

# SCIENTIFIC REPORTS



OPEN

## Rapid replacement by non-vaccine pneumococcal serotypes may mitigate the impact of the pneumococcal conjugate vaccine on nasopharyngeal bacterial ecology

Brenda Kwambana-Adams<sup>1,3</sup>, Blake Hanson<sup>2</sup>, Archibald Worwui<sup>1</sup>, Schadrac Agbla<sup>1,3</sup>, Ebenezer Foster-Nyarko<sup>1</sup>, Fatima Ceasay<sup>1</sup>, Chinelo Ebruke<sup>1,3</sup>, Uzochukwu Egere<sup>1</sup>, Yanjiao Zhou<sup>2</sup>, Maze Ndukum<sup>4</sup>, Erica Sodergren<sup>2</sup>, Michael Barer<sup>5</sup>, Richard Adegbola<sup>6</sup>, George Weinstock<sup>2</sup> & Martin Antonio<sup>1,3,7</sup>

There is growing concern that interventions that alter microbial ecology can adversely affect health. We characterised the impact of the seven-valent pneumococcal conjugate vaccine (PCV7) on pneumococcal carriage and the bacterial component of the nasopharyngeal microbiome during infancy. Newborns were recruited into three groups as follows: Group 1 (n = 33) was the control group and comprised infants who received PCV7 after 6 months and came from unvaccinated communities. Group 2 (n = 30) came from unvaccinated communities and Group 3 (n = 39) came from vaccinated communities. Both group 2 and 3 received PCV7 at 2, 3 and 4 months. Culture and 16S rRNA gene sequencing were performed on nasopharyngeal specimens collected at regular intervals from infants. Nasopharyngeal carriage of PCV7 serotypes in Group 1 was significantly higher than in Group 2 and 3 ( $p < 0.01$ ). However, pneumococcal carriage remained comparable due to an expansion of non-vaccine serotypes in Groups 2 and 3. Determination of phylogenetic dis(similarities) showed that the bacterial community structures were comparable across groups. A mixed effects model showed no difference in community richness ( $p = 0.15$ ) and Shannon  $\alpha$ -diversity ( $p = 0.48$ ) across the groups. Immediate replacement of pneumococcal vaccine serotypes with non-vaccine serotypes may mitigate the impact of PCV7 on nasopharyngeal bacterial community structure and ecology.

*Streptococcus pneumoniae*, the pneumococcus, causes several serious and often life-threatening infections including meningitis, pneumonia and bacteraemia<sup>1</sup>. The 7-valent pneumococcal polysaccharide-diphtheria CRM<sub>197</sub> protein conjugate vaccine (PCV7) was licensed in 2000 and markedly reduced the carriage of vaccine serotypes<sup>2,3</sup> and the incidence of invasive disease<sup>4-6</sup>. Childhood deaths attributed to pneumococcus reduced from 735,000 in 2000 to 476,000 in 2008 following widespread use of PCVs. The 10-valent PCV formulation (PCV10) includes all the serotypes in PCV7 and serotypes 1, 5 and 7F<sup>7,8</sup>. The 13-valent PCV (PCV13) formulation includes all the additional serotypes 3, 6A and 19A. PCV10 and PCV13 have replaced PCV7 and had been introduced in 132 countries globally as of September 2016<sup>7</sup>.

Although PCVs are a remarkable public health success, their long-term utility is threatened by the emergence of serotype replacement and possibly species replacement. Although serotype replacement<sup>9,10</sup> and serotype switching<sup>11-13</sup> associated with widespread use of PCVs have been reported, the impact of vaccination on the

<sup>1</sup>Vaccines and Immunity Theme, Medical Research Council Unit The Gambia (MRCG), Fajara, The Gambia. <sup>2</sup>The Jackson Laboratory for Genomic Medicine, Farmington, CT, USA. <sup>3</sup>London School of Hygiene and Tropical Medicine, London, UK. <sup>4</sup>The Genome Institute (Washington University in St Louis), St. Louis, Missouri, USA. <sup>5</sup>Infection, Immunity and Inflammation, University of Leicester, Leicester, UK. <sup>6</sup>GSK, Wavre, Belgium. <sup>7</sup>Division of Microbiology & Immunity, Warwick Medical School, University Of Warwick, Coventry, UK. Correspondence and requests for materials should be addressed to B.K.-A. (email: [bkwambana@mrc.gm](mailto:bkwambana@mrc.gm))

Characteristics	Category	Group 1 (n = 33)	Group 2 (n = 30)	Group 3 (n = 39)	p-value
Sex, n (%)	Male	20 (60.6)	13 (43.3)	19 (48.7)	0.37 <sup>a</sup>
	Female	13 (39.4)	17 (56.7)	20 (51.3)	
Birthplace, n (%)	Home	9 (27.3)	20 (66.7)	23 (59.0)	<0.01 <sup>b</sup>
	Health Centre	4 (12.1)	2 (6.7)	10 (25.6)	
	Hospital	20 (60.6)	8 (26.7)	6 (15.4)	
Birth type, n (%)	Vaginal	33 (100)	30 (100)	39 (100)	N/A
Mother's age in years, n (%)	<20	3 (9.1)	2 (6.7)	6 (15.4)	0.84 <sup>b</sup>
	20–34	20 (60.6)	20 (66.7)	23 (59.0)	
	≥35	10 (30.3)	8 (26.7)	10 (25.6)	
Gestation, n (%)	Full term	33 (100)	30 (100)	38 (97.4)	0.99 <sup>b</sup>
	Premature	0 (0.0)	0 (0.0)	1 (2.6)	
Siblings, n (%)	0	4 (12.1)	4 (13.3)	6 (15.4)	0.75 <sup>b</sup>
	1–3	16 (48.5)	12 (40.0)	21 (53.9)	
	>3	13 (39.4)	14 (46.7)	12 (30.8)	
Ethnicity, n (%)	Mandinka	6 (18.2)	7 (23.3)	12 (30.8)	0.38 <sup>a</sup>
	Jola	18 (54.6)	18 (60.0)	23 (59.0)	
	Other	9 (27.3)	5 (16.7)	4 (10.3)	
Birth weight in kg*	Median (min, max)	3.1 (2.5–3.9)	3 (2.6–3.4)	3 (1.9–4.3)	0.34 <sup>c</sup>
Mother vaccinated, n (%)	No	33 (100)	30 (100)	18 (46.2)	<0.01 <sup>a</sup>
	Yes	0 (0.0)	0 (0.0)	21 (53.8)	
Vaccinated siblings, n (%)	No	33 (100)	30 (100)	34 (87.2)	0.01 <sup>b</sup>
	Yes	0 (0.0)	0 (0.0)	5 (12.8)	

**Table 1.** Baseline characteristics of newborns and their mothers by study group. <sup>a</sup>Chi square test. <sup>b</sup>Fisher's exact test. <sup>c</sup>Kruskal-Wallis test. \*Birth weight was not recorded for 15, 16 and 11 of Group 1, 2 and 3 newborns respectively.

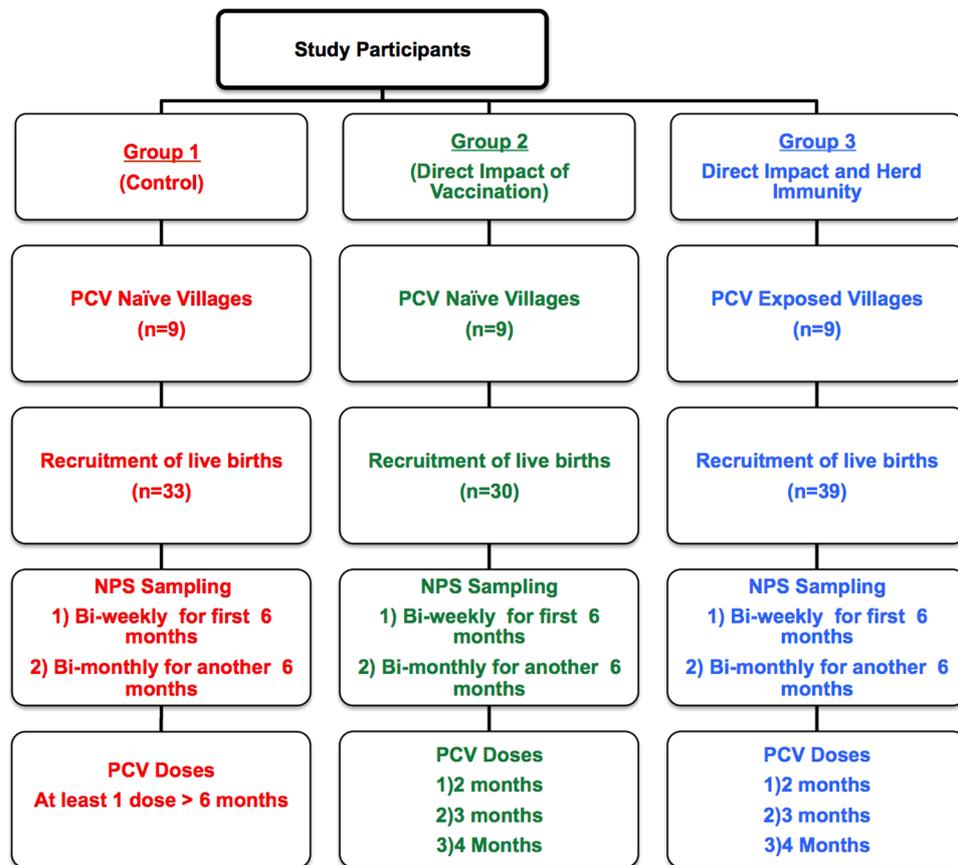
nasopharyngeal microbial ecology, which may influence replacement disease, has not been determined among West African children. Few studies have reported changes in the epidemiology of bacterial diseases associated with widespread use of PCV7<sup>14–16</sup>; however, it is unclear how these changes are related to nasopharyngeal microbial colonisation. The nasopharynx is an important reservoir of commensal and pathogenic microbes, which can migrate to the sinuses, middle ear and lower respiratory tract and invade the blood system. The issue of replacement in carriage and disease is of particular interest within the Gambian context where infant pneumococcal nasopharyngeal carriage rates exceed 90% and PCV7 protects against approximately 63% of previously circulating serotypes<sup>17, 18</sup>. PCV7 was introduced in the routine immunization programme in The Gambia in August 2009, and was replaced by PCV13 in June 2011.

We set out to determine the effects of PCV7 on nasopharyngeal microbial ecology and assess its wider public health impact by asking whether vaccination significantly alters the composition and structure of the infant nasopharyngeal microbiome. A birth cohort of 102 Gambian newborns recruited into three groups with distinct PCV7 vaccine exposures was rigorously sampled for one year to address this question.

## Results

**Characteristics of study participants.** All newborns from participating villages for whom informed parental consent was obtained were recruited into the study until each of the vaccination groups had at least 30 children. A total of 102 children were recruited into each of the study groups as follows, Group 1 (n = 33), Group 2 (n = 30) and Group 3 (n = 39). The baseline characteristics of the newborns recruited into in each of the study groups are summarised in Table 1. There were no significant differences in the baseline characteristics across the study groups except for birthplace, maternal vaccination and sibling vaccination status. Group 3 newborns came from villages with widespread PCV7 vaccination and were expected to have high rates of maternal and sibling vaccination. The study completion rate was 94% (96/102); 6 children did not complete the study, 3 died and 3 dropped out or were lost to follow-up. None of the children that died had been vaccinated with PCV7. 1595 nasopharyngeal swabs were collected from the children, with a mean of 16 specimens (range 1 to 17) per infant over the 12 months of follow-up as outlined in Fig. 1.

**Pneumococcal carriage dynamics.** Pneumococcus was found in 79% (1258/1595) of the nasopharyngeal specimens by culture. Vaccine serotypes were detected in 30% (383/1258) of the pneumococcus positive specimens by the sweep latex agglutination serotyping method<sup>19</sup>. Pneumococcal carriage among Group 1 infants in the first 12 months of life was 79% (95% CI: 76–83), comparable to Group 2 75% (95% CI: 71–79), p = 0.28 and Group 3 82% (95% CI: 79–85), p = 0.29. In contrast, carriage of PCV7 serotypes was 37% (95% CI: 32–41) among Group 1 infants, significantly higher than among Group 2 infants 20% (95% CI: 17–24%), p < 0.01 and Group 3 infants 17% (95% CI: 14–20), p < 0.01. Pneumococcal, vaccine type and non-vaccine type carriage across the 12 months of follow-up are shown in Fig. 2A,B and C. The carriage of vaccine type pneumococci decreased by nearly half



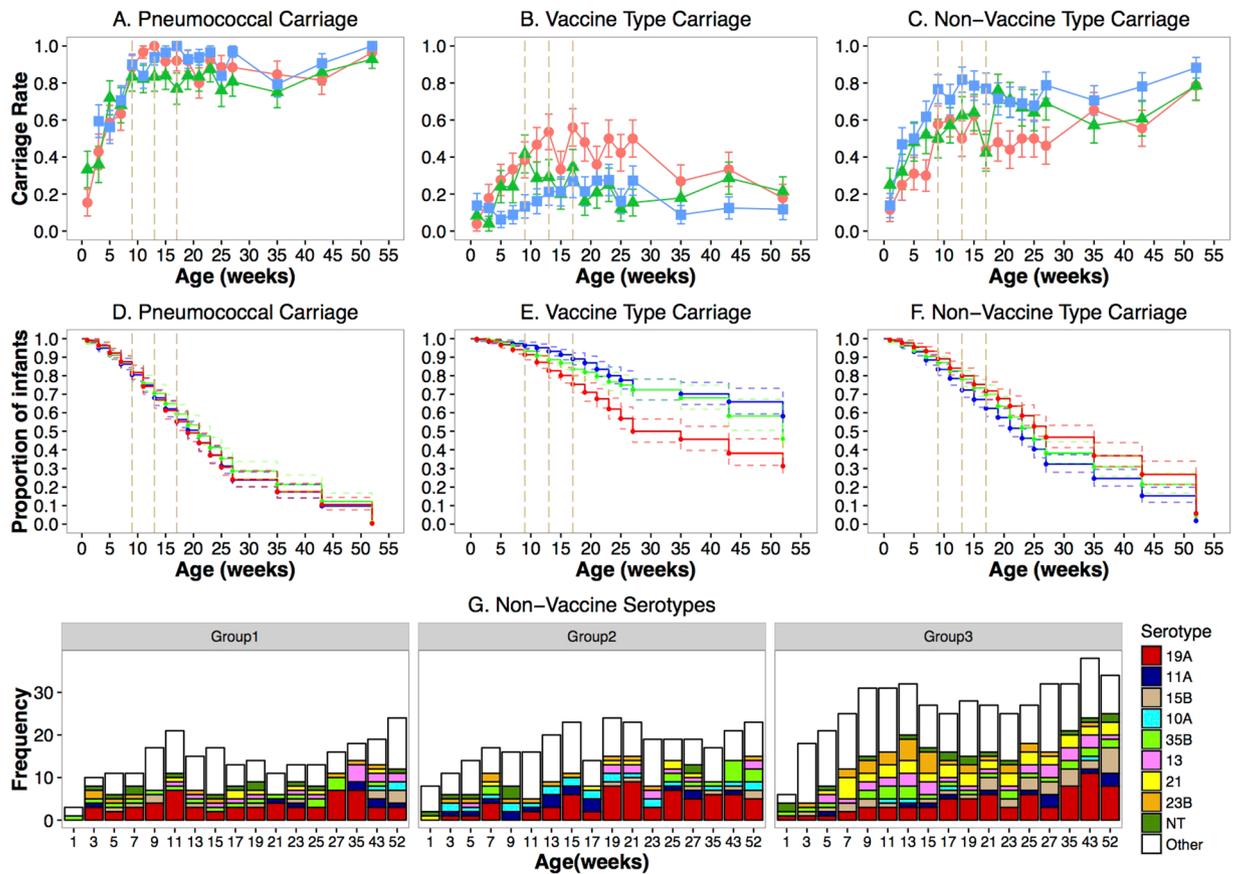
**Figure 1.** Flowchart of infant recruitment and sample collection.

following the first dose of PCV7 given at 9 weeks among Group 2 infants (Fig. 2B). Vaccine type carriage declined between 35 and 52 weeks among Group 1 infants following the nationwide introduction of PCV7. Although pneumococcal acquisition was similar across all three groups (Fig. 2D), Group 1 infants acquired vaccine type pneumococci more rapidly than Group 3 infants (Fig. 2E). The most common non-vaccine serotypes across all the vaccination groups were 19A, 23B and 35B. Carriage of serotype 15B was most common among Group 3 infants who came from heavily vaccinated villages and received PCV at 2, 3 and 4 months (Fig. 2G).

Carriage of pneumococcus among children born to vaccinated mothers was 78% (95% CI: 76–81%), comparable to carriage rates among infants born to unvaccinated mothers 83% (95% CI: 78–87%),  $p < 0.065$ . However, infants with vaccinated mothers had significantly lower carriage of pneumococcal vaccine serotypes, 19% (95% CI: 14–23%) compared to infants with unvaccinated mothers 25% (95% CI: 23–28%),  $p = 0.014$ . Nonetheless, there was no significant difference in carriage of vaccine serotypes among children with at least one vaccinated sibling 23% (95% CI: 14–32%), and those without any vaccinated siblings 24% (95% CI: 22–26%),  $p = 0.920$ . Infants born at a health facility (health center or hospital) had significantly higher carriage of pneumococcal vaccine serotypes 27% (95% CI: 23–30%) than infants born at home 21% (95% CI: 19–24%),  $p = 0.009$ . There was no difference in the carriage of pneumococcus among children born at health facilities 79% (95% CI: 76–82%) and at home 79% (95% CI: 77–82%),  $p = 0.862$ .

**Nasopharyngeal microbial communities.** Of the 1595 nasopharyngeal specimens collected, 16S rDNA deep sequencing was successfully performed on 91% (1460/1595). Initial amplification and/or sequencing with a minimum of 500 reads per sample failed in 9% of (135/1595) NPS specimens. The sequences from this dataset were binned into 814 operational taxonomic units (OTUs) representing at least 297 genera and 15 Phyla. Firmicutes, Proteobacteria, Bacteroidetes and Actinobacteria accounted for at least 90% of the nasopharyngeal microbiota across all time points in each of the groups (Fig. 3A). Likewise, the top ten genera, which included *Streptococcus*, *Staphylococcus*, *Corynebacterium*, *Pseudomonas*, *Moraxella* and *Haemophilus*, accounted for more than 70% of the microbiota throughout the first 12 months of life (Fig. 3B). These microbes form the backbone of the infant nasopharyngeal microbiome while there are numerous low abundance microbes, which may account for most of the diversity within and across individuals.

In the first week of life, *Staphylococcus*, *Corynebacterium* and *Streptococcus* were the dominant genera (Fig. 3B). However, while the initial dominance of *Staphylococcus* decreased in the subsequent weeks, the relative abundance of *Streptococcus* and *Pseudomonas* increased, accounting for over 50% of the microbiota at some visits.

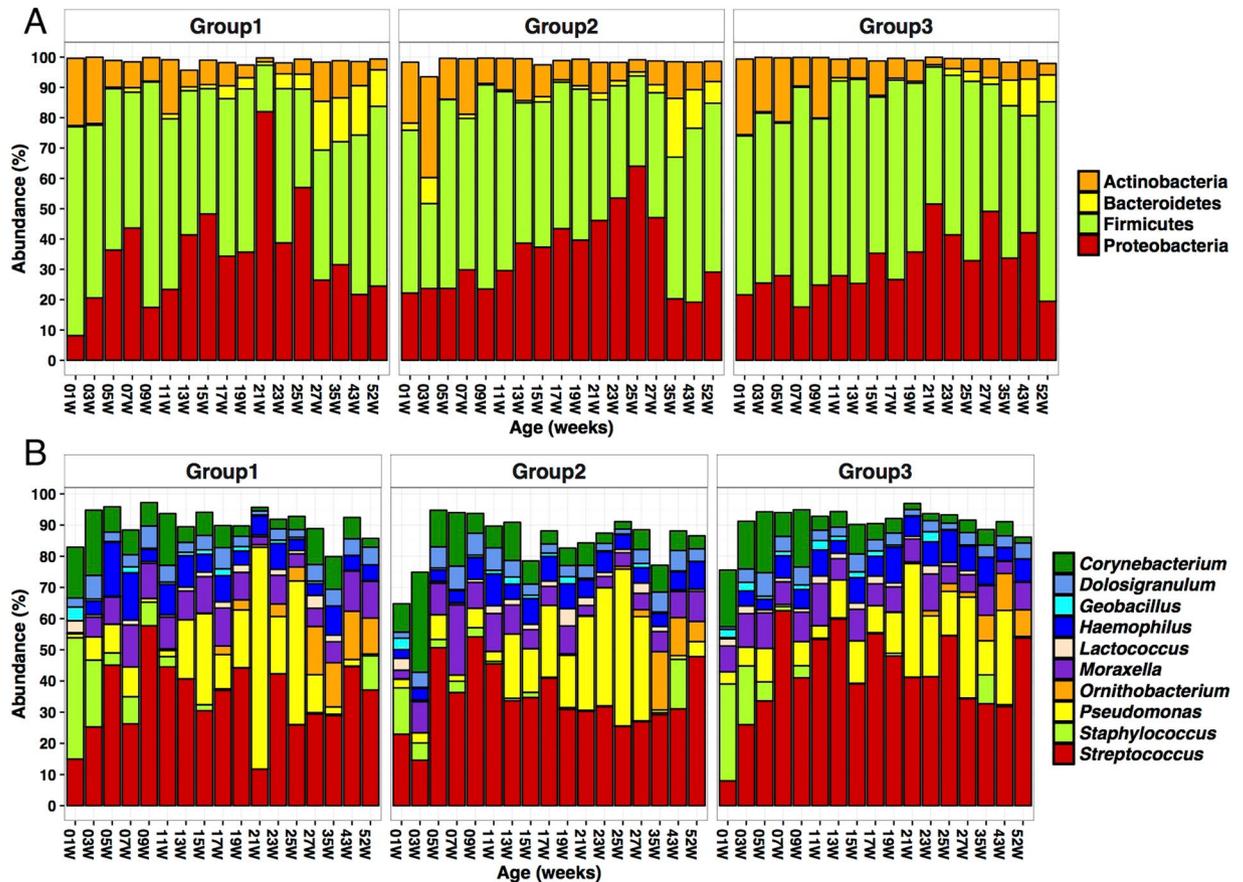


**Figure 2.** Nasopharyngeal carriage dynamics of pneumococcus, vaccine and non-vaccine serotypes in the first year of life. Carriage rates of pneumococcus (A), vaccine (B) and non-vaccine serotypes (C). Kaplan Meier Survival Curve for time to first acquisition of pneumococcus (D), vaccine serotypes (E) and non-vaccine serotypes (F). Distribution of the top ten vaccine serotypes by group (G). The vertical dashed lines are the study 2, 3, and 4 month vaccination time points. P values generated from an adjusted logistic regression model. The line colour denotes the group: red for Group 1, green for group 2 and blue for Group 3.

**Effect of vaccination on community composition and structure.** To evaluate the effect of PCV7 vaccination on overall microbial composition, the phylogenetic dis(similarities) among communities from the three groups were evaluated by MDS/PCoA on weighted-UniFrac distance. There were no distinct shifts in community composition associated with vaccination group as communities from all three groups clustered closely together at each of the 17 age-based visits (Fig. 4). The distribution of the most abundant OTUs across time among the groups appeared very similar at both Phylum and Genus taxonomic levels (Fig. 3). MDS/PCoA on weighted-UniFrac suggested higher variability in community composition between the 15<sup>th</sup> and 27<sup>th</sup> weeks after birth compared to the early and latter weeks (Fig. 4).

A random intercept mixed-effects model was used to investigate the interaction between richness (Supplementary Table 1) and Shannon  $\alpha$ -diversity (Supplementary Table 2) with vaccination group and age adjusting for time-variant and time-invariant explanatory variables. The square root of richness was higher between 11 and 17 weeks after birth (adjusted difference: 0.75 (95% CI: 0.41, 1.10) and between 19 and 27 weeks after birth (adjusted difference: 0.82 (95% CI: 0.30, 1.34) compared to the first 9 weeks of life. No difference in richness was found between the first 9 weeks of life and the last 20 weeks of the first year. These trends are visualized in Fig. 5A. Nasopharyngeal microbial communities harboured a median of 67 (range 1–264) OTUs per specimen and the distribution of richness was similar across all three groups (Fig. 5B). Shannon  $\alpha$ -diversity decreased over time during the first 9 weeks of life (adjusted time effect:  $-0.06$  (95% CI:  $-0.08$ ,  $-0.03$ )) and became steady thereafter (visualized in Fig. 5C). The distribution of Shannon  $\alpha$ -diversity was similar across all three groups (Fig. 5D). There was no evidence of difference in richness ( $p = 0.15$ ) and Shannon  $\alpha$ -diversity across the three vaccination groups ( $p = 0.48$ ) in the mixed effects random intercept model. Other factors such as breastfeeding and mother's age appeared to have significant associations with both richness and Shannon  $\alpha$ -diversity (Supplementary Tables 1 and 2).

**Effect of PCV7 vaccination on individual OTUs.** Differential abundance of individual OTUs across the vaccination groups was evaluated by the use of the negative binomial distribution and shrinkage estimator for the distribution's variance. The variance of most OTUs across all the nasopharyngeal samples was very low ( $< \log_{10}(-4)$ ) as shown in Fig. 6. There were 38, 27, and 33 OTUs that had differential abundance between groups 1 &



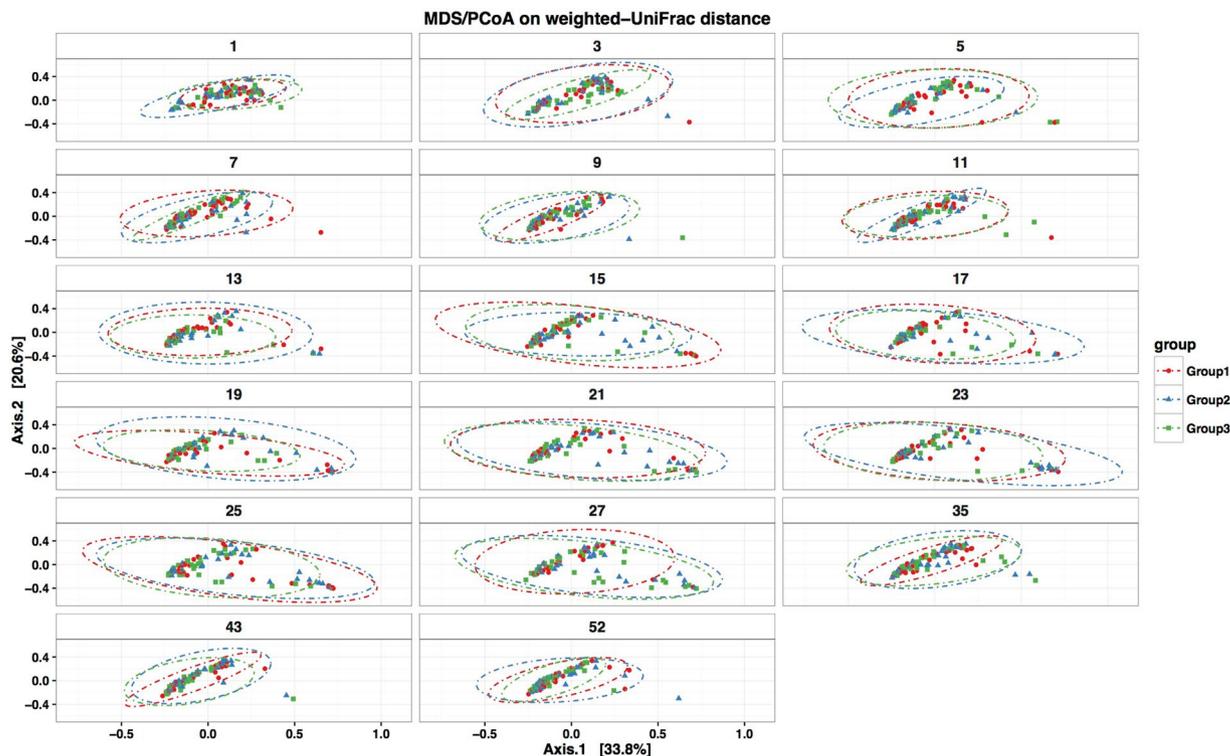
**Figure 3.** Composition of the nasopharyngeal microbiome in the first year of life. The distribution of the top 5 phyla (A) and the top 10 genera (B) in the nasopharyngeal microbiome from birth to 52 weeks for each vaccination group is shown.

2, groups 1 & 3 and groups 2 & 3 respectively (Fig. 6A,B and C). The OTUs with differential abundance across all the groups represented 56 unique OTUs belonging to the Firmicutes, Proteobacteria and Actinobacteria phyla. A total of 9 OTUs representing *Streptococcus*, *Moraxella*, *Dolosigranulum*, *Haemophilus*, *Pseudomonas* and *Pelomonas* genera had significantly different abundance across the vaccination groups after controlling for confounders in a mixed effect model performed as described above (Table 2). Group 2 and Group 3 infants had reduced abundance of 7 of the 9 OTUs compared to Group 1 infants.

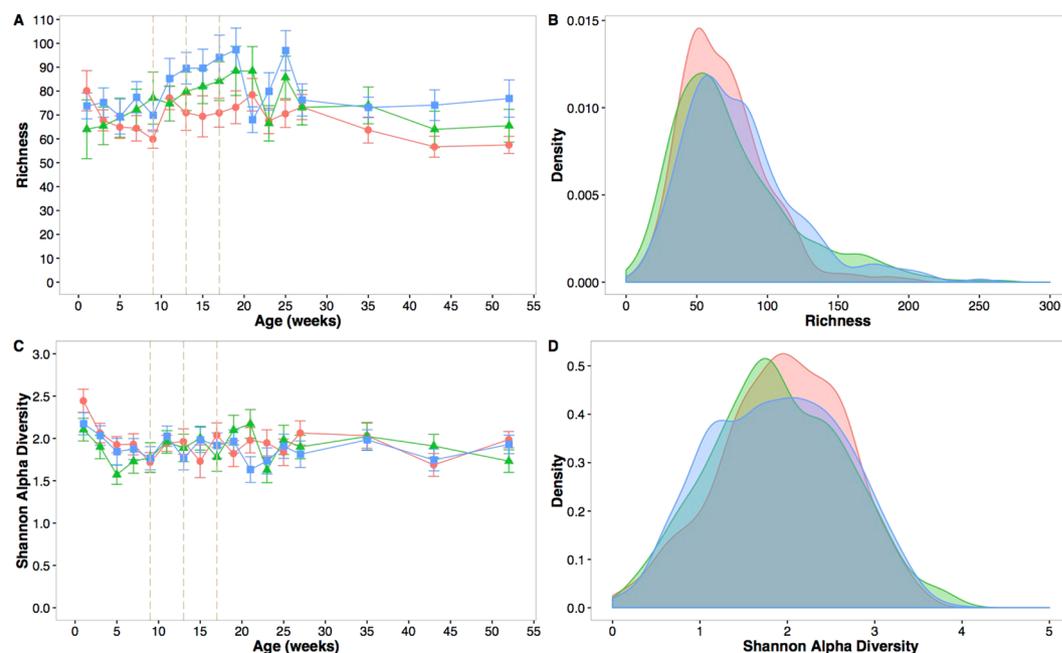
## Discussion

There is growing concern that vaccination and antimicrobial treatment strategies that alter the normal balance of microbial communities could have temporal or pervasive adverse effects on health<sup>20</sup>. In this unique study of infants followed up from birth to 1 year with intensive sampling, we have demonstrated that while PCV7 vaccination effectively reduces the nasopharyngeal carriage of vaccine serotypes, pneumococcal carriage remains high (above 75%) among vaccinated infants due to an immediate expansion of non-vaccine serotypes throughout the first year of life. We have concurrently shown that the overall nasopharyngeal microbial community composition and structure remain largely unaltered by exposure to PCV7, even in the presence of herd immunity. These findings together suggest that exposure to PCV7 does not significantly alter the structure and composition of the infant nasopharyngeal microbiome among Gambian infants, likely due to immediate replacement of vaccine by non-vaccine serotypes.

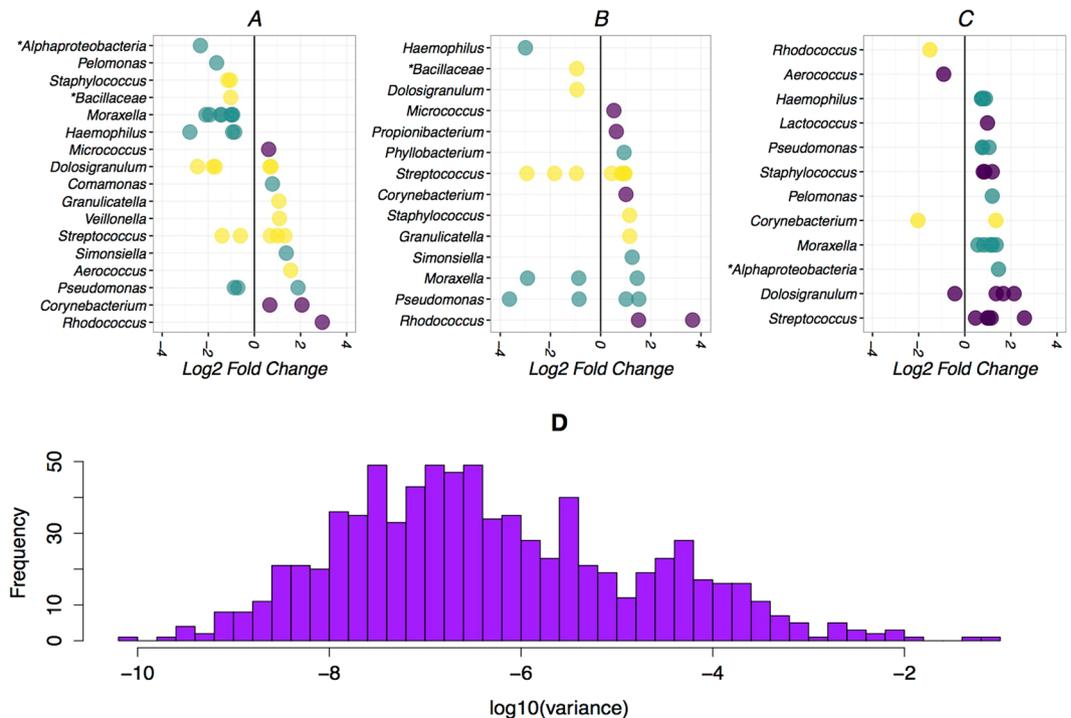
Replacement of vaccine with non-vaccine serotypes associated with PCVs was reported even before the vaccine was first licensed. Following a small study of an experimental pentavalent CRM<sub>197</sub> PCV among children, Obaro and colleagues reported a reduction in the carriage of vaccine types, which was countered by a commensurate increase in the carriage of non-vaccine serotypes<sup>21</sup>. Virtually complete replacement of vaccine with non-vaccine serotypes in carriage associated with PCV vaccination has been reported in several randomized controlled trials in South Africa<sup>22</sup>, the Netherlands<sup>23</sup>, Israel<sup>24</sup>, the United States of America<sup>25</sup> and subsequently in The Gambia<sup>26</sup>. A new insight this study adds is how rapidly this replacement takes place among infants. Nasopharyngeal swabs were collected every two weeks which provided a unique opportunity to show that not only does clearance of vaccine serotypes occur rapidly (<14 days) following the first dose of PCV7 but serotype replacement also occurs concurrently (Fig. 2). This is in contrast to a previous report that suggested that reductions in vaccine-serotype carriage occur within one or two months<sup>27</sup>.



**Figure 4.** Clustering of nasopharyngeal microbial profiles by group in the first year of life. The phylogenetic dis(similarities) among communities from the three vaccination groups were evaluated by MDS/PCoA on weighted-UniFrac distances. The dotted eclipses show the 95% confidence intervals for the ordinations for each group at each time point. Each circle represents the microbial profile of a nasopharyngeal specimen from an infant and the colour denotes the vaccination group; red for Group 1, green for group 2 and blue for Group 3. The distance between circles indicates the variability between nasopharyngeal microbial profiles i.e. longer distances show greater dissimilarity.



**Figure 5.** Nasopharyngeal microbiome richness and Shannon  $\alpha$ -diversity in the first year of life by group. Richness over time (A); Density plot of richness (B); Shannon  $\alpha$ -diversity in the first year of life (C); Shannon  $\alpha$ -diversity in the first year of life (D). The vertical dashed lines are the study vaccination time points. The colour denotes the vaccination group; red for Group 1, green for group 2 and blue for Group 3.



**Figure 6.** Differentially abundant OTUs across vaccination group in the first year of life. Differential abundance of individual OTUs across the vaccination groups was evaluated by the use of the negative binomial distribution and shrinkage estimator for the distribution's variance using the DESeq. 2 extension in the Phyloseq package in R. OTUs that were differentially abundant ( $\alpha < 0.01$ ) between Group 1 and 2 (A), Group 1 and 3 (B), and Group 2 and 3 (C). The log<sub>10</sub> variance of most OTUs across the samples was very low (D). The colours represent the phylum to which the OTUs belong: purple circles represent *Firmicutes*, green circles represent *Proteobacteria* and yellow circles represent *Actinobacteria*. The asterisk (\*) indicates a taxonomic Class while the other labels represent Genera.

The most common non-vaccine serotypes carried by vaccinated children were serotypes 19A, 35B and 15B. Increases in carriage of these serotypes following PCV vaccination have been widely reported across the globe and previously reviewed<sup>9</sup>. Previous pneumococcal surveys conducted in the same region as this study show that all three serotypes were circulating in the population prior to widespread use of PCV7 in the Gambia<sup>17, 28, 29</sup>. A pre-PCV longitudinal study of pneumococcal carriage among Gambian infants showed that serotype 35B was amongst the most commonly isolated serotypes at first acquisition and was amongst the predominant serotypes, accounting for just below 4% of pneumococcal isolates<sup>17</sup>. Both unmasking and increased acquisition are the most probable drivers of the expansion of the non-vaccine serotypes following PCV7 vaccination<sup>30</sup>. The detection of multiple serotypes was enhanced by the use of the sweep method as opposed to the traditional method in which one or two colonies per specimen are picked for serotyping. However, it is important to note that the culture-based detection of pneumococci employed in this study has lower sensitivity than molecular methods<sup>31, 32</sup>. Molecular pneumococcal detection and serotyping tools have the added benefit of enhancing multiple serotype detection and providing quantitative data<sup>19, 32, 33</sup>.

Rapid replacement of vaccine with non-vaccine serotype pneumococci is likely to have important implications for invasive disease. Population-based surveillance in the Upper River Region of The Gambia conducted between May 2008 and December 2014 showed a 47% increase in non-vaccine-type invasive pneumococcal disease, which was coupled with a significant decline in vaccine-type invasive pneumococcal disease following the introduction of PCV7 and PCV13<sup>4</sup>. Interestingly, serotypes 19A and 35B were amongst the leading non-PCV7 serotypes causing invasive pneumococcal disease in The Gambia after the introduction of PCV7. It was also striking that 19A invasive pneumococcal disease persisted between 2013 and 2014 among children less than 5 years old despite widespread use of PCV13 in the country<sup>4</sup>.

At least one other study has investigated the impact of PCV7 on the nasopharyngeal microbiome. Biesbroek and colleagues investigated the impact of PCV7 vaccination on the nasopharyngeal microbiota among 97 vaccinated infants and 103 unvaccinated infants in the Netherlands<sup>34</sup>. The vaccinated infants received 3 doses of PCV7 at 2, 4 and 11 months of age. Deep sequencing of nasopharyngeal specimens collected at 12 and 24 months showed changes in the composition and diversity of the microbial communities associated with PCV7 vaccination. However, these changes were temporary. Of note were the increases in the abundance of *Veillonella*, *Prevotella*, *Fusobacterium*, *Streptococcus* and *Neisseria* observed at 12 months. In this study, we found at least 56 OTUs from the Firmicutes, Proteobacteria and Actinobacteria phyla that were differentially abundant across the vaccination groups. This included members of the *Streptococcus*, *Staphylococcus* and *Moraxella* species. After adjusting for possible confounders such as age, sex, season, birthplace, antibiotic treatment, travel history and

OTU	Taxonomic classification (Phylum; Class; Order; Family; Genus)	Group	Frequency (%) (95% CI)	Mean Abundance% (Range)	Coef. (95% CI)	P-value
169	<i>Proteobacteria</i> ; <i>Gammaproteobacteria</i> ; <i>Pseudomonadales</i> ; <i>Pseudomonadaceae</i> , <i>Pseudomonas</i>	1	43 (39–48)	0.41 (0–24.78)	0.00	—
		2	18 (15–22)	0.01 (0–0.55)	−0.77 (−1.34–−0.21)	0.01
		3	30 (26–34)	0.02 (0–0.43)	−1.01 (−1.51–−0.51)	0.00
2	<i>Firmicutes</i> ; <i>Bacilli</i> ; <i>Lactobacillales</i> ; <i>Streptococcaceae</i> , <i>Streptococcus</i>	1	99 (98–100)	29.74 (0–100)	0.00	—
		2	99 (98–100)	32.09 (0–95.74)	0.14 (−0.15–0.42)	0.35
		3	99 (98–100)	35.82 (0–100)	0.46 (0.19–0.74)	<0.01
203	<i>Firmicutes</i> ; <i>Bacilli</i> ; <i>Lactobacillales</i> ; <i>Streptococcaceae</i> , <i>Streptococcus</i>	1	46 (42–51)	0.21 (0–3.25)	0.00	—
		2	25 (21–29)	0.01 (0–0.32)	−0.71 (−1.32–−0.11)	0.02
		3	37 (33–41)	0.02 (0–0.97)	−0.78 (−1.35–−0.21)	0.01
205	<i>Proteobacteria</i> ; <i>Gammaproteobacteria</i> ; <i>Pseudomonadales</i> ; <i>Moraxellaceae</i> ; <i>Moraxella</i>	1	4 (35–44)	0.46 (0–40.47)	0.00	—
		2	27 (23–31)	0.07 (0–3.36)	−0.36 (−1.07–0.35)	0.32
		3	33 (29–37)	0.03 (0–2.34)	−0.95 (−1.6–−0.3)	<0.01
3	<i>Proteobacteria</i> ; <i>Gammaproteobacteria</i> ; <i>Pseudomonadales</i> ; <i>Moraxellaceae</i> ; <i>Moraxella</i>	1	91 (88–93)	6.32 (0–69.94)	0.00	—
		2	86 (83–89)	5.49 (0–72.88)	0.21 (−0.18–0.61)	0.29
		3	94 (92–96)	7.32 (0–79.12)	0.57 (0.19–0.95)	<0.01
46	<i>Firmicutes</i> ; <i>Bacilli</i> ; <i>Lactobacillales</i> ; <i>Carnobacteriaceae</i> ; <i>Dolosigranulum</i>	1	73 (69–78)	1.65 (0–26.37)	0.00	—
		2	62 (57–67)	0.42 (0–9.78)	−1.02 (−1.67–−0.38)	<0.01
		3	74 (7–77)	0.84 (0–21.34)	−0.94 (−1.55–−0.33)	<0.01
48	<i>Proteobacteria</i> ; <i>Betaproteobacteria</i> ; <i>Burkholderiales</i> ; <i>Comamonadaceae</i> ; <i>Pelomonas</i>	1	34 (29–38)	0.18 (0–6.89)	0.00	—
		2	23 (19–27)	0.09 (0–18.01)	−0.54 (−1.08–0.01)	0.05
		3	34 (31–38)	0.1 (0–3.16)	−0.58 (−1.08–−0.08)	0.02
8	<i>Proteobacteria</i> ; <i>Gammaproteobacteria</i> ; <i>Pasteurellales</i> ; <i>Pasteurellaceae</i> ; <i>Haemophilus</i>	1	77 (73–81)	4.44 (0–86.33)	0.00	—
		2	74 (7–78)	4.4 (0–74.06)	0.66 (0.19–1.12)	0.01
		3	85 (82–88)	5.74 (0–80.8)	0.64 (0.19–1.08)	0.01
89	<i>Proteobacteria</i> ; <i>Gammaproteobacteria</i> ; <i>Pasteurellales</i> ; <i>Pasteurellaceae</i> ; <i>Haemophilus</i>	1	33 (29–37)	0.45 (0–23.73)	0	—
		2	26 (22–3)	0.03 (0–1.37)	−0.27 (−0.91–0.38)	42
		3	25 (22–29)	0.02 (0–0.92)	−0.85 (−1.47–−0.23)	0.01

**Table 2.** OTUs with differential abundance by vaccination group in the first year of life. The mixed effect random intercept model was adjusted for time-invariant explanatory variables (gender, place of birth, number of siblings, maternal age) and time-variant explanatory explanatory (time, vaccination period, season, breast feeding, travelling, antibiotic usage and nutritional status based on weight for height).

malnutrition, only 9 OTUs had differential abundance across the vaccination groups (Table 2). Despite these differences, the overall microbiome structure and composition were indistinguishable across the vaccination groups throughout the first year of life. This is consistent with findings from a small study conducted among 60 children aged 12–59 months in Kenya in which the 10-valent pneumococcal non-typeable *Haemophilus influenzae* protein-D conjugate vaccine (PHiD-CV) was not found to significantly alter the nasopharyngeal microbiome<sup>35</sup>.

Although a few studies have reported changes in the carriage of other pathogens such as *Staphylococcus aureus* and *Haemophilus influenzae* that co-colonise the nasopharynx following PCV vaccination<sup>36–38</sup>, it is not clear if there is a link with nasopharyngeal carriage dynamics. However, this does not rule out the possibility that localised changes in microbial ecology at the sites of infection, such as the middle ear in AOM, may play an important role<sup>39–43</sup>. It may be of some interest that the three OTUs that had significantly higher abundance among infants who received PCV7 and came from heavily vaccinated communities belonged to the *Haemophilus*, *Moraxella* and *Streptococcus* genera. This may warrant further investigation.

A key strength of this study is that the infants were intensively sampled in the first 6 months and followed up at regular intervals up to one year. A trade-off to the regular sampling was that the total number of infants recruited was only 102. Hence, more subtle changes in the nasopharyngeal microbiome composition and structure could have been missed. Group 1 was the control (unvaccinated) group in the initial study design, however, infants recruited into this group received PCV7 once it was introduced in The Gambia in August 2009. PCV7 was replaced by PCV13 in 2011 in The Gambia and it is unclear if the higher valency formulations will have similar impact on the nasopharyngeal microbiome.

The evidence presented here strongly suggests that as vaccine serotypes are cleared from the nasopharyngeal ecological niche, they are rapidly replaced by non-vaccine pneumococcal serotypes and not other bacterial species. Although this may provide some reassurance, diligent surveillance of nasopharyngeal carriage is crucial to ensure that the net public health gains of PCVs are maintained over the long term.

## Materials/Subjects and Methods

**Ethical permission.** Ethical approval to conduct this prospective longitudinal study was granted by the Gambian Government and Medical Research Council Unit The Gambia Joint Ethics Committee. Written,

informed consent was obtained from the participants' parents or guardian at recruitment. Enrolment of participants and all procedures were conducted in accordance with the relevant regulations and guidelines.

**Study population.** The Western Region is representative of rural areas in The Gambia. HIV prevalence is estimated at 2% and the majority of villagers belong to the Jola, Mandinka and Fula ethnic groups<sup>29</sup>. Villages in the Western Region, The Gambia covering an area of approximately 90 km<sup>2</sup> were selected for this study. Study participants were recruited from 27 villages each with estimated birth rates between three and twenty-six per year. The villages were split into 3 groups of 9 villages with estimated population sizes of 2000 and birth rates of approximately 80 per year. Trained village reporters in each village recorded and reported pregnancies, births, deaths and other serious events. Recruitment of subjects was carried out between November 2008 and April 2009 after written informed consent. To avert recruitment bias, all newborns from participating villages for whom informed parental consent was obtained were recruited into the study until each of the vaccination groups had at least 30 children.

**Study vaccination groups.** The study participants were recruited into three groups as outlined in the flow-chart in Fig. 1. Group 1 and 2 participants were born in PCV7 unvaccinated villages while Group 3 participants were born in heavily vaccinated communities following the PCV7 vaccine trials conducted in 2006<sup>2</sup>. Group 2 and 3 infants received three doses of PCV7 at 2, 3 and 4 months whereas Group 1 infants did not receive PCV7 at this age. PCV7 was introduced in the Gambia in August 2009, 11 months after the study was started. Group 1 infants received at least 1 dose of PCV7 during the national catch-up campaigns. All Group 1 infants were at least 6 months old when they received the first dose of PCV. Group 2 and Group 3 infants also received additional doses of PCV7 following the implementation of PCV7 in The Gambia after 6 months of age. Nasopharyngeal swabs were collected within seven days of birth, then biweekly for the first six months and subsequently bi-monthly up to one year from all participants.

**Nasopharyngeal sampling process.** Nasopharyngeal swabs were collected as previously described<sup>29</sup> by trained field workers. The swab was immediately inoculated into a vial containing 1 mL of chilled skim milk–tryptone–glucose–glycerol (STGG) transport medium. The vials were kept on ice, transported to the MRCG Laboratories site in Fajara and stored at –70 °C within eight hours of collection. Pneumococci were isolated as previously described<sup>29</sup>.

**DNA isolation for 454-pyrosequencing.** Nasopharyngeal swabs stored in STGG were thawed on ice and gently vortexed for five seconds. DNA was extracted from each swab using the PowerSoil<sup>®</sup> DNA Isolation Kit (MO BIO Laboratories, Carlsbad, CA, USA). 250 µL of the swab solution was transferred to the PowerBead<sup>®</sup> tubes and DNA was extracted following manufacturer's protocol. DNA was eluted in 100 µL of the kit elution buffer and immediately stored at –20 °C. DNA extractions were carried out in batches of 24 including one extraction control to which 250 µL of sterile DNase free water was added instead.

**16S rDNA sequencing.** The V1–V3 region of the 16S rRNA gene were amplified using primers 5'-AGAGTTTGATCCTGGCTCAG-3' and 5'-ATTACCGCGGCTGCTGG-3' as previously described<sup>44, 45</sup>. The primers contained an adapter sequence and unique barcodes. The positive controls for all runs were purified. *Anaerotruncus colhomini*s DNA, reagent controls and non-template controls were set up for each PCR run. Following amplification and purification, the amplicons were pooled at equimolar concentrations and sequenced on the 454 GS FLX Titanium Sequencing Platform (Roche, USA) as described elsewhere<sup>44, 45</sup> at the Genome Institute (University of Washington in St. Louis, MO, USA).

Quality control and data processing were performed using in-house pipelines at the Jackson Laboratory. Briefly, reads with length <200 bp and/or with more than a single ambiguous base call were discarded. Chimeric sequences and reads without the adapter sequences were also removed. Reads from the same samples were binned based on barcode and then the barcode, adapter and primer sequences at both terminals were trimmed. Samples with read depth less than 500 were discarded based on rarefaction analysis of all the samples. Alignment and taxonomic classification (Phylum to Genus) of the reads was carried out with the Ribosomal Database Project Naïve Bayesian Classifier using a 0.5 filter.

**Data collection.** At recruitment, background data including birth weight, birthplace, maternal age, maternal parity, ethnicity, sex, and vaccination status of mothers and siblings were collected. Socio-demographic and clinical information about the participants was collected at each visit including antibiotic use, anthropometric measurements, travel history, breastfeeding status, and infections (ear and chest). All data were collected on approved study forms and double entered and verified in an OpenClinica Electronic Data Capture system (Waltham, MA, USA) with an Access (Microsoft, Seattle, WA, USA) backend.

**Analysis.** Differences in baseline characteristics of the study participants recruited into the three study groups were tested using the Chi square test, Fisher's exact test or Kruskal–Wallis test where appropriate. A mixed effects model was used to compare richness and Shannon  $\alpha$  diversity across the study groups. Richness was positively skewed and was square root transformed for statistical analysis. Time-invariant risk factors (gender, place of birth, number of siblings, maternal age) and time-variant risk factors (age, vaccination period, season, breast feeding, travelling, antibiotic usage and nutritional status based on weight for height) were used as explanatory variables. Age and vaccination group were the main exposure risk factors and the remaining explanatory risk factors were considered as potential confounders. A random-intercept model was used to handle these two types of explanatory risk factors. All explanatory risk factors mentioned above were included in the fixed effects component of the model as well as pairwise and triple interaction terms between time, group and period. A backward model selection approach using the Wald test was adopted. Explanatory risk factors with a non-significant effect at the 5% significance level were discarded. Model residuals and random effects were checked for reliability. Kaplan Meier survival curves were used

to compare the rates to first acquisition of pneumococcus and PCV7 serotypes in the three study groups. Statistical analyses were performed using STATA/SE 14.1, USA. Microbial community composition, structure and ecology analyses were performed using the Phyloseq package<sup>46</sup> with R version 3.2.1. To reduce noise, low-occurrence, poorly represented OTUs were filtered out i.e. OTUs that did not appear more than once in 10% of the samples. To evaluate the effect of PCV7 vaccination group on overall microbial composition, the phylogenetic dis(similarities) among communities from the three groups were evaluated by metric multidimensional scaling (MDS)/ Principle component analysis (PCoA) on weighted UniFrac distance. Differential abundance of individual OTUs across the vaccination groups was explored using the DESeq package-extensions in the Phyloseq package.

**Data availability.** The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

## References

- O'Brien, K. L. *et al.* Burden of disease caused by *Streptococcus pneumoniae* in children younger than 5 years: global estimates. *Lancet (London, England)* **374**, 893–902, doi:10.1016/S0140-6736(09)61204-6 (2009).
- Roca, A. *et al.* Effects of community-wide vaccination with PCV-7 on pneumococcal nasopharyngeal carriage in the Gambia: a cluster-randomized trial. *PLoS medicine* **8**, e1001107, doi:10.1371/journal.pmed.1001107 (2011).
- Millar, E. V. *et al.* Effect of community-wide conjugate pneumococcal vaccine use in infancy on nasopharyngeal carriage through 3 years of age: a cross-sectional study in a high-risk population. *Clinical infectious diseases: an official publication of the Infectious Diseases Society of America* **43**, 8–15, doi:10.1086/504802 (2006).
- Mackenzie, G. A. *et al.* Effect of the introduction of pneumococcal conjugate vaccination on invasive pneumococcal disease in The Gambia: a population-based surveillance study. *The Lancet. Infectious diseases*, doi:10.1016/S1473-3099(16)00054-2 (2016).
- von Gottberg, A. *et al.* Effects of vaccination on invasive pneumococcal disease in South Africa. *N Engl J Med* **371**, 1889–1899, doi:10.1056/NEJMoa1401914 (2014).
- Flasche, S. *et al.* Effect of pneumococcal conjugate vaccination on serotype-specific carriage and invasive disease in England: a cross-sectional study. *PLoS medicine* **8**, e1001017, doi:10.1371/journal.pmed.1001017 (2011).
- Loharikar, A. *et al.* Status of New Vaccine Introduction - Worldwide, September 2016. *MMWR. Morbidity and mortality weekly report* **65**, 1136–1140, doi:10.15585/mmwr.mm6541a3 (2016).
- Lee, L. A. *et al.* The estimated mortality impact of vaccinations forecast to be administered during 2011–2020 in 73 countries supported by the GAVI Alliance. *Vaccine* **31**(Suppl 2), B61–72, doi:10.1016/j.vaccine.2012.11.035 (2013).
- Dagan, R. Serotype replacement in perspective. *Vaccine* **27**(Suppl 3), C22–24 (2009).
- Hausdorff, W. P. & Hanage, W. P. Interim results of an ecological experiment - Conjugate vaccination against the pneumococcus and serotype replacement. *Hum Vaccin Immunother* **12**, 358–374, doi:10.1080/21645515.2015.1118593 (2016).
- Loman, N. J. *et al.* Clonal expansion within pneumococcal serotype 6C after use of seven-valent vaccine. *PLoS one* **8**, e64731, doi:10.1371/journal.pone.0064731 (2013).
- Brueggemann, A. B., Pai, R., Crook, D. W. & Beall, B. Vaccine escape recombinants emerge after pneumococcal vaccination in the United States. *PLoS pathogens* **3**, e168, doi:10.1371/journal.ppat.0030168 (2007).
- Croucher, N. J. *et al.* Selective and genetic constraints on pneumococcal serotype switching. *PLoS Genet* **11**, e1005095, doi:10.1371/journal.pgen.1005095 (2015).
- Guevara, M. *et al.* Changing epidemiology of invasive pneumococcal disease following increased coverage with the heptavalent conjugate vaccine in Navarre, Spain. *Clinical microbiology and infection: the official publication of the European Society of Clinical Microbiology and Infectious Diseases* **15**, 1013–1019 (2009).
- Laupland, K. B., Gregson, D. B., Vanderkooi, O. G., Ross, T. & Kellner, J. D. The changing burden of pediatric bloodstream infections in Calgary, Canada, 2000–2006. *The Pediatric infectious disease journal* **28**, 114–117, doi:10.1097/INF.0b013e318187ad5a (2009).
- Lacapa, R. *et al.* Changing epidemiology of invasive pneumococcal disease among White Mountain Apache persons in the era of the pneumococcal conjugate vaccine. *Clinical infectious diseases: an official publication of the Infectious Diseases Society of America* **47**, 476–484, doi:10.1086/590001 (2008).
- Hill, P. C. *et al.* Nasopharyngeal carriage of *Streptococcus pneumoniae* in Gambian infants: a longitudinal study. *Clinical infectious diseases: an official publication of the Infectious Diseases Society of America* **46**, 807–814, doi:10.1086/528688 (2008).
- Cutts, F. T. *et al.* Efficacy of nine-valent pneumococcal conjugate vaccine against pneumonia and invasive pneumococcal disease in The Gambia: randomised, double-blind, placebo-controlled trial. *Lancet (London, England)* **365**, 1139–1146, doi:10.1016/S0140-6736(05)71876-6 (2005).
- Turner, P. *et al.* Field evaluation of culture plus latex sweep serotyping for detection of multiple pneumococcal serotype colonisation in infants and young children. *PLoS one* **8**, e67933, doi:10.1371/journal.pone.0067933 (2013).
- Peterson, J. *et al.* The NIH Human Microbiome Project. *Genome Res* (2009).
- Obaro, S. K., Adegbola, R. A., Banya, W. A. & Greenwood, B. M. Carriage of pneumococci after pneumococcal vaccination. *Lancet (London, England)* **348**, 271–272 (1996).
- Mbelle, N. *et al.* Immunogenicity and impact on nasopharyngeal carriage of a nonavalent pneumococcal conjugate vaccine. *J Infect Dis* **180**, 1171–1176, doi:10.1086/315009 (1999).
- van Gils, E. J. *et al.* Effect of reduced-dose schedules with 7-valent pneumococcal conjugate vaccine on nasopharyngeal pneumococcal carriage in children: a randomized controlled trial. *JAMA* **302**, 159–167, doi:10.1001/jama.2009.975 (2009).
- Dagan, R. *et al.* Reduction of nasopharyngeal carriage of *Streptococcus pneumoniae* after administration of a 9-valent pneumococcal conjugate vaccine to toddlers attending day care centers. *J Infect Dis* **185**, 927–936, doi:10.1086/339525 (2002).
- Lipsitch, M. *et al.* Strain characteristics of *Streptococcus pneumoniae* carriage and invasive disease isolates during a cluster-randomized clinical trial of the 7-valent pneumococcal conjugate vaccine. *J Infect Dis* **196**, 1221–1227, doi:10.1086/521831 (2007).
- Cheung, Y. B. *et al.* Nasopharyngeal carriage of *Streptococcus pneumoniae* in Gambian children who participated in a 9-valent pneumococcal conjugate vaccine trial and in their younger siblings. *The Pediatric infectious disease journal* **28**, 990–995, doi:10.1097/INF.0b013e3181a78185 (2009).
- Hammit, L. L. *et al.* Population effect of 10-valent pneumococcal conjugate vaccine on nasopharyngeal carriage of *Streptococcus pneumoniae* and non-typeable *Haemophilus influenzae* in Kilifi, Kenya: findings from cross-sectional carriage studies. *Lancet Glob Health* **2**, e397–405, doi:10.1016/S2214-109X(14)70224-4 (2014).
- Hill, P. C. *et al.* Transmission of *Streptococcus pneumoniae* in rural Gambian villages: a longitudinal study. *Clinical infectious diseases: an official publication of the Infectious Diseases Society of America* **50**, 1468–1476, doi:10.1086/652443 (2010).
- Hill, P. C. *et al.* Nasopharyngeal carriage of *Streptococcus pneumoniae* in Gambian villages. *Clinical infectious diseases: an official publication of the Infectious Diseases Society of America* **43**, 673–679, doi:10.1086/506941 (2006).
- Weinberger, D. M., Malley, R. & Lipsitch, M. Serotype replacement in disease after pneumococcal vaccination. *Lancet (London, England)* **378**, 1962–1973, doi:10.1016/S0140-6736(10)62225-8 (2011).

31. Kwambana, B. A., Barer, M. R., Bottomley, C., Adegbola, R. A. & Antonio, M. Early acquisition and high nasopharyngeal colonisation by *Streptococcus pneumoniae* and three respiratory pathogens amongst Gambian new-borns and infants. *BMC infectious diseases* **11**, 175, doi:10.1186/1471-2334-11-175 (2011).
32. Satzke, C. *et al.* Standard method for detecting upper respiratory carriage of *Streptococcus pneumoniae*: updated recommendations from the World Health Organization Pneumococcal Carriage Working Group. *Vaccine* **32**, 165–179, doi:10.1016/j.vaccine.2013.08.062 (2013).
33. Satzke, C. *et al.* The PneuCarriage Project: A Multi-Centre Comparative Study to Identify the Best Serotyping Methods for Examining Pneumococcal Carriage in Vaccine Evaluation Studies. *PLoS medicine* **12**, e1001903; discussion e1001903, doi:10.1371/journal.pmed.1001903 (2015).
34. Biesbroek, G. *et al.* The impact of breastfeeding on nasopharyngeal microbial communities in infants. *Am J Respir Crit Care Med* **190**, 298–308, doi:10.1164/rccm.201401-0073OC (2014).
35. Feazel, L. M. *et al.* Effects of Vaccination with 10-Valent Pneumococcal Non-Typeable Haemophilus influenzae Protein D Conjugate Vaccine (PHiD-CV) on the Nasopharyngeal Microbiome of Kenyan Toddlers. *PLoS one* **10**, e0128064, doi:10.1371/journal.pone.0128064 (2015).
36. Spijkerman, J. *et al.* Long-term effects of pneumococcal conjugate vaccine on nasopharyngeal carriage of *S. pneumoniae*, *S. aureus*, *H. influenzae* and *M. catarrhalis*. *PLoS one* **7**, e39730, doi:10.1371/journal.pone.0039730 (2012).
37. van Gils, E. J. *et al.* Effect of seven-valent pneumococcal conjugate vaccine on *Staphylococcus aureus* colonisation in a randomised controlled trial. *PLoS one* **6**, e20229, doi:10.1371/journal.pone.0020229 (2011).
38. van Gils, E. J., Veenhoven, R. H., Rodenburg, G. D., Hak, E. & Sanders, E. A. Effect of 7-valent pneumococcal conjugate vaccine on nasopharyngeal carriage with *Haemophilus influenzae* and *Moraxella catarrhalis* in a randomized controlled trial. *Vaccine* **29**, 7595–7598, doi:10.1016/j.vaccine.2011.08.049 (2011).
39. Casey, J. R., Adlowitz, D. G. & Pichichero, M. E. New patterns in the otopathogens causing acute otitis media six to eight years after introduction of pneumococcal conjugate vaccine. *The Pediatric infectious disease journal* **29**, 304–309, doi:10.1097/INF.0b013e3181c1bc48 (2010).
40. Block, S. L. *et al.* Community-wide vaccination with the heptavalent pneumococcal conjugate significantly alters the microbiology of acute otitis media. *The Pediatric infectious disease journal* **23**, 829–833, doi:00006454-200409000-00010 (2004).
41. Eskola, J. *et al.* Efficacy of a pneumococcal conjugate vaccine against acute otitis media. *N Engl J Med* **344**, 403–409, doi:10.1056/NEJM200102083440602 (2001).
42. Dupont, D. *et al.* Evolving microbiology of complicated acute otitis media before and after introduction of the pneumococcal conjugate vaccine in France. *Diagn Microbiol Infect Dis* **68**, 89–92, doi:S0732-8893(10)00173-2/j.diagmicrobio.2010.04.012 (2010).
43. Brunton, S. Current face of acute otitis media: microbiology and prevalence resulting from widespread use of heptavalent pneumococcal conjugate vaccine. *Clin Ther* **28**, 118–123, doi:S0149-2918(06)00025-7/j.clinthera.2006.01.011 (2006).
44. Zhou, Y. *et al.* The conjunctival microbiome in health and trachomatous disease: a case control study. *Genome Med* **6**, 99, doi:10.1186/s13073-014-0099-x (2014).
45. Human Microbiome Project, C. A framework for human microbiome research. *Nature* **486**, 215–221, doi:10.1038/nature11209 (2012).
46. McMurdie, P. J. & Holmes, S. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS one* **8**, e61217, doi:10.1371/journal.pone.0061217 (2013).

## Acknowledgements

We would like to thank the Medical Research Council Unit The Gambia (MRCG) for funding and providing the facilities to conduct this research. We would also like to thank the Jackson Laboratories and The Genome Institute (Washington University in St Louis) for sequencing the samples and providing bioinformatics support. The contributions from the Research Molecular Microbiology group at MRCG were immense and we are grateful. We also thank the dedicated field team that worked tirelessly led by Mr Ma Ansu Kinteh, ensuring the smooth running of the project. This study would not have been possible without the invaluable contributions of the study participants, their families and the communities of the Foni Districts.

## Author Contributions

M.A., R.A., M.B., G.W. and B.K. contributed to the conception and design of the work. B.K., U.E. and M.A. were responsible for the recruitment of study participants and the field data collection. E.F., F.C. and C.E. contributed to the microbiology processing of the samples. A.W. performed data management while S.A. performed the statistical analyses. G.W., E.S., M.N., Y.Z., A.W., B.K. and B.H. contributed to the microbiome sequencing and bioinformatic analyses. B.K. and B.H. oversaw the data analysis and interpretation, and drafted the manuscripts. M.A., R.A., M.B. and G.W. provided critical revision of the article.

## Additional Information

**Supplementary information** accompanies this paper at doi:10.1038/s41598-017-08717-0

**Competing Interests:** There are no conflicts to declare. Richard Adegbola is currently an employee of GSK Vaccines. The opinion presented in this paper is the one of the author and does not reflect GSK's position. This does not alter the authors' adherence to the journal policy.

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



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

© The Author(s) 2017