

Distinct *Streptococcus pneumoniae* cause invasive disease in Papua New Guinea

Kate C. Mellor^{1,*}, Stephanie Lo¹, Mitton Yoannes², Audrey Michael†², Tilda Orami², Andrew R. Greenhill³, Robert F. Breiman⁴, Paulina Hawkins⁵, Lesley McGee⁵, Stephen D. Bentley¹, Rebecca L. Ford^{2,‡} and Deborah Lehmann^{4,‡}

Abstract

Streptococcus pneumoniae is a key contributor to childhood morbidity and mortality in Papua New Guinea (PNG). For the first time, whole genome sequencing of 174 isolates has enabled detailed characterisation of diverse *S. pneumoniae* causing invasive disease in young children in PNG, 1989–2014. This study captures the baseline *S. pneumoniae* population prior to the introduction of 13-valent pneumococcal conjugate vaccine (PCV13) into the national childhood immunisation programme in 2014. Relationships amongst lineages, serotypes and antimicrobial resistance traits were characterised, and the population was viewed in the context of a global collection of isolates. The analyses highlighted a diverse *S. pneumoniae* population associated with invasive disease in PNG, with 45 unique Global Pneumococcal Sequence Clusters (GPSCs) observed amongst the 174 isolates reflecting multiple lineages observed in PNG that have not been identified in other geographic locations. The majority of isolates were from children with meningitis, of which 52% ($n=72$) expressed non-PCV13 serotypes. Over a third of isolates were predicted to be resistant to at least one antimicrobial. PCV13 serotype isolates had 10.1 times the odds of being multidrug-resistant (MDR) compared to non-vaccine serotype isolates, and no isolates with GPSCs unique to PNG were MDR. Serotype 2 was the most commonly identified serotype; we identified a highly clonal cluster of serotype 2 isolates unique to PNG, and a distinct second cluster indicative of long-distance transmission. Ongoing surveillance, including whole-genome sequencing, is needed to ascertain the impact of the national PCV13 programme upon the *S. pneumoniae* population, including serotype replacement and antimicrobial resistance traits.

DATA SUMMARY

Genome sequences were deposited in the European Nucleotide Archive (ENA), the accession number and the sample metadata are available in the supplementary material. The authors confirm all supporting data, code and protocols have been provided within the article or through supplementary data files.

BACKGROUND

Streptococcus pneumoniae is highly endemic and a common cause of pneumonia and meningitis in Papua New Guinea (PNG) [1–3]. Papua New Guinea encompasses the eastern half of the island of New Guinea and its offshore islands. Neighbouring countries include Australia, Indonesia and the Solomon Islands. The predominantly rural population of ~9.1 million people [4]

Received 26 January 2022; Accepted 26 April 2022; Published 11 July 2022

Author affiliations: ¹Parasites and Microbes, Wellcome Sanger Institute, Hinxton, UK; ²Papua New Guinea Institute of Medical Research, Goroka, Papua New Guinea; ³Institute of Innovation, Science and Sustainability, Federation University Australia, Churchill, Australia; ⁴Rollins School of Public Health Emory University, Atlanta, GA, USA; ⁵Centers for Disease Control and Prevention, Atlanta, GA, USA; ⁶Wesfarmers Centre of Vaccines and Infectious Diseases, Telethon Kids Institute, The University of Western Australia, Crawley WA 6009, Australia.

*Correspondence: Kate C. Mellor, km18@sanger.ac.uk

Keywords: *Streptococcus pneumoniae*; Papua New Guinea; invasive disease.

Abbreviations: AMR, antimicrobial resistance; CSF, cerebrospinal fluid; EHP, Eastern Highlands Province; GPSC, Global Pneumococcal Sequence Cluster; IPD, invasive pneumococcal disease; MDR, multi-drug resistant; MIC, minimum inhibitory concentration; MLST, multi-locus sequence typing; PCV, pneumococcal conjugate vaccine; PNG, Papua New Guinea; WGS, whole-genome sequencing.

†Deceased (2014).

‡These authors contributed equally to this work

Data statement: All supporting data, code and protocols have been provided within the article or through supplementary data files. Supplementary material is available with the online version of this article.

000835 © 2022 The Authors



This is an open-access article distributed under the terms of the Creative Commons Attribution License.

Impact Statement

Streptococcus pneumoniae is a major cause of bacteraemic pneumonia and bacterial meningitis in Papua New Guinea (PNG). A 13-valent pneumococcal conjugate vaccine (PCV13) was introduced into the national immunisation programme in 2014. This study analyses samples prior to vaccine introduction, and finds a diverse *S. pneumoniae* population with evidence of the maintenance of multiple lineages that were distinct from those identified in other countries. Concordant with published literature, the results indicate that the majority of cases of invasive pneumococcal disease, in particular meningitis, in PNG are caused by non-PCV13 serotypes, limiting its potential benefit in PNG; however higher valency vaccines in development may be more effective. Over a third of pneumococcal isolates were resistant to one or more antimicrobials and 7.3% of isolates were multidrug resistant (MDR). Vaccine serotype isolates were 10.1 times more likely to be MDR than non-vaccine serotype isolates which suggests potential reduction in AMR with use of PCV13. However, continued monitoring is needed to assess replacement of dominant *S. pneumoniae* serotypes and lineages in PNG following the perturbation caused by introduction of PCV13.

live in highland, coastal and island regions and speak >800 different languages [5]. The mortality rate in children <5 years-of-age is high (57 per 1000 births in 2015) and pneumonia is the most common cause of hospitalisation and death in this age group [6, 7]. Chloramphenicol was the first-line treatment for children with meningitis, but has now been replaced by ceftriaxone due to an increase in cases of meningitis caused by chloramphenicol resistant *Haemophilus influenzae* [8, 9]. Intermediate penicillin resistance of *S. pneumoniae* has long been reported in PNG [10, 11] and was detected in 21.5% of pneumococcal isolates from cerebrospinal fluid (CSF) in children between 1996 and 2005 [2]. A recent increase in resistance of *S. pneumoniae* to cotrimoxazole (40.2% >0.5/9.5 µg ml⁻¹) has been detected in carriage studies (D Lehmann, personal communication). Resistance of carriage isolates to other antimicrobials was consistent with isolates from children with meningitis (1996–2005); approximately 2.1% of isolates from children with meningitis were tetracycline resistant and chloramphenicol and erythromycin resistance were rarely detected [2].

Colonisation with *S. pneumoniae* is a pre-requisite for invasive pneumococcal disease (IPD) [12]. Pneumococcal carriage is exceptionally high in PNG, with studies indicating all infants are colonised by 3 months of age and the median age of first pneumococcal acquisition is approximately 19 days [13, 14]. Consequently, *S. pneumoniae* is a common cause of bacterial meningitis amongst children in PNG [2, 15]. A broad range of serotypes have been reported in both carriage and IPD in PNG, consistent with other highly endemic settings [16–20]. Prior to introduction of 13-valent pneumococcal conjugate vaccine (PCV13), over 50% of IPD cases amongst children in PNG were caused by non-PCV13 serotypes, limiting potential benefit of the vaccine [2, 18]. Interestingly, some serotypes rarely observed in other geographic regions have been common causes of IPD in PNG, such as serotypes 2 and 46 [15, 16]. In PNG, the high rates of pneumococcal carriage [13, 14] combined with a high proportion of disease caused by non-vaccine serotypes highlight the potential for replacement of PCV13 serotypes with non-vaccine serotypes [2].

A GAVI-funded, universal nationwide vaccination programme for infants using the PCV13 began in 2014. However, PCV13 (comprised of serotypes 1, 3, 4, 5, 6A, 6B, 7F, 9V, 14, 18C, 19A, 19F, 23F) was not widely administered in the Eastern Highlands Province (EHP), where the majority of aetiology studies have been conducted, until late 2015 and annual coverage estimates nationally have been estimated to be ≤35% [21]. Encouragingly, PCV13 vaccination was associated with a 57.4% reduction in hospitalisation of children due to pneumonia and a 28.7% reduction of hypoxic pneumonia in PNG [22].

Building upon insights gained through phenotypic isolate characterisation, the present study examines whole genome sequence data for IPD isolates from children in PNG, providing a high-resolution view of the *S. pneumoniae* population structure. The diversity of lineages, serotypes and antimicrobial resistance traits are characterised, providing a baseline of *S. pneumoniae* diversity prior to the national PCV13 introduction.

METHODS

A collection of 174 *S. pneumoniae* isolates from children aged under 8-years-old with IPD in PNG between 1989 and 2014 were whole-genome sequenced using an Illumina HiSeq platform. The isolates originated from multiple formal studies and hospital surveillance (detailed in Table 1) conducted by the PNG Institute of Medical Research (PNGIMR) in collaboration with clinicians at the local hospitals. Isolates were freeze-dried for storage at the PNG Institute of Medical Research, and all viable isolates that were available were sequenced. The majority of isolates were from patients in the Asaro Valley area (which includes Goroka town, the provincial capital, located at 1600 m above-sea-level) and other areas within a 1 h drive of Goroka. The majority of isolates (77.6%, *n*=135) were from a study of the aetiology of suspected meningitis cases of children admitted to Goroka Base Hospital (previously named Goroka Hospital, then Goroka General Hospital and later as Goroka Regional Hospital). These specimens were collected as requested by paediatricians, without pre-defined selection criteria. Whilst most of the studies/surveillance aimed to investigate the aetiology of pneumonia and/or meningitis, in the study investigating the epidemiology of typhoid (Table 1), patients

Table 1. Summary of isolate origins: study, source, clinical manifestation

	Study description and reference(s)	no. of isolates	Years of study	% (no.) isolated from CSF§			% (no.) associated with		
				Meningitis	Pneumonia	Other manifestation	Meningitis and pneumonia	Unknown	Unknown
i	Aetiology of suspected meningitis among children admitted to Goroka Base Hospital*. [2, 19] Samples from children with meningitis, no additional predefined selection criteria	135	1989–1992 1997–2003	92.6 (n=125)	88.9 (n=120)	2.3 (n=3)	7.5 (n=11)	0	0.7% (n=1)
ii	Aetiology study of pneumonia in children aged <5 years admitted to Tari Hospital, Hela Province (unpublished) Hospitalised children were included if antimicrobials were not started prior to blood culture	2	1990–1991	0	0	0	0	0	100 (n=2)
iii	Multicentre Young Infant Study (WHO) – investigation of severe infections in children aged <3 months attending Goroka Base Hospital* Outpatient Department [19] Children were included if illness began at home and rectal temperature >=35°C or <35.5°C or had a cough, high respiratory rate, difficulty breathing, convulsions, fever or difficult to wake.	9	1991–1993	11.1 (n=1)	22.2 (n=2)	44.4 (n=4)	0	22.2 (bacteraemia, n=2)	11.1 (n=1)
iv	Assessment of the epidemiology and transmission of <i>Salmonella typhi</i> in the Asaro Valley, EHP‡ (unpublished) Children with 3+ day history of fever resident in Asaro valley villages who attended the Goroka Base Hospital or Asaro Health Centre.	1	1993	0	0	0	0	0	100 (n=1)
v	Hospital-based invasive disease surveillance† (blood culture), Goroka Hospital*, EHP‡ Samples from children with meningitis/sepsis with no additional predefined selection criteria	21	1995–1998, 2001	23.8 (n=5)	52.4 (n=11)	14.3 (n=3)	14.3 (n=3)	0	19.0 (n=4)
vi	Is routine lumbar puncture indicated in PNG child with febrile convulsion? (Modilon Hospital, Madang Province) [23, 45] Hospitalised children with one or more fever associated seizure	2	2007–2008	50 (n=1)	100 (n=2)	0	0	0	0
vii	Aetiology of moderate or severe pneumonia and meningitis in children aged <5 years enrolled through urban outpatient clinics or admitted to Eastern Highlands Provincial Hospital*, Goroka, EHP‡ [25] Samples from children with pneumonia and meningitis, no additional predefined selection criteria	4	2013–2014	25 (n=1)	25 (n=1)	75 (n=3)	0	0	0

*Goroka Regional Hospital previously known as Goroka General Hospital, Goroka Base Hospital, and Goroka Hospital; Goroka located in the Asaro Valley.

†Opportunistic sampling as requested by paediatrician.

‡EHP Eastern Highlands Province.

§Remainder of isolates were from blood samples.

were enrolled if they had a history of fever for 3 days or more, and therefore cases of bacteraemia due to *S. pneumoniae* and other pathogens were also identified. Invasive isolates were included from a study conducted at Modilon Hospital, Madang Province on the north coast of PNG [23] and from surveillance of pneumonia at Tari Hospital in the remote Hela Province (previously part of Southern Highlands Province) (unpublished data), hospitalised children without antimicrobial treatment prior to blood culture were included. Whole genome sequencing and quality control of paired-end reads was conducted according to the Global Pneumococcal Sequencing (GPS) Project protocol [24].

Specimen collection, blood and CSF culture, phenotypic pneumococcal serotyping and antimicrobial susceptibility have been described in detail elsewhere [1, 2, 15, 19, 25]. Serotyping and antimicrobial susceptibility for all studies were conducted at the PNGIMR laboratory in Goroka.

In silico assignment of serotype was inferred using SeroBA [26] with additional screening of isolates predicted to be 'untypable' by SeroBA using the Centres for Disease Control and Prevention's (CDC) in-house serotyping pipeline [27] and manual characterisation of the *cps* region. Multi-locus sequence typing (MLST) was performed using SRST2 [28]. Prediction of antimicrobial resistance phenotype from whole-genome sequencing (WGS) data was conducted using the CDC antimicrobial resistance pipeline [27, 29] which generates minimum inhibitory concentration (MIC) predictions, and classification of isolates as susceptible, intermediate or resistant based on Clinical and Laboratory Standards Institute (CLSI) guidelines [30], including separate MIC thresholds for meningitis infections. Multidrug resistance (MDR) was defined as *in silico* predicted resistance to three or more antimicrobial classes. Global Pneumococcal Sequencing Cluster (GPSC) assignments were determined using PopPUNK [31]. This method uses nucleotide variation across the whole-genome to cluster isolates, enabling definition of lineages. Lineages local to PNG were defined as GPSCs not identified in other geographic locations in the GPS database (November 2020) and/or MLST not within the PubMLST database (November 2020).

Reads were mapped to a reference *S. pneumoniae* (ATCC 700669, NCBI accession number FM211187) using BWA-MEM (v0.7.17) [32], followed by indel realignment using GATK (v3.7.0) [33], deduplication with Picard Mark Duplicates v1.127 (available at <http://broadinstitute.github.io/picard/>) and variant calling and consensus pseudosequence generation using samtools v1.6 [34] and bcftools v1.5. Regions of recombination were then identified and masked from the alignment using Gubbins v3.0 [35]. Single nucleotide polymorphism data were then extracted from the masked alignment and used to generate a maximum-likelihood phylogeny, using RAxML (v8.2.8) [36]. The metadata and analysis results are available to view interactively using Microreact (<https://microreact.org/project/jhj4cLf93pvVQVp4hPN9Gq>). Lineage-specific phylogenies were generated by mapping to a lineage-specific reference (GPSC96 accession ERR1214695, references for other GPSCs are detailed in supplementary data), then recombination was identified (same approach as previously described). Regions of recombination were masked prior to generation of the GPSC phylogenies using RAxML.

Fisher's Exact test, implemented in R [37], was used to test for differences between PCV13 serotype and non-PCV13 serotype PNG isolates, including clinical presentation, sample source (CSF and blood) and antimicrobial resistance.

Concordance between phenotypic and genotypic AMR profiles was evaluated. Discordance of phenotypic and genotypic AMR was classified according to error type. Major errors occurred where phenotypically susceptible isolates were genotypically predicted to be resistant. Very major errors occurred where isolates were phenotypically resistant but genotypically predicted to be susceptible. Data were compared where phenotypic MIC data were available, with consistent MIC cut-offs (CLSI guidelines) for both the predicted and phenotypic methods. Categorical variables including clinical manifestation, serotype, vaccine status (PCV13 or non-vaccine serotype). Sample source and GPSCs were compared using Fisher's Exact test with the significance level of $P < 0.05$ using R (v 3.0.6).

RESULTS

Sampling

Analyses in this study utilised 174 isolates from children with clinical disease attributed to *S. pneumoniae* collected between September 1989 and January 2014 in PNG, prior to the introduction of PCV13. Metadata are summarised in Table 1, stratified by origin study. The majority of isolates were derived from CSF samples (76.4%, $n=133$), and the remainder were isolated from blood samples. Almost all samples were collected in the Eastern Highlands Province (97.7%, $n=170$), with two isolates from each of Madang and Hela Provinces. The majority of isolates (90.2%, $n=157$) were from children ≤ 2 years of age, 6.3% ($n=11$) aged between 2 and 5 years, and the remaining 1.7% ($n=3$) aged between five and 8 years. Patient age was unknown for 1.7% ($n=3$) of isolates.

Of the 94.8% ($n=165$) of isolates from patients with known clinical diagnosis, the majority were from children with a diagnosis of meningitis (82.4%, $n=136$). A further 8.5% ($n=14$) of isolates were from children with both meningitis and pneumonia, and 7.8% of isolates ($n=13$) were from children with pneumonia as sole diagnosis.

Serotypes

Almost half (45.4%; $n=79$) of the total isolates were non-PCV13 serotype (NVT). The odds of CSF isolates expressing a non-PCV13 serotype was 2.44 times higher than for isolates from blood (95% CI 1.10, 5.72; $P=0.020$). Overall, non-PCV13 serotype 2 was most common, followed by PCV13 serotypes 5, 23F, 7F and 14 (Fig. 1). Clinical manifestation data was available for 96.3% ($n=26$) of serotype 2 isolates, of which most were from children with a diagnosis of meningitis alone (80.8%; $n=21$), 11.5% ($n=3$) pneumonia and meningitis and only 7.7% ($n=2$) pneumonia alone. There was no significant association of clinical manifestation with serotype 2 isolates ($P=0.871$).

Compared with PCV13, the newly licensed PCV15 vaccine would not increase the proportion of vaccine type serotypes. A 10-valent PCV (Pneumosil), which includes serotypes 1, 5, 6A, 6B, 7F, 9V, 14, 19A, 19F, and 23F, would cover 72 (41.4%) isolates and PCV20 (PCV13 plus serotypes 8, 10A, 11A, 12F, 15B, 22F, and 33F) would provide an incremental improvement in serotype coverage (58.1%, $n=101$). More encouragingly, a PCV24 vaccine (PCV13 serotypes plus serotypes 2, 8, 9N, 10A, 11A, 12F, 15B, 17F, 20, 22F, 33F), which is currently under development, would increase the proportion of vaccine serotype isolates in this study to 74.1% ($n=129$). However, of the isolates expressing non-PCV24 serotypes 48.9% ($n=22$) were predicted to be resistant to one or more antimicrobial.

One isolate was characterised as untypable using SeroBA and the CDC in-house serotyping pipeline; therefore, manual assessment of the *cps* region between *dexB* and *aliA* was conducted. *cps* genes were absent from the GPSC18 isolate, with only Δ IS1202 Δ and *pspK* genes between *dexB* and *aliA*, predictive of an unencapsulated phenotype. This isolate was also phenotypically untypable.

GPSCs

Amongst the 174 isolates, 45 GPSCs were identified, of which 15 GPSCs were only detected in PNG within the GPS collection highlighted in Fig. S1 (available in the online version of this article). The STs of isolates within these GPSCs were searched against the PubMLST database, as a proxy for GPSC, and were not identified in other geographic regions. Interestingly, 28.4% of isolates ($n=50$) belonging to 15 GPSCs were of a GPSC not identified in other geographic regions (GPSCs 635 ($n=12$), 595 ($n=9$), 640 ($n=8$), 656 ($n=4$), 658 ($n=4$); the remaining GPSCs were represented by ≤ 2 isolates).

The top five most common GPSCs were 96, 8, 635, 15 and 20 (Table 2). Of the 32 serotypes, 12 were expressed across multiple GPSCs (maximum five GPSCs). For almost all GPSCs there was a single expressed serotype Fig. S2. The exceptions were GPSC149 (serotypes 27 ($n=2$) and 32A ($n=1$)), GPSC635 (serotypes 19F ($n=7$), 14 ($n=3$), 19A ($n=1$), 29 ($n=1$)) and 713 (serotypes 6B ($n=1$), 6C ($n=1$)).

Phylogenies were constructed for GPSCs with >5 isolates from this study, placing isolates from PNG in the context of global GPS collection isolates (Fig. S3). GPSC96 isolates were all serotype 2, the most common serotype in the collection, and formed two distinct phylogenetic clusters which were concordant with MLST designations. The majority ($n=27$) of GPSC96 isolates, spanning 1998–2002, clustered together and were distinct from other GPS collection isolates. Two GPSC96 isolates, from 2000 and 2001, formed a distinct cluster and were most closely related to isolates from Israel (average 82 SNPs) and Morocco (average 86 SNPs) (Fig. 2). Similarly, all GPSC20 isolates were serotype 23F, and formed two distinct clusters within phylogenies of isolates from diverse geographic regions, consistent with MLSTs and AMR profiles. In contrast, within GPSCs 15, 76, 224 and 334 the PNG isolates formed a single, distinct cluster within each global GPSC phylogeny (Fig. S3).

Antimicrobial resistance

Over a third of isolates (36.2%, $n=63$) were predicted to be resistant to one or more antimicrobial and 7.3% ($n=14$) multidrug resistant. For six GPSCs (20, 334, 40, 658, 76) more than one AMR profile was observed. Multidrug resistant isolates included all GPSC658 ($n=4$, serotype 14), GPSC26 ($n=1$, serotype 12F) and 60% ($n=6$) of GPSC20 (serotype 23F) isolates. No isolates collected prior to 1998 were classified as MDR. PCV13 serotype isolates had 10.1 times greater odds of being MDR than non-PCV13-serotype isolates (95% CI [1.41, 444.3], $P=0.007$). No isolates from GPSCs exclusive to PNG were predicted to be MDR.

At the MIC threshold for meningitis ($0.12 \mu\text{g ml}^{-1}$), 24.1% of NVT serotype isolates and 32.6% of PCV13 isolates were classified as resistant to penicillin, with no significant difference in odds of penicillin resistance between NVT and PCV13 serotype isolates (OR 0.66 [0.31, 1.34], $P=0.241$). Most NVT penicillin-resistant strains were serogroup 24 ($n=9$, GPSC595), 19F ($n=8$, GPSC635) or 12F ($n=6$, GPSC334); other serotypes included 29 ($n=2$) and 22A ($n=1$). PCV13 serotype penicillin-resistant strains were 14 ($n=8$), 23F ($n=6$), 6A ($n=2$), 19A ($n=1$). Interestingly, GPSC334 isolates ($n=6$) from 1991 onwards were predicted to be resistant to penicillin which appears to be due to recombination events of both *pbp2B* and *pbp2X*, resulting in import of novel *pbp2B* and *pbp2X* alleles, while the GPSC334 isolate from 1990 was predicted to be susceptible to antimicrobials. Amongst the 50 isolates predicted to be resistant to penicillin at the meningitis MIC threshold, resistance was encoded by seven unique allelic profiles of *pbp1A*, *pbp2B* and *pbp2X* genes, of which six profiles had one or more ‘new’ alleles. An allelic profile containing one or two ‘new’ *pbp* alleles was observed in 12 isolates, each unique to a GPSC. A profile of three novel alleles predicted to confer resistance to penicillin was common to 33 isolates across ten GPSCs. Of the isolates with phenotypic penicillin MIC data (96.6%, $n=168$),

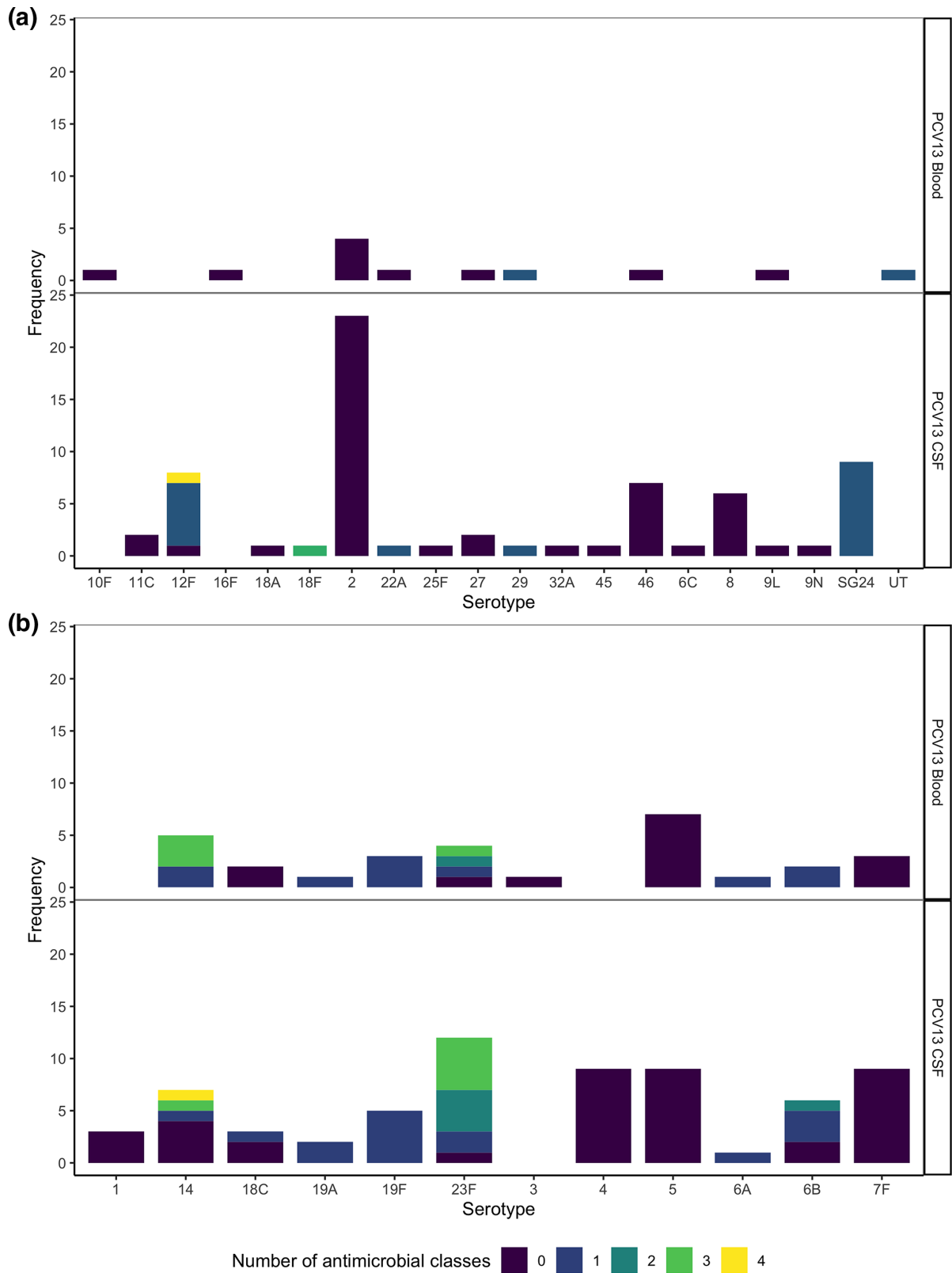


Fig. 1. Summary of serotypes (≥ 3 isolates), grouped by sample source and pneumococcal conjugate vaccine (PCV) serotype status (part A non-vaccine serotype isolates; part B=PCV13 serotype isolates). The frequency bars are coloured by number of antimicrobial classes to which isolates were resistant. UT=untypable, SG24=serogroup 24. NVT=non-PCV13 serotypes. CSF=cerebrospinal fluid.

Table 2. Summary of isolate characteristics for GPSCs with >5 isolates, ranked by number of isolates

GPSC	Total no. of isolates	no. of isolates from CSF†	Years	Serotype(s) (<i>in silico</i> prediction)	Multilocus sequence type(s)	Clonal complex (CC)	Antimicrobial resistance (<i>in silico</i> prediction)
96	27	23	1989–2002	2	1504 (n=24) 16705 (n=3)	74	None
8	16	9	1990–2002	5	16188*	289	None
15	11	9	1991–2003	7F	191 (n=6) 16198* (n=5)	191	None
635*	12	5	1991–1997	19F (n=7) 14 (n=3) 19A (n=1) 29 (n=1)	16204 (n=5) 6904 (n=2) 16210 (n=2) 16244 (n=2) 16068 (n=1)	CC63 (n=1) CC unassigned (n=11)	Resistant to penicillin (meningitis MIC threshold)
20	10	8	1990–2003	23F	802 (n=8) 16189* (n=2)	802 (n=8) CC unassigned (n=2)	ST802 isolates: intermediate susceptibility cotrimoxazole (n=8) and resistant to tetracycline (n=6) ST16189: none
162	8	8	1990–2001	4	14136 (n=7) 4127 (n=1)	4127	None
595*	9	9	1990–2002	Serogroup 24	6755	CC unassigned	Resistant to penicillin at meningitis MIC thresholds
334	7	7	1990–2000	12F	16297* (n=6) 1527 (n=1)	1527 (n=1) CC unassigned (n=6)	Isolates from 1991 onward (n=6) resistant to penicillin at meningitis MIC threshold. Isolates from 1990: none
640*	8	7	1995–2002	46	16207* (n=6) 16246* (n=2)	CC unassigned	None
5	5	4	1990–2002	23F	16193	172	Resistant to penicillin at meningitis threshold and resistant to cotrimoxazole (n=5)
76	6	4	1998–2014	6B	16206 (n=5) 16259 (n=1)	CC unassigned	Resistant to penicillin at meningitis threshold. One ST16206 isolate resistant to cotrimoxazole
224	6	6	1991–2003	8	16229* (n=1) 6748* (n=5)	6022	None

*Not described outside Papua New Guinea (within GPS and PubMLST databases). Data for the remaining 33 GPSCs in supplementary data, antimicrobial susceptibility characteristics summarised in Fig. S1 and serotypes summarised in Fig. S2.

†Remainder of isolates derived from blood samples.

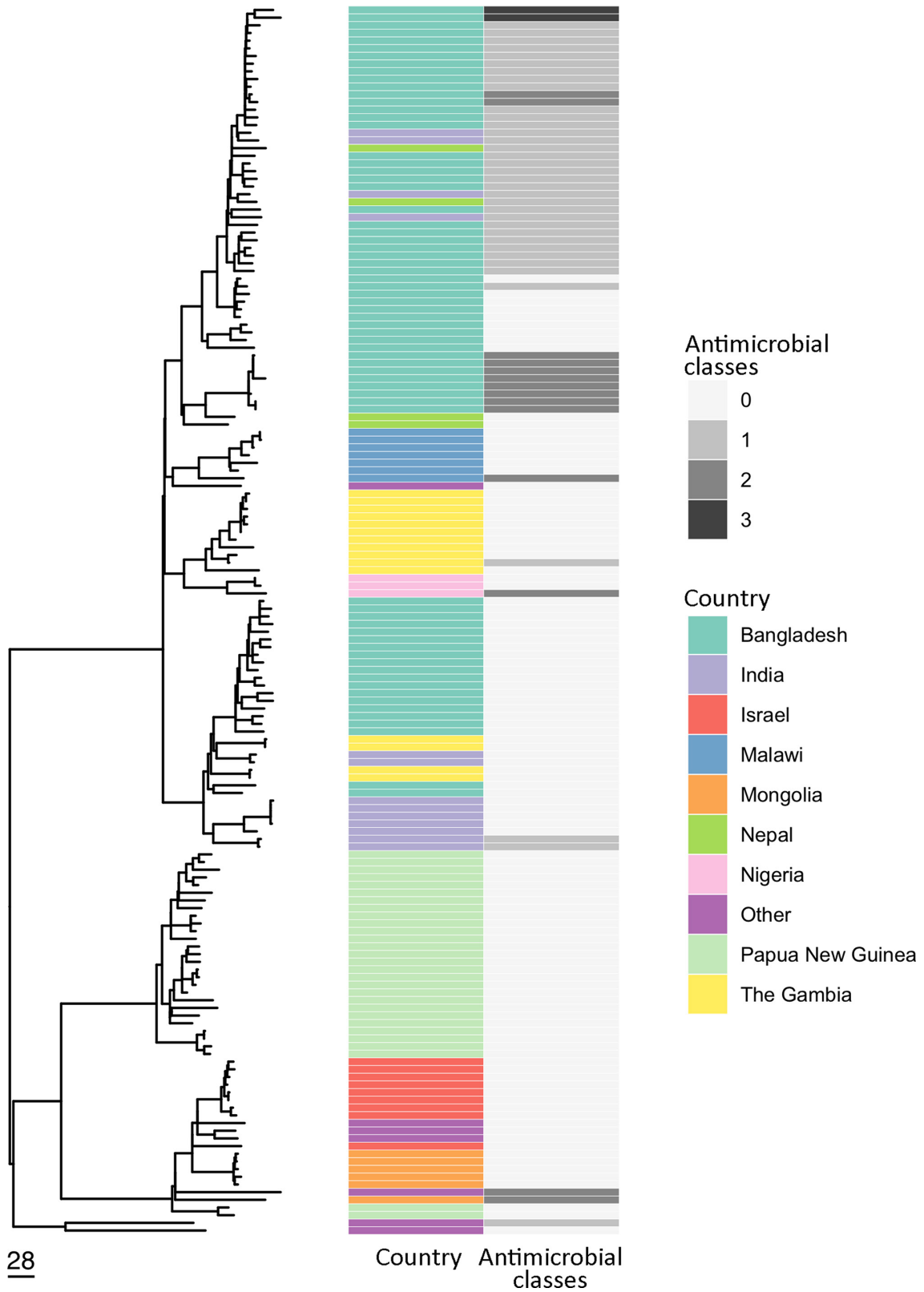


Fig. 2. Phylogeny of Papua New Guinea and published Global Pneumococcal Sequencing collection serotype 2, GPSC 96 isolates. Heat map displays the country of isolate origin and number of antimicrobial classes to which the isolates were predicted to be resistant (using the CDC pipeline). Reference mapped phylogeny (reference: ERR1214695), post-masking of recombination sites using Gubbins and generated using RAxML.

there was 93.5% concordance between phenotypic and genotypic predicted penicillin resistance at the meningitis MIC threshold. Discordance comprised of very major errors for two isolates and major errors for nine isolates.

Cephalosporin-resistant isolates ($n=4$) were all GPSC 658, serotype 14 isolates; these isolates were also cotrimoxazole- and penicillin-resistant (Fig. S1). Cotrimoxazole resistance, encoded by indels in *folP* (five variants) with or without point mutations in *folA*, was predicted for 5.7% ($n=10$) of isolates and a further 7.5% ($n=13$) of isolates were predicted to have intermediate susceptibility to cotrimoxazole. Of the cotrimoxazole-resistant isolates, ten expressed PCV13 serotypes (GPSC5, serotype 23F ($n=5$); GPSC658, serotype 14 ($n=4$)); GPSC76, serotype 6B ($n=1$). There was concordance between phenotypic and genotypic predicted cotrimoxazole resistance for 95.8% ($n=159$) of the 166 isolates with available phenotypic MIC data. Tetracycline-resistant isolates were detected from 1998 onwards, encoded by *tet(M)* in 5.2% of isolates ($n=9$) across both PCV13 serotype isolates (GPSC 9, serotype 14 ($n=1$); GPSC 20, serotype 23F ($n=6$)) and non-PCV13 serotype isolates (GPSC26, serotype 12F ($n=1$); GPSC172, serotype 18F ($n=1$)). For the 60 isolates with phenotypic tetracycline MIC data there was 96.7% concordance between phenotype and genotype.

DISCUSSION

Here we show high serotype and lineage diversity amongst IPD *S. pneumoniae* isolates from PNG. A distinct *S. pneumoniae* population compared to other sampled geographic locations was evident; multiple common *S. pneumoniae* lineages were observed across extended time periods between 1990 and 2014 in PNG and were not observed in other countries based on the GPS or PubMLST databases. The most common serotypes associated with IPD cases in PNG appear to differ from those observed in carriage in neighbouring Indonesia, where serogroups 6 and 23, and untypable isolates have been the most commonly isolated [38–40]. These findings are indicative of diverse serotypes circulating in PNG for an extended time period and limited long-range transmission events to or from PNG; however, a lack of representation of isolates from neighbouring countries limit the assessment of lineage distributions.

The most common predicted antimicrobial resistance trait was resistance to penicillins at the meningitis MIC threshold, including amongst GPSC 635, 640 and 595 isolates which were exclusive to PNG. As there was no significant difference in resistance to penicillins between NVT and PCV13 serotype isolates, the introduction of PCV13 would not be anticipated to reduce proportionate penicillin resistance of *S. pneumoniae* meningitis cases. However, as PCV13 serotype isolates had higher odds of being MDR compared with NVT serotypes, a reduction in MDR strains may occur, though in the future could be anticipated to return to a similar level due to ongoing selection [41].

Almost half of isolates expressed non-PCV13 serotypes, reflective of previous reports from PNG [2; Lehmann *et al.*, 1999 [18, 20]. This contrasts with estimates that PCV13 would cover 74% of IPD in all geographic regions globally, including Oceania [17]. Two GPSCs (635 and 713) expressed both PCV13 and NVT serotypes, suggesting potential for serotype replacement by these lineages. Interestingly, one unencapsulated GPSC18 isolate was identified from a blood sample from a child with meningitis. This strain would not be expected to have been the causative agent of meningitis and the possibility of a mixed pneumococcal infection cannot be excluded [42]. All GPSC635 were penicillin-resistant and all GPSC5 isolates were resistant to both cotrimoxazole and penicillin, reducing treatment options.

GPSC96 was the dominant GPSC represented amongst the PNG dataset utilised for this study. Concerningly, serotype 2 is not included in the PCV13 currently used in PNG, though is included in the higher valency PCV24 Affinivax vaccine. The GPSC96 isolates were all serotype 2, which has been associated with IPD in other geographic regions including Bangladesh and Israel [43, 44]; and reported in PNG for over 30 years [15, 16]. In PNG, the distribution of serotype 2 across age groups has appeared to be similar to other serotypes [2]. However, the epidemiology of serotype 2 can vary between geographic settings; in Bangladesh serotype 2 has been overrepresented, relative to other serotypes, amongst young children [44]. The association of serotype 2 with clinical manifestation also appears to vary with location, though may be influenced by sampling; in Israel the most common presentation was pneumonia (78.5%) followed by septicaemia (9.9%) and meningitis (7.4%) [43]. In stark contrast, in Bangladesh 98% of serotype 2 IPD cases had a diagnosis of meningitis [44]. In PNG serotype 2 *S. pneumoniae* were isolated from children with meningitis and those with pneumonia alone, with no significant association with clinical manifestation. However, sample numbers were limited and cases may not be reflective of the wider population. The global GPSC96 phylogeny revealed two clusters of PNG isolates, with a main clonal cluster comprising of isolates spanning 1989 to 2002 which may be representative of long-term serotype 2 isolates circulating in PNG. These isolates were distantly related to GPSC96 isolates from other geographic regions. In contrast, the cluster of two PNG GPSC96 isolates most closely related to isolates from Israel and Morocco may indicate a long-range transmission event to or from PNG.

The samples available for this study derived from both routine surveillance and multiple studies in PNG, pooled to provide an overview of genomic characteristics of isolates from a location with limited available data. The *S. pneumoniae* captured by this overview will not be fully representative of IPD cases in PNG, however, in the absence of systematically derived surveillance data, the isolates contributing to this study provide a valuable baseline for the evaluation of the impact of PCV13 upon the *S.*

pneumoniae population causing IPD in children in the EHP as the majority of isolates in this study were from this province. The pneumococcal population structure observed in this study may not be representative of other regions of PNG, and comparisons of *S. pneumoniae* pre- and post-perturbation due to PCV13 introduction may therefore be limited to EHP. Additionally, the majority of isolates were from patients with meningitis rather than pneumonia, which is not reflective of hospitalised case numbers [7]. Expansion of surveillance to capture *S. pneumoniae* in carriage, and causing both meningitis and pneumonia across PNG is needed to provide a robust basis for the assessment of the impact of vaccine introductions.

The resolution afforded by genomic data and contextualisation using the GPS collection has highlighted the distinct *S. pneumoniae* population observed to cause IPD in children in PNG. Ongoing standardized surveillance, including WGS of isolates, post-PCV13 distribution is essential to evaluate serotype replacement patterns and impact of PCV13 upon AMR traits. Additionally, surveillance needs to be conducted at several representative sites across the country, perhaps using sampling from healthy carriage alongside pneumococcal disease.

Funding information

This study was co-funded by the Bill and Melinda Gates Foundation (grant code OPP1034556), the Wellcome Sanger Institute (core Wellcome grants 098051 and 206194), and the USA Centres for Disease Control and Prevention. The funding sources had no role in isolate selection, analysis, or data interpretation.

Acknowledgements

We thank the Wellcome Sanger Institute (WSI) Pathogen Informatics team for help with data management. Directors and senior researchers, PNGIMR who identified funds and provided intellectual and technical input/advice for contributory studies detailed in Table 1: Michael Alpers, John Reeder, Peter Siba, William S Pomat. Clinicians and study leaders who were involved in studies listed in Table 1: Peter Richmond and Christopher Blyth (UWA and Wesfarmers Centre of Vaccines and Infectious Diseases, Telethon Kids Institute), Moses Laman (PNGIMR) Christopher Blyth, Fiona Russell (Melbourne Children's Research Institute, University of Melbourne), Megan Passey (University Centre for Rural Health, University of Sydney) Lead microbiologists and heads of microbiology department at PNGIMR: Mike Gratten (deceased), Alison Clegg We thank Geoffrey Rose (deceased), Tony Lupiwa, Mathew Omena, Steven Tiwara, Rex Ivivi Lapule Yuasi, Celestine Aho, Joe Jude (PNGIMR) for their work culturing and phenotypically characterising isolates. We thank Sarah Javati (PNGIMR) for locating and collating isolates for shipment.

Author contributions

Conceptualisation (K.C.M., S.L., S.B., R.L.F., D.L.), methodology (K.C.M., S.L.), investigation (M.Y., T.O., A.M., R.L.F., K.C.M.), formal analysis (K.C.M.), resources (S.D.B., M.Y., A.M., R.L.F.), data curation (K.M., S.L., P.H.), writing - original draft preparation (K.C.M.), writing - review and editing (K.C.M., S.L., M.Y., T.O., A.R.G., R.F.B., P.H., L.M., S.D.B., R.L.F., D.L.), visualisation (K.C.M.), project administration (S.L., S.D.B., L.M.), funding (L.M., R.L.F., D.L., S.D.B.).

Conflicts of interest

S.D.B. reports personal fees from Pfizer and Merck, outside the submitted work. All other authors declare no conflicts of interest. The study "Aetiology of moderate or severe pneumonia and meningitis in children aged <5 years enrolled through urban outpatient clinics or admitted to Eastern Highlands Provincial Hospital, Goroka, EHP (Blyth *et al.*, 2017) was part funded by a Pfizer Global Research Grant (WS1801400).

Ethical statement

Isolates for this study were obtained from a retrospective bacterial collection from the Papua New Guinea Institute of Medical Research. Isolates were obtained as part of routine surveillance or as part of previously published studies (summarised in Table 1), with fully anonymised patient data. No human or animal tissue was obtained.

References

- Gratten M, Montgomery J. The bacteriology of acute pneumonia and meningitis in children in Papua New Guinea: assumptions, facts and technical strategies. *P N G Med J* 1991;34:185–198.
- Greenhill AR, Phuanukoonnon S, Michael A, Yoannes M, Orami T, *et al.* *Streptococcus pneumoniae* and *Haemophilus influenzae* in paediatric meningitis patients at Goroka General Hospital, Papua New Guinea: serotype distribution and antimicrobial susceptibility in the pre-vaccine era. *BMC Infect Dis* 2015;15:485.
- Riley ID. Pneumonia research in Papua New Guinea: 1967–1986. *P N G Med J* 2010;53:106–118.
- UNICEF. Papua New Guinea Country Profile. WWW Document. URL; 2021. <https://data.unicef.org/country/png/> [accessed 3 October 2021].
- Palmer B. (ed.). *The Languages and Linguistics of the New Guinea Area: A Comprehensive Guide (Vol. 4)*. Walter de Gruyter GmbH & Co KG, 2017.
- GAVI. Papua New Guinea Country Information; 2015. <https://www.gavi.org/programmes-impact/country-hub/western-pacific/papua-new-guinea> [accessed 10 March 2021].
- Paediatric society of PNG. *Child Morbidity and Mortality. 10th Annual Report*. PNG department of health, 2019.
- Manning L, Laman M, Greenhill AR, Michael A, Siba P, *et al.* Increasing chloramphenicol resistance in *Streptococcus pneumoniae* isolates from Papua New Guinean children with acute bacterial meningitis. *Antimicrob Agents Chemother* 2011;55:4454–4456.
- Paediatrics society of Papua New Guinea. *Standard Treatment for Common Illnesses of Children in Papua New Guinea. A Manual for Nurses, Community Health Workers, Health Extension Officers and Doctors*. Port Moresby: National Department of Health; 2011.
- Douglas RM. Pneumonia in Papua New Guinea, from the past to the future. *P N G Med J* 2010;53:8.
- Gratten M, Naraqi S, Hansman D. High prevalence of penicillin-insensitive pneumococci in Port Moresby, Papua New Guinea. *Lancet* 1980;2:192–195.
- Simell B, Auranen K, Käyhty H, Goldblatt D, Dagan R, *et al.* The fundamental link between pneumococcal carriage and disease. *Expert Rev Vaccines* 2012;11:841–855.
- Francis JP, Richmond PC, Pomat WS, Michael A, Keno H, *et al.* Maternal antibodies to pneumolysin but not to pneumococcal surface protein A delay early pneumococcal carriage in high-risk Papua New Guinean infants. *Clin Vaccine Immunol* 2009;16:1633–1638.
- Riley ID, Lehmann D, Alpers MP, Marshall TF, Gratten H, *et al.* Pneumococcal vaccine prevents death from acute lower-respiratory-tract infections in Papua New Guinean children. *Lancet* 1986;2:877–881.
- Lehmann D, Yeka W, Rongap T, Javati A, Saleu G, *et al.* Aetiology and clinical signs of bacterial meningitis in children admitted to

- Goroka Base Hospital, Papua New Guinea, 1989-1992. *Ann Trop Paediatr* 1999;19:21-32.
16. Gratten M, Barker J, Shann F, Gerega G, Montgomery J, et al. The aetiology of purulent meningitis in highland children: a bacteriological study. 1985;28:233-240.
 17. Johnson HL, Deloria-Knoll M, Levine OS, Stoszek SK, Freimanis Hance L, et al. Systematic evaluation of serotypes causing invasive pneumococcal disease among children under five: the pneumococcal global serotype project. *PLoS Med* 2010;7:e1000348.
 18. Lehmann D, Gratten M, Montgomery J. Susceptibility of pneumococcal carriage isolates to penicillin provides a conservative estimate of susceptibility of invasive pneumococci. *Pediatr Infect Dis J* 1997;16:297-305.
 19. Lehmann D, Michael A, Omena M, Clegg A, Lupiwa T, et al. Bacterial and viral etiology of severe infection in children less than three months old in the highlands of Papua New Guinea. *Pediatr Infect Dis J* 1999;18:S42-9.
 20. Yuasi L, Sapura J, Kave J, Nivio B, Ford R, et al. The aetiology of pneumonia and meningitis in hospitalised children from the Eastern Highlands Province, Papua New Guinea. 55th Annual medical symposium of the medical society of Papua New Guinea, 1st-6th Sep, 2019. Papua New Guinea: Port Moresby, 2019.
 21. WHO. WHO and UNICEF Estimates of National immunization coverage. Papua New Guinea; 2020
 22. Blyth CC, Ford R, Sapura J, Kumani T, Masiria G, et al. Childhood pneumonia and meningitis in the Eastern Highlands Province, Papua New Guinea in the era of conjugate vaccines: study methods and challenges. *Pneumonia (Nathan)* 2017;9:5.
 23. Laman M, Manning L, Hwaiwhange I, Vince J, Aipit S, et al. Lumbar puncture in children from an area of malaria endemicity who present with a febrile seizure. *Clin Infect Dis* 2010;51:534-540.
 24. Gladstone RA, Lo SW, Lees JA, Croucher NJ, van Tonder AJ, et al. International genomic definition of pneumococcal lineages, to contextualise disease, antibiotic resistance and vaccine impact. *EBioMedicine* 2019;43:338-346.
 25. Blyth CC, Britton KJ, Nguyen CD, Sapura J, Kave J, et al. Effectiveness of 13-valent *Pneumococcal conjugate* vaccine against hypoxic pneumonia and hospitalisation in Eastern Highlands Province, Papua New Guinea: an observational cohort study. *Lancet Reg Health West Pac* 2022;22:100432.
 26. Epping L, van Tonder AJ, Gladstone RA, Bentley SD, et al. SeroBA: rapid high-throughput serotyping of *Streptococcus pneumoniae* from whole genome sequence data. *Microb Genom* 2018;4.
 27. Metcalf BJ, Gertz RE, Gladstone RA, Walker H, Sherwood LK, et al. Strain features and distributions in pneumococci from children with invasive disease before and after 13-valent conjugate vaccine implementation in the USA. *Clin Microbiol Infect* 2016;22:60.
 28. Inouye M, Dashnow H, Raven L-A, Schultz MB, Pope BJ, et al. SRST2: Rapid genomic surveillance for public health and hospital microbiology labs. *Genome Med* 2014;6:90.
 29. Li Y, Metcalf BJ, Chochua S, Li Z, Gertz RE, et al. Penicillin-binding protein transpeptidase signatures for tracking and predicting β -lactam resistance levels in *Streptococcus pneumoniae*. *mBio* 2016;7:e00756-16.
 30. CLSI. *Performance Standards for Antimicrobial Susceptibility Testing*. Clinical and Laboratory Standards Institute, 2014.
 31. Lees JA, Harris SR, Tonkin-Hill G, Gladstone RA, Lo SW, et al. Fast and flexible bacterial genomic epidemiology with PopPUNK. *Genomics* 2018. DOI: 10.1101/360917.
 32. Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM; 2013. <https://arxiv.org/abs/1303.3997>
 33. Van der Auwera GA, Carneiro MO, Hartl C, Poplin R, Del Angel G, et al. From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Curr Protoc Bioinformatics* 2013;43:11.
 34. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, et al. Genome project data processing subgroup genome project data processing subgroup, 2009 the sequence alignment/map format and samtools. *Bioinformatics* 2009;25:2078-2079.
 35. Croucher NJ, Page AJ, Connor TR, Delaney AJ, Keane JA, et al. Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using Gubbins. *Nucleic Acids Res* 2015;43:e15.
 36. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 2014;30:1312-1313.
 37. R Core Team. *R: A Language and Environment for Statistical Computing Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing; 2019.
 38. Dunne EM, Murad C, Sudigdoadi S, Fadlyana E, Tarigan R, et al. Carriage of *Streptococcus pneumoniae*, *Haemophilus influenzae*, *Moraxella catarrhalis*, and *Staphylococcus aureus* in Indonesian children: a cross-sectional study. *PLoS One* 2018;13:e0195098.
 39. Murad C, Dunne EM, Sudigdoadi S, Fadlyana E, Tarigan R, et al. Pneumococcal carriage, density, and co-colonization dynamics: a longitudinal study in Indonesian infants. *Int J Infect Dis* 2019;86:73-81.
 40. Soewignjo S, Gessner BD, Sutanto A, Steinhoff M, Prijanto M, et al. *Streptococcus pneumoniae* nasopharyngeal carriage prevalence, serotype distribution, and resistance patterns among children on Lombok Island, Indonesia. *Clin Infect Dis* 2001;32:1039-1043.
 41. Corander J, Fraser C, Gutmann MU, Arnold B, Hanage WP, et al. Frequency-dependent selection in vaccine-associated pneumococcal population dynamics. *Nat Ecol Evol* 2017;1:1950-1960.
 42. Ndlangisa K, du Plessis M, Allam M, Wolter N, de Gouveia L, et al. Invasive disease caused simultaneously by dual serotypes of *Streptococcus pneumoniae*. *J Clin Microbiol* 2018;56:e01149-17.
 43. Dagan R, Ben-Shimol S, Benisty R, Regev-Yochay G, Lo SW, et al. A nationwide outbreak of invasive pneumococcal disease in Israel caused by *Streptococcus pneumoniae* serotype 2. *Clin Infect Dis* 2021;73:e3768-e3777.
 44. Saha SK, Al Emran HM, Hossain B, Darmstadt GL, Saha S, et al. *Streptococcus pneumoniae* serotype-2 childhood meningitis in Bangladesh: a newly recognized pneumococcal infection threat. *PLoS One* 2012;7:e32134.
 45. Lehmann D. Celebrating ARI research at PNGIMR. In: *Papua New Guinea Institute of Medical Research Fortieth Anniversary Colloquium*; Jun 29. Goroka, PNG, 2008.

Five reasons to publish your next article with a Microbiology Society journal

1. When you submit to our journals, you are supporting Society activities for your community.
2. Experience a fair, transparent process and critical, constructive review.
3. If you are at a Publish and Read institution, you'll enjoy the benefits of Open Access across our journal portfolio.
4. Author feedback says our Editors are 'thorough and fair' and 'patient and caring'.
5. Increase your reach and impact and share your research more widely.

Find out more and submit your article at microbiologyresearch.org.