Clonal Expansion of a *Streptococcus pneumoniae* Serotype 3 Capsule Variant Sequence Type 700 With Enhanced Vaccine Escape Potential After 13-Valent Pneumococcal Conjugate Vaccine Introduction

Akuzike Kalizang’oma,1,2 Todd D. Swarthout,1,2,3 Thandie S. Mwalukomo,4 Arox Kamng’ona,5,6 Comfort Brown,5,6 Jacqueline Msefula,5,6 Hayley Demetriou,6 Jia Mun Chan,1 Lucy Roalfe,6 Uri Obolski,7 Jose Lourenço,6 David Goldblatt,1,6 Chrispin Chaguza,1,9,10,11 Neil French,12,a, and Robert S. Heyderman12,a

1NIHR Mucosal Pathogens Research Unit, Research Department of Infection, Division of Infection and Immunity, University College London, London, United Kingdom; 2Pneumonia and Meningitis Pathogens Associate Research Group, Malawi-Liverpool-Wellcome Research Programme, Blantyre, Malawi; 3Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht University, Utrecht, Netherlands; 4School of Medicine and Oral Health, Kamuzu University of Health Sciences, Blantyre, Malawi; 5School of Life Sciences and Allied Health Professionals, Kamuzu University of Health Sciences, Blantyre, Malawi; 6Great Ormond Street Institute of Child Health, University College London, London, United Kingdom; 7Porter School of the Environment and Earth Science, Tel-Aviv University, Tel-Aviv, Israel; 8Faculdade de Ciências, BioISI, Universidade de Lisboa, Lisboa, Portugal; 9Parasites and Microbes Programme, Wellcome Sanger Institute, Hinxton, United Kingdom; 10Department of Epidemiology of Microbial Diseases, Yale School of Public Health, Yale University, New Haven, Connecticut, USA; 11Yale Institute for Global Health, Yale University, New Haven, Connecticut, USA; and 12Institute of Infection, Veterinary and Ecological Sciences, Department of Clinical Infection, Microbiology and Immunology, University of Liverpool, Liverpool, United Kingdom

**Background.** *Streptococcus pneumoniae* serotype 3 remains a problem globally. Malawi introduced 13-valent pneumococcal conjugate vaccine (PCV13) in 2011, but there has been no direct protection against serotype 3 carriage. We explored whether vaccine escape by serotype 3 is due to clonal expansion of a lineage with a competitive advantage.

**Methods.** The distribution of serotype 3 Global Pneumococcal Sequence Clusters (GPSCs) and sequence types (STs) globally was assessed using sequences from the Global Pneumococcal Sequencing Project. Whole-genome sequences of 135 serotype 3 carriage isolates from Blantyre, Malawi (2015–2019) were analyzed. Comparative analysis of the capsule locus, entire genomes, antimicrobial resistance, and phylogenetic reconstructions were undertaken. Opsonophagocytosis was evaluated using serum samples from vaccinated adults and children.

**Results.** Serotype 3 GPSC10-ST700 isolates were most prominent in Malawi. Compared with the prototypical serotype 3 capsular polysaccharide locus sequence, 6 genes are absent, with retention of capsule polysaccharide biosynthesis. This lineage is characterized by increased antimicrobial resistance and lower susceptibility to opsonophagocytic killing.

**Conclusions.** A serotype 3 variant in Malawi has genotypic and phenotypic characteristics that could enhance vaccine escape and clonal expansion after post-PCV13 introduction. Genomic surveillance among high-burden populations is essential to improve the effectiveness of next-generation pneumococcal vaccines.

**Keywords.** *Streptococcus pneumoniae*; serotype 3; pneumococcal conjugate vaccine; capsule; Africa.
termed GPSC12-CC180, is the most dominant serotype 3 strain worldwide comprising 3 lineages (Iα, Iβ, and II) [8]. Compared with other GPSC12-CC180 lineages, the clade II lineage is characterized by a higher prevalence of antimicrobial resistance (AMR), greater diversity of surface protein antigens, and a higher rate of recombination mediated by transposons, phages, and integrative conjugative elements carrying genes associated with competence and virulence [8].

Malawi introduced PCV13 into the national Expanded Programme on Immunisation in 2011, using a 3 + 0 schedule (1 dose at 6, 10, and 14 weeks of age) with a catch-up campaign limited to children aged <1 year. There has been high PCV13 coverage, exceeding 90%, with direct protection against vaccine serotype IPD among vaccinated infants [11]. However, in line with reports from Kenya and The Gambia [12, 13], indirect protection of older unvaccinated children and adults has been less effective than in high-income countries, with relatively high residual carriage of all 13 vaccine serotypes among PCV13-vaccinated children, up to 8 years after PCV13 introduction in Malawi [1, 14]. In particular, there has been no apparent direct protection against serotype 3 carriage [14], and population-level serosurveillance suggests limited vaccine-induced serotype 3-specific immunity [15]. Because severe disease is a dead end for S. pneumoniae, virulence and pathogenicity largely develop as a consequence of active pneumococcal adaptation to colonize the host and transmit from person to person. We therefore hypothesized that the apparent vaccine escape by serotype 3 in Malawi could be due, at least in part, to the clonal expansion of a lineage with a competitive advantage in the nasopharyngeal niche.

Through an analysis of the genome sequences of 135 serotype 3 isolates from sequential carriage surveys [14, 16], we describe the clonal expansion of a sequence type (ST) 700–GPSC10 serotype 3 lineage in Malawi following PCV13 introduction, a lineage distinct from the GPSC12-CC180 lineage described elsewhere [8]. ST700-GPSC10 is characterized by the absence of 6 genes in its serotype 3 CPS locus, a different AMR profile from other serotype 3 lineages in Malawi, and lower susceptibility to opsonophagocytosis that may explain the spread of this lineage with enhanced evasion of vaccine-induced immunity.

**METHODS**

**Genome Sequences**

The genome sequences of 135 serotype 3 carriage isolates were used in this analysis (Supplementary Table 1) from pneumococcal carriage surveillance between 2015 and 2019 in Blantyre, Malawi [14, 16]. An additional 592 genome sequences were accessed from the Global Pneumococcal Sequencing (GPS) project spanning >2 decades (1993–2018) in 32 countries worldwide, including Malawi [10, 17].

**Genome Analysis**

GPSCs and STs based on the pneumococcal multilocus sequence typing scheme [18] were determined using PopPUNK software (version 2.6.0) [19] and PubMLST, respectively [20]. We scanned each genome for the presence of AMR genes (tetM, cat, mefA, and ermB), using Abricate software (version 1.0.1) with the ResFinder database (https://github.com/tseemann/abricate) and PathogenWatch (https://pathogen.watch/) with ResFinder and the National Center for Biotechnology Information databases; minimum inhibitory concentration (MICs) to penicillin were genotypically predicted using an analysis pipeline developed by the US Centers for Disease Control and Prevention [21]. The MICs were interpreted using Clinical and Laboratory Standards Institute guidelines and clinical breakpoints (https://clsi.org/). Genome assemblies were generated using SPAdes software (version 2.5) [22]. CONTIGuator software (version 2.0) [23] was used to order the contigs in the assemblies of the Malawian serotype 3 isolates, and serotype 3 (strain 670-6B; GenBank accession no. NC_014498) was used as a reference.

The CPS locus is known to be physically located between aliA and dexpB genes in the pneumococcus [24]. The prototypical serotype 3 CPS reference locus [24] has the following genes, dexpB, aliB, tsp, wzg, wzf, wzd, wze, tsp, wgo, wcE, galU, and pgm, and our analysis is based on the gene content comparisons at the physical locations with the serotype 3 reference CPS locus. Although only the core CPS genes (wzg, wzf, wzd, and wze) have known capsule biosynthesis functions [24, 25], others have unknown or potential auxiliary functions. Therefore, to avoid confusion, we refer to any genes located in the entire region known to harbor genes involved in capsule biosynthesis for serotype 3 as CPS genes. To locate the CPS locus and manually extract the corresponding sequences, each ordered genome was first compared against the serotype 3 reference genome described by Bentley et al [24] using blastn software (version 2.14.0) [26] and the Artemis Comparison Tool (version 18.0.3) [25]. The presence of serotype 3 capsule locus genes was also assessed by aligning the sequencing reads of each strain to the sequence of the CPS locus described by Bentley et al in 2006 [24] (using Bowtie2 [27] and Samtools [28] software).

The presence and absence of the CPS locus genes was done by mapping each genome against the serotype 3 reference CPS locus sequence (GenBank accession no. CR931634) from Bentley et al [24]. We used Snippy software (version 4.3.6) to map reads of each isolate against the reference CPS locus sequences to identify missing genes. The contiguity of genes present was assessed using the Artemis Comparison Tool (version 18.1.0) [25]. Genome comparisons between different serotype 3 lineages were carried out using whole-genome alignment and maximum-likelihood was estimated [27, 28]. The phylogeny was assessed in a subset of isolates (RaxML [29]; GTRGAMMA, 100 iterations), which included all 135 serotype 3 isolates from Malawi [14], as well as all 18 non-serotype 3 strains belonging to GPSC10 isolated in the same Malawi study.
Bacterial Growth
Representative strains from ST700 and ST5435 were grown overnight on Columbia blood agar supplemented with 5% horse blood at 37°C and 5% carbon dioxide (microaerophilic condition) or at 37°C anaerobically using the AnaeroGen W-Zip Compact system (Oxoid), per the manufacturer’s instructions. Image brightness and contrast were adjusted using Fiji software (version 2.14.0).

Opsonophagocytosis Assay
We undertook opsonophagocytosis analysis to evaluate bacterial resistance to antibody plus complement mediated killing, as described elsewhere [30]. Briefly, 3 ST700 and 3 ST5435 serotype 3 strains were selected at random from the isolate collection. These were compared to killing reduction against a reference serotype 3 strain (an optochin resistant variant of Wu2), in serum samples from 5 healthy adults vaccinated with 23-valent pneumococcal polysaccharide vaccine and from 6 PCV13-vaccinated children. Todd-Hewitt broth with yeast extract and filtered fetal bovine serum (4%) were used for the assay. The incubations were carried out in a 96-well plate with serial dilutions of human serum before further incubation with HL-60 cells and baby rabbit complement. Wells were plated out onto agar, and bacterial colonies were quantified with a plate reader. Opsonophagocytosis titers were expressed as the opsonic index [30].

Statistical Analysis
Statistical tests and associated diagrams were generated using R software (version 2.11.1) (R Core Team 2014; https://www.R-project.org/) and GraphPad Prism (version 8.0) (GraphPad Software). Nonparametric data, which included opsonic indices, are presented as individual data points and median values, and Mann-Whitney tests were used to compare theses indices. Differences were considered significant at \( P < .05 \).

Data Availability
Pneumococcal CPS locus de novo assemblies for representative STs are available from the National Center for Biotechnology Information (accession nos. OR805039–OR805049).

RESULTS

**GPSC10-ST700 as the Most Prominent *S. pneumoniae* Serotype 3 Genotype in Malawi**

We first compiled a global collection of whole-genome–sequenced serotype 3 isolates from the GPS data set to understand the geographic distribution of GPSC10-ST700. The carriage and invasive isolates included in this analysis originated from 32 countries worldwide, including Malawi. The proportion of the most common serotype 3 GPSCs and STs within each country is shown in Figure 1A and 1B (detailed in Supplementary Table 1) and displayed geographically in Figure 1C. Among the sequenced genomes, we identified serotype 3 GPSC10 in 6 countries (Figure 1A–1C): Kenya (14 of 14 isolates), India (1 of 20), Malawi (33 of 43), South Africa (7 of 110), The Gambia (2 of 49), and Togo (4 of 6). In contrast, although as reported previously [8], most of the serotype 3 isolates globally were associated with the GPSC12 lineage, including strains from the CC180 clone, we did not find the GPSC12 lineage among the 135 serotype 3 carriage isolates from Malawi or in our previous genomic analysis of invasive and carriage isolates from the same population [31, 32]. Instead, among the 135 carriage serotype 3 isolates from Malawi collected after the introduction of PCV13, GPSC10-ST700 was most prominent (72 of 135 isolates). Pneumococcal carriage studies have shown that serotype 3 has expanded after PCV13 introduction in Malawi [14], with ST700 now the most prominent clone [10].

**Multiple Absent Genes in the CPS Locus of GPSC10-ST700**

We then assessed the genetic diversity of the serotype 3 isolates from Malawi to understand the genetic attributes of GPSC10-ST700 serotype 3 isolates that may have contributed to their dominance compared with the other serotype 3 lineages in Malawi. First, mapping the sequences of each serotype 3 isolate genome against the prototypical serotype 3 reference CPS locus sequence described by Bentley et al [24], we observed an absence of 6 genes in the CPS locus of the GPSC10-ST700 lineage (Figure 1D and Table 1). We confirmed this finding after performing de novo assemblies of the whole-genome sequences to reconstruct the CPS locus. We considered all the genes located between *dexB* and *pgm* in the serotype 3 reference CPS locus sequence described by Bentley et al [24] as part of the CPS locus (see Methods). We found that this CPS locus variant characterized by the absent genes was not exclusive to the GPSC10-ST700 lineage; rather, it is polyphyletic, appearing independently in ≥3 distinct phylogenetic clusters expressing serotype 3, including ST700 and ST3214. Despite the absent genes in the CPS locus, the resulting shorter CPS locus sequence still included the core CPS genes (*wchE*, *ugd*, and *galU*) necessary to produce the serotype 3 polysaccharide [24].

Our literature search revealed that the observed serotype 3 CPS locus variant associated with GPSC10-ST700 was analogous to a serotype 3 variant isolated from deceased chimpanzees in 2011 in Kenya [33]. Although the authors described the missing genes as pseudogenes, our annotation highlights their potential involvement in polysaccharide production and export (Table 1). The GPSC10-ST700 isolates from Malawi produce a functional serotype 3 polysaccharide capsule, as assessed by latex agglutination [34] and typical colony morphology (Figure 2). We observed mucoid colonies with no
morphological differences between representative strains showing the full (ST5435) and the variant (ST700) CPS locus when grown anaerobically or in a microaerophilic environment, although colony size and alpha hemolysis were reduced for both strains in anaerobic conditions (Figure 2). Owing to the absent 6 genes, ST700 strains lack homologues of \textit{cpsB} (\textit{wzh}), \textit{cpsC} (\textit{wzd}), and \textit{cpsD} (\textit{wze}), which encode for a bacterial tyrosine kinase system important for the regulation of capsule production in D39 derivative strains (serotype 2), particularly at the septum and in low oxygen conditions [3, 35, 36]. As only \textit{wchE}, \textit{ugd}, and \textit{galU} are involved in CPS-3 synthesis [24], the role that nonessential genes such as \textit{cpsB} play in regulating capsule thickness or function in serotype 3 remains to be determined.

GPSC10-ST700 Characterized by AMR and a Unique Genetic Profile

Next, we assessed the distribution of the AMR and virulence genes in the GPSC10-ST700 and other lineages associated with serotype 3 isolates. Figure 1C shows the maximum-likelihood phylogeny of all Malawian serotype 3 isolates (multiple GPSCs) together with GPSC10 strains associated with other serotypes. All except the isolates belonging to GPSC9

Figure 1.  Countries of origin and genomic profiles of \textit{Streptococcus pneumoniae} serotype 3 isolates. A, Global Pneumococcal Sequence Clusters (GPSCs) isolated in each country. B, Sequence types (STs) isolated in each country. C, Countries (in red) from which the serotype 3 \textit{S. pneumoniae} isolates described in this analysis originated. D, Maximum-likelihood tree including all serotype 3 pneumococci isolated from Blantyre, Malawi [14], and the non–serotype 3 strains belonging to the GPSC10 lineage. The tree is annotated with metadata columns showing (from left to right) the serotype, ST, GPSC, penicillin minimum inhibitory concentration (MIC) (predicted genetically), the presence and absence of antimicrobial resistance (AMR) genes (\textit{cat}, \textit{mefA}, \textit{ermB}, \textit{tetM}), and the presence and absence of each gene annotated on the serotype 3 capsular polysaccharide (CPS) locus [24]. In white are the absent 6 genes, and in red are the core CPS-3 genes (in reference to the serotype 3 locus defined by Bentley et al [24], shown above the gene locus panel).
are characterized by the 6-gene absence seen in the serotype 3 cluster described above. Only GPSC10-ST700 genomes were characterized by the presence of a tetracycline resistance gene and by a predicted MIC to penicillin >0.5 μg/mL (penicillin nonsusceptible).

The maximum-likelihood phylogeny of all pneumococcal isolates from Blantyre, Malawi, collected between 2015 and 2019, showed that GPSC10 included isolates belonging to serotype 3 and other serotypes, including 19A, 19F, and 14. This observation suggests that the GPSC10 lineage tends to undergo capsule switching. The GPSC10 serotype 3 and the other capsule-switched isolates—GPSC10-19A, GPSC10-19F, GPSC10-10A, and GPSC10-14—were distinguished by a minimum of 4599, 3264, 7774, and 6232 single-nucleotide polymorphisms, respectively. Therefore, considering that the serotype 3 isolates in the GPSC10-ST700 lineage differed by up to 7120 single-nucleotide polymorphisms (similar to the genetic divergence between serotype 3 and the other serotypes in the GPSC10 lineage), we suggest that GPSC10-ST700 serotype 3 strains may have originated through multiple capsular switching events from the same original genomic background.

To identify genotypic traits that explain the clonal expansion of serotype 3 GPSC10-ST700, full-genome comparison was carried out between isolates belonging to the different serotypes within GPSC10. This analysis highlighted 2 gene clusters present in ST700 serotype 3 strains (Table 2). The first cluster

---

### Table 1. Details of the 6 Absent Genes in the Sequence Type 700 Serotype 3 Capsular Polysaccharide Locus

<table>
<thead>
<tr>
<th>Feature Name</th>
<th>Locus Tag</th>
<th>Start</th>
<th>End</th>
<th>Feature Length</th>
<th>Feature Product</th>
</tr>
</thead>
<tbody>
<tr>
<td>aliB</td>
<td>SPC03_0002</td>
<td>321</td>
<td>494</td>
<td>174</td>
<td>Putative oligopeptide-binding protein AliB (pseudogene)</td>
</tr>
<tr>
<td>tnp</td>
<td>SPC03_0003</td>
<td>487</td>
<td>1336</td>
<td>850</td>
<td>Putative IS630-Spn1 transposase</td>
</tr>
<tr>
<td>wzg (cpsA)</td>
<td>SPC03_0004</td>
<td>1524</td>
<td>2748</td>
<td>1225</td>
<td>Integral membrane regulatory protein Wzg (pseudogene)</td>
</tr>
<tr>
<td>wzh (cpsB)</td>
<td>SPC03_0005</td>
<td>2750</td>
<td>3480</td>
<td>731</td>
<td>Protein-tyrosine phosphatase Wzh (pseudogene)</td>
</tr>
<tr>
<td>wzd (cpsC)</td>
<td>SPC03_0006</td>
<td>3489</td>
<td>4181</td>
<td>693</td>
<td>Capsular polysaccharide biosynthesis protein Wzd</td>
</tr>
<tr>
<td>wzg (cpsD)</td>
<td>SPC03_0007</td>
<td>4191</td>
<td>4604</td>
<td>414</td>
<td>Tyrosine-protein kinase Wze (pseudogene)</td>
</tr>
</tbody>
</table>

*Indicated in blue in Figure 1D (capsular polysaccharide reference sequence described by Bentley et al [24]; GenBank accession no. CR931634).
TABLE 2. Gene Clusters Potentially Associated With Pathogenicity—dipeptide permease Adenosine Triphosphate–Binding Cassette Transporter Cluster and Lantibiotic Cluster Present in Sequence Type 700 Serotype 3 Isolates

<table>
<thead>
<tr>
<th>Gene Cluster</th>
<th>Feature Name</th>
<th>Feature Length</th>
<th>Feature Product</th>
<th>Accession No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>dpp-ABC transporter cluster</td>
<td>Cyd</td>
<td>407</td>
<td>Cytidine deaminase</td>
<td>c00269</td>
</tr>
<tr>
<td></td>
<td>pggpA</td>
<td>488</td>
<td>Phosphatidyglycerophosphatase A (lipid transport and metabolism)</td>
<td>COG1267</td>
</tr>
<tr>
<td></td>
<td>shp3</td>
<td>1409</td>
<td>Hyphothetical protein</td>
<td>c14647</td>
</tr>
<tr>
<td></td>
<td>gisA</td>
<td>1982</td>
<td>ABC-type glutathione transport system ATPase component</td>
<td>COG1123</td>
</tr>
<tr>
<td></td>
<td>dppC</td>
<td>851</td>
<td>ABC-type dipeptide/oligopeptide/nickel transport system (permease)</td>
<td>COG1173</td>
</tr>
<tr>
<td></td>
<td>dppB</td>
<td>950</td>
<td>ABC-type dipeptide/oligopeptide/nickel transport system (permease)</td>
<td>COG00601</td>
</tr>
<tr>
<td></td>
<td>dppA</td>
<td>1628</td>
<td>ABC-type dipeptide/oligopeptide/nickel transport system (ATPase)</td>
<td>cd00999</td>
</tr>
<tr>
<td></td>
<td>nanM</td>
<td>923</td>
<td>N-acetylneuraminate (sialic acid) epimerase</td>
<td>c57290</td>
</tr>
<tr>
<td>Lantibiotic cluster</td>
<td>lanC</td>
<td>407</td>
<td>Lantionine synthetase C–like domain associated with serine/threonine kinases</td>
<td>cd04791</td>
</tr>
<tr>
<td></td>
<td>ptrB</td>
<td>488</td>
<td>Protease II</td>
<td>COG1770</td>
</tr>
<tr>
<td></td>
<td>mdlB</td>
<td>1409</td>
<td>ABC-type multidrug transport system, ATPase and permease component, ABC-type bacteriocin transporter</td>
<td>COG1132</td>
</tr>
<tr>
<td></td>
<td>lhaD</td>
<td>1982</td>
<td>ABC domain of the transporters involved in the export of lipoprotein and macrolide, putative bacteriocin export ABC transporter, ABC-type antimicrobial peptide transport system</td>
<td>cd03255</td>
</tr>
<tr>
<td></td>
<td>lhp5</td>
<td>851</td>
<td>Hyphothetical protein</td>
<td>...</td>
</tr>
</tbody>
</table>

Abbreviations: ABC, adenosine triphosphate–binding cassette; ATPase, adenosine triphosphatase; dpp, dipeptide permease.

*No conserved domains were found in the frame.

The dpp-ABC cluster carries a dipeptide permease (dpp) adenosine triphosphate–binding cassette (ABC) transporter, and the second cluster (the lantibiotic cluster) carries the genes for the production and metabolism of lantibiotics, a class of bacteriocins expressed by gram-positive bacteria [29].

Reduced Opsonophagocytic Killing of GPSC10-ST700 Strains Highlighted by Phenotypic Analysis

Three serotype 3 ST700 isolates (CPS-3 gene cluster absent) and 3 serotype 3 ST5435 isolates (CPS-3 gene cluster present) from Blantyre, obtained between 2015 and 2019. Were assessed for their ability to survive opsonophagocytosis with immune serum samples. When tested both against serum samples from adults (vaccinated with 23-valent pneumococcal polysaccharide vaccine) and serum samples from children (PCV13 vaccinated), ST700 was less susceptible to opsonophagocytic killing than other serotype 3 lineages from Malawi carrying a full-length CPS locus, specifically ST5435 (GPSC9). Taken together, these characteristics may explain the apparent competitive advantage of the GPSC10-ST700 lineage.

**DISCUSSION**

The control of *S. pneumoniae* serotype 3 carriage and disease has become a prominent challenge among vaccine developers and public health agencies [4, 14]. In contrast to other serotypes included in the PCV13 formulation, there is increasing evidence that PCV13 introduction has achieved no or only partial immunity to serotype 3 [4]. Moreover, as demonstrated by the CC180 lineage, serotype 3 has an increased tendency to recombine, which has resulted in the emergence of several clones associated with this serotype worldwide through capsule switching due to recombination-mediated swapping of the capsule locus between pneumococcal strains, as well as related streptococcal species [8]. However, detailed genomic analysis and phenotypic characterization of these serotype 3 lineages, especially those cocirculating in the same setting, have not been conducted thus far to determine why some of the serotype 3-associated lineages are more prevalent than others.

In the current study, we performed these genomic and phenotypic analyses to gain insights into the clonal expansion of the GPSC10-ST700 serotype 3 lineage in Malawi, the now-predominant serotype 3 lineage in Malawi, which is distinct from the CC180 lineage (GPSC12) more widespread globally [8]. We show that the GPSC10-ST700 serotype 3 lineage is common in Africa and has been identified among carriage and IPD isolates from Malawi and South Africa [17]. The serotype 3 strains in the GPSC10 lineage are characterized by the absence of 6 genes in the CPS locus, compared with the reference serotype 3 CPS locus sequence described by Bentley et al [24] and other serotype 3 lineages circulating in Malawi. The lineage is also tetracycline resistant, has a higher MIC for penicillin than other serotype 3 lineages isolated in Malawi, appears to have a propensity for capsule switching, and possesses virulence genes involved in pathogenicity and colonization. Furthermore, the GPSC10-ST700 lineage is less susceptible to opsonophagocytic killing than other serotype 3 lineages from Malawi carrying a full-length CPS locus, specifically ST5435 (GPSC9). Taken together, these characteristics may explain the apparent competitive advantage of the GPSC10-ST700 lineage.
lineage compared with the other serotype 3 lineages identified after the introduction of PCV13 vaccine in Malawi [14, 16]. Serotype and genotype specific factors may have contributed to GPSC10-ST700 becoming the most prominent S. pneumoniae serotype 3 genotype in Malawi. It has been well documented that the PCV13 rollout has had minimal effect on serotype 3 carriage globally [9, 35, 37, 38], and this has been supported by post-PCV13 studies that have reported that PCV13 is less immunogenic for serotype 3 [4, 15, 39, 40]. Such evasion of vaccine-induced immunity could promote transmission of serotype 3 strains with certain virulence traits in both vaccinated and unvaccinated individuals, which may lead to the clonal expansion of serotype 3–associated lineages, including GPSC10-ST700. Together these factors may improve the population-level fitness and transmissibility of the GPSC10-ST700, especially following the elimination of competing vaccine type serotypes after PCV13 introduction.

The GPSC10-ST700 serotype 3 lineage appears to have emerged initially through capsular switching events, from serotype 19A/19F to serotype 3. Such acquisition of the serotype 3 capsule is not unexpected, as the GPSC10 genetic background is known to be highly permissive to capsule switching and expresses ≥16 serotypes [41]. This capsule switching has been associated with the acquisition of a shorter capsular locus that lacks the wzg (cpsA), wzh (cpsB), wzd (cpsC), and wze (cpsD) genes, also known as cpsA-D genes [24]. We propose that the absence of the genes in the GPSC-ST700 serotype 3 cps locus is more likely due to gene loss rather than to these genes having never been present, as this CPS locus structure was not widely evident before PCV13 introduction in Malawi or elsewhere. Indeed, there is increasing evidence supporting bacterial evolution through gene loss and genome reduction [42], including among Streptococcus species and between strains [43, 44]. It remains unknown whether this shorter version of the serotype 3 CPS locus has an impact on the structure of the serotype 3 capsule. However, we hypothesize that a shorter, yet functional serotype 3 CPS locus may be acquired relatively easily by non–serotype 3 strains, including vaccine serotypes, leading to capsule switching. Further studies are required to fully understand the potential ecological benefit of the shorter serotype 3 CPS locus.

We show that the absence of the 6 genes in the GPSC10-ST700 serotype 3 CPS locus does not affect capsule production as detected by serotyping and colony morphology. We show that this lineage is less susceptible to opsonophagocytic killing with immune serum samples, which we postulate may be due to a modified capsule, contributing to immune evasion among PCV13-vaccinated individuals. Serotype 3, along with serotype 37, produces capsule using the synthase pathway rather than the Wzx/Wzy-dependent pathway [45–48]. As a result, only wchE, ugd, and galU are required [24]. However, we acknowledge that it is also possible that other genes coding subcapsular

Figure 3. Opsonophagocytic killing of Streptococcus pneumoniae serotype 3 sequence type (ST) 700 and ST5435 isolates by serum samples from 5 healthy adults vaccinated with 23-valent pneumococcal polysaccharide vaccine (PPV23) (A) and 6 children vaccinated with 13-valent pneumococcal conjugate vaccine (PCV13) (B). Symbols differentiate serum samples from individual donors, and colors differentiate isolates.
In conclusion, in the context of considerable global concern over the efficacy of current PCVs against serotype 3 [6, 7], we highlight the clonal expansion of a GPSC10-ST700 serotype 3 lineage after PCV13 introduction, which exhibits elevated penicillin MICs, possesses clusters of virulence genes involved in pathogenicity and colonization, is resistant to opsonophagocytosis, and has a propensity for capsule switching. These findings demonstrate the importance of continued genomic surveillance among vulnerable populations to improve pneumococcal vaccine effectiveness. This globally important pathogen is naturally competent, which may promote the acquisition of further AMR [48] and emergence of novel lineages through capsule switching. Therefore, we suggest that these characteristics are likely to enhance vaccine escape and should be considered when designing the next generation of pneumococcal vaccines. The S. pneumoniae serotype 3 human challenge model, which appears to be safe and achieves experimental colonization rates similar to those previously seen with S. pneumoniae serotype 6B [49], may offer the opportunity to better understand serotype 3 colonization and test the efficacy of novel vaccines.

Supplementary Data
Supplementary materials are available at The Journal of Infectious Diseases online (http://jid.oxfordjournals.org/). Supplementary materials consist of data provided by the author that are published to benefit the reader. The posted materials are not copyedited. The contents of all supplementary data are the sole responsibility of the authors. Questions or messages regarding errors should be addressed to the author.

Notes
Acknowledgments. We thank the individuals who participated in this study and the local schools and authorities for their support. We are grateful for the study field teams, the Malawi-Liverpool-Wellcome (MLW) laboratory team, and the hospitality of the QECH ART Clinic, led by Ken Malisita. Our thanks also extend to the MLW laboratory management team (led by Brigitte Denis) and the MLW data management team (led by the late Clemens Masesa). We thank Andrea Gori for his expert input into the analysis. Finally, we thank the High-Throughput Genomics Group at the Wellcome Trust Centre for Human Genetics (funded by Wellcome Trust) for funding the sequencing data.

Disclaimer. The views expressed in this publication are those of the authors and not necessarily those of the National Institute for Health and Care Research or the Department of Health and Social Care.

Financial support. This work was supported by the Bill & Melinda Gates Foundation (grant OPP1117653 to R. S. H.), the Medical Research Council (MRC grant MR/N023129/1 and grant MR/T016329/1 to J. M. C.), the National Institute for Health and Care Research (NIHR) (using UK Aid from the UK government; grant 16/136/46 to R. S. H.), the Wellcome Trust (strategic award 206545/Z/17/Z supporting the Malawi-Liverpool-Wellcome Clinical Research Programme), and the NIHR Global Health Research Unit on Mucosal Pathogens (using UK Aid from the UK government; support to A.K., T. D. S., N. F., and R. S. H.)

Potential conflicts of interest. All authors: No reported conflicts.

All authors have submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Conflicts that the editors consider relevant to the content of the manuscript have been disclosed.

References


43. Kilian M, Riley DR, Jensen A, Brüggemann H, Tettelin H. Parallel evolution of *Streptococcus pneumoniae* and *Streptococcus mitis* to pathogenic and mutualistic lifestyles. mBio 2014; 5:e01490-01414.


47. Waite RD, Penfold DW, Struthers JK, Dowson CG. Spontaneous sequence duplications within capsule genes cap8E and tts control phase variation in *Streptococcus pneumoniae* serotypes 8 and 37. Microbiology (Reading) 2003; 149(pt 2):497–504.

