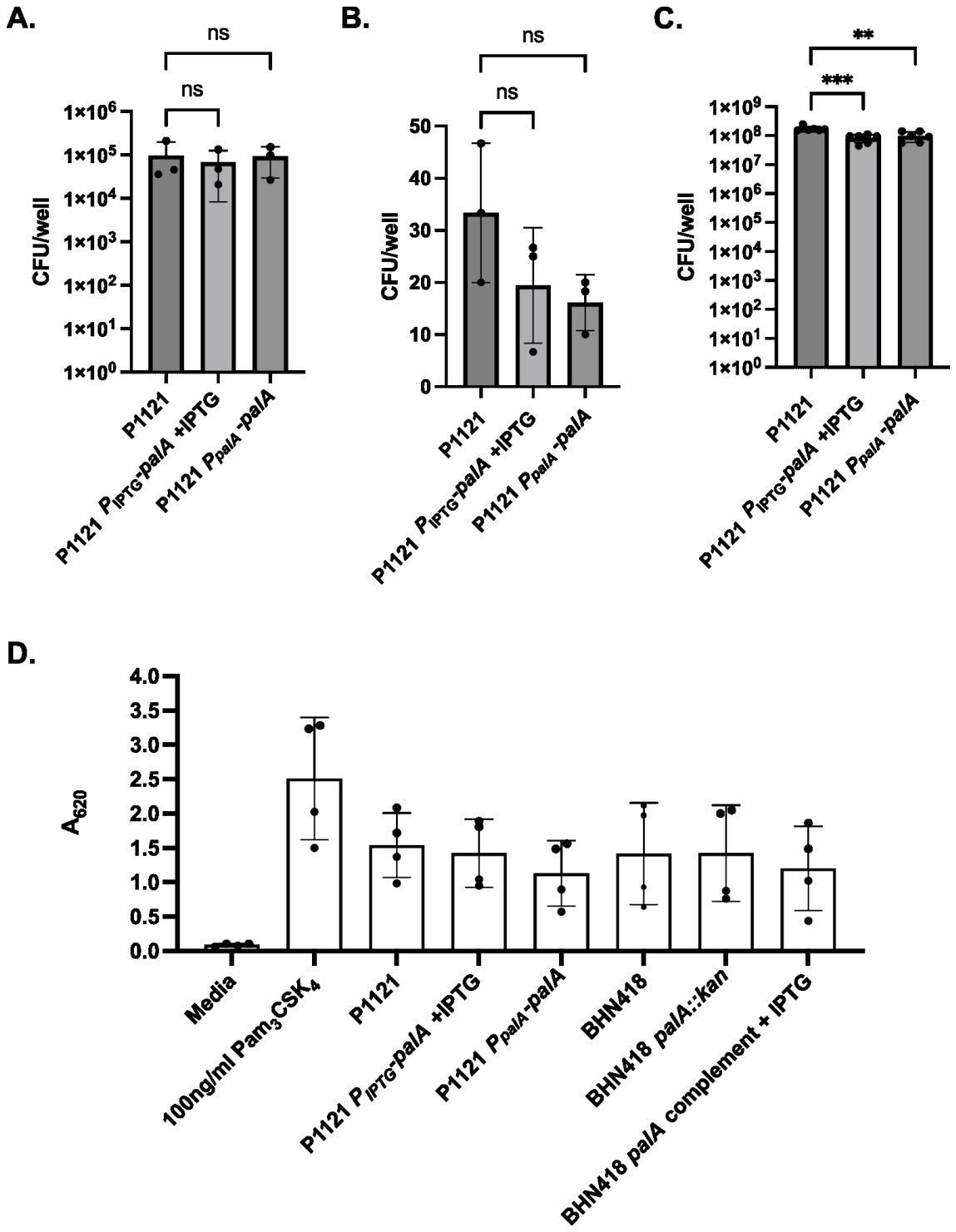

FIG 6 Heterologous expression of *palA* in P1121 (serotype 23F) did not increase epithelial microinvasion or TLR2 signaling. (A–C) NPE microinvasion by WT P1121 and *palA* expression mutants, measured as NPE-associated bacteria (A), internalized bacteria (B), and planktonic bacteria growing in proximity with Detroit 562 NPE cells 3 hours post infection (C). (D) SEAP reporter readout from HEK-Blue hTLR2 reporter cells treated with pneumococcal strains at MOI 10 for 16 hours.

inflammation and retain PiuA and other lipoproteins in whole-cell lysate, but does not restore WT levels of microinvasion, further suggests that regulation of lipoprotein processing may be more complex than previously thought. The assumption that Lgt and

other lipoprotein modification proteins are constitutively expressed and active and all 30+ lipoproteins are processed at similar rates may not be entirely correct.

We, therefore, speculate that unlike Gram-negative bacteria, for which lipoprotein processing is essential, pneumococci and other Gram-positive bacteria may compensate for the loss of Lgt by differentially regulating the expression of lipoprotein encoding genes, which in turn is regulated by external stimuli in nasopharyngeal niche (64, 65). Proteomics and immunoblotting analyses showed that the abundance of specific lipoproteins increase, decrease, or show no change when lipoprotein processing is disrupted in *S. pneumoniae*, although no clear patterns were apparent (5, 13). Additionally, we and others have observed that mutation of *lgt* does not lead to complete loss/release of preprolipoprotein from the cell surface (Fig. S3) (5). Future investigation into whether pneumococci vary *lgt* expression or Lgt activity to thrive in different tissue niches, in tandem with compensatory regulation of lipoprotein expression, may be warranted. Such a study may reveal novel mechanisms of innate immune evasion. It would also be informative to determine whether there are strain-dependent differences in regulating the generation of TLR2 agonists (mature lipoproteins) and how this may affect the invasiveness and virulence of different pneumococcal lineages.

While investigating the intraspecies variation in the role of *lgt* in microinvasion, we discovered a previously uncharacterized lipoprotein encoding gene (*palA*) and its associated genetic island. Although *palA* presence is enriched in the carriage and otitis media isolates, we have not been able to demonstrate a clear role for PalA in epithelial microinvasion or in a murine model of colonization or disease. We evaluated bacterial burden in mice 7 days post infection and thus cannot eliminate the possibility that PalA plays a role in the early phases of colonization or in persistence. PalA is predicted by sequence and structure homology to be involved in carbohydrate or sugar transport, although we have yet to identify a substrate for PalA. Raffinose metabolism has been shown to contribute to lung vs ear tropism in serotype 3 and serotype 14 strains (48). It is, therefore, possible that *palA* functions in promoting niche specialization by facilitating the uptake and metabolism of uncommon sugars, such as raffinose, found in the nasopharynx or middle ear.

In conclusion, we demonstrated a role for pneumococcal surface lipoproteins in triggering epithelial inflammation and augmenting interferon signaling in response to pneumococcal-epithelial interactions. We show that pneumococcal lipoproteins mediate microinvasion in a strain-dependent manner, which may explain the significant attenuation in carriage duration and disease with *lgt* mutants reported by others (8, 14, 53). Additionally, we have characterized a novel accessory lipoprotein likely acquired through horizontal gene transfer but rejected the hypothesis that this lipoprotein contributes to strain differences in pneumococcal epithelial microinvasion. Instead, we postulate that differential regulation of lipoprotein gene expression responding to the nasopharyngeal niche regulates this microinvasion process.

MATERIALS AND METHODS

Bacterial growth and maintenance

S. pneumoniae strains were grown on Columbia agar base with 5% defibrinated horse blood (CBA plates; EO Labs, Oxoid) statically in Todd-Hewitt broth supplemented with 0.5% yeast extract (THY; Oxoid) or in brain heart infusion broth (BHI; Oxoid) at 37°C, 5% CO_2. Where appropriate, growth medium was supplemented with antibiotics at the following concentrations: chloramphenicol (10 µg/mL), erythromycin (0.5 µg/mL), and kanamycin (250 µg/mL). Working stocks for infections were prepared by freezing THY cultures at OD_{600} 0.3–0.4 with 10% glycerol. NEB Stable competent *Escherichia coli*-derived strains were grown in Luria-Bertani (LB) broth or LB agar (Difco) supplemented with ampicillin (200 µg/mL) where appropriate. Bacterial strains used in this paper are listed in Table 3.

TABLE 3 Bacterial strains and plasmids used in this study

Designation	Genotype/description	Source
Strains		
TIGR4	WT serotype 4 isolate	(49)
BHN418	WT serotype 6B isolate	(16)
P1121	WT serotype 23F isolate	(2)
ECSPN100	TIGR4 *lgt::cm*	This work
ECSPN106	TIGR4 *lgt::cm* P_{IPTG}-*lgt-erm* (TIGR4 *lgt* complementation)	This work
ECSPN200	BHN418 *lgt::cm*	This work
ECSPN210	BHN418 *lgt::cm* P_{IPTG}-*lgt-erm* (BHN418 *lgt* complementation)	This work
ECSPN211	BHN418 *palA::kan*	This work
ECSPN213	BHN418 *palA::kan* P_{IPTG}-*palA-erm* (BHN418 *palA* complementation)	This work
ECSPN400	P1121 P_{IPTG}-*palA-erm*	This work
ECSPN401	P1121 P_{palA}-*palA-kan*	This work
Plasmids		
pASR103	Complementation construct with an IPTG inducible promoter and *erm* selectable marker	(66)
pPEPY	Complementation construct with a *kan* selectable marker	(66)
pEMcat	Minitransposon plasmid; source of cm^R cassette	(67)
pABG5	Cloning plasmid; source of kan^R cassette	(68)
pEC210	TIGR4 *lgt* coding region cloned into pASR103	This work
pEC211	BHN418 *lgt* coding region cloned into pASR103	This work
pEC213	BHN418 *palA* coding region cloned into pASR103	This work
pEC213	BHN418 *palA* promoter and coding region cloned into pPEPY	This work

Bacterial genetic manipulation

S. pneumoniae were genetically manipulated using a competence-stimulating peptide (CSP)-mediated transformation assay (69). Briefly, pneumococci were grown in THY pH 6.8 supplemented with 1 mM $CaCl_2$ and 0.02% bovine serum albumin (BSA) at 37°C, 5% CO_2 to OD_{600} 0.01–0.03, pelleted and resuspended in 1/12 vol THY pH 8.0 supplemented with 1 mM $CaCl_2$ and 0.2% BSA. A total of 400 ng CSP (Cambridge Biosciences; CSP-2 for TIGR4; 1:1 ratio of CSP-1:CSP-2 for BHN418; CSP-1 for P1121) was added to the bacterial suspension and incubated at RT for 5 min. The suspensions were then mixed with ~300 ng transforming DNA, incubated at 37°C, 5% CO_2 for 2 hours, and plated on CBA plates supplemented with relevant antibiotics. Antibiotic-resistant transformants were screened using colony PCR and confirmed by sequencing.

Transforming DNA for generating *lgt::cm* and *palA::kan* mutants was generated using overlap-extension PCR. Complementation and expression constructs were generated by inserting the target gene into the complementation plasmid pASR103 or pPEPY (66), which allows for integration of the construct at a chromosomal ectopic site. Plasmids used are listed in Table 3, while primers are listed in Table S3.

Cell culture

Detroit 562 (ATCC CCL-138; human pharyngeal carcinoma epithelial cells) was expanded and maintained in MEMα (minimum essential medium; Gibco 22561021) supplemented with 10% heat-inactivated fetal bovine serum (HI-FBS; LabTech FB-1001/500 or Gibco 10438–026) at 37°C, 5% CO_2. HEK-Blue hTLR2 reporter cells (Invivogen, hkb-htlr2) were expanded and maintained in Dulbecco's modified Eagle medium (DMEM; 4.5 g/L glucose, 2 mM glutamine, and sodium pyruvate) supplemented with 10% HI-FBS at 37°C, 5% CO_2. Per the manufacturer's instructions, DMEM growth medium was supplemented with 100 µg/mL normocin and/or 1× HEK-Blue Selection (Invivogen) where appropriate.

NPE infections

Adherence-invasion infections of confluent Detroit 562 cells with *S. pneumoniae* strains were performed at multiplicity of infection (MOI) 20 (P1121/23F derived strains) or MOI 10 (all others) for 3 hours. Working bacterial stocks were thawed, centrifuged to remove the freezing medium, and resuspended in the infection medium (MEMα with 1% HI-FBS) to the appropriate colony forming unit (CFU). One milliliter bacterial suspension was added to each well containing confluent Detroit 562 cells. Plates were incubated statically at 37°C, 5% CO_2 for 3 hours, after which 10 µL of the supernatant was removed for CFU enumeration. For adherence assays, cells were washed thrice with phosphate buffered saline (PBS), lysed with cold 1% saponin (10 min incubation at 37°C, followed by vigorous pipetting), and 10 µL cell lysate removed for CFU enumeration. For invasion assays, cells were washed thrice with PBS, incubated with 0.5 mL infection medium supplemented with 200 µg/mL gentamicin at 37°C, 5% CO_2 for 1 hour to kill extracellular bacteria, followed by 3× PBS wash, lysis with 1% saponin and CFU enumeration. Experiments were performed at least thrice on different days ($n \geq 3$ biological replicates) with technical duplicates. Statistical significance was determined using one-way analysis of variance (ANOVA) with Bonferroni's multiple comparison test.

To harvest RNA for qPCR, confluent Detroit 562 cells were treated with synthetic agonists or infected with *S. pneumoniae* strains at MOI 10 for 6 hours. Briefly, working bacterial stocks were thawed, centrifuged to remove the freezing medium, and resuspended in infection medium (MEMα with 1% HI-FBS) to the appropriate CFU. Bacterial suspensions, infection medium (negative control), or infection medium supplemented with synthetic agonists [20 µg/mL Poly(I:C); TLR3 agonist, Bio-Techne] were added to each flask. Flasks were incubated statically at 37°C, 5% CO_2 for 6 hours, after which 10 µL were removed for CFU enumeration. Detroit 562 cells were washed thrice with PBS and harvested by scraping into 300 µL RNA*later* (ThermoFisher). For each treatment condition, RNA harvesting was performed at least thrice on different days ($n \geq 3$ biological replicates) without technical replicates.

For growth curve experiments, *S. pneumoniae* strains were seeded into 1 mL THY or 1 mL infection medium (MEMα with 1% HI-FBS, LabTech) with and without confluent Detroit 562 cells in 12-well plates at a similar CFU number as used in infection experiments. Plates were incubated at 37°C, 5% CO_2 for 7 hours, with aliquots taken for CFU enumeration every hour. CFU growth curves were performed at least thrice on different days ($n \geq 3$ biological replicates) without technical replicates. Statistical significance was determined using Student's *t* test assuming equal variance.

Quantitative and semi-quantitative PCR

RNA from epithelial cells stored in RNA*later* was extracted using RNeasy Mini kit (Qiagen) according to manufacturer instructions. Carryover DNA was removed with TURBO DNA-*free* kit (Ambion), and cDNA was generated using LunaScript RT Supermix kit (NEB). Quantitative PCR (qPCR) was performed using Luna Universal qPCR Master Mix (NEB) in technical triplicates with primers specific for *GAPDH*, *CXCL10*, *IFNB1*, *IFNL1*, and *IFNL3* (Table S3). Whenever possible, qPCR primers were designed to span exon-exon junctions. Cycling conditions are as follows: 95°C for 5 min, 40 cycles of 95°C for 15 s, and 60.5°C for 45 s, with a plate read at the end of each cycle. Data were analyzed using the $2^{\Delta\Delta Ct}$ method, with media-only control and *GAPDH* levels for normalization. Statistical significance was determined using Student's *t* test assuming equal variance.

RNA from *S. pneumoniae* stored in RNA*later* was extracted using a modified TRIzol (Ambion) protocol. Briefly, pneumococcal strains were grown in BHI to $OD_{600} \sim 0.5$, harvested by centrifugation at 8,000 × *g* for 8 min, resuspended in 300 µL RNA*later*, and saved at −70°C. On the extraction day, the suspensions were thawed and subjected to centrifugation at 8,000 × *g* for 8 min, followed by removal of RNA*later* and resuspension of the bacterial pellet in 1 mL TRIzol reagent. The whole 1 mL suspension was transferred to pre-chilled VK01 Precellys lysing tubes containing glass beads (Stretton Scientific) and subjected to lysis by bead beating (Precellys Evolution; 6,200 RPM, 4 × 45 s cycles

with 20 s rest in between cycle). The TRIzol lysates were then centrifuged at 5,000 × *g* for 10 min at 4°C, and ~800 µL supernatant was transferred to a new, pre-chilled centrifuge tube. RNA extraction and cDNA generation were performed as described above. Semi-quantitative PCR was performed using OneTaq Quick-Load Master Mix (NEB) using primers against *lgt*, *palA*, and 16S rRNA with cycling conditions: 95°C for 5 min, 40 cycles of 95°C for 15 s, and 60.5°C for 45 s. Amplification products were analyzed using DNA gel electrophoresis.

Immunoblotting

Five milliliter cultures of pneumococcal strains grown in BHI to OD ~ 0.5 were harvested by centrifugation at 3,900 × *g* for 15 min, resuspended in 1 mL 1× PBS with 0.1% NP-40, and transferred into pre-chilled VK01 Precellys lysing tubes containing glass beads (Stretton Scientific). Pneumococcal suspensions were lysed by bead beating (Precellys Evolution; 6,200 RPM, 4 × 45 s cycles with 20 s rest in between cycle), followed by centrifugation at 5,000 × *g* for 10 min at 4°C to pellet debris. Approximately 800 µL supernatant (whole-cell lysates) was transferred to a fresh centrifuge tube and saved at −70°C until further use. Protein concentration was determined using Bradford reagent (Thermo Scientific) per the manufacturer's instructions and used to normalize the amount of whole-cell lysate used in immunoblotting. Lysates were mixed with loading buffer, incubated at 70°C for 10 min, and chilled on ice prior to gel loading.

Approximately 3.5 µg and 2 µg whole-cell lysate were loaded on NuPage 4%–12% Bis-Tris protein gels (Invitrogen) and subjected to gel electrophoresis and transferred onto nitrocellulose membranes. Membranes were blocked in 1× Tris-buffered saline (TBS) pH 7.4 with 0.05% Tween-20 and 5% skim milk (hereafter blocking buffer) at RT for 1 hour, washed thrice in 1× TBS pH 7.4 with 0.05% Tween-20 (wash buffer; 5 min incubation at RT per wash), and probed with antisera from mice inoculated with polysaccharide conjugated-PiuA (1:1,000) or human intravenous immunoglobulin (1:1,000) (Vigam Liquid) in blocking buffer overnight at 4°C (70). Membranes were washed thrice, incubated with IRDye 800W-conjugated polyclonal goat α-mouse or goat α-human antibody in blocking buffer (1:10,000; Abcam) at RT for 1 hour, washed thrice, and imaged using LiCor Odyssey CLx. Equal loading was checked using PonceauS staining.

HEK-Blue hTLR2 reporter assay

HEK-Blue hTLR2 secreted alkaline phosphatase (SEAP) reporter assays were performed according to manufacturer instructions (Invivogen, hkb-htlr2). Briefly, HEK-Blue hTLR2 cells, *S. pneumoniae*, and control reagents were resuspended or diluted in pre-warmed HEK-Blue Detection medium (Invivogen). 5×10^4 HEK-Blue hTLR2 cells were mixed with 5×10^5 CFU *S. pneumoniae* (MOI 10) and incubated for 16 hours at 37°C, 5% CO_2. SEAP activity was then measured spectroscopically at A_{620}. One hundred nanograms per milliliter of Pam_2CSK_4 and Pam_3CSK_4 (TLR2 agonist, Bio-Techne) was used as positive controls, while a bacterium-free medium was used as a negative control. Experiments were performed at least thrice on different days ($n \geq 3$ biological replicates) with technical triplicates. Statistical significance was determined using one-way ANOVA with Bonferroni's multiple comparison test.

Lipoprotein prediction using MEME suite

Amino acid sequences of 39 published D39 lipoproteins were used with the motif discovery tool MEME to identify pneumococcal lipoprotein motif(s) (5, 13, 22). The top two MEME results were combined to obtain motif: L[LA][AS][AL]LXL[AV]A**C**[SG][NQS], a modified extension of the minimal lipobox motif LAG**C** (5).

The obtained motif was used with the motif scanning tool FIMO to identify lipoproteins in the genomes of *S. pneumoniae* TIGR4, BHN418, and D39, with the latter used for quality control (71). The match *P*-value was set to 0.001. FIMO results were further

filtered with the following criteria: (i) presence of the lipidated cysteine residue in the motif, (ii) presence of motif in the first 70 a.a. of the sequence, and iii) positive prediction as lipoprotein by SignalP-6.0 (72).

Genomic analysis

The presence of *palA* and its associated genetic island was determined using Local-BLAST (BLASTN and TBLASTN) for the Malawian carriage data set (n = 51,379) and serotype 23F strain P1121 (73). The built-in BLAST tool on pubmlst.org was used for the analysis of the BIGSdb data set (52). BLASTN and TBLASTN tools on the NCBI database were used to identify *palA* and PalA homolog in non-pneumococcal species (73, 74). BLAST results were exported in csv format and further analyzed using R (v3.6.0) in RStudio (http://www.rstudio.com/). The presence/absence of *palA* was annotated onto a Newick tree showing the phylogeny of the Malawian carriage strains by metabolic type and visualized using iTOL (51, 75). Potential gene functions were inferred through the results of BLASTP and NCBI Conserved Domain Database searches (73, 74, 76).

The BHN418 genome assembly was generated by combining long-read sequencing (PacBio) and short-read sequencing (Illumina) methods which resulted in a single contiguous chromosome of BHN418 of length 2,107,426 bp. *De novo* assembly was performed using the Unicycler v0.4.8 pipeline in bold mode, quality assessed using QUAST v5.1.0rc1, and annotated using Bakta v1.8.2 as described previously (77–80). TIGR4 sequencing reads were aligned to the BHN418 genome using Samtools v1.14 and visualized using IGV v2.16.1 (49, 74).

Genomes of TIGR4 and/or BHN418 *lgt::cm*, *palA::kan*, and respective complementation strains were assembled *de novo* using SPAdes 3.15.5 with standard parameters and subjected to Local-BLAST (TBLASTN) to determine the presence or absence of *lgt*, *palA*, *cm*, and *kan* in the respective genomes (74, 81).

TLR2 transcriptional module analysis

TLR2-mediated transcriptional activity in Detroit 562 cells infected with TIGR4 and BHN418 for 3 hours was determined using published RNAseq data (2). We generated a transcriptional module reflective of TLR2 activity derived from genes overexpressed in fibroblasts stimulated with TLR2 agonists Pam_2CSK_4 and/or FSL-1 for 6 hours relative to unstimulated controls [>1.5-fold; paired t test with α of $P < 0.05$ without multiple testing correction; Gene Expression Omnibus (GEO) data set GSE92466; Fig. S1A] (19). Module expression was determined by calculating the geometric mean expression of all constituent genes found in the analyzed RNAseq data set. Performance was validated using data derived from Acute Myeloid Leukemia cells (GEO data sets GSE92744) and CD14+ monocytes stimulated with Pam_3CSK_4 (GEO data set GSE78699; Fig. S1B and C) (82, 83).

Murine experiments

Outbred female CD1 mice (Charles River Laboratories) were inoculated intranasally under anesthetic (isoflurane) with 1×10^7 CFU bacteria (n = 6 for single inoculation colonization and pneumonia model, n = 7 for competition experiment). For colonization experiments, nasal washes were performed 7 days post infection using 1 mL PBS. For pneumonia model, mice were sacrificed 24 hpi, and bacteria were recovered from the blood and homogenized lungs. CFU numbers were enumerated using CBA supplemented with 4 µg/mL gentamicin, with an additional 250 µg/mL kanamycin where appropriate. All animal procedures were approved by the local ethical review process and conducted in accordance with the relevant UK Home Office-approved project license (PPL70/6510). Mice were housed for at least 1 week under standard conditions before use. Randomization or blinding was not performed for these experiments. Statistical significance was determined using Mann-Whitney test.

ACKNOWLEDGMENTS

We would like to thank A. Gori and CJ Yang for bioinformatics support and the Jolly, Tomlinson, and Towers groups at UCL for sharing supplies and equipment.

This study was funded by a Medical Research Council grant (MR/T016329/1) awarded to R.S.H. and J.S.B. which supported J.M.C. and C.M.W. J.M.C. and M.B. were also supported by funding from NIHR Global Health Research Unit on Mucosal Pathogens at UCL, commissioned by the National Institute for Health Research using Official Development Assistance (ODA) funding. G.P. is supported by funding from the UCLH NIHR Biomedical Research Center. R.S.H. is a NIHR Senior Investigator. The views expressed are those of the authors and not necessarily those of the NIHR.

AUTHOR AFFILIATIONS

[1]Research Department of Infection, Division of Infection and Immunity, University College London, London, United Kingdom
[2]UCL Respiratory, Division of Medicine, University College London, London, United Kingdom

PRESENT ADDRESS

Modupeh Betts, Institute of Infection, Veterinary and Ecological Sciences, University of Liverpool, Liverpool, United Kingdom
Caroline M. Weight, Department of Biomedical and Life Sciences, Lancaster University, Lancaster, United Kingdom

AUTHOR ORCIDs

Jia Mun Chan http://orcid.org/0000-0003-4223-7527
Modupeh Betts http://orcid.org/0000-0003-0083-7487
Robert S. Heyderman http://orcid.org/0000-0003-4573-449X

FUNDING

Funder	Grant(s)	Author(s)
UKRI | Medical Research Council (MRC)	MR/T016329/1	Jia Mun Chan Caroline M. Weight Jeremy S. Brown Robert S. Heyderman
National Institute for Health and Care Research (NIHR)	16/136/46	Jia Mun Chan Modupeh Betts Robert S. Heyderman

DATA AVAILABILITY

BHN418, ECSPN100, ECSPN106, ECSPN200, ECSPN210, ECSPN211 and ECSPN213 genomes were deposited to NCBI with BioProject accession number PRJNA1022026 (BHN418) and PRJNA1087740 (everything else). TIGR4 sequencing reads were downloaded from the NCBI Sequence Read Archive (accession SRX6259281), while P1121 reads were downloaded from the EMBL-EBI database (accession ERS1072059) (85, 86). D39 and TIGR4 whole genome assemblies were downloaded from the NCBI GenBank database (accession numbers CP000410.2 and AE005672.3, respectively) (87). All other genomic sequences used were hosted on the PubMLST Pneumococcal Genome Library (https://pubmlst.org/organisms/streptococcus-pneumoniae/pgl) or the Global Pneumococcal Sequencing project database (https://www.pneumogen.net/gps/) (52, 75). RNAseq data used in the TLR2 transcriptional module expression analysis were obtained from the ArrayExpress database (accession E-MTAB-7841) (6).

ADDITIONAL FILES

The following material is available online.

Supplemental Material

Supplemental material (IAI00447-23-s0001.pdf). Figures S1 to S6; Tables S1 to S3.

REFERENCES

1. Weiser JN, Ferreira DM, Paton JC. 2018. *Streptococcus pneumoniae*: transmission, colonization and invasion. Nat Rev Microbiol 16:355–367. https://doi.org/10.1038/s41579-018-0001-8
2. Weight CM, Venturini C, Pojar S, Jochems SP, Reiné J, Nikolaou E, Solórzano C, Noursadeghi M, Brown JS, Ferreira DM, Heyderman RS. 2019. Microinvasion by *Streptococcus pneumoniae* induces epithelial innate immunity during colonisation at the human mucosal surface. Nat Commun 10:3060. https://doi.org/10.1038/s41467-019-11005-2
3. Jochems SP, de Ruiter K, Solórzano C, Voskamp A, Mitsi E, Nikolaou E, Carniel BF, Pojar S, German EL, Reiné J, et al. 2019. Innate and adaptive nasal mucosal immune responses following experimental human pneumococcal colonization. J Clin Invest 129:4523–4538. https://doi.org/10.1172/JCI128865
4. Trimble A, Connor V, Robinson RE, McLenaghan D, Hancock CA, Wang D, Gordon SB, Ferreira DM, Wright AD, Collins AM. 2020. Pneumococcal colonisation is an asymptomatic event in healthy adults using an experimental human colonisation model. PLoS One 15:e0229558. https://doi.org/10.1371/journal.pone.0229558
5. Kohler S, Voß F, Gómez Mejia A, Brown JS, Hammerschmidt S. 2016. Pneumococcal lipoproteins involved in bacterial fitness, virulence, and immune evasion. FEBS Lett 590:3820–3839. https://doi.org/10.1002/1873-3468.12352
6. Mitchell AM, Mitchell TJ. 2010. *Streptococcus pneumoniae*: virulence factors and variation. Clin Microbiol Infect 16:411–418. https://doi.org/10.1111/j.1469-0691.2010.03183.x
7. Moffitt K, Howard A, Martin S, Cheung E, Herd M, Basset A, Malley R. 2015. T_H17-mediated protection against pneumococcal carriage by a whole-cell vaccine is dependent on toll-like receptor 2 and surface lipoproteins. Clin Vaccine Immunol 22:909–916. https://doi.org/10.1128/CVI.00118-15
8. Jang A-Y, Ahn KB, Zhi Y, Ji H-J, Zhang J, Han SH, Guo H, Lim S, Song JY, Lim JH, Seo HS. 2019. Serotype-independent protection against invasive pneumococcal infections conferred by live vaccine with *lgt* deletion. Front Immunol 10:1212. https://doi.org/10.3389/fimmu.2019.01212
9. Tomlinson G, Chimalapati S, Pollard T, Lapp T, Cohen J, Camberlein E, Stafford S, Periselneris J, Aldridge C, Vollmer W, Picard C, Casanova J-L, Noursadeghi M, Brown J. 2014. TLR-mediated inflammatory responses to *Streptococcus pneumoniae* are highly dependent on surface expression of bacterial lipoproteins. J Immunol 193:3736–3745. https://doi.org/10.4049/jimmunol.1401413
10. Barendt SM, Sham L-T, Winkler ME. 2011. Characterization of mutants deficient in the L,D-carboxypeptidase (DacB) and WalRK (VicRK) regulon, involved in peptidoglycan maturation of *Streptococcus pneumoniae* serotype 2 strain D39. J Bacteriol 193:2290–2300. https://doi.org/10.1128/JB.01555-10
11. Russell H, Tharpe JA, Wells DE, White EH, Johnson JE. 1990. Monoclonal antibody recognizing a species-specific protein from *Streptococcus pneumoniae*. J Clin Microbiol 28:2191–2195. https://doi.org/10.1128/jcm.28.10.2191-2195.1990
12. Lawrence MC, Pilling PA, Epa VC, Berry AM, Ogunniyi AD, Paton JC. 1998. The crystal structure of pneumococcal surface antigen PsaA reveals a metal-binding site and a novel structure for a putative ABC-type binding protein. Structure 6:1553–1561. https://doi.org/10.1016/S0969-2126(98)00153-1
13. Pribyl T, Moche M, Dreisbach A, Bijlsma JJE, Saleh M, Abdullah MR, Hecker M, van Dijl JM, Becher D, Hammerschmidt S. 2014. Influence of impaired lipoprotein biogenesis on surface and exoproteome of *Streptococcus pneumoniae*. J Proteome Res 13:650–667. https://doi.org/10.1021/pr400768v
14. Chimalapati S, Cohen JM, Camberlein E, MacDonald N, Durmort C, Vernet T, Hermans PWM, Mitchell T, Brown JS. 2012. Effects of deletion of the *Streptococcus pneumoniae* lipoprotein diacylglyceryl transferase gene *lgt* on ABC transporter function and on growth *in vivo*. PLoS One 7:e41393. https://doi.org/10.1371/journal.pone.0041393
15. Aaberge IS, Eng J, Lermark G, Løvik M. 1995. Virulence of *Streptococcus pneumoniae* in mice: a standardized method for preparation and frozen storage of the experimental bacterial inoculum. Microb Pathog 18:141–152. https://doi.org/10.1016/s0882-4010(95)90125-6
16. Browall S, Norman M, Tångrot J, Galanis I, Sjöström K, Dagerhamn J, Hellberg C, Pathak A, Spadafina T, Sandgren A, Bättig P, Franzén O, Andersson B, Örtqvist Å, Normark S, Henriques-Normark B. 2014. Intraclonal variations among *Streptococcus pneumoniae* isolates influence the likelihood of invasive disease in children. J Infect Dis 209:377–388. https://doi.org/10.1093/infdis/jit481
17. Sleeman KL, Griffiths D, Shackley F, Diggle L, Gupta S, Maiden MC, Moxon ER, Crook DW, Peto TEA. 2006. Capsular serotype-specific attack rates and duration of carriage of *Streptococcus pneumoniae* in a population of children. J Infect Dis 194:682–688. https://doi.org/10.1086/505710
18. Richard AL, Siegel SJ, Erikson J, Weiser JN. 2014. TLR2 signaling decreases transmission of *Streptococcus pneumoniae* by limiting bacterial shedding in an infant mouse influenza A co-infection model. PLoS Pathog 10:e1004339. https://doi.org/10.1371/journal.ppat.1004339
19. Della Mina E, Borghesi A, Zhou H, Bougarn S, Boughorbel S, Israel L, Meloni I, Chrabieh M, Ling Y, Itan Y, et al. 2017. Inherited human IRAK-1 deficiency selectively impairs TLR signaling in fibroblasts. Proc Natl Acad Sci U S A 114:E514–E523. https://doi.org/10.1073/pnas.1620139114
20. Lee KS, Scanga CA, Bachelder EM, Chen Q, Snapper CM. 2007. TLR2 synergizes with both TLR4 and TLR9 for induction of the MyD88-dependent splenic cytokine and chemokine response to *Streptococcus pneumoniae*. Cell Immunol 245:103–110. https://doi.org/10.1016/j.cellimm.2007.04.003
21. Dessing MC, Hirst RA, de Vos AF, van der Poll T. 2009. Role of toll-like receptors 2 and 4 in pulmonary inflammation and injury induced by pneumolysin in mice. PLoS One 4:e7993. https://doi.org/10.1371/journal.pone.0007993
22. Bailey TL, Johnson J, Grant CE, Noble WS. 2015. The MEME suite. Nucleic Acids Res 43:W39–W49. https://doi.org/10.1093/nar/gkv416
23. Khandavilli S, Homer KA, Yuste J, Basavanna S, Mitchell T, Brown JS. 2008. Maturation of *Streptococcus pneumoniae* lipoproteins by a type II signal peptidase is required for ABC transporter function and full virulence. Mol Microbiol 67:541–557. https://doi.org/10.1111/j.1365-2958.2007.06065.x
24. Dintilhac A, Claverys J-P. 1997. The *adc* locus, which affects competence for genetic transformation in *Streptococcus pneumoniae*, encodes an ABC transporter with a putative lipoprotein homologous to a family of streptococcal adhesins. Res Microbiol 148:119–131. https://doi.org/10.1016/S0923-2508(97)87643-7
25. Loisel E, Jacquamet L, Serre L, Bauvois C, Ferrer JL, Vernet T, Di Guilmi AM, Durmort C. 2008. AdcAII, a new pneumococcal Zn-binding protein homologous with ABC transporters: biochemical and structural analysis. J Mol Biol 381:594–606. https://doi.org/10.1016/j.jmb.2008.05.068
26. Alloing G, de Philip P, Claverys JP. 1994. Three highly homologous membrane-bound lipoproteins participate in oligopeptide transport by the Ami system of the Gram-positive *Streptococcus pneumoniae*. J Mol Biol 241:44–58. https://doi.org/10.1006/jmbi.1994.1472
27. Alloing G, Trombe MC, Claverys JP. 1990. The *ami* locus of the Gram-positive bacterium *Streptococcus pneumoniae* is similar to binding protein-dependent transport operons of Gram-negative bacteria. Mol Microbiol 4:633–644. https://doi.org/10.1111/j.1365-2958.1990.tb00632.x

28. Farshchi Andisi V, Hinojosa CA, de Jong A, Kuipers OP, Orihuela CJ, Bijlsma JJE. 2012. Pneumococcal gene complex involved in resistance to extracellular oxidative stress. Infect Immun 80:1037–1049. https://doi.org/10.1128/IAI.05563-11
29. Saleh M, Bartual SG, Abdullah MR, Jensch I, Asmat TM, Petruschka L, Pribyl T, Gellert M, Lillig CH, Antelmann H, Hermoso JA, Hammerschmidt S. 2013. Molecular architecture of *Streptococcus pneumoniae* surface thioredoxin - fold lipoproteins crucial for extracellular oxidative stress resistance and maintenance of virulence. EMBO Mol Med 5:1852–1870. https://doi.org/10.1002/emmm.201202435
30. Härtel T, Klein M, Koedel U, Rohde M, Petruschka L, Hammerschmidt S. 2011. Impact of glutamine transporters on pneumococcal fitness under infection-related conditions. Infect Immun 79:44–58. https://doi.org/10.1128/IAI.00855-10
31. Potter AJ, Trappetti C, Paton JC. 2012. *Streptococcus pneumoniae* uses glutathione to defend against oxidative stress and metal ion toxicity. J Bacteriol 194:6248–6254. https://doi.org/10.1128/JB.01393-12
32. Basavanna S, Khandavilli S, Yuste J, Cohen JM, Hosie AHF, Webb AJ, Thomas GH, Brown JS. 2009. Screening of *Streptococcus pneumoniae* ABC transporter mutants demonstrates that LivJHMGF, a branched-chain amino acid ABC transporter, is necessary for disease pathogenesis. Infect Immun 77:3412–3423. https://doi.org/10.1128/IAI.01543-08
33. Weinrauch Y, Lacks SA. 1981. Nonsense mutations in the amylomaltase gene and other loci in *Streptococcus pneumoniae*. Mol Gen Genet 183:7–12. https://doi.org/10.1007/BF00270130
34. Basavanna S, Chimalapati S, Maqbool A, Rubbo B, Yuste J, Wilson RJ, Hosie A, Ogunniyi AD, Paton JC, Thomas G, Brown JS. 2013. The effects of methionine acquisition and synthesis on *Streptococcus pneumoniae* growth and virulence. PLoS One 8:e49638. https://doi.org/10.1371/journal.pone.0049638
35. Adamou JE, Heinrichs JH, Erwin AL, Walsh W, Gayle T, Dormitzer M, Dagan R, Brewah YA, Barren P, Lathigra R, Langermann S, Koenig S, Johnson S. 2001. Identification and characterization of a novel family of pneumococcal proteins that are protective against sepsis. Infect Immun 69:949–958. https://doi.org/10.1128/IAI.69.2.949-958.2001
36. Brown JS, Gilliland SM, Holden DW. 2001. A *Streptococcus pneumoniae* pathogenicity island encoding an ABC transporter involved in iron uptake and virulence. Mol Microbiol 40:572–585. https://doi.org/10.1046/j.1365-2958.2001.02414.x
37. Brown JS, Gilliland SM, Ruiz-Albert J, Holden DW. 2002. Characterization of pit, a *Streptococcus pneumoniae* iron uptake ABC transporter. Infect Immun 70:4389–4398. https://doi.org/10.1128/IAI.70.8.4389-4398.2002
38. Bidossi A, Mulas L, Decorosi F, Colomba L, Ricci S, Pozzi G, Deutscher J, Viti C, Oggioni MR. 2012. A functional genomics approach to establish the complement of carbohydrate transporters in *Streptococcus pneumoniae*. PLoS One 7:e33320. https://doi.org/10.1371/journal.pone.0033320
39. Saxena S, Khan N, Dehinwal R, Kumar A, Sehgal D. 2015. Conserved surface accessible nucleoside ABC transporter component SP0845 is essential for pneumococcal virulence and confers protection *in vivo*. PLoS One 10:e0118154. https://doi.org/10.1371/journal.pone.0118154
40. Cron LE, Bootsma HJ, Noske N, Burghout P, Hammerschmidt S, Hermans PWM. 2009. Surface-associated lipoprotein PpmA of *Streptococcus pneumoniae* is involved in colonization in a strain-specific manner. Microbiology (Reading) 155:2401–2410. https://doi.org/10.1099/mic.0.026765-0
41. Orihuela CJ, Mills J, Robb CW, Wilson CJ, Watson DA, Niesel DW. 2001. *Streptococcus pneumoniae* PstS production is phosphate responsive and enhanced during growth in the murine peritoneal cavity. Infect Immun 69:7565–7571. https://doi.org/10.1128/IAI.69.12.7565-7571.2001
42. Rosenow C, Maniar M, Trias J. 1999. Regulation of the α-galactosidase activity in *Streptococcus pneumoniae*: characterization of the raffinose utilization system. Genome Res 9:1189–1197. https://doi.org/10.1101/gr.9.12.1189
43. Hermans PWM, Adrian PV, Albert C, Estevão S, Hoogenboezem T, Luijendijk IHT, Kamphausen T, Hammerschmidt S. 2006. The streptococcal lipoprotein rotamase A (SlrA) is a functional peptidyl-prolyl isomerase involved in pneumococcal colonization. J Biol Chem 281:968–976. https://doi.org/10.1074/jbc.M510014200
44. Dwyer MA, Hellinga HW. 2004. Periplasmic binding proteins: a versatile superfamily for protein engineering. Curr Opin Struct Biol 14:495–504. https://doi.org/10.1016/j.sbi.2004.07.004
45. Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJE. 2015. The Phyre2 web portal for protein modelling, prediction, and analysis. Nat Protoc 10:845–858. https://doi.org/10.1038/nprot.2015.053
46. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Žídek A, Potapenko A, et al. 2021. Highly accurate protein structure prediction with AlphaFold. Nature 596:583–589. https://doi.org/10.1038/s41586-021-03819-2
47. Wilkens S. 2015. Structure and mechanism of ABC transporters. F1000Prime Rep 7:14. https://doi.org/10.12703/P7-14
48. Minhas V, Harvey RM, McAllister LJ, Seemann T, Syme AE, Baines SL, Paton JC, Trappetti C. 2019. Capacity to utilize raffinose dictates pneumococcal disease phenotype. mBio 10:e02596-18. https://doi.org/10.1128/mBio.02596-18
49. Tettelin H, Nelson KE, Paulsen IT, Eisen JA, Read TD, Peterson S, Heidelberg J, DeBoy RT, Haft DH, Dodson RJ, et al. 2001. Complete genome sequence of a virulent isolate of *Streptococcus pneumoniae*. Science 293:498–506. https://doi.org/10.1126/science.1061217
50. Spellerberg B, Cundell DR, Sandros J, Pearce BJ, Idanpaan-Heikkila I, Rosenow C, Masure HR. 1996. Pyruvate oxidase, as a determinant of virulence in *Streptococcus pneumoniae*. Mol Microbiol 19:803–813. https://doi.org/10.1046/j.1365-2958.1996.425954.x
51. Obolski U, Swarthout TD, Kalizang'oma A, Mwalukomo TS, Chan JM, Weight CM, Brown C, Cave R, Cornick J, Kamng'ona AW, Msefula J, Ercoli G, Brown JS, Lourenço J, Maiden MC, French N, Gupta S, Heyderman RS. 2023. The metabolic, virulence and antimicrobial resistance profiles of colonising *Streptococcus pneumoniae* shift after PCV13 introduction in urban Malawi. Nat Commun 14:7477. https://doi.org/10.1038/s41467-023-43160-y
52. Jolley KA, Maiden MCJ. 2010. BIGSdb: scalable analysis of bacterial genome variation at the population level. BMC Bioinformatics 11:595. https://doi.org/10.1186/1471-2105-11-595
53. Petit CM, Brown JR, Ingraham K, Bryant AP, Holmes DJ. 2001. Lipid modification of prelipoproteins is dispensable for growth *in vitro* but essential for virulence in *Streptococcus pneumoniae*. FEMS Microbiol Lett 200:229–233. https://doi.org/10.1111/j.1574-6968.2001.tb10720.x
54. Joyce EA, Popper SJ, Falkow S. 2009. *Streptococcus pneumoniae* nasopharyngeal colonization induces type I interferons and interferon-induced gene expression. BMC Genomics 10:404. https://doi.org/10.1186/1471-2164-10-404
55. Skovbjerg S, Nordén R, Martner A, Samuelsson E, Hynsjö L, Wold AE. 2017. Intact pneumococci trigger transcription of interferon-related genes in human monocytes, while fragmented, autolyzed bacteria subvert this response. Infect Immun 85:e00960-16. https://doi.org/10.1128/IAI.00960-16
56. Parker D, Martin FJ, Soong G, Harfenist BS, Aguilar JL, Ratner AJ, Fitzgerald KA, Schindler C, Prince A. 2011. *Streptococcus pneumoniae* DNA initiates type I interferon signaling in the respiratory tract. mBio 2:e00016-11. https://doi.org/10.1128/mBio.00016-11
57. D'Mello A, Riegler AN, Martínez E, Beno SM, Ricketts TD, Foxman EF, Orihuela CJ, Tettelin H. 2020. An *in vivo* atlas of host–pathogen transcriptomes during *Streptococcus pneumoniae* colonization and disease. Proc Natl Acad Sci U S A 117:33507–33518. https://doi.org/10.1073/pnas.2010428117
58. Koppe U, Högner K, Doehn J-M, Müller HC, Witzenrath M, Gutbier B, Bauer S, Pribyl T, Hammerschmidt S, Lohmeyer J, Suttorp N, Herold S, Opitz B. 2012. *Streptococcus pneumoniae* stimulate a STING- and IFN regulatory factor 3- dependent type I IFN production in macrophages, which regulates RANTES production in macrophages, cocultured alveolar epithelial cells, and mouse lungs. J Immunol 188:811–817. https://doi.org/10.4049/jimmunol.1004143
59. Ruiz-Moreno JS, Hamann L, Jin L, Sander LE, Puzianowska-Kuznicka M, Cambier J, Witzenrath M, Schumann RR, Suttorp N, Opitz B. 2018. The cGAS/STING pathway detects *Streptococcus pneumoniae* but appears dispensable for antipneumococcal defense in mice and humans. Infect Immun 86:e00849-17. https://doi.org/10.1128/IAI.00849-17
60. LeMessurier KS, Häcker H, Chi L, Tuomanen E, Redecke V. 2013. Type I interferon protects against pneumococcal invasive disease by inhibiting

bacterial transmigration across the lung. PLoS Pathog 9:e1003727. https://doi.org/10.1371/journal.ppat.1003727
61. Zangari T, Ortigoza MB, Lokken-Toyli KL, Weiser JN. 2021. Type I interferon signaling is a common factor driving *Streptococcus pneumoniae* and influenza A virus shedding and transmission. mBio 12:e03589-20. https://doi.org/10.1128/mBio.03589-20
62. Anderton JM, Rajam G, Romero-Steiner S, Summer S, Kowalczyk AP, Carlone GM, Sampson JS, Ades EW. 2007. E-cadherin is a receptor for the common protein pneumococcal surface adhesin A (PsaA) of *Streptococcus pneumoniae*. Microb Pathog 42:225–236. https://doi.org/10.1016/j.micpath.2007.02.003
63. Abdullah MR, Gutiérrez-Fernández J, Pribyl T, Gisch N, Saleh M, Rohde M, Petruschka L, Burchhardt G, Schwudke D, Hermoso JA, Hammerschmidt S. 2014. Structure of the pneumococcal L,D-carboxypeptidase DacB and pathophysiological effects of disabled cell wall hydrolases DacA and DacB. Mol Microbiol 93:1183–1206. https://doi.org/10.1111/mmi.12729
64. Smithers L, Olatunji S, Caffrey M. 2021. Bacterial lipoprotein posttranslational modifications. New insights and opportunities for antibiotic and vaccine development. Front Microbiol 12:788445. https://doi.org/10.3389/fmicb.2021.788445
65. Nguyen MT, Matsuo M, Niemann S, Herrmann M, Götz F. 2020. Lipoproteins in Gram-positive bacteria: abundance, function, fitness. Front Microbiol 11:582582. https://doi.org/10.3389/fmicb.2020.582582
66. Keller LE, Robinson DA, McDaniel LS. 2016. Nonencapsulated *Streptococcus pneumoniae*: emergence and pathogenesis. mBio 7:e01792. https://doi.org/10.1128/mBio.01792-15
67. Moscoso M, García E, López R. 2006. Biofilm formation by *Streptococcus pneumoniae*: role of choline, extracellular DNA, and capsular polysaccharide in microbial accretion. J Bacteriol 188:7785–7795. https://doi.org/10.1128/JB.00673-06
68. Granok AB, Parsonage D, Ross RP, Caparon MG. 2000. The RofA binding site in *Streptococcus pyogenes* is utilized in multiple transcriptional pathways. J Bacteriol 182:1529–1540. https://doi.org/10.1128/JB.182.6.1529-1540.2000
69. Zhu L, Lau GW. 2011. Inhibition of competence development, horizontal gene transfer and virulence in *Streptococcus pneumoniae* by a modified competence stimulating peptide. PLoS Pathog 7:e1002241. https://doi.org/10.1371/journal.ppat.1002241
70. Reglinski M, Ercoli G, Plumptre C, Kay E, Petersen FC, Paton JC, Wren BW, Brown JS. 2018. A recombinant conjugated pneumococcal vaccine that protects against murine infections with a similar efficacy to Prevnar-13. NPJ Vaccines 3:53. https://doi.org/10.1038/s41541-018-0090-4
71. Grant CE, Bailey TL, Noble WS. 2011. FIMO: scanning for occurrences of a given motif. Bioinformatics 27:1017–1018. https://doi.org/10.1093/bioinformatics/btr064
72. Teufel F, Almagro Armenteros JJ, Johansen AR, Gíslason MH, Pihl SI, Tsirigos KD, Winther O, Brunak S, von Heijne G, Nielsen H. 2022. SignalP 6.0 predicts all five types of signal peptides using protein language models. Nat Biotechnol 40:1023–1025. https://doi.org/10.1038/s41587-021-01156-3
73. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25:3389–3402. https://doi.org/10.1093/nar/25.17.3389
74. Sayers EW, Barrett T, Benson DA, Bryant SH, Canese K, Chetvernin V, Church DM, DiCuccio M, Edgar R, Federhen S, et al. 2009. Database resources of the national center for biotechnology information. Nucleic Acids Res 37:D5–D15. https://doi.org/10.1093/nar/gkn741
75. Letunic I, Bork P. 2021. Interactive tree of life (iTOL) v5: an online tool for phylogenetic tree display and annotation. Nucleic Acids Res 49:W293–W296. https://doi.org/10.1093/nar/gkab301
76. Marchler-Bauer A, Bo Y, Han L, He J, Lanczycki CJ, Lu S, Chitsaz F, Derbyshire MK, Geer RC, Gonzales NR, Gwadz M, Hurwitz DI, Lu F, Marchler GH, Song JS, Thanki N, Wang Z, Yamashita RA, Zhang D, Zheng C, Geer LY, Bryant SH. 2017. CDD/SPARCLE: functional classification of proteins via subfamily domain architectures. Nucleic Acids Res 45:D200–D203. https://doi.org/10.1093/nar/gkw1129
77. Wick RR, Judd LM, Gorrie CL, Holt KE. 2017. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. PLOS Comput Biol 13:e1005595. https://doi.org/10.1371/journal.pcbi.1005595
78. Gurevich A, Saveliev V, Vyahhi N, Tesler G. 2013. QUAST: quality assessment tool for genome assemblies. Bioinformatics 29:1072–1075. https://doi.org/10.1093/bioinformatics/btt086
79. Betts M, Jarvis S, Jeffries A, Gori A, Chaguza C, Msefula J, Weight CM, Kwambana-Adams B, French N, Swarthout TD, Brown JS, Heyderman RS. 2021. Complete genome sequence of *Streptococcus pneumoniae* strain BVJ1JL, a serotype 1 carriage isolate from Malawi. Microbiol Resour Announc 10:e0071521. https://doi.org/10.1128/MRA.00715-21
80. Schwengers O, Jelonek L, Dieckmann MA, Beyvers S, Blom J, Goesmann A. 2021. Bakta: rapid and standardized annotation of bacterial genomes via alignment-free sequence identification. Microbial Genomics 7:000685. https://doi.org/10.1099/mgen.0.000685
81. Prjibelski A, Antipov D, Meleshko D, Lapidus A, Korobeynikov A. 2020. Using SPAdes *de novo* assembler. Curr Protoc Bioinformatics 70:e102. https://doi.org/10.1002/cpbi.102
82. Lachmandas E, Boutens L, Ratter JM, Hijmans A, Hooiveld GJ, Joosten LAB, Rodenburg RJ, Fransen JAM, Houtkooper RH, van Crevel R, Netea MG, Stienstra R. 2016. Microbial stimulation of different toll-like receptor signalling pathways induces diverse metabolic programmes in human monocytes. Nat Microbiol 2:16246. https://doi.org/10.1038/nmicrobiol.2016.246
83. Eriksson M, Peña-Martínez P, Ramakrishnan R, Chapellier M, Högberg C, Glowacki G, Orsmark-Pietras C, Velasco-Hernández T, Lazarević VL, Juliusson G, Cammenga J, Mulloy JC, Richter J, Fioretos T, Ebert BL, Järås M. 2017. Agonistic targeting of TLR1/TLR2 induces p38 MAPK-dependent apoptosis and NFκB-dependent differentiation of AML cells. Blood Adv 1:2046–2057. https://doi.org/10.1182/bloodadvances.2017006148
84. Tettelin H, Masignani V, Cieslewicz MJ, Donati C, Medini D, Ward NL, Angiuoli SV, Crabtree J, Jones AL, Durkin AS, et al. 2005. Genome analysis of multiple pathogenic isolates of *Streptococcus agalactiae*: implications for the microbial "pan-genome". Proc Natl Acad Sci U S A 102:13950–13955. https://doi.org/10.1073/pnas.0506758102
85. Pojar S, Basset A, Gritzfeld JF, Nikolaou E, van Selm S, Eleveld MJ, Gladstone RA, Solórzano C, Dalia AB, German E, et al. 2020. Isolate differences in colonization efficiency during experimental human pneumococcal challenge. medRxiv. https://doi.org/10.1101/2020.04.20.20066399
86. Lanie JA, Ng W-L, Kazmierczak KM, Andrzejewski TM, Davidsen TM, Wayne KJ, Tettelin H, Glass JI, Winkler ME. 2007. Genome sequence of Avery's virulent serotype 2 strain D39 of *Streptococcus pneumoniae* and comparison with that of unencapsulated laboratory strain R6. J Bacteriol 189:38–51. https://doi.org/10.1128/JB.01148-06