Bacterial surface lipoproteins mediate epithelial microinvasion by *Streptococcus pneumoniae*

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

[a] Research Department of Infection, Division of Infection and Immunity, University College London, London, UK.
[b] Department of Respiratory Medicine, Centre for Inflammation and Tissue Repair, University College London, London, United Kingdom

Running title: Pneumococcal lipoproteins in epithelial microinvasion

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

* Present address: Institute of Infection, Veterinary and Ecological Sciences, University of Liverpool, Liverpool, UK.
** Present address: Department of Biomedical and Life Sciences, Lancaster University, Lancaster, UK.

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

Abstract word count: 236 words.
Importance word count: 150 words.
Manuscript word count: 4838 words.

ABSTRACT

Streptococcus pneumoniae, a common coloniser of the upper respiratory tract, invades nasopharyngeal epithelial cells without causing disease in healthy people. We hypothesised that surface expression of pneumococcal lipoproteins, recognised by the innate immune receptor TLR2, mediate epithelial microinvasion. Mutation of *lgt* in serotype 4 (TIGR4) and serotype 6B (BHN418) pneumococcal strains abolishes the ability of the mutants to activate TLR2 signalling. Loss of *lgt* also led to concomitant decrease in interferon signalling 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 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 colonisation is considered a necessary prerequisite for disease, yet many people are transiently and asymptomatically colonised by pneumococci without becoming unwell. It is therefore important to better understand how the colonisation 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. Our results highlight the potential importance of the post-translational modification of lipoproteins in the mediation of epithelial invasion during pneumococcal colonisation. Targeting mucosal innate immunity and epithelial microinvasion alongside the induction of an adaptive immune response may be effective in preventing pneumococcal colonisation and disease.

INTRODUCTION

Streptococcus pneumoniae (pneumococcus) is a versatile pathobiont capable of asymptomatically colonising the nasopharynx, causing localised infections of the middle ear, respiratory tract and lungs, and causing disseminated invasive disease (e.g. bacteraemic pneumonia and meningitis) with high mortality rates (1). *S. pneumoniae* has long been considered an extracellular pathogen despite 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 characterise 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 mediate 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, are 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 signalling (8, 9, 13). Mutating *lgt* also attenuates pneumococcal virulence and shortens colonisation 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 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 colonise as well as cause invasive disease in susceptible hosts, albeit to different extents (17). Our findings indicate that lipoproteins contributed to epithelial inflammation by activating the TLR2 pathway and augmenting an interferon response. While attenuation of inflammatory responses were 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 uncharacterised 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. We hypothesized *that palA* mediate the intraspecies differences in microinvasion, however, cell culture and murine colonisation experiments did not show an essential role for *palA* in microinvasion, colonisation or disease. Our results suggest that the contribution of lipoproteins and

lipoprotein processing to microinvasion may be more complex and we propose may occur through differential post-translational lipidation.

RESULTS

Pneumococcal *lgt* mutants induce lower levels of TLR2 and interferon signalling compared to than 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 signalling in HEK-Blue™ hTLR2 reporter cells, while genetic complementation of *lgt* at a chromosomal ectopic site restored WT-like ability to stimulate the TLR2 pathway (Figure 1A) (9). Although macrophages respond to pneumococcal infections by activating TLR2 signalling pathways, it is unknown if nasopharyngeal epithelial cells respond in the same way (8, 9, 18). Using a transcriptional module reflective of TLR2 signalling and previously published transcriptomic datasets (2, 19), we found evidence of elevated TLR2-mediated transcriptional activity in nasopharyngeal epithelial cells infected with TIGR4 and BHN418 (Figure 1B; Supplementary Figure 1).

TLR2 activation is necessary for 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 signalling (6, 7). We therefore hypothesize that TLR2 activation potentiates interferon signalling 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 (Figure 1C-E), while cells infected with BHN418 *lgt::cm* have reduced expression of *CXCL10* and *IFNL3* compared to those infected with WT BHN418 (Figure 1C,1F).

160 Our results suggest that lipoprotein-mediated TLR2 activation augments the

161 epithelial interferon response during pneumococcal microinvasion.

162

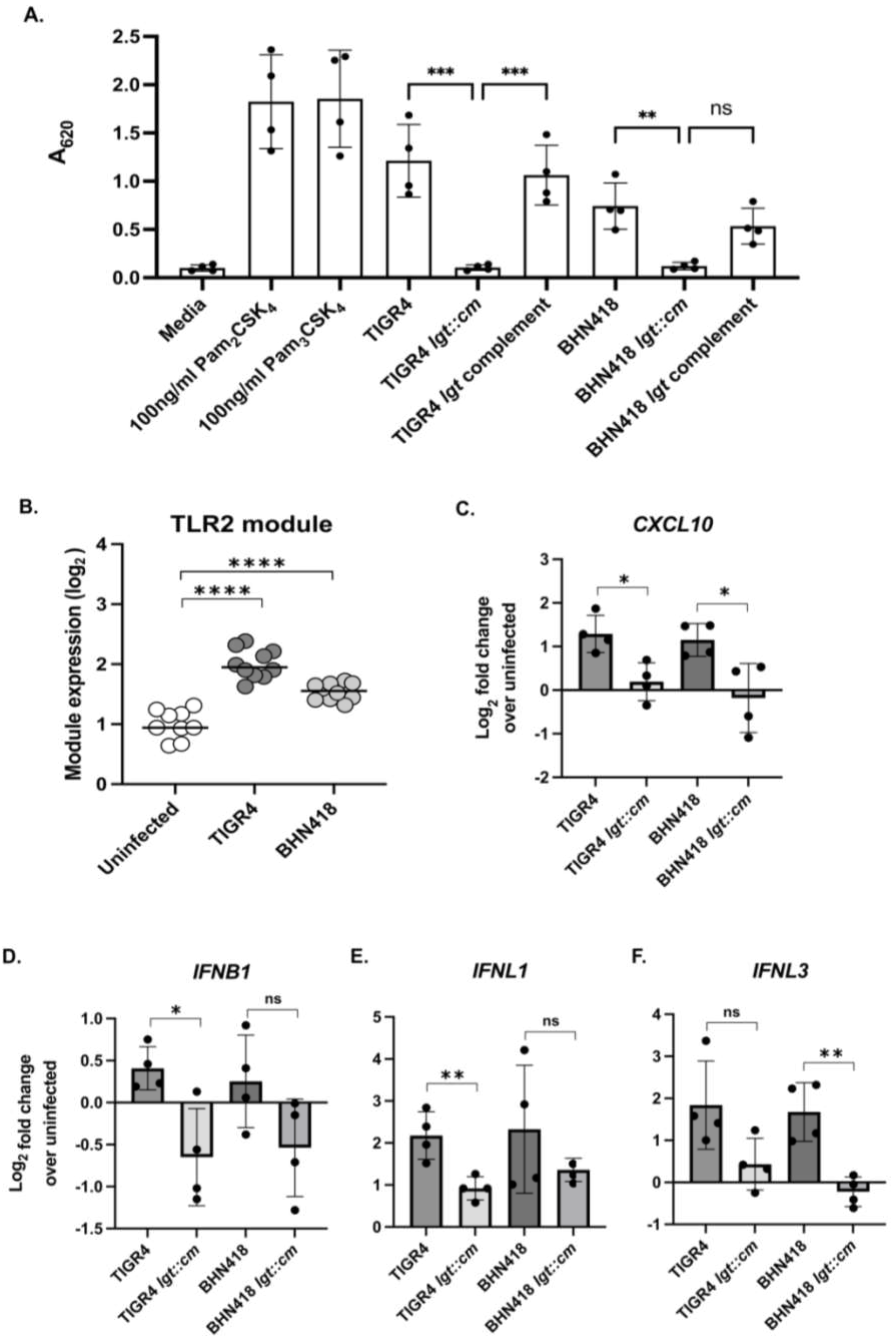

Figure 1. Pneumococcal *lgt* mutants were less inflammatory compared to WT strains. (A) SEAP reporter readout from HEK-Blue™ hTLR2 reporter cells treated with pneumococcal strains at MOI 10 for 16 hours. (B) Expression of a transcriptional module reflective of TLR2-mediated activity in Detroit 562 cells 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 $p < 0.001$, *** indicates $p < 0.0001$.

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 signalling impact on pneumococcal microinvasion, we infected confluent Detroit 562 nasopharyngeal cells (NPE) with the serotype 6B (BHN418) and 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 internalised into Detroit 562 cells (Figure 2A-B, 2D-2E). 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 (Figure 2B, 2D) (15, 17). The *lgt* mutation also significantly reduced the number of planktonic BHN418 but not TIGR4 (Figure 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 (Figure 1A, Figure 2A-F).

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 (Figure 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 (Figure 3A-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 cell-free medium and is therefore not dependent on the presence of NPE cells (Figure 3C-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 signalling neither promotes nor inhibits epithelial microinvasion by *S. pneumoniae*.

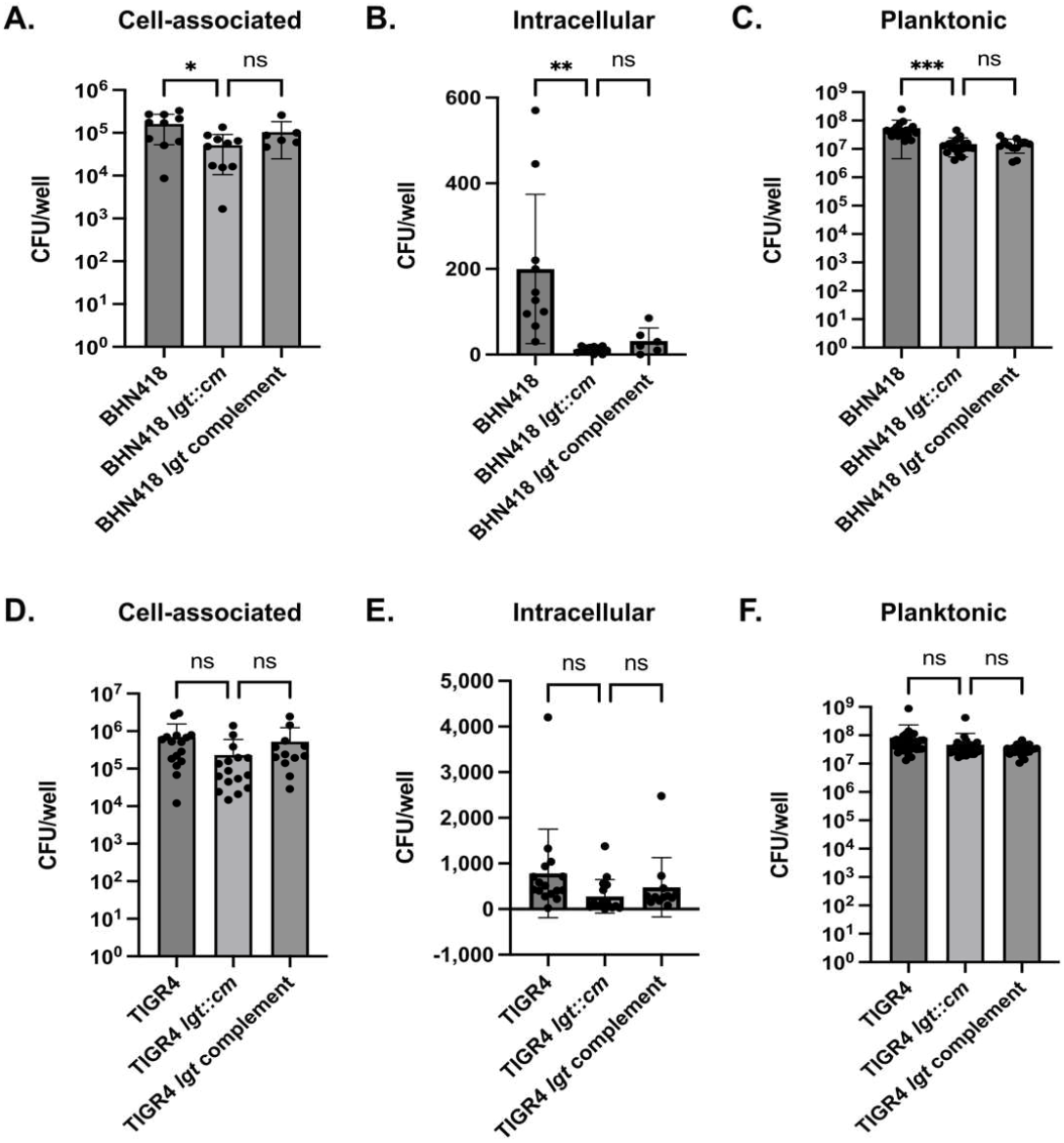

Figure 2. Mutation of *lgt* impaired epithelial microinvasion by *Streptococcus pneumoniae*, with greater defects for BHN418. (A-F) NPE microinvasion by WT and *lgt* mutants. Graphs show CFU numbers for BHN418-derived (A-C) and TIGR4-derived strains (D-F) associated with (A,D), internalised into (B,E),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

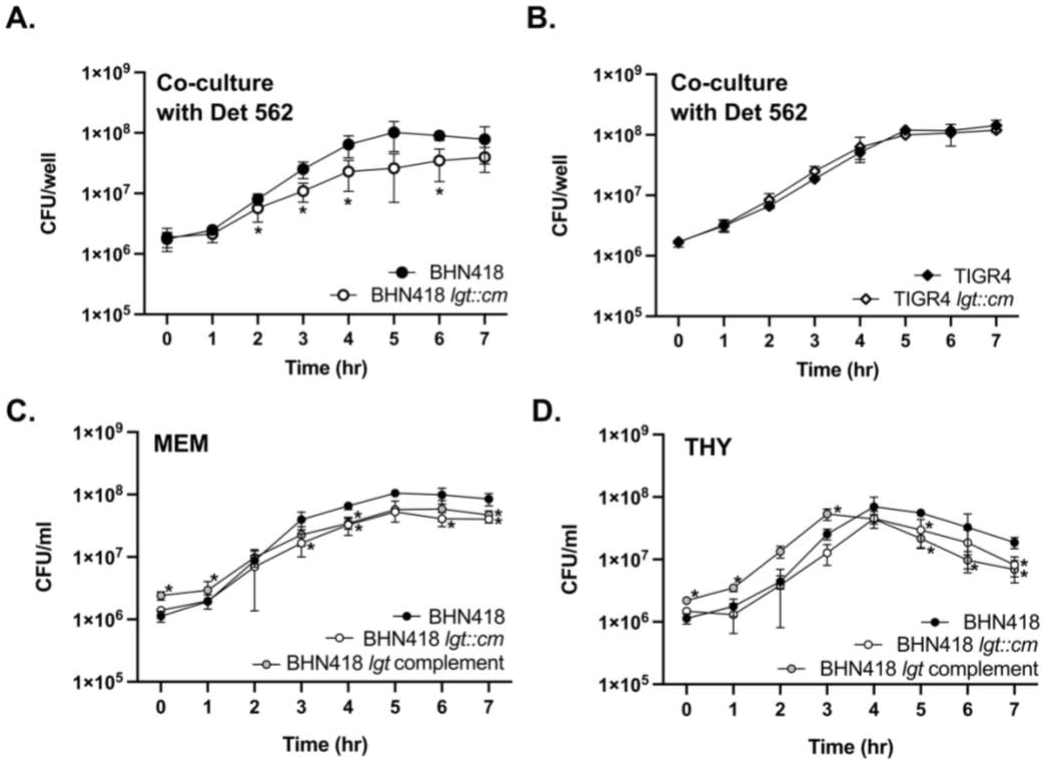

Figure 3. Growth of BHN418 *lgt::cm* compared to WT. (A-B) Growth of BHN418-derived (A) and TIGR4-derived (B) strains in infection medium (MEM + 1% FBS) in the presence of Detroit 562 NPE cells. (C-D) Growth of BHN418-derived strains in infection medium (C) and in the rich growth medium THY (D) in the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

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 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.

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 (Figure 4A-B). ABC transporters are multi subunit proteins comprising of two transmembrane permease domains and two cytoplasmic ATPase domains (23). Two genes encoding ABC transporter permease domain proteins, annotated as *yteP* and *araQ*, were found ~3.3kb and ~2.4kb upstream of *palA.* We were unable to locate ORF(s) encoding for the ATPase domain proteins in the 10kb 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 genome, in a similar strategy as the raffinose utilisation system, or no longer function as transporters (24, 25).

This genetic context suggests that *palA* is the fifth gene in an operon encoding for carbohydrate import and utilisation genes, which includes *yteP* and *araQ* (Figure 4A). The operon sits 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% (26). This putative genetic island is directly downstream of *spxB*, which encodes an important pneumococcal virulence factor involved in the production of H_2O_2 (27). Alignment of TIGR4 whole genome sequencing reads to the BHN418 genome revealed the absence of the entire putative island in the TIGR4 genome (Figure 4C) (26).

260 **Table 1. Lipoproteins encoded by TIGR4 and BHN418, identified**
261 **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	(28)
RSS80_10715	SP_2169	*adcA*	Adhesin competence protein A; zinc transporter	(29)
RSS80_04890	SP_1002	*adcAII*	Adhesin competence protein AII; zinc transporter	(30)
RSS80_01875	SP_0366	*aliA*	AmiA-like protein A; oligopeptide transporter	(31)
RSS80_07270	SP_1527	*aliB*	AmiA-like protein B; oligopeptide transporter	(31)
RSS80_09070	SP_1891	*amiA*	Aminopterin resistance locus protein A; oligopeptide transporter	(32)
RSS80_03115	SP_0629	*dacB*	L,D-carboxypeptidase	(10)
RSS80_03240	SP_0659	*etrx1*	Extracellular thioredoxin-like protein 1; thiol-disulfide oxidoreductase	(33)
RSS80_04875	SP_1000	*etrx2*	Extracellular thioredoxin-like protein 2; thiol-disulfide oxidoreductase	(34)
RSS80_06715	SP_1394	*glnH*	GlnH glutamine/polar amino acid ABC transporter substrate-binding protein	(35)
RSS80_00785	SP_0148	*gshT*	Glutathione transporter	(36)
RSS80_03685	SP_0749	*livJ*	Branched chain amino-acid transporter	(37)
RSS80_10385	SP_2108	*malX*	Maltosaccharide transporter	(38)
RSS80_00790	SP_0149	*metQ*	Methionine-binding lipoprotein Q	(39)
RSS80_05810	SP_1175	*phtA*	Pneumococcal histidine triad protein A	(40)
RSS80_05040	SP_1032	*piaA*	Pneumococcal iron acquisition protein A	(41)
RSS80_01305	SP_0243	*pitA*	Pneumococcal iron transporter protein A	(42)
RSS80_08955	SP_1872	*piuA*	Pneumococcal iron uptake protein A	(41)
RSS80_04120	SP_0845	*pnrA*	Nucleoside transporter	(43, 44)
RSS80_04795	SP_0981	*ppmA*	Putative proteinase maturation protein A; peptidyl-prolyl cis–trans isomerase	(45)
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	(46)
RSS80_09095	SP_1897	*rafE*	Raffinose transporter	(24)
RSS80_03790	SP_0771	*slrA*	Streptococcal lipoprotein rotamase A;cyclophilin-type peptidyl-prolyl cis–trans isomerase	(47)
RSS80_04895§	SP_1003	*phtB*	Pneumococcal histidine triad protein B	(40)
RSS80_04895§	SP_1174	*phtD*	Pneumococcal histidine triad protein D	(40)
RSS80_06745	SP_1400	*pstS2*	phosphate binding protein	-
RSS80_10885	SP_2197	-	ABC transporter binding protein	-
RSS80_00530	SP_0112	-	Amino acid binding protein	-
RSS80_09960	SP_2041	-	Membrane protein insertase	-

RSS80_04185	SP_0857*	-	ABC transporter substrate binding protein	-
RSS80_03070	SP_0620	-	Amino acid ABC transporter binding protein	-
RSS80_09570	SP_1975	-	Membrane protein insertase	-
RSS80_03435*	SP_0708*	-	ABC transporter substrate binding protein (truncated)	-
RSS80_03595	**Not present**	**-**	**Extracellular solute binding protein**	**-**
RSS80_04430	SP_0899	-	Hypothetical protein	-
RSS80_05800*	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	-

[bold] ORF present in BHN418 but not TIGR4.
* Annotated as pseudogene, contained premature stop codon, or interrupted by insertion sequence
§ RSS80_04895 was the best match BLAST result for more than one TIGR4 CDS
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).

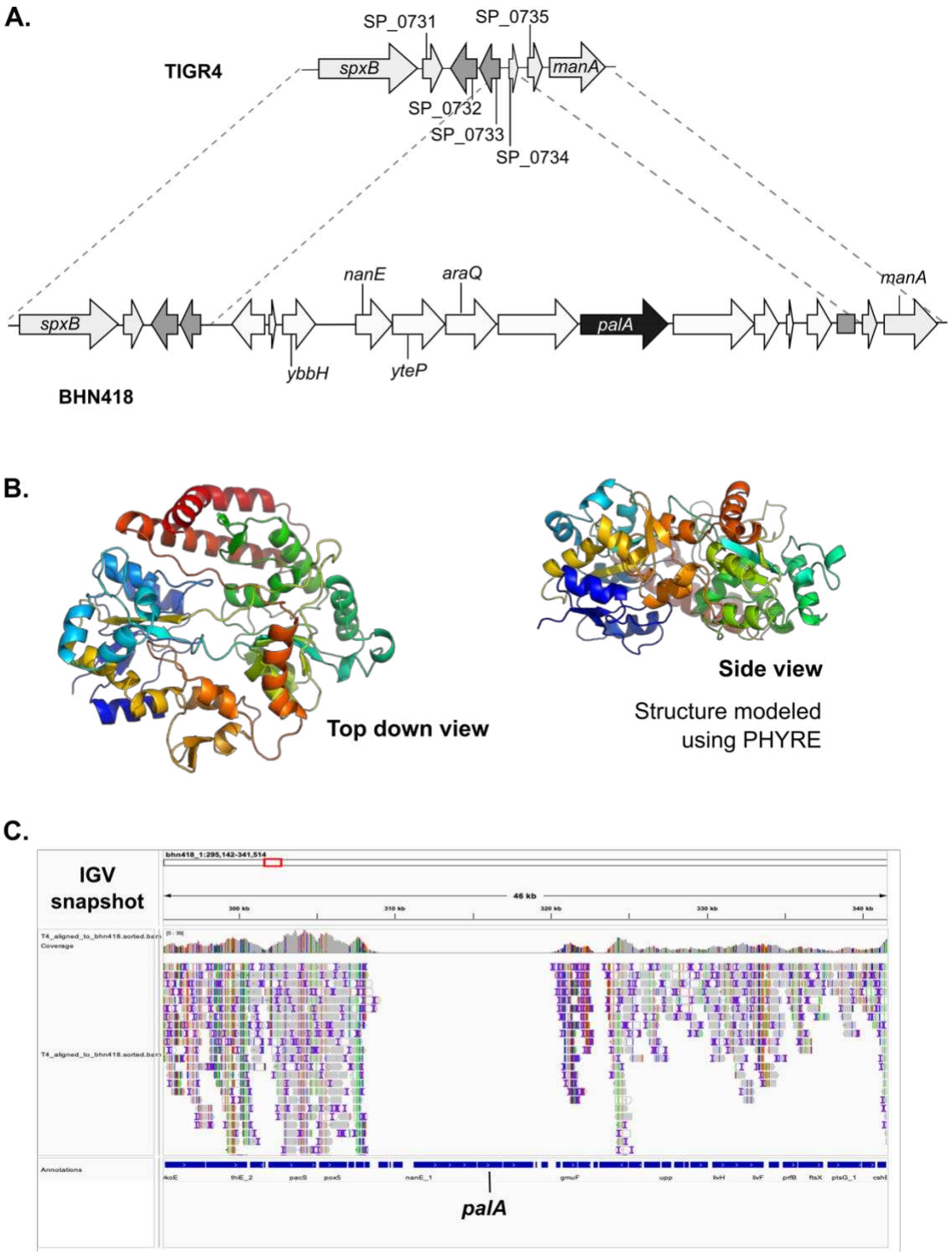

Figure 4. PalA is a lipoprotein present in BHN418 but not in TIGR4. (A) Genetic context of the *palA* gene (black arrow), which is within a multi-gene operon that is part of a putative genetic island. ORFs with homology to previously described genes are labelled with the gene name. (B) Predicted structure of PalA, modelled using PHYRE, which shows a barrel-like structure with a possible central ligand binding pocket. (C) IGV snapshot demonstrating lack of TIGR4 sequencing reads mapping to the putative *palA* genetic island.

279 To determine if *palA* is predominantly present in more carriage-type
280 serotypes, such as serotype 6B, we examined the presence of *palA* in a well-curated
281 dataset of 2806 carriage isolates from Malawi (48). 567 of these carriage isolates
282 (20.5%) carry *palA* in their chromosome. Mapping the analysis results onto a
283 hierarchal clustering (Newick) tree showed that *palA* is present in specific lineages,
284 with no clear association to capsular serotypes or sequence types (genetic
285 relatedness, visualised as neighbouring branches on a Newick tree) (Supplementary
286 Figure 2) (48). However, presence of *palA* is enriched in certain serotypes,
287 particularly serotype 6A (39/93, 41.9%), 6B (12/31, 38.7%), 10A (29/34, 85%), 15B
288 (52.76, 68.4%), 16F (60/94, 63.8%), 23B (45/103, 43.7%), 35A (28/28, 100%) and
289 35B (60/114, 52.6%) (Supplementary Table 1). Additionally, the branching patterns
290 of the phylogenetic tree for the Malawi carriage isolates supports the inference that
291 *palA* and its associated genetic island were acquired via horizontal gene transfer and
292 expanded in specific lineages (Supplementary Figure 2).
293

294 **PalA presence is enriched in carriage and ear isolates.** Maintenance of this
295 11.5kb genetic island is potentially costly and suggests that the island confers some
296 form of advantage to isolates that carry it. *S. pneumoniae* is capable of colonising
297 and infecting multiple body sites including the nasopharynx, lungs, blood, CSF,
298 meninges, and middle ear. We therefore examined 51,379 genomes in the BIGSdb
299 database to determine if there is an association between the presence of *palA* and
300 the isolation site of the strain ("source") (49).

301 The *palA* gene presence is enriched in carriage isolates and in strains isolated
302 from ear infections compared to strains isolated from IPD or lower respiratory tract
303 disease (Table 2). More than half of serotype 22F and 6A strains isolated from the
304 ear carried *palA* and approximately 28% of all serotype 22F and 6A genomes in the
305 database carry *palA*, in contrast to the overall *palA* prevalence rate of 9.97% (Table
306 2). In the BIGSdb database, the only serotype 4 strain isolated from the ear carried
307 *palA* in its genome. These observations suggest that *palA* and/or its putative genetic
308 island may facilitate spread to and cause infection of the ear, although *palA*'s
309 presence is not necessary for colonisation of the ear.

310

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

Category	Source (BIGSdb label)	Proportion (%)
Carriage	"nasopharynx", "pharynx", "sputum"	14.57 (3114/21369)
Otitis	"ear swab", "middle ear fluid"	13.72 (129/940)
Pneumonia	"lung aspirate", "sinus aspirate", "bronchoalveolar lavage", "bronchi"	8.96 (25/279)
Invasive	"blood", "cerebrospinal fluid", "joint fluid", "pleural fluid	6.54 (1837/28075)
Eye/pus/others	"eye swab", "pus", "other"	2.65 (19/716)
Overall		9.97 (5124/51379)

313

Mutation of *palA* does not alter pneumococcal colonisation 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. 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 to that of WT BHN418 (Figure 5A-C). Additionally, we did not observe a growth defect when BHN418 *palA::kan* was grown in THY or MEM (Figure 5D-E). Heterologous expression of *palA* in a serotype 23F strain naturally lacking the island (P1121) did not increase the microinvasion potential of the resulting strains and reduced the number of planktonic bacteria in the cell culture supernatant (Figure 6A-C). Moreover, the BHN418 *palA* knockout strains and the P1121 *palA* knock-in strains activated TLR2 signalling to similar levels as their respective wild-type strains (Figure 6D). We therefore conclude that 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.

Mutation of *palA* attenuates nasopharyngeal colonisation density and duration in mice (14). We next asked if presence of *palA* confer 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* mutant either singly or in a 1:1 competitive mix. After 7 days of colonisation, similar CFU numbers for WT BHN418 and the *palA::kan* were recovered from nasal washes (Figure 5F-G). Similar CFU numbers for BHN418 and *palA::kan* were also recovered in homogenised lungs and blood 24 hours post inoculation in a murine pneumonia model (Figure 5H-I). We conclude that presence of *palA* does not confer

339 a colonisation advantage in the murine nasopharynx or in the progression to

340 bacteraemic pneumonia.

341

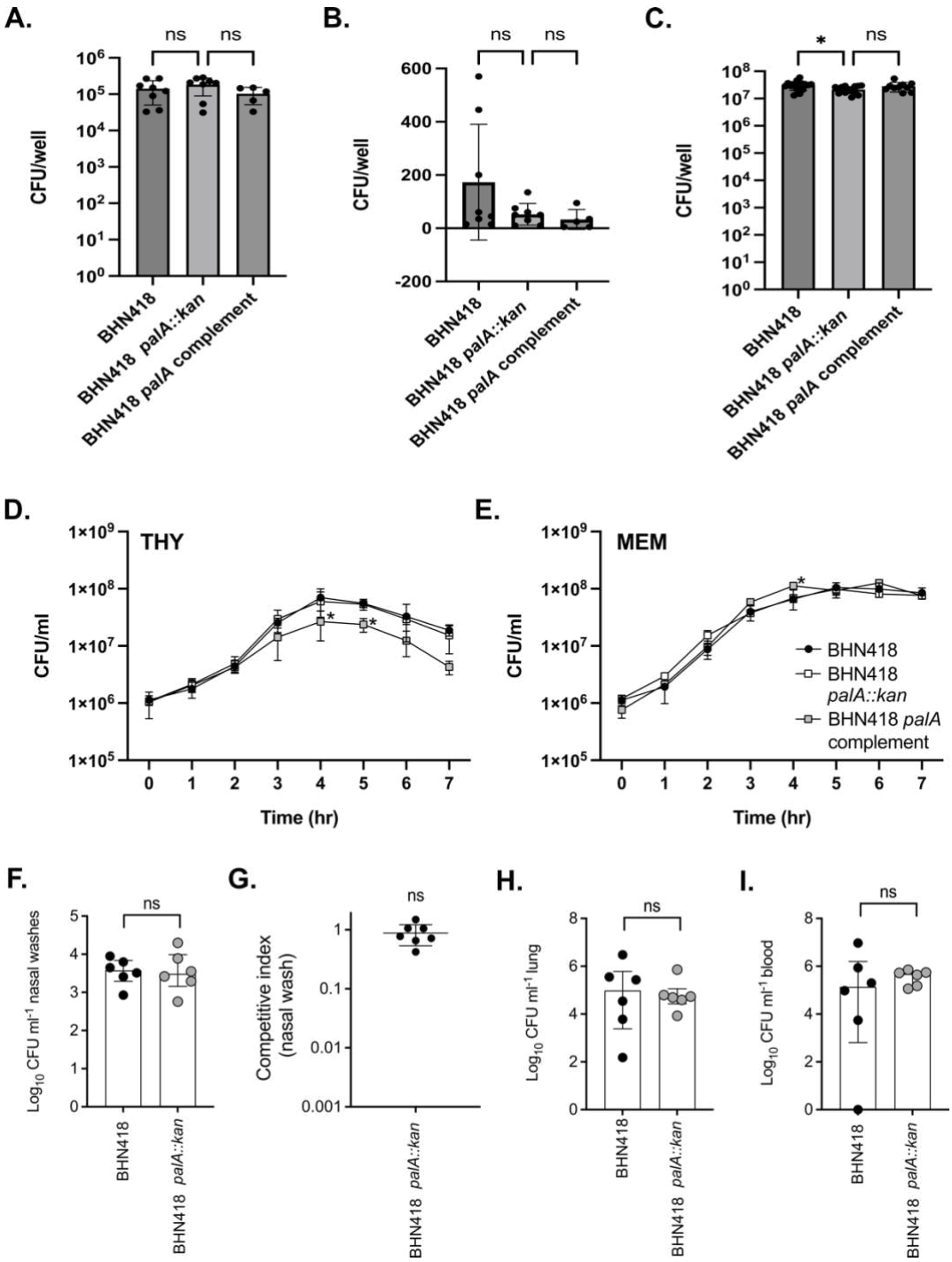

Figure 5. PalA was not essential for NPE microinvasion, murine colonisation and progression to disease. (A-C) NPE microinvasion by WT BHN418 and *palA* mutants, measured as NPE-associated bacteria (A), internalised bacteria (B) and planktonic bacteria growing in proximity with Detroit 562 NPE cells 3 hours post infection. (D-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 $p < 0.05$.

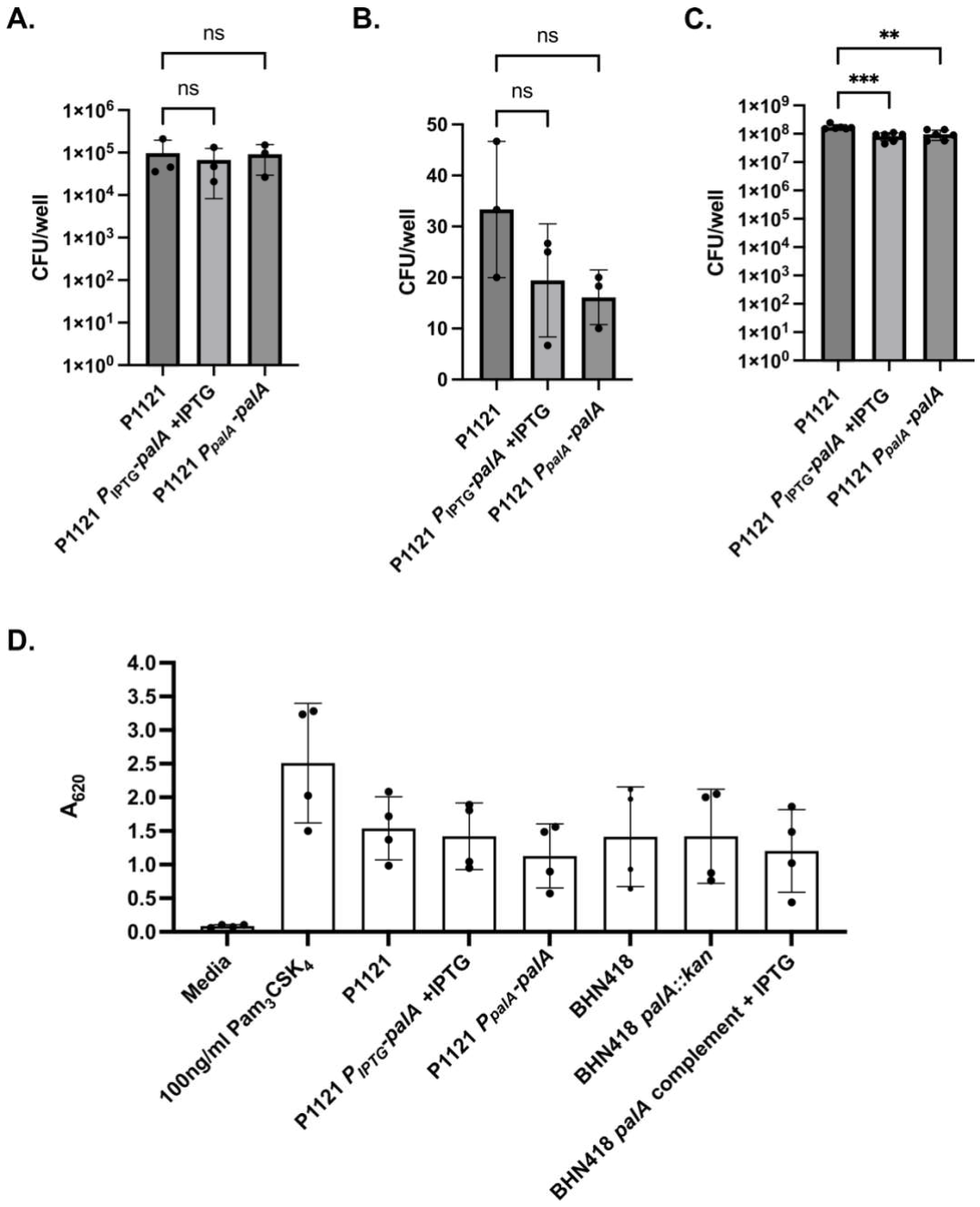

Figure 6. Heterologous expression of *palA* in P1121 (serotype 23F) did not increase epithelial microinvasion or TLR2 signalling. (A-C) NPE microinvasion by WT P1121 and *palA* expression mutants, measured as NPE-associated bacteria (A), internalised bacteria (B) and planktonic bacteria growing in proximity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

DISCUSSION

In this study, we demonstrated that pneumococcal lipoproteins trigger inflammation during epithelial colonisation at least partially via the TLR2-dependent pathway. We have previously shown that epithelial microinvasion occurs in the absence of disease and that 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, 50). However, 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 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 (51–56). In infant mice co-infected with pneumococci and influenzae, interferon signalling increases bacterial shedding while protecting against invasive disease, while TLR2 signalling limits bacterial shedding and transmission (18, 57, 58). Since TLR2 signalling 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, colonisation or progression to disease in people. Nonetheless, our results suggest that mucosal innate immunity

386 could be targeted alongside the induction of an adaptive immune response to
387 prevent pneumococcal colonisation, transmission and invasive disease.
388 We have implicated lipoprotein expression in intraspecies differences in NPE
389 cell microinvasion but our genetic and mutational analysis lead us to suggest that
390 this is not due to differences in lipoprotein repertoire, but rather due to differential
391 post-translational lipidation. The lipoprotein PsaA functions as a bacterial adhesin
392 which binds to host cell E-cadherin (59). In theory, we should have observed a
393 significant reduction in the number of epithelial-associated bacteria for both the
394 TIGR4 and BHN418 *lgt* mutants compared to wild-type due to the loss of PsaA
395 surface presentation; however, our data did not reflect this hypothesis. Similar subtle
396 effects have been seen for example when mutating the lipoprotein encoding gene
397 *dacB* but not *lgt* alters murein sacculus composition, and when mutating *lgt* results in
398 strain-dependent variable effects on growth in rich medium (5, 9, 14, 50, 60). The
399 observation that genetic complementation of *lgt* in BHN418 restores pneumococcal
400 ability to trigger epithelial inflammation but not WT levels of microinvasion further
401 suggests that regulation of lipoprotein processing may be more complex than
402 previously thought. The assumption that Lgt and other lipoprotein modification
403 proteins are constitutively expressed and active, and all 30+ lipoproteins are
404 processed at similar rates may not be entirely correct. Proteomics and
405 immunoblotting analyses showed that abundance of specific lipoproteins may
406 increase, decrease or show no change when lipoprotein processing is disrupted in *S.*
407 *pneumoniae*, although no clear patterns were apparent (5, 13). We therefore
408 speculate that unlike Gram-negative bacteria, for which lipoprotein processing is
409 essential, pneumococci and other Gram-positive bacteria compensate for the loss of

Lgt by differentially regulating expression of lipoprotein encoding genes, which in turn is regulated by external stimuli in nasopharyngeal niche (61, 62).

While investigating the intraspecies variation in the role of *lgt* in microinvasion, we discovered a previously uncharacterised lipoprotein encoding gene (*palA*) and its associated genetic island. Although *palA* presence is enriched in 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, colonisation or disease. 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 versus ear tropism in serotype 3 and serotype 14 strains (25). It is therefore possible that *palA* functions in promoting niche specialisation by facilitating 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 signalling 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, 50). Additionally, we have characterised a novel accessory lipoprotein likely acquired through horizontal gene transfer but rejected the hypothesis that this lipoprotein contribute to strain differences in pneumococcal epithelial microinvasion. Instead, we postulate that differential regulation of lipoprotein gene expression responding to the nasopharyngeal niche regulate this microinvasion process.

MATERIALS AND METHODS

Bacterial growth and maintenance. *Streptococcus pneumoniae* strains were grown on Columbia agar base with 5% defibrinated horse blood (CBA plates; EO Labs, Oxoid) or statically in Todd-Hewitt broth supplemented with 0.5% yeast extract (THY; 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), 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 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.

Bacterial genetic manipulation. *S. pneumoniae* were genetically manipulated using a competence stimulating peptide (CSP)-mediated transformation assay (63). Briefly, pneumococci were grown in THY pH 6.8 supplemented with 1 mM $CaCl_2$ and 0.02% BSA at 37°C, 5% CO_2 to OD_{600} 0.01- 0.03, pelleted and resuspended in 1/12 volume 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) was added to the bacterial suspension and incubated at RT for 5 mins. The suspensions were then mixed with ~300 ng transforming DNA, incubated at 37°C, 5% CO_2 for two 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 were generated using overlap-extension PCR. Complementation and expression constructs were generated by inserting the target gene into the complementation plasmid pASR103 or pPEPY (64), which allows for integration of the construct at a chromosomal ectopic site. Plasmids used are listed in Table 3, while primers are listed in Supplementary Table 2.

Cell culture. Detroit 562 (ATCC® CCL-138™; human pharyngeal carcinoma epithelial cells) were expanded and maintained in MEMα (Gibco™ 22561021) supplemented with 10% heat-inactivated FBS (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 DMEM (4.5 g/L glucose, 2mM glutamine, sodium pyruvate) supplemented with 10% HI-FBS at 37°C, 5% CO_2. Per manufacturer's instructions, DMEM growth medium was supplemented with 100 µg/ml normocin™ and/or 1X HEK-Blue™ Selection (Invivogen) where appropriate.

NPE infections. Adherence-invasion infections of confluent Detroit 562 cells with *S. pneumoniae* strains were performed at MOI 20 (P1121/23F derived strains) or MOI 10 (all others) for 3 hours. Working bacterial stocks were thawed, centrifuged to remove freezing medium and resuspended in infection medium (MEMα with 1% HI-FBS) to the appropriate CFU. 1 ml bacterial suspension were 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 were removed for CFU enumeration. For adherence assays, cells were washed thrice with 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 3x PBS wash, lysis with 1% saponin and CFU enumeration. Experiments were performed at least thrice on different days (n≥3 biological replicates) with technical duplicates. Statistical significance was determined using one-way 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 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≥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

509 on different days (n≥3 biological replicates) without technical replicates. Statistical
510 significance was determined using Student's *t*-test assuming equal variance.
511
512 **qPCR.** RNA from epithelial cells stored in RNA*later* were extracted using RNeasy
513 Mini kit (Qiagen) according to manufacturer instructions. Carryover DNA was
514 removed with TURBO DNA-*free* kit (Ambion), and cDNA generated using
515 LunaScript® RT Supermix kit (NEB). qPCR was performed using Luna® Universal
516 qPCR Master Mix (NEB) in technical triplicates with primers specific for *GAPDH*,
517 *CXCL10*, *IFNB1*, *IFNL1,* and *IFNL3* (Supplementary Table 2). Whenever possible
518 qPCR primers were designed to span exon-exon junctions. Cycling conditions are as
519 follows: 95°C for 5 mins, 40 cycles of 95°C for 15 secs and 60.5°C for 45 secs, with
520 a plate read at the end of each cycle. Data was analysed using the $2^{\Delta\Delta Ct}$ method,
521 with media only control and *GAPDH* levels for normalization. Statistical significance
522 was determined using Student's *t*-test assuming equal variance.
523
524 **HEK-Blue hTLR2 reporter assay.** HEK-Blue™ hTLR2 secreted alkaline
525 phosphatase (SEAP) reporter assays were performed according to manufacturer
526 instructions (Invivogen, hkb-htlr2). Briefly, HEK-Blue™ hTLR2 cells, *S. pneumoniae*
527 and control reagents were resuspended or diluted in pre-warmed HEK-Blue™
528 Detection medium (Invivogen). 5 x 10^4 HEK-Blue™ hTLR2 cells were mixed with 5 x
529 10^5 CFU *S. pneumoniae* (MOI 10) and incubated for 16 hours at 37°C, 5% CO_2.
530 SEAP activity was then measured spectroscopically at A_{620}. 100 ng/ml of Pam_2CSK_4
531 and Pam_3CSK_4 (TLR2 agonist, Bio-Techne) were used as positive controls, while
532 bacterial-free medium was used as negative control. Experiments were performed at
533 least thrice on different days (n≥3 biological replicates) with technical triplicates.

534 Statistical significance was determined using one-way ANOVA with Bonferroni's
535 multiple comparison test.
536
537 **Lipoprotein prediction using MEME suite.** Amino acid sequences of thirty-nine
538 published D39 lipoproteins were used with the motif discovery tool MEME to identify
539 pneumococcal lipoprotein motif(s) (5, 13, 22). The top two MEME results were
540 combined to obtain motif: L[LA][AS][AL]LXL[AV]A**C**[SG][NQS], a modified extension
541 of the minimal lipobox motif LAG**C** (5).
542
543 The obtained motif was used with the motif scanning tool FIMO to identify
544 lipoproteins in the genomes of *S. pneumoniae* TIGR4, BHN418 and D39, with the
545 latter used for quality control (65). Match *p*-value was set to 0.001. FIMO results
546 were further filtered with the following criteria: (i) presence of the lipidated cysteine
547 residue in the motif, (ii) presence of motif in the first 70 a.a. of the sequence, iii)
548 positive prediction as lipoprotein by SignalP-6.0 (66).
549
550 **Genomic analysis.** Presence of *palA* and its associated genetic island were
551 determined using Local-BLAST (BLASTN, TBLASTN) for the Malawian carriage
552 dataset (n=51,379) and serotype 23F strain P1121 (67). The built-in BLAST tool on
553 pubmlst.org was used for analysis of the BIGSdb dataset (49). BLASTN and
554 TBLASTN tools on the NCBI database were used to identify *palA* and PalA
555 homologues in non-pneumococcal species (67, 68). BLAST results were exported in
556 csv format and further analysed using R (v3.6.0) in RStudio
557 (http://www.rstudio.com/). Presence/absence of *palA* was annotated onto a Newick
558 tree showing phylogeny of the Malawian carriage strains by metabolic type and

559 visualized using iTOL (48, 69). Potential gene functions were inferred through the
560 results of BLASTP and NCBI Conserved Domain Database searches (67, 68, 70).
561
562 The BHN418 genome assembly was generated by combining long read sequencing
563 (PacBio) and short read sequencing (Illumina) methods which resulted in a single
564 contiguous chromosome of BHN418 of length 2,107,426 bp. *De novo* assembly was
565 performing using the Unicycler v0.4.8 pipeline in bold mode, quality assessed using
566 QUAST v5.1.0rc1 and annotated using Bakta v1.8.2 as described previously (71–
567 74). TIGR4 sequencing reads were aligned to the BHN418 genome using Samtools
568 v1.14 and visualised using IGV v2.16.1 (26, 68).
569
570 **TLR2 transcriptional module analysis.** TLR2-mediated transcriptional activity in
571 Detroit 562 cells infected with TIGR4 and BHN418 for 3 hours were determined
572 using published RNAseq data (2). We generated a transcriptional module reflective
573 of TLR2 activity derived from genes overexpressed in fibroblasts stimulated with
574 TLR2 agonists Pam_2CSK_4 and/or FSL-1 for 6 hours relative to unstimulated controls
575 (>1.5 fold; paired *t*-test with α of $p<0.05$ without multiple testing correction) (Gene
576 Expression Omnibus (GEO) dataset GSE92466) (Supplementary Table 1) (19).
577 Module expression was determined by calculating the geometric mean expression of
578 all constituent genes found in the analysed RNAseq dataset. Performance was
579 validated using data derived from Acute Myeloid Leukemia cells (GEO datasets
580 GSE92744) and CD14+ monocytes stimulated with Pam_3CSK_4 (GEO dataset
581 GSE78699) (Supplementary Figure 1) (75, 76).
582

583 **Murine experiments.** Outbred female CD1 mice (Charles River Laboratories) were
584 inoculated intranasally under anaesthetic (isoflurane) with 1 x 10^7 CFU bacteria (n=6
585 for single inoculation colonisation and pneumonia model, n=7 for competition
586 experiment). For colonisation experiments, nasal washes were performed 7 days
587 post infection using 1 ml PBS. For pneumonia model, mice were sacrificed 24 hpi
588 and bacteria recovered from the blood and homogenized lungs. CFU numbers were
589 enumerated using CBA supplemented with 4 µg/ml gentamicin, with additional 250
590 µg/ml kanamycin where appropriate. All animal procedures were approved by the
591 local ethical review process and conducted in accordance with the relevant UK
592 Home Office approved project license (PPL70/6510). Mice were housed for at least
593 one week under standard conditions before use. Randomisation or blinding was not
594 performed for these experiments. Statistical significance was determined using
595 Mann-Whitney test.
596
597 **Data availability.** BHN418 genome was deposited to NCBI with accession number
598 PRJNA1022026. TIGR4 sequencing reads were downloaded from NCBI Sequence
599 Reads Archive (accession SRX6259281), while P1121 reads were downloaded from
600 the EMBL-EBI database (accession ERS1072059) (77, 78). D39 and TIGR4 whole
601 genome assemblies were downloaded from NCBI GenBank database (accession
602 numbers CP000410.2 and AE005672.3, respectively) (79). All other genomic
603 sequences used were hosted on the PubMLST Pneumococcal Genome Library
604 (https://pubmlst.org/organisms/streptococcus-pneumoniae/pgl) (48, 49). RNAseq
605 data used in the TLR2 transcriptional module expression analysis were obtained
606 from the ArrayExpress database (accession E-MTAB-7841) (6).
607

608 **Table 3. Bacterial strains and plasmids used in this study.**

Designation	Genotype/ Description	Source
Strains		
TIGR4	WT Serotype 4 isolate	(26)
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	(64)
pPEPY	Complementation construct with a *kan* selectable marker	(64)
pEMcat	Minitransposon plasmid; source of cm^R cassette	(80)
pABG5	Cloning plasmid; source of kan^R cassette	(81)
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

609

ACKNOWLEDGEMENTS

This study was funded by a Medical Research Council grant (MR/T016329/1) awarded to RSH and JSB which supported JMC and CMW. JMC and MB 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. GP is supported by funding from the UCLH NIHR Biomedical Research Centre. RSH is a NIHR Senior Investigator. The views expressed are those of the authors and not necessarily those of the NIHR.

REFERENCES

1. Weiser JN, Ferreira DM, Paton JC. 2018. *Streptococcus pneumoniae*: Transmission, colonization and invasion. Nat Rev Microbiol 16:355–367.
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.
3. Jochems SP, de Ruiter K, Solórzano C, Voskamp A, Mitsi E, Nikolaou E, Carniel BF, Pojar S, German EL, Reiné J, Soares-Schanoski A, Hill H, Robinson R, Hyder-Wright AD, Weight CM, Durrenberger PF, Heyderman RS, Gordon SB, Smits HH, Urban BC, Rylance J, Collins AM, Wilkie MD, Lazarova L, Leong SC, Yazdanbakhsh M, Ferreira DM. 2019. Innate and adaptive nasal mucosal immune responses following experimental human pneumococcal colonization. J Clin Invest 129:4523–4538.
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.
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.
6. Mitchell AM, Mitchell TJ. 2010. *Streptococcus pneumoniae*: Virulence factors and variation. Clin Microbiol Infect 16:411–418.
7. Moffitt K, Howard A, Martin S, Cheung E, Herd M, Basset A, Malley R. 2015. Th17-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.
8. Jang A-Y, Ahn KB, Zhi Y, Ji H, 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.
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. 2020. TLR-mediated inflammatory responses to *Streptococcus pneumoniae* are highly dependent on surface expression of bacterial lipoproteins. J Immunol 193:3736–3745.
10. Barendt SM, Sham L, 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.
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.
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. Curr Biol 6:1553–1561.
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.
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.
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.
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.
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.
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.
19. Della Mina E, Borghesi A, Zhou H, Bougarn S, Boughorbel S, Israel L, Meloni I, Chrabieh M, Ling Y, Itan Y, Renieri A, Mazzucchelli I, Basso S, Pavone P, Falsaperla R, Ciccone R, Cerbo RM, Stronati M, Picard C, Zuffardi O, Abel L, Chaussabel D, Marr N, Li X, Casanova J-L, Puel A. 2017. Inherited human IRAK-1 deficiency selectively impairs TLR signaling in fibroblasts. Proc Natl Acad Sci 114:E514–E523.

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.
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.
22. Bailey TL, Johnson J, Grant CE, Noble WS. 2015. The MEME Suite. Nucleic Acids Res 43:W39–W49.
23. Wilkens S. 2015. Structure and mechanism of ABC transporters. F1000Prime Rep 7:14.
24. Rosenow C, Maniar M, Trias J. 1999. Regulation of the alpha-galactosidase activity in *Streptococcus pneumoniae*: Characterization of the raffinose utilization system. Genome Res 9:1189–1197.
25. 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.
26. Tettelin H, Nelson KE, Paulsen IT, Eisen JA, Read TD, Peterson S, Heidelberg J, DeBoy RT, Haft DH, Dodson RJ, Durkin AS, Gwinn M, Kolonay JF, Nelson WC, Peterson JD, Umayam LA, White O, Salzberg SL, Lewis MR, Radune D, Holtzapple E, Khouri H, Wolf A le. M, Utterback TR, Hansen CL, McDonald LA, Feldblyum T V., Angiuoli S, Dickinson T, Hickey EK, Holt IE, Loftus BJ, Yang F, Smith HO, Venter JC, Dougherty BA, Morrison DA, Hollingshead SK, Fraser CM. 2001. Complete genome sequence of a virulent isolate of *Streptococcus pneumoniae*. Science (80-) 293:498–506.
27. Spellerberg B, Cundell DR, Sandros J, Pearce BJ, Idänpään-Heikkilä I, Rosenow C, Masure HR. 1996. Pyruvate oxidase, as a determinant of virulence in Streptococcus pneumoniae. Mol Microbiol 19:803–813.
28. 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.
29. 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.
30. 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.
31. Alloing G, de Philip P, Claverys J-P. 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.
32. Alloing G, Trombe M, Claverys J-P. 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.
33. Farshchi Andisi V, Hinojosa CA, de Jong A, Kuipers OP, Orihuela CJ, Bijlsma JJE. 2011. Pneumococcal gene complex involved in resistance to extracellular oxidative stress. Infect Immun 80:1037–1049.

34. 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.
35. 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.
36. 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.
37. 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.
38. Weinrauch Y, Lacks SA. 1981. Nonsense mutations in the amylomaltase gene and other loci in *Streptococcus pneumoniae*. Mol Gen Genet 183:7–12.
39. 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.
40. 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.
41. 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.
42. 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.
43. 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.
44. 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.
45. 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 155:2401–2410.
46. 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.
47. Hermans PWM, Adrian P V, 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.

48. Gori A, Obolski U, Swarthout TD, Lourenço J, Weight CM, Cornick J, Kamng'ona A, Mwalukomo TS, Msefula J, Brown C, Maiden MC, French N, Gupta S, Heyderman RS. 2021. The metabolic, virulence and antimicrobial resistance profiles of colonizing Streptococcus pneumoniae shift after pneumococcal vaccine introduction in urban Malawi. medRxiv 2021.07.21.21260914.
49. Jolley KA, Maiden MCJ. 2010. BIGSdb: Scalable analysis of bacterial genome variation at the population level. BMC Bioinformatics 11:595.
50. 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.
51. Joyce EA, Popper SJ, Falkow S. 2009. *Streptococcus pneumoniae* nasopharyngeal colonization induces type I interferons and interferon-induced gene expression. BMC Genomics 10:404.
52. Skovbjerg S, Nordén R, Martner A, Samuelsson E, Hynsjö L, Wold AE. 2017. Intact pneumococci trigger transcription of interferon-related genes in human moncytes, while fragmented, autolyzed bacteria subvert this response. Infect Immun 85:e00960-16.
53. 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.
54. 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.
55. 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.
56. Ruiz-Moreno JS, Hamann L, Jin L, Sander LE, Puzianowska-Kuznicka M, Cambier J, Witzenrath M, Schumann RR, Suttorp N, Opitz B, Group CS. 2018. The cGAS/STING pathway detects *Streptococcus pneumoniae* but appears dispensable for antipneumococcal defense in mice and humans. Infect Immun 86:e00849-17.
57. 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.
58. 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.
59. 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.
60. 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.
61. Smithers L, Olatunji S, Caffrey M. 2021. Bacterial Lipoprotein Posttranslational Modifications. New Insights and Opportunities for Antibiotic and Vaccine Development. Front Microbiol 12:788445.
62. Nguyen MT, Matsuo M, Niemann S, Herrmann M, Götz F. 2020. Lipoproteins in Gram-Positive Bacteria: Abundance, Function, Fitness. Front Microbiol 11:582582.
63. 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.
64. Keller LE, Robinson DA, McDaniel LS. 2016. Nonencapsulated *Streptococcus pneumoniae*: Emergence and pathogenesis. MBio 7:e01792-15.
65. Grant CE, Bailey TL, Noble WS. 2011. FIMO: scanning for occurrences of a given motif. Bioinformatics 27:1017–1018.
66. 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.
67. 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.
68. Sayers EW, Barrett T, Benson DA, Bryant SH, Canese K, Chetvernin. V, Church DM, Dicuccio M, Edgar R, Federhen S, Feolo M, Geer LY, Helmberg W, Kapustin Y, Landsman D, Lipman DJ, Madden TL, Maglott DR, Miller V, Mizrachi I, Ostell J, Pruitt KD, Schuler GD, Sequeira E, Sherry ST, Shumway M, Sirotkin K, Souvorov A, Starchenko G, Tatusova TA, Wagner L, Yaschenko E, Ye J. 2009. Database resources of the National Center for Biotechnology Information. Nucleic Acids Res 37:5–15.
69. 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.
70. Machler-Bauer A, Bo Y, Han L, He J, Lanczycki C, Lu S, Chitsaz F, Derbyshire M, Geer R, Gonzales N, Gwadz M, Hurwitz D, Lu F, Machler G, Song J, Thanki N, Wang Z, Yamashita R, Zhang D, Zheng C, Geer L, Bryant S. 2017. CDD/SPARCLE: functional classification of proteins via subfamily domain architectures. Nucleic Acids Res 45:D200–D203.
71. 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.
72. Gurevich A, Saveliev V, Vyahhi N, Tesler G. 2013. QUAST: quality assessment tool for genome assemblies. Bioinformatics 29:1072–1075.
73. 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 10:10.1128/mra.00715-21.
74. 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. Microb Genomics 7:000685.

75. 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.
76. 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.
77. Tettelin H, Masignani V, Cieslewicz MJ, Donati C, Medini D, Ward NL, Angiuoli S V, Crabtree J, Jones AL, Durkin AS, Deboy RT, Davidsen TM, Mora M, Scarselli M, Ros IM y, Peterson JD, Hauser CR, Sundaram JP, Nelson WC, Madupu R, Brinkac LM, Dodson RJ, Rosovitz MJ, Sullivan SA, Daugherty SC, Haft DH, Selengut J, Gwinn ML, Zhou L, Zafar N, Khouri H, Radune D, Dimitrov G, Watkins K, O'Connor KJB, Smith S, Utterback TR, White O, Rubens CE, Grandi G, Madoff LC, Kasper DL, Telford JL, Wessels MR, Rappuoli R, Fraser CM. 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.
78. Pojar S, Basset A, Gritzfeld JF, Nikolaou E, Selm S van, Eleveld MJ, Gladstone RA, Solórzano C, Dalia AB, German E, Mitsi E, Connor V, Hyder-Wright AD, Hill H, Hales C, Chen T, Camilli A, Collins AM, Rylance J, Bentley SD, Jochems SP, Jonge MI de, Weiser JN, Cleary DW, Clarke S, Malley R, Gordon SB, Ferreira DM. 2020. Isolate differences in colonization efficiency during experimental human pneumococcal challenge. medRxiv 2020.04.20.20066399.
79. Lanie JA, Ng W-L, Kazmierczak KM, Andrzejewski TM, Davidsen TM, Wayne KJ, Hervé T, 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.
80. 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.
81. 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.
82. Ramos I, Smith G, Ruf-Zamojski F, Martínez-Romero C, Fribourg M, Carbajal EA, Hartmann BM, Nair VD, Marjanovic N, Monteagudo PL, DeJesus VA, Mutetwa T, Zamojski M, Tan GS, Jayaprakash C, Zaslavsky E, Albrecht RA, Sealfon SC, García-Sastre A, Fernandez-Sesma A. 2019. Innate Immune Response to Influenza Virus at Single-Cell Resolution in Human Epithelial Cells Revealed Paracrine Induction of Interferon Lambda 1. J Virol 93:10.1128/jvi.00559-19.