Article

Bifidobacteria support optimal infant vaccine responses

https://doi.org/10.1038/s41586-025-08796-4

Received: 21 November 2023

Accepted: 17 February 2025

Published online: 2 April 2025

Open access

Check for updates

Feargal J. Ryan[1,2,14], Michelle Clarke[3,4,5,14], Miriam A. Lynn[1,2,14], Saoirse C. Benson[1,2], Sonia McAlister[6,7], Lynne C. Giles[8], Jocelyn M. Choo[1,2], Charné Rossouw[1,2], Yan Yung Ng[9], Evgeny A. Semchenko[10], Alyson Richard[2], Lex E. X. Leong[2], Steven L. Taylor[1,2], Stephen J. Blake[1,2], Joyce I. Mugabushaka[1,2], Mary Walker[3], Steve L. Wesselingh[1,2], Paul V. Licciardi[9,11], Kate L. Seib[10], Damon J. Tumes[1,12], Peter Richmond[6,7,13], Geraint B. Rogers[1,2], Helen S. Marshall[3,4,5,15] & David J. Lynn[1,2,15] ✉

Accumulating evidence indicates that antibiotic exposure may lead to impaired vaccine responses[1–4]; however, the mechanisms underlying this association remain poorly understood. Here we prospectively followed 191 healthy, vaginally born, term infants from birth to 15 months, using a systems vaccinology approach to assess the effects of antibiotic exposure on immune responses to vaccination. Exposure to direct neonatal but not intrapartum antibiotics was associated with significantly lower antibody titres against various polysaccharides in the 13-valent pneumococcal conjugate vaccine and the *Haemophilus influenzae* type b polyribosylribitol phosphate and diphtheria toxoid antigens in the combined 6-in-1 Infanrix Hexa vaccine at 7 months of age. Blood from infants exposed to neonatal antibiotics had an inflammatory transcriptional profile before vaccination; in addition, faecal metagenomics showed reduced abundance of *Bifidobacterium* species in these infants at the time of vaccination, which was correlated with reduced vaccine antibody titres 6 months later. In preclinical models, responses to the 13-valent pneumococcal conjugate vaccine were strongly dependent on an intact microbiota but could be restored in germ-free mice by administering a consortium of *Bifidobacterium* species or a probiotic already widely used in neonatal units. Our data suggest that microbiota-targeted interventions could mitigate the detrimental effects of early-life antibiotics on vaccine immunogenicity.

Global immunization programmes prevent millions of deaths per year[5,6]; however, for reasons that are unclear, immune responses to vaccination exhibit substantial interindividual variation and are frequently suboptimal in the populations at most risk from infectious disease[7]. Evidence increasingly indicates that the gut microbiota is an important and targetable factor that may, at least in part, explain these interindividual and interpopulation differences in immune responses to vaccination[7,8]. In mice, antibiotic exposure in early life has been shown to lead to impaired antibody responses to five different adjuvanted and live-attenuated parenteral vaccines[1]. Antibiotic-treated and germ-free (GF) adult mice have also been shown to have impaired antibody responses to an unadjuvanted influenza vaccine but not to the live yellow fever vaccine[9] or a COVID-19 messenger RNA (mRNA) vaccine[10]. Furthermore, adult participants randomized to antibiotics before immunization with unadjuvanted trivalent influenza vaccine had significantly impaired neutralizing and binding antibody titres; however, this was not observed in adults with pre-existing immunity to influenza[11]. On the other hand, oral rotavirus vaccine immunogenicity was not found to be reduced in adults randomized to antibiotics[12]. Consistent with our previous preclinical data suggesting that antibiotics may have a more profound effect on antibody responses to adjuvanted vaccines in early life[1], a recent retrospective study found that vaccine antibody titres were lower in antibiotic-exposed infants[2]. Cumulative antibiotic exposure during the first year of life has also been associated with reduced vaccine response at 12 months[3,4].

To prospectively assess whether antibiotic exposure in early life affects antibody responses to vaccination and to better understand the mechanisms involved, we established a prospective observational study that recruited healthy, full-term, vaginally delivered infants. Infants were defined as not exposed to antibiotics (no-ABX, $n = 80$) or

[1]Precision Medicine, South Australian Health and Medical Research Institute (SAHMRI), Adelaide, South Australia, Australia. [2]Flinders Health and Medical Research Institute, Flinders University, Bedford Park, South Australia, Australia. [3]Women's and Children's Health Network, North Adelaide, South Australia, Australia. [4]Adelaide Medical School and The Robinson Research Institute, The University of Adelaide, Adelaide, South Australia, Australia. [5]Vaccinology and Immunology Research Trials Unit, Women's and Children's Health Network, Adelaide, South Australia, Australia. [6]Wesfarmers Centre for Vaccines and Infectious Diseases, The Kids Institute, Perth, Western Australia, Australia. [7]School of Medicine, University of Western Australia, Perth, Western Australia, Australia. [8]School of Public Health, The University of Adelaide, Adelaide, South Australia, Australia. [9]Vaccine Immunology, Murdoch Children's Research Institute, Melbourne, Victoria, Australia. [10]Institute for Biomedicine and Glycomics, Griffith University, Southport, Queensland, Australia. [11]Department of Paediatrics, University of Melbourne, Melbourne, Victoria, Australia. [12]Centre for Cancer Biology, SA Pathology and University of South Australia, Adelaide, South Australia, Australia. [13]Department of Immunology and General Paediatrics, Perth Children's Hospital, Nedlands, Western Australia, Australia. [14]These authors contributed equally: Feargal J. Ryan, Michelle Clarke, Miriam A. Lynn. [15]These authors jointly supervised this work: Helen S. Marshall, David J. Lynn. ✉e-mail: david.lynn@sahmri.com

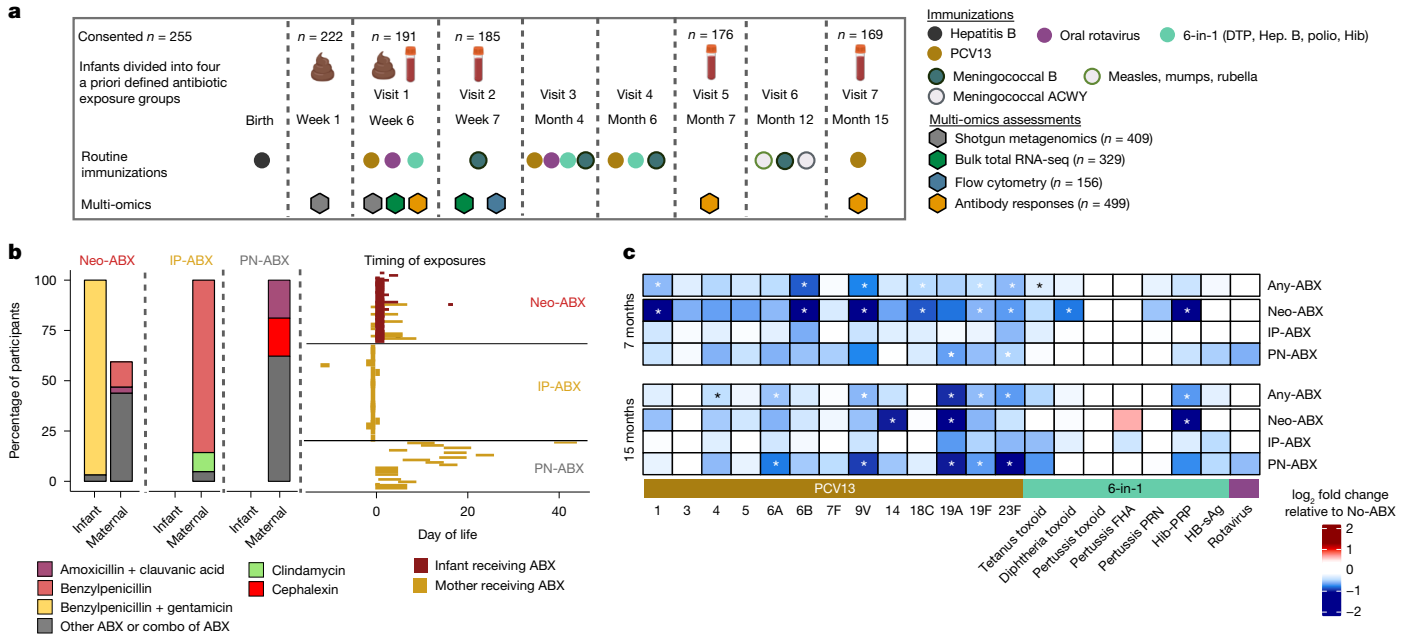


Fig. 1 | Vaccine-specific antibody responses are frequently impaired in infants directly exposed to antibiotics in the neonatal period. a, Overview of the AIR study design. Responses to the polio, measles, mumps, rubella and meningococcal ACWY vaccines were not measured in this study. **b**, Antibiotics administered to infants and their mothers in the direct neonatal antibiotic exposure group (Neo-ABX, $n = 32$ infants), intrapartum antibiotic exposure group (IP-ABX, $n = 49$ infants) and postnatal antibiotic exposure group (PN-ABX, $n = 30$ infants). **c**, Heatmap showing fold change in the geometric mean concentration (GMC) of vaccine antigen-specific antibodies in serum collected from Neo-ABX ($n = 27$ at 7 months, $n = 26$ at 15 months), IP-ABX ($n = 43$ at 7 months, $n = 42$ at 15 months) and PN-ABX ($n = 26$ at 7 months, $n = 27$ at 15 months) infants, relative to No-ABX infants ($n = 64$ at 7 months, $n = 65$ at 15 months). The blue–red gradient indicates decreased and increased fold change, respectively. Statistical significance in **c** was assessed using a generalized linear model adjusting for sex and baseline antibody titres. *$P < 0.05$ (unadjusted). All statistical tests were two-sided. Exact P values are provided in Supplementary Table 3a. Icons in **a** were created using BioRender (https://biorender.com). DTP, diphtheria, tetanus and pertussis; Hep., hepatitis.

exposed to antibiotics (ABX, $n = 111$) in the peripartum period (Fig. 1a and Extended Data Fig. 1). ABX infants were further subdivided (Fig. 1b) into three antibiotic exposure groups defined a priori (Methods). Thirty of 32 infants (94%) in the direct neonatal antibiotic exposure group (neo-ABX) started antibiotics in the 48 h after birth, and 31 (97%) of these infants received intravenous (i.v.) benzylpenicillin and gentamicin for at least 48 h. Of these infants, 19 were also exposed or potentially exposed to maternal antibiotics intrapartum or postnatally (Supplementary Table 1). In the intrapartum antibiotic exposure (IP-ABX) group, 36 of 49 (73%) of mothers received i.v. benzylpenicillin within 48 h before delivery, seven (14%) received i.v. benzylpenicillin in combination with at least one other oral or i.v. antibiotic, four (8%) received i.v. clindamycin, one received i.v. clindamycin plus oral cephalexin and—consistent with the a priori definition of this group, which included mothers treated with antibiotics in the last 28 days of pregnancy—one received oral cephalexin from day 12 to day 10 before birth. Mothers of infants in the postnatal antibiotic exposure group (PN-ABX, $n = 30$) received a variety of different antibiotics orally in the postnatal period from birth to day 45, and seven also received i.v. antibiotics (Supplementary Table 1). Maternal and infant demographics, as well as breastfeeding rates, were comparable across exposure groups, except for probiotic usage, which was higher in the ABX group (Extended Data Table 1).

Eighty-six per cent of infants in this study received hepatitis B vaccine at birth. At approximately week 6 of life, the first doses of the other primary childhood vaccines were administered as recommended by the Australian National Immunisation Program (NIP) schedule as of 2017. Further primary doses were administered at approximately 4 and 6 months of age (Fig. 1a). Serum IgG titres against 20 different vaccine antigens and anti-rotavirus vaccine IgA titres were measured at baseline and at 7 and 15 months (Extended Data Fig. 2a–u), using previously validated immunoassays[13–18]. To assess the relationship between ABX group and antibody responses to vaccination, we used multivariate regression adjusting for sex and baseline antibody titres; this was because baseline titres, as expected[19], were frequently negatively correlated with subsequent antibody titres (Extended Data Fig. 2v). Compared with No-ABX infants, infants exposed to any antibiotics before the first dose of the 13-valent pneumococcal conjugate vaccine (PCV13), 6-in-1 Infanrix Hexa and rotavirus vaccines had significantly ($P < 0.05$) decreased IgG titres against six of the capsular pneumococcal polysaccharides (PPS1, 6B, 9V, 18C, 19F and 23F) in the PCV13 vaccine and significantly lower anti-tetanus toxoid IgG titres at age 7 months (Fig. 1c and Supplementary Table 2a). After subdivision by exposure group, this decrease in anti-PPS IgG titres was evident in Neo-ABX infants (PPS1, 6B, 9V, 18C, 19F and 23F) but not IP-ABX infants (Fig. 1c, Extended Data Fig. 3a–f and Supplementary Table 3a). A similar trend towards reduced anti-PPS4 and PPS5 titres (but not those for other polysaccharides) was observed (Extended Data Fig. 3g–m). IgG titres against diphtheria toxoid and the *Haemophilus influenzae* type b (Hib) polyribosylribitol phosphate (PRP) antigen (antigens in the 6-in-1 Infanrix Hexa vaccine) were also significantly lower in Neo-ABX infants at 7 months (Fig. 1c and Extended Data Fig. 3n,o). A similar trend was observed for anti-tetanus toxoid IgG titres ($P = 0.09$) (Extended Data Fig. 3p). There was no detectable association of ABX group with IgG titres against pertussis toxoid (PT), hepatitis B surface antigen (HB-sAg), pertactin (PRN) or filamentous haemagglutinin (FHA), or with serum oral rotavirus vaccine IgA titres (Extended Data Fig. 3q–u). PN-ABX infants had significantly reduced anti-PPS19A and 23F IgG titres at 7 months (Fig. 1c), indicating that maternal postnatal antibiotic exposure may have an impact on subsequent infant responses to vaccination but to a lesser extent than direct administration of antibiotics to neonates.

Antibody titres were significantly lower at 15 months compared with 7 months for all of the antigens assessed, except anti-tetanus toxoid IgG titres, consistent with waning of immunity over time[20] (Extended Data Fig. 2a–u). Any antibiotics before the first dose of the PCV13,

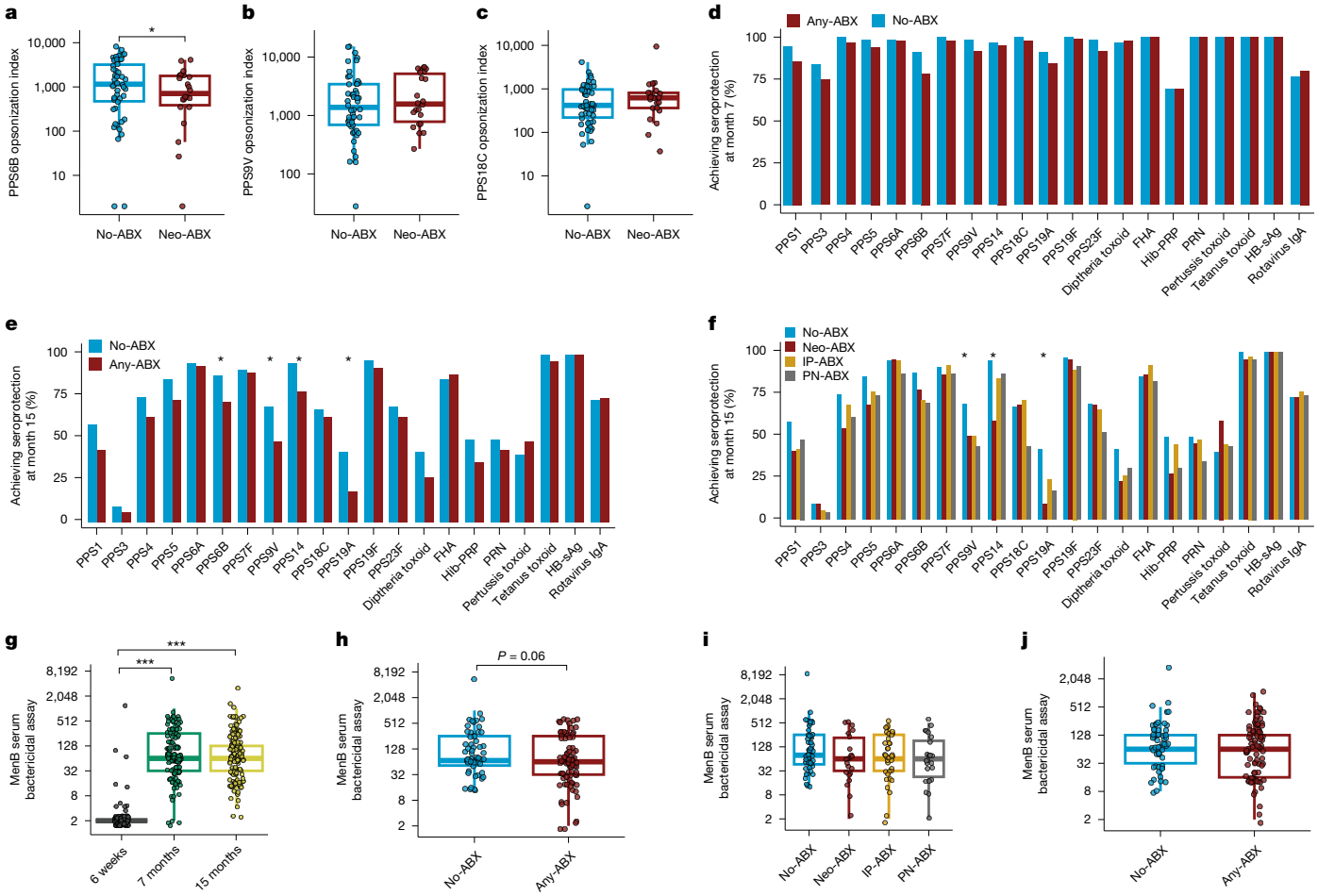


Fig. 2 | Antibiotic exposure in early life is associated with impaired antibody functionality and lower seroprotective responses against some vaccine antigens. a–**c**, Opsonophagocytic activity of serum collected from Neo-ABX ($n = 26$) and No-ABX ($n = 59$) infants at 7 months against PPS6B (**a**), PPS9V (**b**) and PPS18C (**c**). **d**,**e**, Proportions of No-ABX ($n = 55$) and Any-ABX ($n = 83$) infants achieving a seroprotective response against each vaccine antigen at 7 (**d**) and 15 (**e**) months. **f**, Proportions of No-ABX ($n = 55$), Neo-ABX ($n = 22$), IP-ABX ($n = 38$) and PN-ABX ($n = 23$) infants achieving a seroprotective response against each vaccine antigen at 15 months. **g**, Serum bactericidal activity against *N. meningitidis* NZ98/254 at week 7 (prevaccination, $n = 140$) and at 7 ($n = 139$) and 15 ($n = 138$) months. **h**, Serum bactericidal activity geometric mean titres (GMTs) subdivided by no or any antibiotic exposure at 7 months (No-ABX, $n = 57$; Any-ABX, $n = 82$). **i**, Serum bactericidal activity GMTs at 7 months subdivided by the indicated antibiotic exposure groups (No-ABX, $n = 57$; Neo-ABX, $n = 22$; IP-ABX, $n = 37$; PN-ABX, $n = 23$). **j**, Serum bactericidal activity GMTs subdivided by no or any antibiotic exposure at 15 months (No-ABX, $n = 56$; Any-ABX, $n = 82$). *n* represents the sample size of infants in each case. For the box plots in **a**–**c** and **g**–**j**, the box denotes the 25th and 75th percentiles, the whiskers the 5th and 95th percentiles, and the middle bar the median. Bar charts in **d**–**f** represent data as percentages. Statistical significance was assessed in **a**–**c** and **g**–**j** using a generalized linear model (Gaussian distribution) and in **d**–**f** using logistic regression, adjusting for sex. $^*P < 0.05$, $^{***}P < 0.001$ (unadjusted). All statistical tests were two-sided. Exact *P* values are provided in the source data.

6-in-1 and rotavirus vaccines was associated with significantly lower IgG titres at 15 months against six different serotypes in the PCV13 vaccine (PPS4, 6A, 9V, 19A, 19F and 23F) and Hib-PRP (Fig. 1c and Supplementary Table 2a). After subdivision by exposure group, significantly reduced IgG titres were observed only in Neo-ABX (PPS14, PPS19A, Hib-PRP) and PN-ABX (PPS6A, 9V, 19A, 19F, 23F) infants at this time point (Extended Data Fig. 4a–u and Supplementary Table 3a). There were no significant differences in IP-ABX infants. We also analysed the antibody data longitudinally (rather than separately at 7 and 15 months) and found that IgG titres against nine of 13 serotypes were significantly lower in Neo-ABX infants (Supplementary Table 3b). Taken together, our data indicate that antibiotic-exposed infants, particularly those exposed directly to antibiotics in the neonatal period, have impaired immunogenicity against several different vaccine antigens, particularly those in conjugate polysaccharide vaccines.

We next used opsonophagocytic assays to assess the functionality of antibodies against three PCV13 serotypes (6B, 9V and 18C) that had significantly reduced binding antibody titres in Neo-ABX infants. Compared with No-ABX infants, the opsonophagocytic activity of serum from Neo-ABX infants at 7 months was significantly lower ($P < 0.05$) against serotype 6B but not 9V or 18C (Fig. 2a–c). These data indicate that neonatal antibiotic exposure can not only impair vaccine-induced binding antibody titres months later but also, in some cases, reduce the functionality of those antibodies.

Reduced PCV13 seroprotective responses

Next, we assessed the association between ABX group and the proportion of infants with antibody responses above the seroprotective threshold defined for each vaccine antigen (Supplementary Table 2b). At 7 months, 1 month after the three primary doses of PCV13 and 6-in-1 vaccines had been administered, most of the infants (69–100%, dependent on the vaccine antigen) achieved a seroprotective response (Fig. 2d and Supplementary Table 2b), although there was a significantly reduced proportion of Neo-ABX infants above the seroprotective threshold for one PCV13 polysaccharide, PPS6B (Supplementary Table 3c).

By 15 months, the percentage of infants in any of the exposure groups maintaining antibody titres above the seroprotective thresholds had dropped considerably, below 50% for several antigens (Fig. 2e). Despite this waning, the proportion above the seroprotective threshold for infants exposed to any antibiotics was lower for all 13 PCV13 serotypes at 15 months compared with No-ABX infants. These reductions were statistically significant for four serotypes (PPS6B, 9V, 14 and 19A), and there was a trend for anti-diphtheria toxoid titres ($P = 0.06$). The proportions of Neo-ABX and PN-ABX infants above the seroprotective threshold were significantly lower ($P < 0.05$) for PPS14 and 19A, and PPS9V and 19A, respectively (Fig. 2f and Supplementary Table 3c). No significantly impaired seroprotective responses were observed in IP-ABX infants. These analyses used the seroprotective threshold accepted by the World Health Organization for PCV13 (0.35 μg ml^{-1}). Reanalysis of the data using serotype-specific seroprotective thresholds[21] showed the same trend (Supplementary Table 3d). These results indicate that there may be more rapid waning of vaccine-induced protection against certain pneumococcal serotypes in antibiotic-exposed infants, particularly those directly treated with antibiotics in the neonatal period.

Responses to meningococcal B vaccination

Infants in our study also received four doses of the four-component meningococcal B vaccine (4CMenB) (Fig. 1a). Sera from infants with matched samples at 7 weeks (pre-4CMenB vaccination), 7 months and 15 months were tested for bactericidal activity against *Neisseria meningitidis* NZ98/254, an internationally accepted surrogate of protection for meningococcal serogroup B vaccines, for which a titre greater than or equal to 1:4 is considered to be protective[22]. There was a significant increase in serum bactericidal activity titres postvaccination in the 7- and 15-month samples compared with baseline (Fig. 2g). Despite these very high seroprotective responses, we observed a trend ($P = 0.06$) towards reduced serum bactericidal activity titres at 7 months in infants exposed to any antibiotics (Fig. 2h). This was not observed when infants were divided further by antibiotic exposure group (Fig. 2i), possibly owing to reduced statistical power. No significant differences were observed at 15 months (Fig. 2j), around 3 months after the fourth dose of 4CMenB had been administered.

Antibiotics alter the infant transcriptome

In adults, transcriptional responses in blood at baseline or induced postvaccination have been shown to be correlated with subsequent vaccine-specific antibody responses[23]. Here we profiled the peripheral blood transcriptome in 329 infant blood samples collected at baseline and approximately 1 week after first-dose vaccination (Supplementary Table 4). Gene set enrichment analysis, a powerful approach that can be used to detect subtle but coordinated changes in gene expression[23–25], showed upregulation of several inflammation-related blood transcriptional modules (BTMs) in Neo-ABX infants at baseline (Fig. 3a and Supplementary Table 5). Few of these BTMs were found to be enriched among genes upregulated in IP-ABX infants, which may be functionally relevant, given that impaired vaccine antibody responses were observed in Neo-ABX infants and not in IP-ABX infants. Upregulation of inflammatory modules has also been previously reported in adults treated with a cocktail of antibiotics[11]. By contrast, B cell-related and plasma-cell-related modules were frequently downregulated in Neo-, IP- and PN-ABX infants. The activity of upregulated monocyte-related BTMs was negatively correlated with antibody responses at 7 months, whereas the activity of B cell-related BTMs was positively correlated (Extended Data Fig. 5a). Taken together, these data indicate that BTMs associated with antibody responses to vaccination are altered at baseline, particularly in Neo-ABX infants. Notably, several BTMs related to natural killer cells were enriched among genes downregulated in IP-ABX infants, indicating that although IP-ABX may not be associated with impaired humoral responses to vaccination, it may alter the immune system in early life in other ways.

Compared with the prevaccination baseline, there was a clear difference in peripheral blood transcriptomes approximately 1 week after the first dose of vaccines had been administered (Fig. 3b), with 3,089 differentially expressed genes (false discovery rate (FDR) < 0.05) identified postvaccination (Fig. 3c and Supplementary Table 6). These differences were driven by vaccination and not the 1-week difference in the age of the infants between the two time points (Extended Data Fig. 5b,c). BTMs upregulated postvaccination (Fig. 3d) included several that have been associated with response to vaccination in adults[23] (Extended Data Fig. 5d and Supplementary Table 7). BTMs related to B cells, monocytes and neutrophils were enriched among genes downregulated postvaccination. Broadly, the magnitude of the transcriptional response to vaccination was similar across the different ABX groups (Extended Data Fig. 5e,f). However, because Neo-ABX infants had altered transcriptional activity at baseline, these infants had notably different BTM activity postvaccination compared with that of No-ABX infants (Extended Data Fig. 5g and Supplementary Table 5).

Circulating immune cell populations

We performed flow cytometry analysis of more than 40 different immune cell populations in fresh peripheral blood samples collected concurrently with the week-7 blood samples used for RNA sequencing (Supplementary Table 8 and Supplementary Figs. 1–3). Although there was a modest reduction in the number of CD45RA$^-$ regulatory T cells in IP-ABX infants (Fig. 3e and Extended Data Fig. 5h,i), the numbers and frequencies of all other immune cell populations were not significantly different in any of the ABX groups compared with unexposed infants (Extended Data Fig. 5j–o and Supplementary Table 9). In addition, we explored whether any circulating immune cell populations differed between high and low vaccine responders, but we did not detect any significant differences (data not shown). These results indicate that the transcriptional differences observed in ABX infants are not primarily driven by differences in the frequency of major circulating immune cell populations.

Impact of antibiotics on microbiota

We collected stool specimens ($n = 409$) at weeks 1 and 6, with 89% of the week-6 samples being collected within 1 day of the first dose of the PCV13, 6-in-1 and oral rotavirus vaccines being administered. Shotgun metagenomic sequencing of these samples therefore enabled unique insight into the composition of the faecal microbiota in these infants at the time of vaccination (Supplementary Table 10). More than 12 million scaffolds were assigned into 3,774 metagenome-assembled genomes (MAGs), encompassing more than 300 species from 43 bacterial families. Actinomycetota and Pseudomonadota were the most abundant phyla at both time points, consistent with findings of previous studies[26]. Bacterial load did not significantly differ between ABX groups at either time point (Extended Data Fig. 6a). Alpha diversity (both evenness and richness) was significantly lower ($P \leq 0.05$) in Neo-ABX and IP-ABX infants at week 1 (Extended Data Fig. 6b,c). From week 1 to week 6, both bacterial load and alpha diversity significantly increased ($P < 0.001$) (Extended Data Fig. 6a–c). We also assessed antimicrobial-resistant-gene carriage and observed increased carriage in Neo-ABX and IP-ABX infants (Extended Data Fig. 6d,e).

Using Dirichlet multinomial mixtures modelling to cluster metagenomic samples, we identified six community state types (CSTs) (Fig. 4a,b). CST-2 had a high centred log ratio (CLR) abundance of *Bifidobacterium* species, particularly *Bifidobacterium breve* but also included *Bifidobacterium longum, Bifidobacterium bifidum* and *Bifidobacterium infantis*, and was significantly overrepresented among No-ABX infants at both week 1 and week 6 (Fig. 4c and Extended Data Fig. 6f). Notably, this was not observed for CST-3, which had a higher

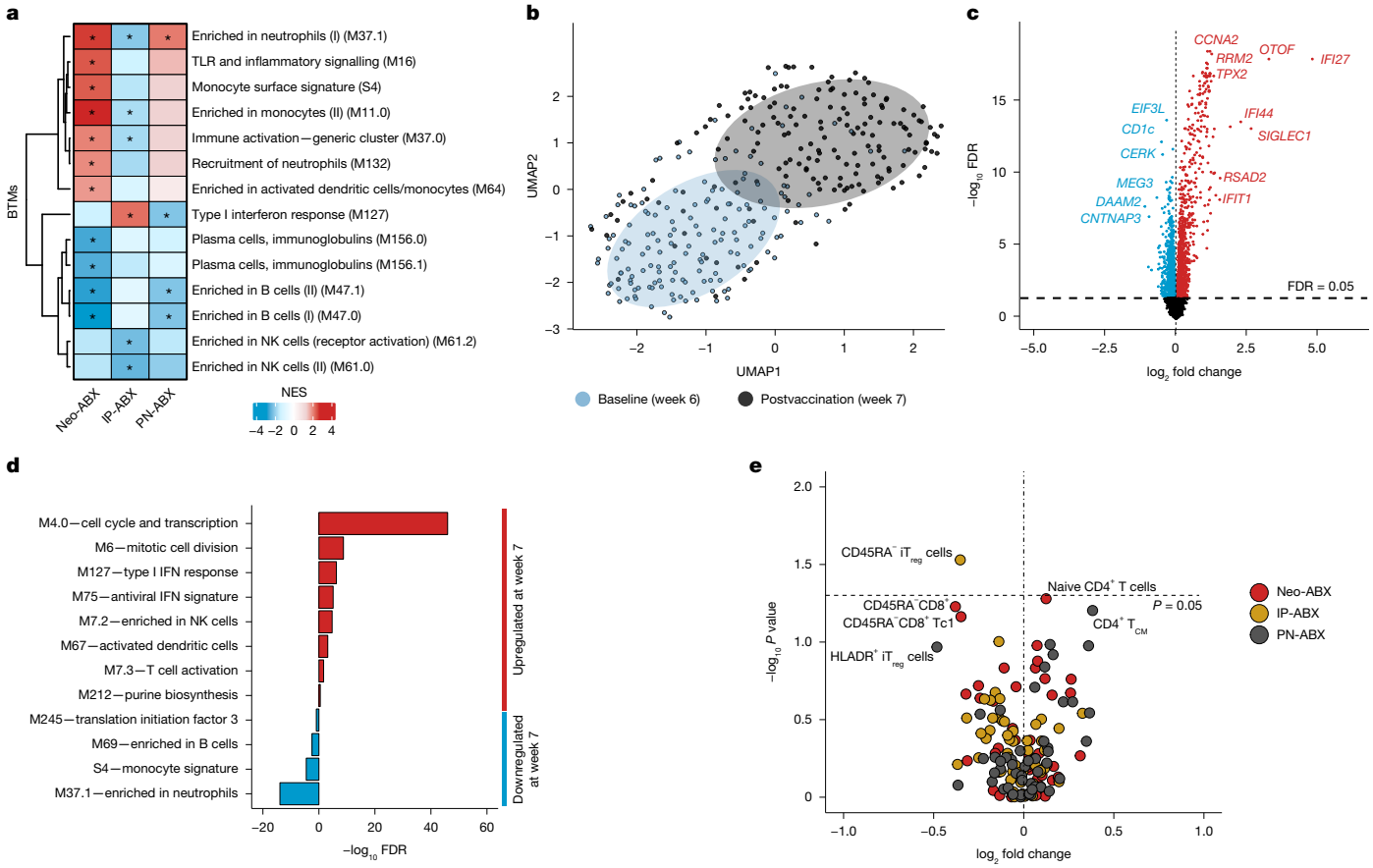


Fig. 3 | Antibiotic-exposed infants have altered blood transcriptional profiles pre- and postvaccination. a, Heatmap showing normalized enrichment scores (NES) for selected BTMs before vaccination in Neo-ABX ($n = 24$), IP-ABX ($n = 44$) and PN-ABX ($n = 25$) infants compared with No-ABX infants ($n = 66$). n represents the number of infants in each group prevaccination when RNA sequencing was performed. The blue–red gradient indicates decreased and increased BTM activity, respectively, relative to that of No-ABX infants. **b**, UMAP projection of whole-blood gene expression data pre- and postvaccination ($n = 329$ infant blood samples), adjusted for sex and batch using SVAseq[38]. **c**, Volcano plot of differentially expressed genes postvaccination. **d**, Selected BTMs that were statistically enriched among differentially expressed genes. **e**, Volcano plot showing fold change in circulating immune cell populations in Neo-ABX ($n = 27$), IP-ABX ($n = 40$) and PN-ABX ($n = 26$) infants relative to No-ABX ($n = 63$) infants. Statistical significance was assessed using the fgsea[39] R package in **a** and **d** and using the edgeR[40] package in **c**. The Benjamini–Hochberg method was used to adjust for multiple comparisons in **a**, **c** and **d**, with an FDR < 0.05 considered to indicate statistical significance. Statistical significance in **e** was assessed using two-tailed Wilcoxon signed-rank tests without adjustment for multiple comparisons. Asterisk indicates FDR < 0.05. Exact P values can be found in the supplementary tables or source data. NK, natural killer; T_{CM}, central memory T cell; T_{reg}, regulatory T cell.

CLR abundance of *B. longum* compared with CST-2. We next assessed whether specific bacterial taxa were differentially abundant among the ABX groups. Compared with No-ABX infants, numerous taxa, including several from the genera *Enterobacter*, *Klebsiella*, *Haemophilus, Citrobacter* and *Streptococcus*, had significantly higher (FDR < 0.05) CLR abundance in the IP-ABX infants at week 1 but not at week 6 (Fig. 4d,e, Extended Data Fig. 6g,h and Supplementary Table 11). Correspondingly, the CLR abundances of several different *Bifidobacterium* species were significantly lower (FDR < 0.05) in IP-ABX infants at week 1 but not at week 6 (Fig. 4d,e, Extended Data Fig. 6i and Supplementary Table 11). Consistent with the changes in the composition of the microbiota in IP-ABX infants, we detected 298 metagenomically encoded pathways that were differentially abundant (FDR < 0.05) at week 1 but none at week 6 (Supplementary Table 12). These data indicate that the effects of intrapartum antibiotics on the composition and functional capacity of the faecal microbiota were transient and had resolved by the time the first doses of vaccines were administered.

By contrast, the CLR abundances of various *Bifidobacterium* species including *B. breve, Bifidobacterium pseudocatenulatum* and *B. bifidum* (Fig. 4d,e and Extended Data Fig. 6j–l) and several *Collinsella* species were significantly decreased (FDR < 0.05) in Neo-ABX infants at week 6 (Supplementary Table 11), the time that the first doses of PCV13, 6-in-1 and oral rotavirus vaccines were administered. Reduced *Bifidobacterium* relative abundance has been reported in some but not all previous studies assessing the impact of antibiotics on the gut microbiota in early life[27]. Consistent with the persistent changes in the composition of the faecal microbiota of Neo-ABX infants at week 6, we identified 406 metagenomically encoded pathways that were differentially abundant (FDR < 0.05) in Neo-ABX infants at this time point compared with No-ABX infants (Supplementary Table 12). These data indicate that the persistent changes induced in the composition and encoded function of the gut microbiota at the time of vaccination by neonatal antibiotic exposure—in particular, reduced abundance of *Bifidobacterium* species—could play a part in the suboptimal antibody responses to vaccination observed in these infants. Consistent with this, the CLR abundances of several *Bifidobacterium* species at week 6 were positively correlated with IgG titres against several different vaccine antigens at 7 months (Fig. 4f and Supplementary Table 13a). Correlations between abundances of taxa at week 1 and antibody responses at month 7 were weaker, although there was positive correlation ($P < 0.05$) between the CLR abundance of *Bifidobacterium* species and anti-PPS6B IgG titres (Extended Data Fig. 7a and Supplementary Table 13b). Given these associations, we next assessed whether *Bifidobacterium* at 6 weeks was increased in high versus low vaccine responders. The CLR abundances of

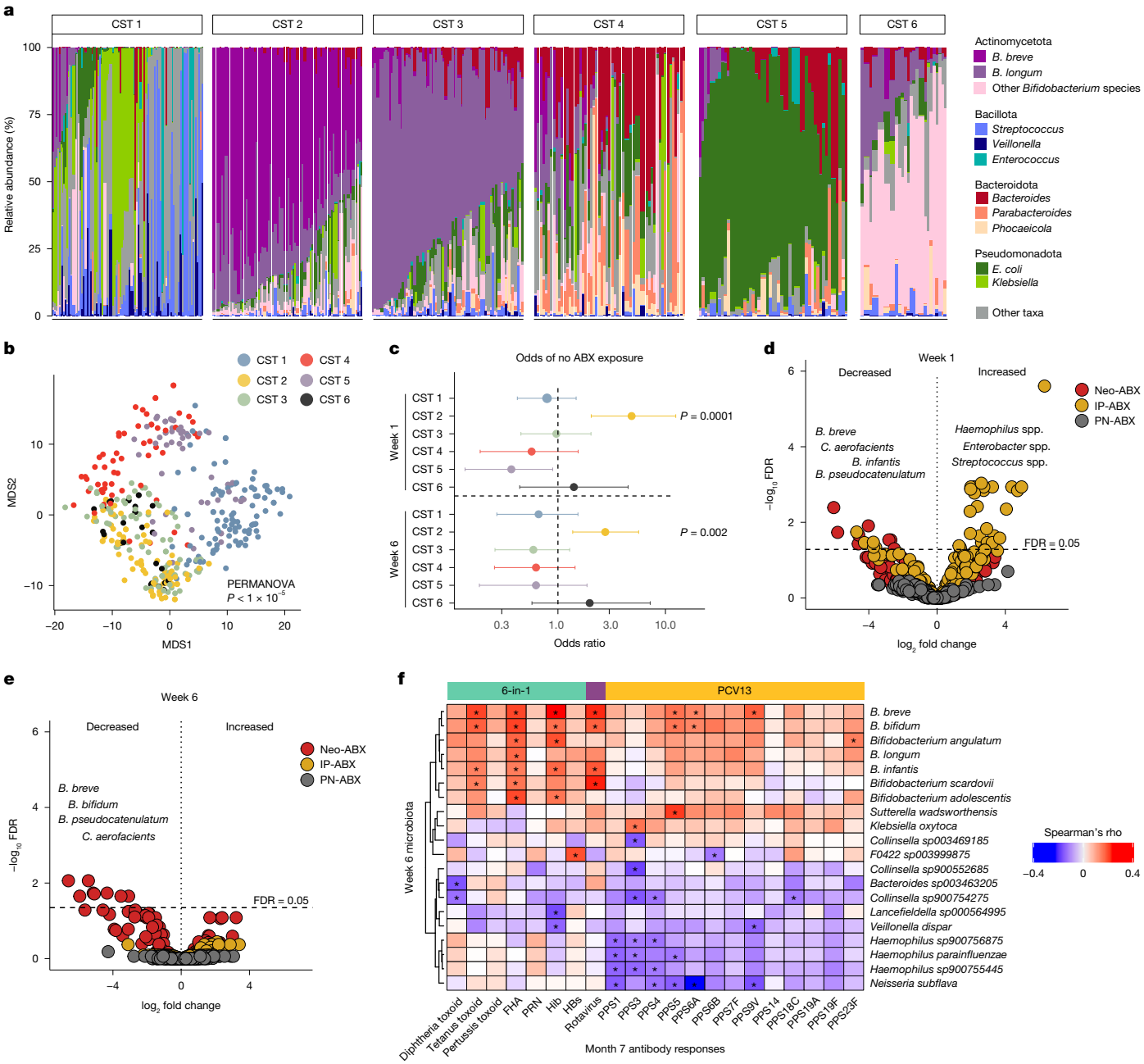


Fig. 4 | Depletion of *Bifidobacterium* species at the time of first-dose immunization is associated with impaired antibody responses to vaccination. a, Taxonomic composition of stool samples ($n = 409$) collected from study infants, as determined by shotgun metagenomics sequencing. Samples are grouped by CST. **b**, Multidimensional scaling (MDS) analysis of the six CSTs. **c**, Odds ratios of a sample ($n = 396$) in each CST at week 1 and week 6 of life being from an infant not exposed to antibiotics. **d**,**e**, Volcano plots showing fold change in the CLR abundance of taxa in the faecal microbiota of Neo-ABX, IP-ABX and PN-ABX infants, relative to No-ABX infants, at week 1 (No-ABX, $n = 79$; Neo-ABX, $n = 36$; IP-ABX, $n = 55$; PN-ABX $n = 33$) (**d**) and week 6 (No-ABX, $n = 80$; Neo-ABX, $n = 33$; IP-ABX, $n = 49$; PN-ABX, $n = 31$) (**e**). n represents the number of infants. **f**, Heatmap representing correlations between the CLR abundance of the top ten most positively correlated and top ten most negatively correlated taxa in the infant faecal microbiota at week 6 and antibody responses to vaccination (adjusted for sex and baseline titres) at month 7. The blue–red gradient represents negative and positive Spearman correlations, respectively ($n = 139$ with matched data across the three time points). See Supplementary Table 13 for exact correlation coefficients. Error bars in **c** indicate 95% confidence intervals. Statistical significance was assessed using permutational multivariate analysis of variance in **b**, Fisher's exact test in **c**, the LinDA method in the MicrobiomeState R package in **d** and **e**, and the cor.test R function in **f**. All statistical tests were two-sided. The Benjamini–Hochberg method was used to adjust for multiple comparisons in **d** and **e**, with FDR < 0.05 considered to indicate statistical significance. *$P < 0.05$ (unadjusted). Exact FDR values for **d** and **e** can be found in Supplementary Table 11.

B. breve, *B. bifidum* and *B. longum*—but not *B. pseudocatenulatum*—were significantly increased in high responders to several different vaccine antigens, particularly PPS6B (Extended Data Fig. 7b and Supplementary Table 13). Causal mediation analysis[28,29] also indicated significant positive direct relationships of both *B. breve* and *B. bifidum* with PPS6B IgG titres ($P < 0.05$) that were mediated by inflammation-related (BTM M11.0, $P < 0.05$) but not B cell-related transcriptional activity. These data indicate that antibiotic exposure may trigger a pro-inflammatory state that impedes *Bifidobacterium* colonization and/or growth and subsequently affects optimal vaccine responses.

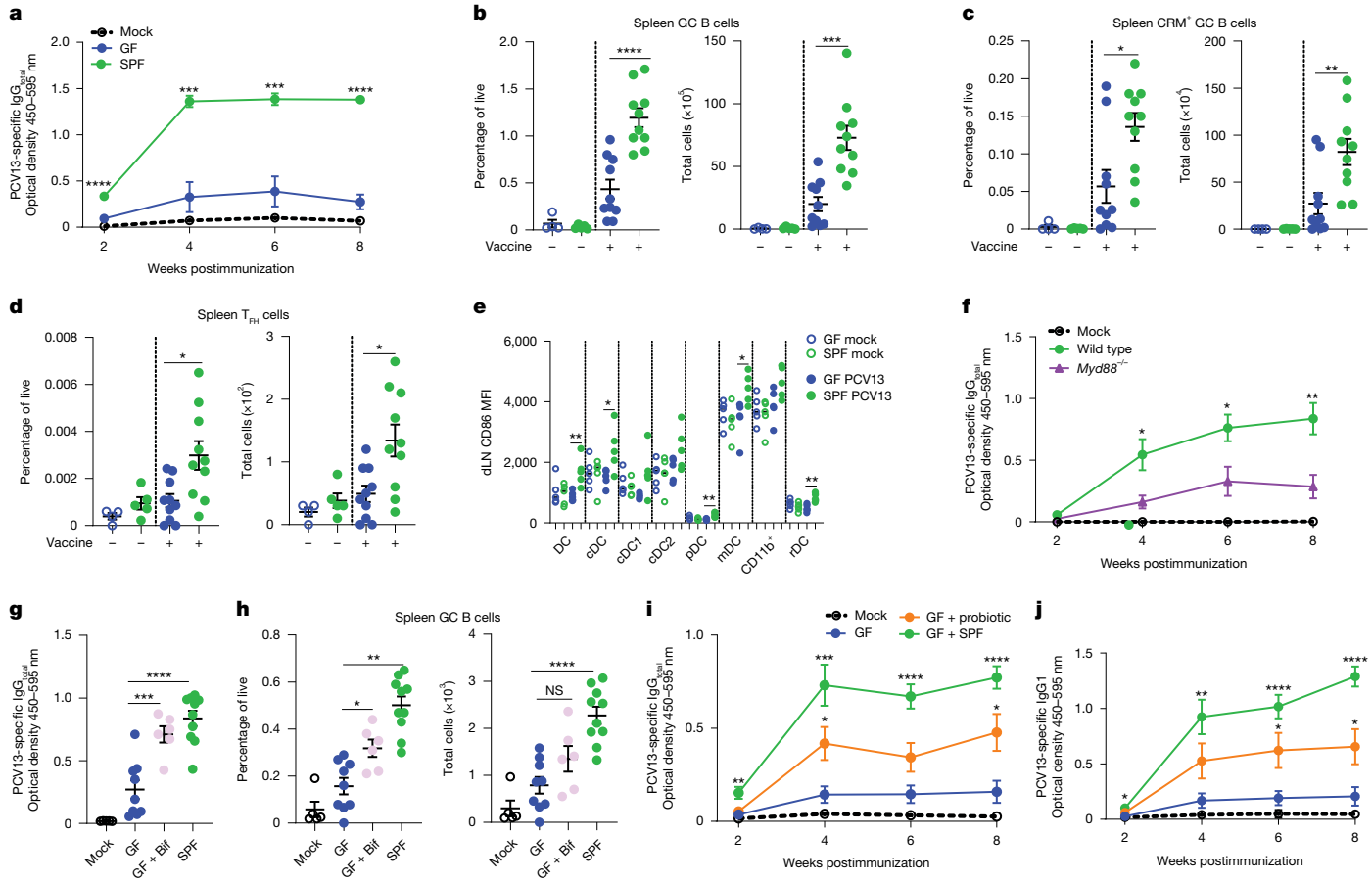


Fig. 5 | Immune responses to PCV13 in mice strongly depend on the gut microbiota and signalling through a MyD88-dependent pathway. a, PCV13-specific IgG_{total} in sera from GF ($n = 9$) and SPF ($n = 10$) mice vaccinated with two doses of PCV13. Mock, PBS-vaccinated mice ($n = 3$). **b**–**d**, Frequencies and total numbers of total GC B cells (**b**) or CRM_{197}^{+} GC B cells (**c**) in the spleens of mock-vaccinated mice ($n = 9$) and PCV13-vaccinated GF ($n = 10$) and SPF ($n = 10$) mice at 2 weeks postboost; and frequencies and total numbers of T_{FH} cells (**d**) in the spleens of these mice. **e**, CD86 mean fluorescence intensity (MFI) on indicated myeloid cell populations in the dLN of GF and SPF mice 24 h following PCV13 or mock vaccination ($n = 5$ mice per group). **f**, PCV13-specific IgG_{total} antibodies in the serum of mock-vaccinated mice ($n = 5$) and PCV13-vaccinated *Myd88*$^{-/-}$ ($n = 10$) and littermate wild-type ($n = 10$) mice. **g**, PCV13-specific IgG_{total} in the serum of mock-vaccinated mice ($n = 5$) and PCV13-vaccinated (two doses intraperitoneally, 2 weeks apart) SPF ($n = 10$) and GF ($n = 9$) mice, and GF mice born to dams colonized with a consortium of *Bifidobacterium* species in pregnancy and administered these strains again at days 7 and 14 postbirth (GF + Bif, $n = 6$). Data shown are at 2 weeks postboost. **h**, Frequencies and total numbers of GC B cells in the spleens of mice from **g** at 2 weeks postboost. **i**,**j**, PCV13-specific IgG_{total} (**i**) and IgG1 (**j**) in the sera of mock-vaccinated mice ($n = 3$), SPF mice ($n = 8$), GF mice ($n = 8$ mice), and GF mice colonized with probiotic strains *L. acidophilus* and *B. bifidum* at day 21 of life (GF + Probiotic, $n = 7$). Mice were vaccinated at day 28 of life. Data in **a**–**j** represent the mean ± s.e.m. Data in **a**–**d** are representative of three or more independent experiments; data in **e**–**j** are representative of two independent experiments. Statistical significance was assessed in **a**, **f**, **i** and **j** using two-way analysis of variance with Tukey's posttest analysis for multiple comparisons, in **b**–**e** using two-tailed Student's *t*-tests, and in **g** and **h** using one-way analysis of variance with Dunnett's posttest analysis for multiple comparisons. $^{*}P < 0.05$; $^{**}P < 0.01$; $^{***}P < 0.001$; $^{****}P < 0.001$; NS, not significant. Exact *P* values are provided in the source data.

Impaired PCV13 responses in GF mice

Our clinical data indicated that antibody responses to the PCV13 vaccine might be particularly dependent on signals from the gut microbiota. To assess whether this was also reflected in preclinical models, we compared antibody responses in sera collected from PCV13-vaccinated conventional specific-pathogen-free (SPF) and GF mice. PCV13-specific IgG, IgG1, IgG2c, IgG3 and IgM responses in sera were markedly impaired in GF mice (Fig. 5a and Extended Data Fig. 8a–d), as were IgG titres against PPS1, PPS3, PPS6B, PPS9V and PCV13 conjugate protein CRM_{197} (Extended Data Fig. 8e–i). GF mice immunized with the 6-in-1 vaccine also had significantly impaired anti-diphtheria toxoid and anti-tetanus toxoid IgG responses; however, analogous to our observations in Neo-ABX infants, their anti-pertussis toxoid responses were not significantly different to those of SPF mice (Extended Data Fig. 8j–l). GF mice also had antibody responses comparable with those of SPF mice following immunization with the 4CMenB vaccine (Extended Data Fig. 8m). We recently reported that GF mice have B and T cell responses to the Pfizer/BioNTech COVID-19 mRNA vaccine that are comparable with those of SPF mice[10]. Taken together, these data indicate that as in antibiotic-treated infants, responses to PCV13 in mice are particularly dependent on signals from the gut microbiota.

To investigate why, we next assessed whether germinal centre (GC) responses, which have a critical role in optimal T-dependent humoral immune responses, were suboptimal in PCV13-immunized GF mice. Compared with SPF mice, GF mice immunized with two doses of PCV13 had significantly lower total and CRM_{197}-specific GC B cells in their spleens (Fig. 5b,c) and draining lymph nodes (dLN) (Extended Data Fig. 8n–q). The proportion of T follicular helper (T_{FH}) cells in the spleen (Fig. 5d) but not in the dLN (Extended Data Fig. 8r,s) was also significantly lower in GF mice. Confocal microscopy confirmed impaired GC formation in PCV13-immunized GF mice (Extended Data Fig. 9). These data indicate that PCV13-immunized GF mice have impaired

T-dependent antibody responses to vaccination because of a failure of optimal GC formation in these mice.

We next proposed that GF mice might fail to mount optimal GC responses to PCV13 because of impaired innate immune activation postvaccination. Consistent with this hypothesis, 24 h after the first dose of PCV13, we observed significant increases in surface expression of activation marker CD86 on cells of several different myeloid populations, including dendritic cells (DC), conventional dendritic cells (cDC), plasmacytoid dendritic cells (pDC), migratory dendritic cells (mDC) and resident dendritic cells (rDC) in the dLN of SPF but not GF mice (Fig. 5e). Expression of MHCII or CD80 was not induced in either GF or SPF mice at 24 h postvaccination (Extended Data Fig. 10a,b). Notably, given the associations we observed between increased CLR abundance of *Bifidobacterium* species and higher antibody titres in infants, mice born to GF dams colonized with *B. breve* during pregnancy (to ensure colonization from birth) had significantly higher CD86 expression at 24 h after PCV13 vaccination compared with mice that remained GF (Extended Data Fig. 10c).

Given that GF mice seemed to have impaired innate immune responses to PCV13 immunization, we proposed a further hypothesis that signals from the microbiota might act through an MyD88-dependent pathway. Consistent with this, serum IgG, IgG1, IgG2c and IgM responses to PCV13 were severely impaired in *Myd88*$^{-/-}$ mice (Fig. 5f and Extended Data Fig. 10d–f). A previous study found that TLR5-mediated sensing of the microbiota was required for antibody responses to a non-adjuvanted influenza vaccine[9]; however, we have previously shown that antibody responses to the alum-adjuvanted PCV13 vaccine are not impaired in *Tlr5*$^{-/-}$ mice[1]. Given this, we instead investigated responses to PCV13 in *Tlr4*$^{-/-}$ and *Tlr2*$^{-/-}$ mice and found a trend suggestive of impaired PCV13-specific antibody responses in *Tlr4*$^{-/-}$ but not *Tlr2*$^{-/-}$ mice (Extended Data Fig. 10g,h). These data indicate that sensing of the microbiota through an MyD88-dependent pathways may be required for optimal antibody responses to the PCV13 vaccine; however, further work is needed to determine the upstream pattern recognition receptor(s) involved, noting that redundancy between different receptors is possible.

Bifidobacteria enhance PCV13 responses

Our clinical data suggested that loss of *Bifidobacterium* species in infants at the time of immunization leads to impaired antibody responses to PCV13 and other vaccines. To assess whether administering a consortium of *Bifidobacterium* species to GF mice could restore PCV13-specific antibody responses, we administered a mixture of *B. breve*, *B. longum* subsp. *longum* and *B. longum* subsp. *infantis* to pregnant dams to ensure colonization of their offspring from birth. These pups (GF + Bif) and pups born to GF or SPF mice were vaccinated with PCV13 at around day 21 of life and again 2 weeks later. Consistent with our hypothesis, PCV13-specific IgG, IgG1 and IgM antibody responses in GF + Bif pups were significantly higher 2 weeks following the second dose of PCV13 compared with those in GF mice (Fig. 5g and Extended Data Fig. 10i,j). Compared with GF mice, GF + Bif mice also had significantly higher proportions or total numbers of GC B cells in their spleens (Fig. 5h) and dLN (Extended Data Fig. 10k). Proportions or numbers of T_{FH} cells were also higher in GF + Bif mice (Extended Data Fig. 10l,m). Notably, we observed no increase in antibody responses to PCV13 in GF + Bif mice after only a single dose of vaccine (Extended Data Fig. 10n–p).

Next, we wanted to establish whether responses in GF mice to PCV13 could be restored by directly administering a probiotic intervention to the GF mice closer to the time of vaccination (rather than through their mothers). To assess this, we compared responses to PCV13 in SPF mice, GF mice and GF mice colonized with strains from commercial probiotic Infloran, 1 week before vaccination with PCV13 (GF + Probiotic). We selected Infloran as it contains a mixture of *B. bifidum* and *Lactobacillus acidophilus* and also because it is commonly used in neonatal intensive care units to prevent necrotizing enterocolitis[30,31] and thus has an established clinical safety profile with clear translational potential. In addition, *B. bifidum* was among the *Bifidobacterium* species that we observed to show decreased abundance in Neo-ABX infants and increased abundance in infants that were high vaccine responders (Extended Data Fig. 7b). PCV13-specific IgG and IgG1 responses in GF mice were significantly enhanced by colonization of the mice with the Infloran probiotic strains (Fig. 5i,j). Notably, colonization of GF mice with either strain alone was insufficient to enhance antibody responses to PCV13 (Extended Data Fig. 10q–s). *Bifidobacterium* spp. are known to use a variety of cross-feeding synergistic strategies involving other members of the gut microbiota to access carbohydrates that they cannot metabolise;[32] this potentially explains why neither culture alone enhanced antibody responses. Consistent with this, the bacterial load in mice colonized with both the *B. bifidum* and *L. acidophilus* strains simultaneously was two log-fold higher than that of mice colonized with either strain alone (Extended Data Fig. 10t). Colonizing GF mice with a range of other commensal isolates before PCV13 vaccination did not enhance responses to PCV13 vaccination (Extended Data Fig. 10u–w). In conclusion, our data indicate that restoring a healthy *Bifidobacterium*-rich microbiota in antibiotic-exposed infants, before vaccination, may lead to enhanced antibody responses to vaccination and increased protection against infectious diseases.

Discussion

We have previously shown that early-life antibiotic exposure in mice leads to impaired responses to a range of different infant vaccines[1]; this motivated us to conduct the prospective clinical study we report here. Consistent with these previous preclinical data, we found that infants exposed to antibiotics, particularly direct antibiotic treatment in the neonatal period, had significantly reduced serum IgG titres against several vaccine antigens at 7 and 15 months of age. Our results, which add to the findings of several important previous studies (Supplementary Discussion), indicate that immune responses to certain types of vaccine, particularly protein-polysaccharide conjugate vaccines such as PCV13, may be more dependent on signals from the gut microbiota than others. Further investigation is needed to determine why this is the case; however, it may be related to the lower immunogenicity of polysaccharide conjugate vaccines compared with adjuvanted protein vaccines and other platforms such as mRNA vaccines[10].

Our clinical and preclinical data suggest that *Bifidobacterium* species in the infant gut microbiota have an important role in mediating optimal innate, GC and serum antibody responses to PCV13 and other vaccines and that administering probiotics to infants exposed to neonatal antibiotics may be a feasible, cheap and safe intervention to enhance responses to vaccination. This needs to be rigorously tested in randomized controlled trials in appropriate populations (Supplementary Discussion). Consistent with our data, some previous correlative studies have reported an association between the relative abundance of *Bifidobacterium* species and responses to the Bacillus Calmette–Guérin (BCG), oral polio and tetanus toxoid vaccines[33,34]. In addition, GF mice colonized with *B. longum* subp. infantis have recently been found to have increased T cell responses to the BCG vaccine[35]. Notably, intrapartum antibiotic exposure was not associated with reduced antibody responses to vaccination in our study (Supplementary Discussion). A possible explanation is that at the time of first-dose vaccination, abundances of *Bifidobacterium* species were not significantly reduced in infants exposed to intrapartum antibiotics, whereas they were significantly reduced in infants who were directly treated with antibiotics.

Despite our important findings, our study had several limitations. The sample size was relatively modest, especially when the antibiotic exposure group was broken down further according to the type of antibiotic exposure. Further studies in larger cohorts are now warranted. Encouragingly, our results in infants were strongly concordant with our

observations in animal models. Our study excluded infants born by caesarean section and those born to mothers with a body mass index greater than 30. Future research should investigate whether these and other factors that alter the infant gut microbiota[36] can also influence immune responses to vaccination. A recent study reported that infants born by caesarean section had significantly reduced IgG titres against pneumococcal capsular polysaccharides in saliva (no serum was collected in that study)[37]. Furthermore, assessment of gene expression responses at more time points postvaccination and using single-cell technologies could further refine our understanding of the link between altered transcriptional responses in infants exposed to antibiotics and suboptimal responses to vaccination (Supplementary Discussion). Finally, our study focused predominantly on binding antibody titres, given that these are the established correlates of protection for most vaccines; however, we found that the opsonophagocytic activity of serum from infants treated with antibiotics in the neonatal period against one of the three PCV13 polysaccharides, PPS6B, was significantly reduced. Furthermore, there was a modest reduction in the bactericidal activity (against *N. meningitidis*) of serum collected from infants vaccinated with the 4CMenB meningococcal B vaccine that had been exposed to any antibiotics, compared with those that had not been exposed to antibiotics. Future studies should also assess the impact of antibiotic exposure on functional antibody responses in more detail, as well as T cell responses to vaccination. In conclusion, we found that early-life antibiotic exposure was associated with significantly impaired vaccine immunogenicity, with a particular impact on immune responses to the PCV13 vaccine. These effects seemed to be mediated by a reduction in abundance of *Bifidobacterium* species, which in turn led to altered baseline immune status. Our findings emphasize the need to broaden the rationale for antibiotic stewardship in early life beyond antimicrobial resistance, as well as the potential of precision vaccination approaches to optimize vaccine-mediated protection from infectious diseases.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-025-08796-4.

1. Lynn, M. A. et al. Early-life antibiotic-driven dysbiosis leads to dysregulated vaccine immune responses in mice. *Cell Host Microbe* **23**, 653–660 (2018).
2. Chapman, T. J., Pham, M., Bajorski, P. & Pichichero, M. E. Antibiotic use and vaccine antibody levels. *Pediatrics* **149**, e2021052061 (2022).
3. Shaffer, M. et al. Very early life microbiome and metabolome correlates with primary vaccination variability in children. *mSystems* **8**, e0066123 (2023).
4. Pichichero, M. E., Xu, L., Gonzalez, E., Pham, M. & Kaur, R. Variability of vaccine responsiveness in young children. *J. Infect. Dis.* **229**, 1856–1865 (2024).
5. Andre, F. E. et al. Vaccination greatly reduces disease, disability, death and inequity worldwide. *Bull. World Health Organ.* **86**, 140–146 (2008).
6. Watson, O. J. et al. Global impact of the first year of COVID-19 vaccination: a mathematical modelling study. *Lancet Infect. Dis.* **22**, 1293–1302 (2022).
7. Lynn, D. J., Benson, S. C., Lynn, M. A. & Pulendran, B. Modulation of immune responses to vaccination by the microbiota: implications and potential mechanisms. *Nat. Rev. Immunol.* **22**, 33–46 (2022).
8. Rossouw, C., Ryan, F. J. & Lynn, D. J. The role of the gut microbiota in regulating responses to vaccination: current knowledge and future directions. *FEBS J.* https://doi.org/10.1111/febs.17241 (2024).
9. Oh, J. Z. et al. TLR5-mediated sensing of gut microbiota is necessary for antibody responses to seasonal influenza vaccination. *Immunity* **41**, 478–492 (2014).
10. Norton, T. et al. B and T cell responses to the BNT162b2 COVID-19 mRNA vaccine are not impaired in germ-free or antibiotic-treated mice. *Gut* **73**, 1222–1224 (2024).
11. Hagan, T. et al. Antibiotics-driven gut microbiome perturbation alters immunity to vaccines in humans. *Cell* **178**, 1313–1328 (2019).
12. Harris, V. C. et al. Effect of antibiotic-mediated microbiome modulation on rotavirus vaccine immunogenicity: a human, randomized-control proof-of-concept trial. *Cell Host Microbe* **24**, 197–207.e194 (2018).
13. Elberse, K. E., Tcherniaeva, I., Berbers, G. A. & Schouls, L. M. Optimization and application of a multiplex bead-based assay to quantify serotype-specific IgG against *Streptococcus pneumoniae* polysaccharides: response to the booster vaccine after immunization with the pneumococcal 7-valent conjugate vaccine. *Clin. Vaccine Immunol.* **17**, 674–682 (2010).
14. McAlister, S. M., van den Biggelaar, A. H. J., Thornton, R. B. & Richmond, P. C. Optimising a 6-plex tetanus-diphtheria-pertussis fluorescent bead-based immunoassay. *MethodsX* **8**, 101360 (2021).
15. Pavliakova, D. et al. Development and validation of 13-plex Luminex-based assay for measuring human serum antibodies to *Streptococcus pneumoniae* capsular polysaccharides. *mSphere* **3**, e00128-18 (2018).
16. Pickering, J. W. et al. A multiplexed fluorescent microsphere immunoassay for antibodies to pneumococcal capsular polysaccharides. *Am. J. Clin. Pathol.* **117**, 589–596 (2002).
17. Bernstein, D. I. et al. Safety and immunogenicity of live, attenuated human rotavirus vaccine 89-12. *Vaccine* **16**, 381–387 (1998).
18. Lee, B. et al. The effect of increased inoculum on oral rotavirus vaccine take among infants in Dhaka, Bangladesh: a double-blind, parallel group, randomized, controlled trial. *Vaccine* **38**, 90–99 (2020).
19. Vono, M. et al. Maternal antibodies inhibit neonatal and infant responses to vaccination by shaping the early-life B cell repertoire within germinal centers. *Cell Rep.* **28**, 1773–1784.e1775 (2019).
20. Swarthout, T. D. et al. Waning of antibody levels induced by a 13-valent pneumococcal conjugate vaccine, using a 3+0 schedule, within the first year of life among children younger than 5 years in Blantyre, Malawi: an observational, population-level, serosurveillance study. *Lancet Infect. Dis.* **22**, 1737–1747 (2022).
21. Andrews, N. J. et al. Serotype-specific effectiveness and correlates of protection for the 13-valent pneumococcal conjugate vaccine: a postlicensure indirect cohort study. *Lancet Infect. Dis.* **14**, 839–846 (2014).
22. Findlow, J., Lucidarme, J., Taha, M. K., Burman, C. & Balmer, P. Correlates of protection for meningococcal surface protein vaccines: lessons from the past. *Expert Rev. Vaccines* **21**, 739–751 (2022).
23. Hagan, T. et al. Transcriptional atlas of the human immune response to 13 vaccines reveals a common predictor of vaccine-induced antibody responses. *Nat. Immunol.* **23**, 1788–1798 (2022).
24. Li, S. et al. Molecular signatures of antibody responses derived from a systems biology study of five human vaccines. *Nat. Immunol.* **15**, 195–204 (2014).
25. Ryan, F. J. et al. A systems immunology study comparing innate and adaptive immune responses in adults to COVID-19 mRNA and adenovirus vectored vaccines. *Cell Rep. Med.* **4**, 100971 (2023).
26. Derrien, M., Alvarez, A. S. & de Vos, W. M. The gut microbiota in the first decade of life. *Trends Microbiol.* **27**, 997–1010 (2019).
27. McDonnell, L. et al. Association between antibiotics and gut microbiome dysbiosis in children: systematic review and meta-analysis. *Gut Microbes* **13**, 1870402 (2021).
28. Tingley, D., Yamamoto, T., Hirose, K., Keele, L. & Imai, K. Mediation: R package for causal mediation analysis. *J. Stat. Softw.* **59**, 1–38 (2014).
29. Imai, K., Keele, L., Tingley, D. & Yamamoto, T. in *Advances in Social Science Research Using R*. (ed. Vinod, H. D.) Ch. 8 (Springer, 2010).
30. Hoyos, A. B. Reduced incidence of necrotizing enterocolitis associated with enteral administration of *Lactobacillus acidophilus* and *Bifidobacterium infantis* to neonates in an intensive care unit. *Int. J. Infect. Dis.* **3**, 197–202 (1999).
31. Cripps, E. K., Dargaville, P. A. & De Paoli, A. G. Impact of probiotic administration on the incidence of necrotising enterocolitis: a single-centre cohort study. *J. Paediatr. Child Health* **59**, 760–765 (2023).
32. Turroni, F. et al. Glycan utilization and cross-feeding activities by Bifidobacteria. *Trends Microbiol.* **26**, 339–350 (2018).
33. Huda, M. N. et al. *Bifidobacterium* abundance in early infancy and vaccine response at 2 years of age. *Pediatrics* **143**, e20181489 (2019).
34. Huda, M. N. et al. Stool microbiota and vaccine responses of infants. *Pediatrics* **134**, e362–e372 (2014).
35. Nyangahu, D. D. et al. Bifidobacterium infantis associates with T cell immunity in human infants and is sufficient to enhance antigen-specific T cells in mice. *Sci. Adv.* **9**, eade1370 (2023).
36. Lundgren, S. N. et al. Maternal diet during pregnancy is related with the infant stool microbiome in a delivery mode-dependent manner. *Microbiome* **6**, 109 (2018).
37. de Koff, E. M. et al. Mode of delivery modulates the intestinal microbiota and impacts the response to vaccination. *Nat. Commun.* **13**, 6638 (2022).
38. Leek, J. T. svaseq: removing batch effects and other unwanted noise from sequencing data. *Nucleic Acids Res.* **42**, e161 (2014).
39. Korotkevich, G. et al. Fast gene set enrichment analysis. Preprint at *bioRxiv* https://doi.org/10.1101/060012 (2016).
40. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).

Methods

Study design

Infants were enrolled in the study at the Women's and Children's Hospital Adelaide, Australia, between April 2017 and March 2021 as part of the Antibiotics and Immune Responses (AIR) study, a prospective, observational clinical study (ACTRN12617000856314). Potential participants were approached by study research nurses in postnatal wards and invited to participate if the mother and their infant met the study eligibility criteria. Study information had also been previously provided to women delivering at the Women's and Children's Hospital through pamphlets in antenatal care packs. Following written informed consent, stool collection packs were provided to families, and medical case notes were reviewed to confirm eligibility criteria and antibiotic exposure status. Ethics approval for the study was obtained from the Human Research Ethics Committee of the Women's and Children's Health Network (approval number HREC/17/WCHN/19), and the authors affirm that they have complied with all relevant ethical regulations.

Exclusion criteria were gestation of less than 37 weeks; maternal body mass index greater than 30 (at first-trimester antenatal visit); maternal sepsis (defined by laboratory-confirmed bacterial infection in blood cultures or cerebrospinal fluid) during pregnancy; infant delivered by caesarean section; confirmed sepsis in infant (defined by laboratory-confirmed bacterial infection in blood cultures or cerebrospinal fluid); known or suspected disorder of the immune system that would prevent an immune response to the vaccines, such as congenital or acquired immunodeficiency, or receiving systemic immunosuppressive therapy; infant with suspected or confirmed HIV, major congenital abnormality or serious illness; maternal or infant participation in a clinical study that could interfere with participation in this study; and anything that would place the individual at increased risk or preclude their full compliance with or completion of the study. For immunogenicity analyses, only participants who received at least one vaccine at study visit 1 were included.

Participants were allocated to an antibiotic exposure group for immunogenicity analyses on the basis of medical record review and discussion with the mother to ascertain any antibiotic exposure before vaccination at study visit 1 (approximately 6 weeks of age). Antibiotic exposure groups were defined a priori as follows: infants with no direct or maternal antibiotic exposure ($n = 80$, No-ABX); (1) infants who received at least 48 h of antibiotic treatment in the neonatal period (the first 28 days after birth), with or without maternal antibiotic exposure ($n = 32$, Neo-ABX); (2) infants whose mothers received intrapartum antibiotics (within 28 days before or during delivery) and no direct infant antibiotic exposure up to 6 weeks of age ($n = 49$, IP-ABX); and (3) infants who did not meet the criteria for inclusion in the No-ABX, Neo-ABX or IP-ABX group. Upon study completion, all infants in group 3 were those who could have been exposed to maternal antibiotics during breastfeeding in the postnatal prevaccination period, so we called this group the postnatal exposure group (PN-ABX, $n = 30$). Researchers conducting the serological assays, omics analyses and other data analyses were blinded to antibiotic exposure group allocations. Participant data were recorded in a secure REDCap (https://project-redcap.org/) database.

Sample collection

Faecal samples for metagenomics were collected from the nappies (diapers) of each infant by parents at week 1 (days 2–10 of life) and week 6 (days 38–56 of life) and placed into Zymo DNA/RNA Shield faecal collection tubes (Zymo Research) before being stored at −80 °C. Blood (3–5 ml) was obtained from infants at approximately 6 weeks, 7 weeks, 7 months and 15 months. At 6 weeks, 1–3 ml of whole blood was placed into BD Vacutainer serum tubes, and 2 ml was placed in a PAXgene blood RNA tube (BD) for RNA preservation and stored at −80 °C. At 7 weeks, up to 1 ml of blood was collected in a BD Vacutainer K3 EDTA tube, 2 ml was collected for flow cytometry analysis, and the remainder (up to 2 ml) was placed into a PAXgene (BD) tube. At 7 months and 15 months, 3–5 ml of blood was collected into BD Vacutainer serum tubes.

Vaccine schedule

Infants were vaccinated with PCV13 (Prevenar 13, Pfizer) and the combination 6-in-1 vaccine (Infanrix Hexa, GSK) at 6 weeks and 4 and 6 months of age and with the oral rotavirus vaccine (Rotarix, GSK) at 6 weeks and 4 months of age as per the NIP Schedule for South Australia as of 2017. Most of the infants (86%) were also vaccinated with the hepatitis B vaccine at birth (either H-B-Vax II Paediatric or Engerix B Paediatric). Infants in this study were also vaccinated with the meningococcal B vaccine (4CMenB/Bexsero, GSK) at 7 weeks, 4 months and 6 months of age, with a booster dose at 12 months of age. The meningococcal ACWY (Nimenrix, Pfizer) and the measles, mumps, rubella (M-M-R II, Merck Sharp & Dohme; or Priorix, GSK) vaccines were administered at 12 months of age as part of the standard NIP schedule (although antibody responses were not evaluated for these vaccines in this study). In this study, the booster dose of PCV13 was provided at 15 months rather than the current NIP scheduled time point of 12 months of age. All vaccines were administered intramuscularly except for the oral rotavirus vaccine. All vaccines were administered in line with Australian Immunisation Handbook recommendations.

Pneumococcal polysaccharide-specific IgG

Antibodies against PCV13 serotypes were measured as previously described[13,15,16], with modifications. In brief, Bio-Plex COOH-microspheres (6.25×10^6, 500 µl, Bio-Rad) were conjugated overnight with optimized doses of DMTMM-modified pneumococcal polysaccharides (ATCC) in 1× PBS pH 7.2 (Life Technologies) as per Schlottmann et al.[41]. The optimal coating doses were a 1:1 dose, 500 µl microspheres with 500 µl DMTMM-modified polysaccharides for serotypes 1, 3, 5, 7F and 9V; and a 1:0.5 dose, 500 µl microspheres with 250 µl DMTMM-modified polysaccharides for serotypes 4, 6A, 6B, 14, 18C, 19A, 19F, 23F and 11A.

A ten-step, three-fold serial dilution of the international reference standard for human antipneumococcal capsule, 007sp (NIBSC), two in-house quality controls and test samples were diluted in adsorbent buffer containing 0.05% Tween-20, 2% *N*-butyl benzenesulfonamide (Sigma-Aldrich) and 2.5 µg ml^{-1} CWPS Multi (SSI Diagnostica) for 30 min to remove non-specific antibodies before incubation with polysaccharide-conjugated microspheres. Equal volumes (25 µl) of diluted samples and bead-mix containing 4,000 microspheres per region were combined on a MultiScreen filter plate (Merck) prewetted with 50 µl of 0.05% PBS with Tween-20 (PBS-T). Plates were incubated at room temperature in the dark on an orbital shaker at 500 rpm for 20 min. They were then washed twice with 100 µl 0.05% PBS-T, followed by addition of 100 µl (1:150) RPE-conjugated goat anti-human IgG Fc secondary antibody (Jackson ImmunoResearch Laboratories Inc.) and incubation for a further 20 min. Following washing, the beads were resuspended in 125 µl 0.05% PBS-T and read on a Bio-Plex-200 machine (Bio-Rad).

Diphtheria, tetanus and pertussis-specific IgG

Antibody concentrations specific to pertussis toxoid, pertactin, filamentous haemagglutinin, tetanus toxoid and diphtheria toxoid were measured as previously described[14,42].

Hib-PRP- and HB-sAg-specific IgG

Antibodies against Hib-PRP and HB-sAg were measured as previously described for diphtheria, tetanus and pertussis (DTP)[14,42], changing only the antigen-coated microspheres, quality controls and reference sera used. In brief, DMTMM-modified[41] Hib-PRP 12/306 (NIBSC) was conjugated to Bio-Plex COOH-microspheres (Bio-Rad) as outlined above for pneumococcal polysaccharides, with an optimal dose of 1:1, 500 µl microspheres with 500 µl DMTMM-modified polysaccharide.

Recombinant HB-sAg subtype adr (Prospec Bio) was conjugated as previously described for DTP antigens, with an optimal coating concentration of 10 μg ml^{-1}. A ten-step, three-fold serial dilution of the international standard for anti-*H. influenzae* type b, 09/222 (NIBSC), was used as the reference serum, after anti-HB-sAg titres had been determined in that serum. Anti-HB-sAg titres in 09/222 were quantified against the second international standard for anti-HB-sAg, 07/164 (NIBSC), using the final Bio-Plex protocol; they were further validated using the current gold-standard method, the Alinity i Anti-HBs kit (PathWest), which showed excellent concordance (coefficient of variation less than 5%).

IgA Serology

Rotavirus IgA antibodies were measured by validated enzyme-linked immunoassay, as previously described[17,18].

High and low vaccine responders

High and low vaccine responders were defined as infants with IgG titres in the 75th percentile and 25th percentile, respectively, for each of the antigens assessed.

Multiplex opsonophagocytosis method

Multiplex opsonophagocytosis assays were performed on three serotypes (6B, 9V and 18C) using a previously published method[43,44]. Serum samples were heat-inactivated at 56 °C for 30 min; then, serial dilutions were made, and 20 μl per well of each dilution was tested in duplicate in 96-well u-bottomed microtitre plates (Interpath Services). Serial three-fold dilutions were used. Frozen aliquots of target pneumococci were thawed, washed with opsonization buffer (HBSS with magnesium and calcium, and 0.1% gelatin), and diluted to approximately 5×10^4 CFU ml^{-1} of each serotype. All serotypes were then pooled in one bacterial suspension, and 10 μl of the bacterial suspension was added to each well. Plates were incubated at room temperature with shaking (mini Orbital Shaker; Bellco Biotechnology) at 700 rpm for 30 min. Then, 10 μl of baby rabbit complement and 40 μl of HL60 cells (approximately 1×10^7 cells ml^{-1}) were then added to each well (the HL60 cells had been washed twice with HBSS before use). Plates were incubated in a tissue culture incubator (37 °C in a 5% CO_2 atmosphere) with shaking at 700 rpm. After 45 min, plates were placed on ice to stop the reaction, and 10 μl of the final reaction mixture was spotted onto Todd-Hewitt agar plates supplemented with 0.5% yeast extract. An overlay agar containing one of the selective antibiotics (selective markers for target bacteria) was added to one of each of the Todd-Hewitt agar plates, which was then incubated overnight at 37 °C with 5% CO_2. The number of bacterial colonies was enumerated using a ProtoCol3 automatic colony counter (Synbiosis). Results were expressed as opsonization indices, defined as the interpolated dilutions of serum that killed 50% of bacteria. The lower limit of detection in the assay was 4. The opsonization indices of samples that did not kill 50% of bacteria were reported as 2 for purposes of analysis. A cut-off of 8 was used to define a positive response.

Human serum bactericidal activity assays

Human serum bactericidal activity assays were performed on 140 matched study participant serum samples (from blood draws at 6 weeks, 7 months and 15 months of the study) as described previously[45]. Briefly, heat-inactivated study participant sera were serially diluted two-fold in assay buffer and inoculated with approximately 1×10^3 CFU of *N. meningitidis* strain NZ98/254. Exogenous human complement (25% v/v, Pel-Freez Biologicals) was added, and assay microtitre plates were incubated at 37 °C with 5% CO_2 for 60 min; then, samples were plated on GC agar base plates prepared using the tilt method and incubated at 37 °C with 5% CO_2 overnight to enable CFU enumeration. Bactericidal titres are reported as the reciprocal of the lowest dilution of participant serum that induced at least 50% killing relative to the no-serum control. The serologic correlate of protection for *N. meningitidis* is considered to be a human serum bactericidal activity titre greater than or equal to 4 (ref. 22). Statistical analyses were performed using one-way analysis of variance and two-tailed paired Student's *t*-test.

Infant blood preparation for flow cytometry

For flow cytometry analysis, 450 μl of infant blood was transferred from a BD Vacutainer K3 EDTA tube into a sterile 10-ml tube within 3 h of collection, and red blood cells were lysed by addition of 4.5 ml of 1× BD Pharm Lyse for 15 min at room temperature. Cells were then pelleted by centrifugation at 200*g* for 5 min, the supernatant was aspirated, and the cell pellet was resuspended in 10 ml flow cytometry buffer and centrifuged again. Cells were subsequently resuspended in 150 μl flow cytometry buffer and divided equally into three tubes for staining. Fc receptors were first blocked using Human TruStain FcX (BioLegend) in Brilliant Stain Buffer (BD) for 10 min on ice. Three panels were designed for same-day flow cytometry analysis (see Supplementary Table 8 for a complete list of cell subsets identified in each panel). Panel 1 was a pan-leukocyte panel enabling identification of neutrophils, monocytes, dendritic cells, eosinophils, basophils, T cells, B cells and natural killer cells. Cells were stained with a cocktail of anti-CD3 FITC (HIT3a, 1:20, BD Biosciences), anti-CD11c PE (B-ly6, 1:20, BD Biosciences), anti-HLA-DR APC-H7 (G46-6, 1:40, BD Biosciences), anti-CD15 V500 (HI98, 1:80, BD Biosciences), anti-CD45 BUV395 (HI30, 1:80, BD Biosciences), anti-CD16 BV421 (3G8, 1:143, BD Biosciences), anti-CD14 AF647 (Mop-9, 1:143, BD Biosciences), anti-CD123 BV711 (9F5, 1:143, BD Biosciences), anti-CD56 PE-Cy7 (5.1H11, 1:40, BioLegend), anti-CD19 BV605 (HIB19, 1:80, BioLegend), anti-CD20 BV786 (2H7, 1:80, BioLegend), anti-FcεRIα PerCP Cy5.5 (AER-37, 1:80, BioLegend) and anti-Siglec8 PE Dazzle (7C9, 1:143, BioLegend). Panel 2 enabled characterization of around 20 functional subsets of B and T cells including naive, central memory and effector memory $CD4^+$ and $CD8^+$ T cells; regulatory T cells; and naive, memory and plasma B cells. The antibody cocktail consisted of anti-CD19 PE (HIB19, 1:10, BD Biosciences), anti-CCR7/CD197 AF647 (150503, 1:10, BD Biosciences), anti-CD127 BV421 (A019D5, 1:40, BioLegend), anti-CD38 BV510 (HIT2, 1:40, BD Biosciences), anti-HLA-DR APC-H7 (G46-6, 1:40, BD Biosciences), anti-CD3 BUV395 (UCHT1, 1:40, BD Biosciences), anti-CD4 BV605 (RPA-T4, 1:80, BD Biosciences), anti-CD8 Cy7 (RPA-T8, 1:80, BD Biosciences), anti-IgD PerCP Cy5.5 (IA6-2, 1:80, BD Biosciences), anti-CD27 BV711 (L128, 1:143, BD Biosciences), anti-CD25 PE Dazzle594 (M-A51, 1:80, BioLegend), anti-CD20 BV786 (2H7, 1:80, BioLegend), and anti-CD45RA FITC (HI100, 1:143, BioLegend). Panel 3 was designed to assess specific T helper (T_H) subsets, including T_H1, T_H2, T_H9, T_H17, T_H22 and natural killer T cells, using anti-CD3 BUV395 (UCHT1, 1:50, BD Biosciences), anti-CD4 BV605 (RPA-T4, 1:100, BD Biosciences), anti-CD8 PE Cy7 (RPA-T8, 1:100, BD Biosciences), anti-CXCR3 BV421 (1C6/CXCR3, 1:100, BD Biosciences), anti-CCR10 APC (1B5, 1:100, BD Biosciences), anti-CCCR6 BV711 (11A9, 1:50, BD Biosciences), anti-CCR4 PECF594 (1G1, 1:50, BD Biosciences), anti-CD45RA FITC (HI100, 1:100, BioLegend) and CD1d-Tet PBS57 PE (1:5,000, NIH). See Supplementary Table 8 for a comprehensive list of antibody sources, catalogue numbers and fluorophores.

After each antibody cocktail had been added, samples were incubated in the dark on ice for 30 min before being washed and resuspended in 300 μl of flow cytometry buffer. Panel 1 was added to a BD Trucount FACS tube (BD), whereas panels 2 and 3 were added to round-bottomed 5-ml FACS tubes alongside 10 μl liquid counting beads (BD). Counts were generated as previously described[25]. Dead cells were stained by addition of DNA binding dye DAPI and were excluded from analysis. Samples were run on a FACSymphony (BD) flow cytometer using FACSDiva software, and results were analysed using FlowJo v.10.6.1 (Tree Star).

Extraction of RNA from blood

RNA extraction and elimination of genomic DNA were carried out using a PAXgene Blood RNA kit (Qiagen) per the manufacturer's instructions, with final elution into 80 μl RNase-free water. A further RNA

precipitation reaction was carried out. Briefly, RNA was resuspended in 2.5 × 100% ethanol and 10% sodium acetate and spun at 12,000*g* for 30 min at 4 °C. Samples were washed in 75% ethanol. Pellets were air-dried and resuspended in 40 µl of RNase-free water. The total RNA yield was determined using a Bioanalyzer 2100/TapeStation (Agilent) and Qubit (Thermo Fisher Scientific). Before library preparation, 0.2 ng of synthetic reference RNA standards (Sequins, Garvan Institute) were added to 250 ng of RNA.

RNA library preparation

Total RNA was converted to strand-specific Illumina-compatible sequencing libraries using a Nugen Universal Plus mRNA-Seq library kit from Tecan, with Any Deplete probes and custom gamma globin depletion per the manufacturer's instructions (MO1442 v.2). Briefly, 250 ng of total RNA was polyA-selected, and the mRNA was fragmented before reverse transcription. Any deplete and custom gamma globin depletion was performed before strand selection II complementary DNA (cDNA) synthesis with pooled gamma depletion oligos, per the manufacturer's instructions. The resulting cDNA was end-repaired, followed by ligation of Illumina-compatible barcoded sequencing adaptors. The cDNA libraries were strand-selected and PCR amplified for 12 cycles before quality control assessment using a TapeStation (Agilent Technologies). Library concentrations were assessed using a Qubit fluorescence assay (Thermo Fisher Scientific). Sequencing pools were generated by mixing equimolar amounts of compatible sample libraries on the basis of the Qubit measurements. The library pool was sequenced using an Illumina NovaSeq 6000 S1 flow cell with 2 × 100-bp paired-end reads.

RNA sequencing analysis

Sequence read quality was assessed using FastQC v.0.11.4 (ref. 46) and summarized with MultiQC v.1.8 (ref. 47), followed by quality control with Trimmomatic v.0.38 (ref. 48) with a window size of 4 and an average quality score of 25. Reads of fewer than 50 nucleotides after trimming were discarded. Reads that passed all quality control steps were then aligned to the human genome (GRCh38) using HISAT2 v.2.1.0 (ref. 49). The gene count matrix was generated with FeatureCounts version 1.5.0-p2 (ref. 50) using the union model with Ensembl v.93 annotation. Data were then imported into R v.4.2.0 for further analysis. Counts were normalized using the trimmed mean of M values method in EdgeR v.3.38 (ref. 40), followed by multidimensional scaling analysis and differential gene expression analysis performed with the glmLRT function. For differential expression analysis comparing week 6 and week 7, a paired design was used, fitting each participant into the model as well as each time point. As each participant's samples were processed together, this intrinsically controlled for batch effects. For non-paired comparisons, processing batch and infant sex were fitted into the model to adjust for those effects. Unknown sources of variation were identified using svaseq v.3.46.0 (ref. 38) and adjusted for before multidimensional scaling and gene set variation analysis. The gene count matrix was filtered to remove genes with a count per million of less than 5 in 15 samples before differential expression analysis. The Benjamini–Hochberg method was used to correct for multiple comparisons. Gene set enrichment analysis was carried out using the fgsea R package v.1.22.0 (ref. 39) with predefined BTMs[24], with input ranks as fold changes calculated by edgeR. Gene set variation analysis was used to calculate a per-sample activity score for each of the BTMs (excluding unannotated modules labelled 'TBA') using R Bioconductor package GSVA v.1.44.1 (ref. 51). Spearman correlation analysis was performed using the Hmisc v.4.7-0 package in R v.4.2.

DNA extraction from faecal samples

Two millilitres of the faecal suspension from the Zymo DNA/RNA Shield faecal collection tube (Zymo Research) was transferred into a sterile microcentrifuge tube and pelleted by centrifugation at 16,200*g* for 20 min at 4 °C. The supernatant was then removed, and DNA extraction was performed on the faecal pellet using the PowerLyzer PowerSoil Kit protocol (Qiagen) as described previously[52].

Quantitative real-time PCR analysis to determine bacterial load in infant stool samples

Total bacterial load in the samples was determined by quantitative PCR (qPCR) of the 16S ribosomal RNA (rRNA) gene using a SYBR-based assay, as described previously[53]. Each qPCR reaction comprised 1× PowerUP SYBR Green Master Mix (Applied Biosystems by Thermo Fisher Scientific), 0.2 µM of each forward and reverse primer, and 1 µl of DNA in a 35-µl total reaction volume. Each reaction was then aliquoted into three technical replicates of 10 µl for qPCR analysis. Numbers of 16S rRNA gene copies were calculated on the basis of a standard curve generated by serial dilutions of *Escherichia coli* genomic DNA and normalized against the faecal weight (in mg) of each sample.

Metagenomic library preparation

Shotgun metagenomics libraries were generated as previously described with modifications[54]. Briefly, libraries were generated for samples with a DNA concentration of at least 0.2 ng μl^{-1} using a Nextera DNA XT Library Preparation Kit and indexed using a Nextera XT Index Kit v.2 (Illumina). Quality control checks for DNA fragment size following PCR enrichment were performed using a DNA 1000 chip on a Bioanalyzer (Agilent Technologies). All libraries with a concentration of 5 nM and above were pair-end sequenced (2 × 151 bp) on the NovaSeq 6000 platform (Illumina).

Metagenomic data analysis

Read quality was assessed using FastQC v.0.11.4 (ref. 46) and summarized with MultiQC v.1.8 (ref. 47). Adaptors were removed with cutadapt v.1.18, and quality-filtering and trimming were performed using Trimmomatic v.0.38 (ref. 48) with a window size of 4, average quality score of 25 and minimum length of 50 nucleotides. The percentage of human DNA per sample was estimated using Kraken2 v.2.1.1 (ref. 55). MAG bins were generated by assembling each of the 409 samples individually using metaSPAdes v.3.15.2 (ref. 56). Reads were then aligned back to this assembly catalogue, and the mapped output was used for binning with MetaBAT2 v.2.15 (ref. 57). Counts per bin were then calculated on the basis of SAM files output from bowtie2 v.2.5.0. Each MAG bin was taxonomically classified using GTDB-Tk v.2.1.0 and reference data release 207 (ref. 58). This MAG bin count matrix was then imported into R for further analyses. Count data were transformed to relative abundances before being visualized using bar plots and were normalized by CLR elsewhere as implemented in R library PhyloSeq v.1.28. Differential abundance analysis was performed with LinDA as implemented in R library MicrobiomeStat v.1.18 with raw count data as input. Counts were filtered for taxa detected in at least 5% of samples before testing with the prev.filter flag in LinDA[59]. Spearman correlation analyses were performed using functions in base R with the CLR-normalized data as input. Alpha and beta diversity analyses were performed in R using PhyloSeq v.1.28 (ref. 60). Metagenomic samples were clustered into CSTs with Dirichlet multinomial mixture modelling (R package DirichletMultinomial v.1.38.0), with the optimal number of clusters calculated using the lowest LaPlace approximation. Quality-filtered reads were also used as input to HUMAnN v.3.0 (ref. 61) for generation of microbial pathway data and summed to represent counts for each MetaCyc pathway and gene ontology term (biological processes only). Pathway count data were also CLR-normalized. To account for variation in 16S rRNA copy numbers between samples, which could have biased qPCR estimates of bacterial load, a mean copy number per sample was calculated on the basis of the species-level relative abundance profiles from the shotgun metagenomic data and the copy number from rrnDB[62]. Antimicrobial resistance genes were analysed by prediction of the encoded protein sequences from MAGs using prodigal v.2.6.3,

clustering at 95% identity with cd-hit v.4.8.1, and conversion to amino acid sequences using transseq to generate a non-redundant predicted protein catalogue. The non-redundant catalogue was aligned to the Comprehensive Antibiotics Resistance Database (v.3.2.7) using the Resistance Gene Identifier (v.5.2.1) with default parameters except the addition of '–exclude_nudge'. Per-gene counts were generated by aligning reads using bowtie2 against the gene catalogue and generating a per-sample count with SAMtools v.1.17. Reads per kilobase per million mapped reads values were generated in R, summed by antibiotic class and compared between groups using a Wilcoxon rank sum test, controlling for multiple comparisons with the Benjamini–Hochberg method. Causal mediation analysis was performed in R with the mediation R package (v.4.5.0)[28].

Preclinical mouse experiments

C57BL/6J mice (The Jackson Laboratory; RRID: IMSR_JAX:000664) were bred and maintained under specific and opportunistic pathogen free (SPF) conditions or gnotobiotic conditions at the South Australian Health and Medical Research Institute (SAHMRI). SPF mice were kept in cages with free access to commercial pelleted food and water and were housed under standardized conditions with regulated daylight, humidity and temperature. Gnotobiotic (GF) C57BL/6 mice (Translational Research Institute) were housed in positively pressurized HEPA-filtered isolators at the SAHMRI Preclinical, Imaging and Research Laboratories (PIRL) with access to autoclaved commercially pelleted food and sterilized water ad libitum, with regulated 12 h day/night and regulated humidity (45–75%) and temperature (18–24 °C). All experiments were carried out in accordance with protocols approved by the SAHMRI Animal Ethics Committee. Each vaccine experiment was conducted as an independent experiment. Equal numbers of male and female mice were used. Young mice (aged approximately 3 weeks) or dams (aged 8–16 weeks) and pups were randomized to the treatment groups. Investigators were not blinded to the mouse treatment groups. The appropriate sample size for the murine experiments was determined on the basis of data from our previous mouse vaccine experiments.

GF mouse dams and their pups remained GF (as determined by culturing of faecal samples and 16S rRNA gene qPCR) or were colonized as previously described[63] by means of oral gavage with a consortium of *Bifidobacterium* species consisting of *B. breve* (DSM 20213, Leibniz Institute), *B. longum* subspecies *longum* (DSM 20219) and *B. longum* subspecies *infantis* (DSM 20088) at approximately 14 days postcoitus (E14) and at day 7 and day 14 postbirth. Successful colonization was confirmed by culturing and Sanger sequencing. For culturing, *Bifidobacterium* isolates were plated on MRS agar (Thermo Fisher Scientific) enriched with 1 mg ml^{-1} cysteine, and single colonies were isolated, confirmed using a MALDI Biotyper (Bruker) and cultured for 48 h at 37 °C using BD GasPak pouch systems (BD). To assess whether administration of a commercially available probiotic could restore antibody responses to PCV13, GF mice were colonized by means of oral gavage at day 21 of life with 2×10^7 CFU *L. acidophilus* and *B. bifidum* isolated from commercial probiotic Infloran (Laboratorio Farmaceutico). Further GF mice were similarly colonized with 2×10^7 CFU with each of the strains individually. *L. acidophilus* and *B. bifidum* cultures were prepared by serially diluting Infloran on to MRS agar plates (Thermo Fisher Scientific) enriched with cysteine for *L. acidophilus* or reinforced clostridial agar (Thermo Fisher Scientific) for *B. bifidum*, under anaerobic conditions. Single colonies were identified using a MALDI Biotyper (Bruker). Further GF mice, aged 21 days, were colonized with *Lactobacillus murinus* (DSM 20452), *Enterococcus gallinarum* (DSM 24841), *Akkermansia muciniphila* (DSM 22959), *Bacteriodes acidifaciens* (DSM 15896), *Blautia producta* (DSM 2950), *Lactobacillus plantarum* (ATCC 202195, The Global Bioresource Centre) or *Enterobacter cloacae*[1].

To produce littermate wild-type and knockout mice for the experiments in $Myd88^{-/-}$, $Tlr2^{-/-}$ and $Tlr4^{-/-}$ mice, pairs of heterozygous $Myd88^{+/-}$, $Tlr2^{+/-}$ and $Tlr4^{+/-}$ dams (8–12 weeks of age) were bred with heterozygous studs. Equal numbers of wild-type and knockout offspring were cohoused for studies.

Vaccination of mice

Mice were vaccinated with 3 µg PCV13 (Pfizer), 8 µg Infanrix Hexa combination vaccine (GSK) or 17.5 µg Bexsero meningococcal serogroup B vaccine (GSK) intraperitoneally at day 28 postbirth and boosted with the same dose 2 weeks later. Control mice were mock-vaccinated with 100 µl of PBS. Four to eight sex-matched mice per group were vaccinated or mock-vaccinated in each experiment. Sera were collected 2 and 4 weeks postimmunization in each experiment and also at 6 and 8 weeks in selected experiments.

Murine faecal sampling

Faecal samples were collected from mice at the start of each experiment and at each experimental time point to confirm the GF or colonization status of the mice. Faecal pellets were aseptically collected, immediately snap-frozen and stored at −80 °C.

DNA extraction from murine faecal samples

Faecal samples were individually weighed, resuspended in 1 ml of phosphate-buffered saline (PBS) (pH 7.2) by vortexing and pelleted by centrifugation at 13,000*g* for 5 min. Microbial DNA extraction was performed on the faecal pellets using a DNeasy PowerSoil Kit (Qiagen) with the following minor modifications: faecal pellets were resuspended in 750 µl of PowerSoil bead solution and 60 µl solution C1 in a PowerSoil bead plate and incubated at 65 °C for 10 min before bead beating as per previously described protocols[63].

16S rRNA gene sequence analysis

Murine faecal DNA extracts were used to generate amplicons of the V4 hypervariable region of the 16S rRNA gene as described previously[63]. Sequencing of the amplicon library was performed using an Illumina MiSeq system (2 × 300-bp run). Paired-end 16S rRNA gene sequences were demultiplexed and imported into QIIME2 (release 2023.2) for processing[64]. Sequences were error-corrected, and counts of error-corrected reads per sample, which we refer to herein as exact sequence variants, were generated with DADA2 (ref. 65). Taxonomy was assigned to sequences with the sklearn plugin for QIIME2 with an 80% confidence threshold, using the SILVA v.138 database[66]. Exact sequence variant counts were converted to relative abundance, which was plotted in R v.4.2.0 with ggplot2 v.3.3.6.

Quantitative real-time PCR analysis to determine bacterial load

Real-time PCR was performed on a Quant Studio 7 Flex Real-time PCR system (Thermo Fisher Scientific) using primers that targeted conserved regions of the 16S rRNA gene as described in ref. 67. Each reaction comprised 5 µl of SYBR Green qPCR (Thermo Fisher Scientific), 0.2 µM of each forward and reverse primer, 1 µl of DNA template and sterile water to a total reaction volume of 22 µl. A no-template DNA control, with the DNA substituted with 1 µl sterile water, was included in each run. Real-time PCR reactions for each sample were performed in duplicate, consisting of 10 µl per reaction. The amplification program was 50 °C for 2 min, 95 °C for 10 min, followed by 40 cycles of 95 °C for 15 s and 60 °C for 1 min. Melting curve analysis was performed at the end of the program. The total bacterial load in each sample was quantified against a standard curve of serial dilutions of *E. coli* genomic DNA, which were performed in the same run, and was detectable within a range of 3 to 3×10^6-equivalent *E. coli* cells. Numbers of cells were then calculated against the faecal weight used for each sample. A Mann–Whitney test was used to assess statistical significance.

Assessment of antibody responses in mice

Levels of antigen-specific IgG_{total} IgG1, IgG2c, IgG3 and IgM were measured by enzyme-linked immunosorbent assay (ELISA) in serum

collected from each vaccine experiment. Briefly, Nunc MaxiSorp ELISA plates (Sigma) were coated either with CRM_{197} (0.5 μg ml^{-1}, Finabio), PCV13 vaccine (0.05 μg ml^{-1}, Pfizer), pneumococcal polysaccharide (4 μg ml^{-1} PPS1, 161-X, ATCC), PPS3 (4 μg ml^{-1} 169-X, ATCC), PPS6B (4 μg ml^{-1} 105-X, ATCC), PPS9V (4 μg ml^{-1} 305-X, ATCC) or diphtheria toxoid (1 μg ml^{-1}, List Biological Laboratories), tetanus toxoid (1 μg ml^{-1}, List Biological Laboratories), pertussis toxoid (1 μg ml^{-1}, List Biological Laboratories) or Bexsero vaccine (2 μg ml^{-1}, GSK). Plates were washed with 0.5% Tween-20/PBS and blocked with 1× ELISA buffer (eBioscience). Mouse serum samples were diluted in 1× ELISA buffer and incubated on blocked plates. Antigen-specific serum antibodies were detected using horseradish peroxidase (HRP)-conjugated antibodies: anti-mouse IgG (Novex) or anti-mouse IgG1 (Abcam) at a dilution of 1:1,000 or IgG2c, IgG3 or IgM (Abcam) at 1:500. HRP activity was detected using tetramethylbenzidine substrate (Thermo Fisher Scientific) and stopped with 2 N H_2SO_4. Developed plates were recorded using a Synergy HTX microplate absorbance reader (BioTek) at 450 nm with correction at 595 nm by subtraction. Data were analysed using Prism 10 (GraphPad).

Tissue dissociation and flow cytometry analysis

Single-cell suspensions of murine spleen or lymph nodes were prepared by mechanical dissociation and filtration through a 70-μm nylon filter (Merck Millipore). Cells were stained with Fixable Viability Stain 780 (BD Biosciences) and murine Fc-blocking antibody clone 2.4G2 (1:100, BD Biosciences) in PBS for 15 min. Cells were washed and then stained with fluorochrome-conjugated antibodies in FACS buffer (PBS, 0.1% bovine serum albumin, 2 mM EDTA) against CD3 (142-C11, 1:300, BD Biosciences), CD4 (RM4, 1:300, BD Biosciences), CD8 (53-6.7, 1:600, BD Biosciences), CD19 (1D3, 1:300, BD Biosciences), CD38 (90, 1:600, BioLegend), CD44 (IM7, 1:600, BD Biosciences), CD62L (MEL-14, 1:600, BD Biosciences), CD95 (J02, 1:300,BD Biosciences), CD138 (281-2, 1:600, BD Biosciences), B220 (RA3-6B2, 1:600, BD Biosciences), CXCR5 (L138D7, 1:200, BioLegend), GL7 (GL7, 1:600, BD Biosciences), IgD (11-26 c.2a, 1:300, BioLegend), IgM (R6-60.2, 1:300, BD Biosciences), PD-1 (RMP1-30, 1:300, BD Biosciences) and streptavidin-PECF549 (1:400, BioLegend).

For assessment of CRM^+ GC B cells, cells were stained with Fixable Viability Stain 780 (1:1000, BD) and murine Fc block (1:100, BD) in PBS for 15 min; then, they were washed and incubated with biotinylated CRM (1:100, Fina BioSolutions) in FACS buffer. Surface antibody cocktails were added directly after staining for 20 min at 4 °C. Data were acquired on a BD FACS Symphony and analysed using FlowJo v.10. For assessment of CD86, CD80 and MCHII expression, single-cell suspensions of spleen or lymph nodes were stained with Zombie Aqua Fixable Viability Stain 780 (1:1000, BioLegend) and murine Fc-blocking antibody clone 2.4G2 (1:100, BD Biosciences) in PBS for 15 min. Cells were washed and then stained with fluorochrome-conjugated antibodies in FACS buffer (PBS, 0.1% bovine serum albumin, 2 mM EDTA) against CD19 (ID3, 1:200, BD), CD3 (145-2C11, 1:150, BD Biosciences), CD8 (53-6.7, 1:200, BD), NK1.1 (PK136, 1:200, BD), Ly6G (1A8, 1:600, BD), B220 (RA3-6B2, 1:3000, BD), CD11b (M1/70, 1:500, BD), CD11c (HL3, 1:400, BD Biosciences), MHCII (M5/114.15.2, 1:150, Miltenyi), CD86 (PO3.3, 1:150, Miltenyi) and CD80 (16-10A1, 1:150, Tonbo Biosciences).

Immunofluorescence staining and imaging

Mediastinal lymph nodes were fixed in 1 ml 4% paraformaldehyde (Sigma) for 4 h and then incubated in 1 ml 30% sucrose solution overnight at 4 °C before being embedded in OCT (Tissue Tek) as described previously[68]. Slides were blocked in blocking buffer (normal goat serum 4% v/v (Southern Cross Science), normal donkey serum 4% v/v (Stratech Scientific APAC) and bovine serum albumin 15% v/v (AusGeneX)) for 2 h at room temperature. Slides were permeabilized with Triton X-100 2% v/v (Sigma) and stained with CD16/32 (1:100, BD Biosciences) and CD21/35 clone 7E9 (1:200), GL7 (1:200), CD3 clone 17A2 (1:200) and IgD clone 11-26c.2a (1:200) (all from BioLegend). Images were acquired using a Leica TCS SP8X confocal microscope and analysed using Fiji[69] in ImageJ.

Statistical analysis

The primary objective of the AIR study was to determine whether antibody responses to the PCV13 vaccine at 7 months of age were different between antibiotic-exposed and unexposed groups, specifically comparing the following:

1. The percentage of participants achieving a seroprotective antibody response (greater than or equal to 0.35 μg ml^{-1}) against pneumococcal polysaccharide serotypes in the PCV13 vaccine at approximately 7 months of age.
2. Geometric mean concentrations in μg ml^{-1} of anti-PPS IgG as measured by serum assay at approximately 7 months of age.

As secondary outcomes, we compared PCV13 vaccine responses at 15 months of age and antibody responses to the Infanrix Hexa 6-in-1 vaccine (6-in-1, DTPa-HepB-IPV-Hib) and the Rotarix oral rotavirus vaccine in antibody-exposed and unexposed infants at both time points. Multivariable linear regression adjusting for baseline prevaccination titres and sex was used to test for any association between antibiotic exposure group and IgG geometric mean concentrations for each of the vaccine antigens assessed. Further longitudinal analyses were conducted by fitting a mixed effects multivariable linear regression, including a random effect for participant ID, to assess the interaction between time point and antibiotic exposure group on IgG geometric mean concentrations for each of the vaccine antigens. Multivariable logistic regression was used to test for any association between antibiotic exposure group and the proportion of infants achieving a seroprotective response. Statistical significance was defined as $P < 0.05$. Further sensitivity analyses were conducted to assess the robustness of associations to adjustment for probiotic consumption, formula consumption, baseline IgG/IgA concentration, infant sex and gestation at delivery. The AIR study also included a range of exploratory systems immunology assessments for which analyses were undertaken in line with best practices for each data type and as described in the sections above. Measurements were taken from distinct samples. All statistical tests throughout were two-sided; the specific statistical analyses used are specified in the figure legends.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

Data that support the findings of this study have been deposited in the Gene Expression Omnibus under accession code GSE198276 and in the Sequence Read Archive under BioProject accession code PRJNA807448. All other data supporting the findings of this study are available within this paper and its Supplementary Information. Source data are provided with this paper.

Code availability

No custom code was developed as part of this manuscript. R scripts used for omics data analysis are available from the Lynn laboratory BitBucket (https://bitbucket.org/lynnlab/air).

41. Schlottmann, S. A., Jain, N., Chirmule, N. & Esser, M. T. A novel chemistry for conjugating pneumococcal polysaccharides to Luminex microspheres. *J. Immunol. Methods* **309**, 75–85 (2006).
42. McAlister, S. M. et al. An observational study of antibody responses to a primary or subsequent pertussis booster vaccination in Australian healthcare workers. *Vaccine* **39**, 1642–1651 (2021).
43. Balloch, A. et al. Interlaboratory comparison of the pneumococcal multiplex opsonophagocytic assays and their level of agreement for determination of antibody function in pediatric sera. *mSphere* **3**, e00070-18 (2018).
44. Temple, B. et al. Efficacy against pneumococcal carriage and the immunogenicity of reduced-dose (0+1 and 1+1) PCV10 and PCV13 schedules in Ho Chi Minh City, Viet Nam: a parallel, single-blind, randomised controlled trial. *Lancet Infect. Dis.* **23**, 933–944 (2023).

45. Lucidarme, J., Louth, J., Townsend-Payne, K. & Borrow, R. Meningococcal serogroup A, B, C, W, X, and Y serum bactericidal antibody assays. *Methods Mol. Biol.* **1969**, 169–179 (2019).
46. Andrews, S. FastQC: a quality control tool for high throughput sequence data. *Babraham Institute* www.bioinformatics.babraham.ac.uk/projects/fastqc/ (2010).
47. Ewels, P., Magnusson, M., Lundin, S. & Kaller, M. MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* **32**, 3047–3048 (2016).
48. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
49. Kim, D., Langmead, B. & Salzberg, S. L. HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods* **12**, 357–360 (2015).
50. Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2014).
51. Hänzelmann, S., Castelo, R. & Guinney, J. GSVA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinformatics* **14**, 7 (2013).
52. Choo, J. M., Leong, L. E. & Rogers, G. B. Sample storage conditions significantly influence faecal microbiome profiles. *Sci. Rep.* **5**, 16350 (2015).
53. Choo, J. M. et al. Impact of long-term erythromycin therapy on the oropharyngeal microbiome and resistance gene reservoir in non-cystic fibrosis bronchiectasis. *mSphere* **3**, e00103-18 (2018).
54. Mobegi, F. M. et al. Intestinal microbiology shapes population health impacts of diet and lifestyle risk exposures in Torres Strait Islander communities. *eLife* **9**, e58407 (2020).
55. Wood, D. E., Lu, J. & Langmead, B. Improved metagenomic analysis with Kraken 2. *Genome Biol.* **20**, 257 (2019).
56. Nurk, S., Meleshko, D., Korobeynikov, A. & Pevzner, P. A. metaSPAdes: a new versatile metagenomic assembler. *Genome Res.* **27**, 824–834 (2017).
57. Kang, D. D. et al. MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. *PeerJ* **7**, e7359 (2019).
58. Chaumeil, P.-A., Mussig, A. J., Hugenholtz, P. & Parks, D. H. GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. *Bioinformatics* **36**, 1925–1927 (2019).
59. Zhou, H., He, K., Chen, J. & Zhang, X. LinDA: linear models for differential abundance analysis of microbiome compositional data. *Genome Biol.* **23**, 95 (2022).
60. McMurdie, P. J. & Holmes, S. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS ONE* **8**, e61217 (2013).
61. Franzosa, E. A. et al. Species-level functional profiling of metagenomes and metatranscriptomes. *Nat. Methods* **15**, 962–968 (2018).
62. Stoddard, S. F., Smith, B. J., Hein, R., Roller, B. R. & Schmidt, T. M. rrnDB: improved tools for interpreting rRNA gene abundance in bacteria and archaea and a new foundation for future development. *Nucleic Acids Res.* **43**, D593–D598 (2015).
63. Lynn, M. A., Ryan, F. J., Tee, Y. C. & Lynn, D. J. Protocol to colonize gnotobiotic mice in early life and assess the impact on early life immune programming. *STAR Protoc.* **3**, 101914 (2022).
64. Bolyen, E. et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotechnol.* **37**, 852–857 (2019).
65. Callahan, B. J. et al. DADA2: high-resolution sample inference from Illumina amplicon data. *Nat. Methods* **13**, 581–583 (2016).
66. Quast, C. et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* **41**, D590–D596 (2013).
67. Nadkarni, M. A., Martin, F. E., Jacques, N. A. & Hunter, N. Determination of bacterial load by real-time PCR using a broad-range (universal) probe and primers set. *Microbiology* **148**, 257–266 (2002).
68. Fra-Bido, S., Walker, S. A., Innocentin, S. & Linterman, M. A. Optimized immunofluorescence staining protocol for imaging germinal centers in secondary lymphoid tissues of vaccinated mice. *STAR Protoc.* **2**, 100499 (2021).
69. Schindelin, J. et al. Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012).

Acknowledgements We thank all the infants and their parents or guardians for their participation in this study. This study was financially supported by grants from the National Health and Medical Research Council of Australia (APP1156415 and APP2017404) and the Women's and Children's Hospital Foundation. K.L.S. and H.S.M. are supported by an NHMRC Leadership Investigator Grant (APP2017383, APP2034644). We thank R. van der Steere for assistance with isolating bacterial strains from the Infloran probiotic; T. Norton for assistance with flow cytometry; S. Forster, P. Hertzog and J. McVernon for input on study design and grant applications that supported the study; L. Quah and R. Higgins at MCRI for their assistance with the opsonophagocytosis assays; M. De Virgilio, S. Read and the SAHMRI Bioresources/PIRL facility staff for assistance with mouse husbandry and breeding; and the Garvan Institute for providing the sequin controls for RNA sequencing. Flow cytometry was performed at the ACRF Cellular Imaging and Cytometry Core Facility in SAHMRI. We also thank the South Australian Genomics Center (SAGC) for help with RNA sequencing. The SAGC is supported by the National Collaborative Research Infrastructure Strategy (NCRIS) through Bioplatforms Australia and by the SAGC partner institutes. We acknowledge the UK Health Security Agency for providing *N. meningitidis* strain NZ98/254.

Author contributions The study was conceived and designed by D.J.L., H.S.M. and S.L.W. The documentation for human research ethics committee approval for the clinical study were written by M.C., D.J.L. and H.S.M. Sample processing protocols were written by M.C., M.A.L. and D.J.T. Recruitment and sample collection were managed by M.C., M.W. and H.S.M. Blood samples were processed for RNA sequencing by M.A.L. and S.J.B. Flow cytometry was performed by S.C.B., S.J.B. and D.J.T. Faecal metagenomic and 16S rRNA qPCR data were generated by J.M.C, A.R. and L.E.X.L under the supervision of G.B.R. Resistome analysis was performed by S.L.T and F.J.R. Serology data were generated by S.M. under the supervision of P.R. Y.Y.N. performed the opsonophagocytosis assays under the supervision of P.V.L. E.A.S. performed the serum bactericidal assays under the supervision of K.L.S. Data analysis, interpretation and visualization were performed by F.J.R. and M.C. under the supervision of D.J.L., H.S.M. and L.C.G. Preclinical animal experiments were performed by M.A.L., S.C.B., C.R. and J.I.M. under the supervision of D.J.L. The manuscript was written by F.J.R., M.C., M.A.L. and D.J.L. with contributions and approval from all authors.

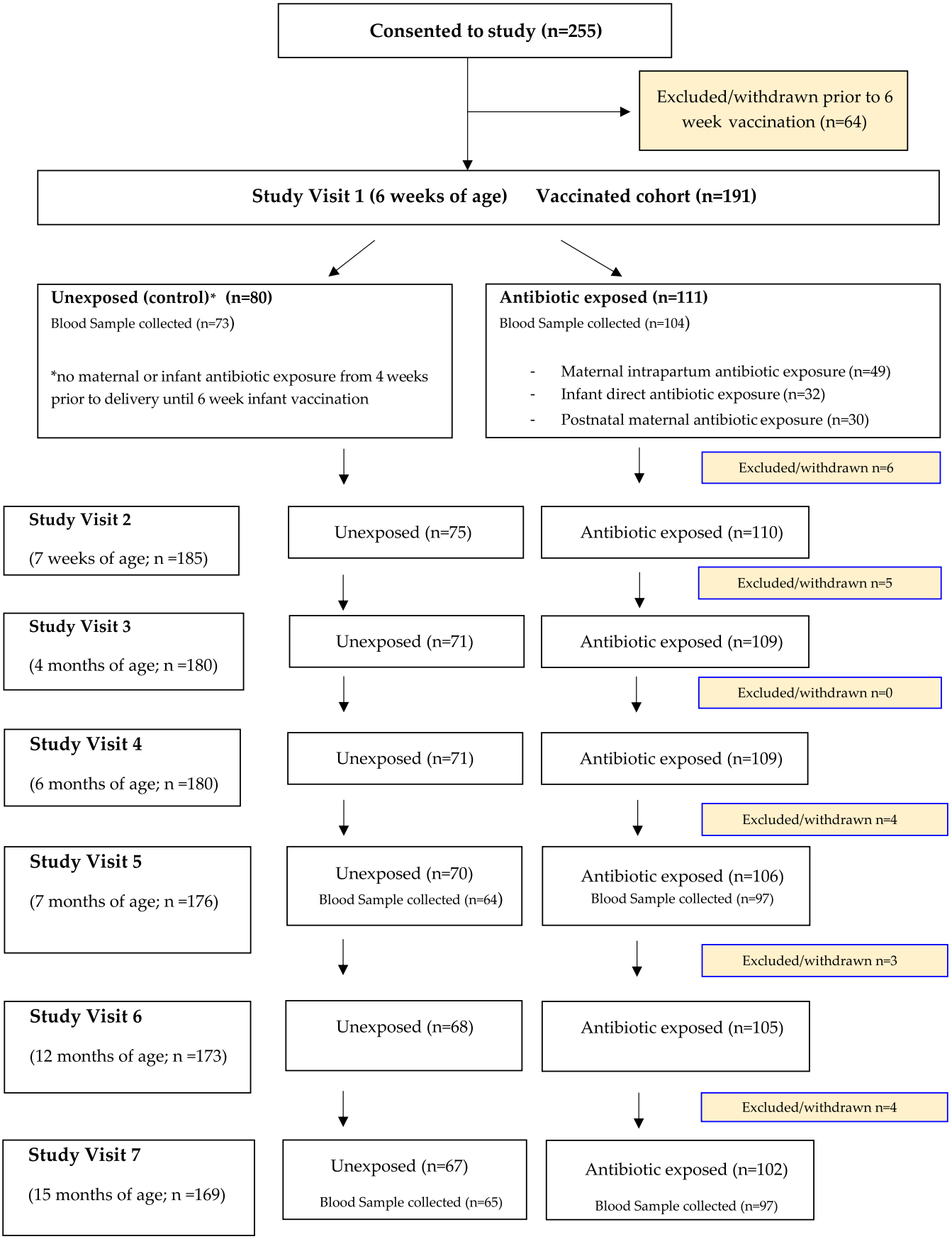
Competing interests D.J.L., S.J.B., F.J.R. and M.A.L. have received funding from GSK related to this area of research. D.J.L. and H.S.M. are consultants for GPN Vaccines (Australia). H.S.M. is an investigator on sponsored vaccine trials funded by Iliad and Pfizer. H.S.M. has received institutional funding for investigator-led studies from Pfizer, Sanofi-Aventis and Seqiris. P.R. has received institutional funding for investigator-led grants from Merck and Sanofi outside this work. His institution receives funding for his participation in vaccine scientific advisory boards from Pfizer, Merck, GSK, GPN Vaccines (Australia) and Sanofi-Aventis unrelated to this work. He is an investigator on sponsored vaccine trials funded by Iliad, GSK, Merck, Sanofi, Dynavax and Pfizer. He is coinventor on a patent for a new bacteriotherapy for treating or preventing respiratory infection unrelated to this work. The other authors declare no competing interests.

Additional information

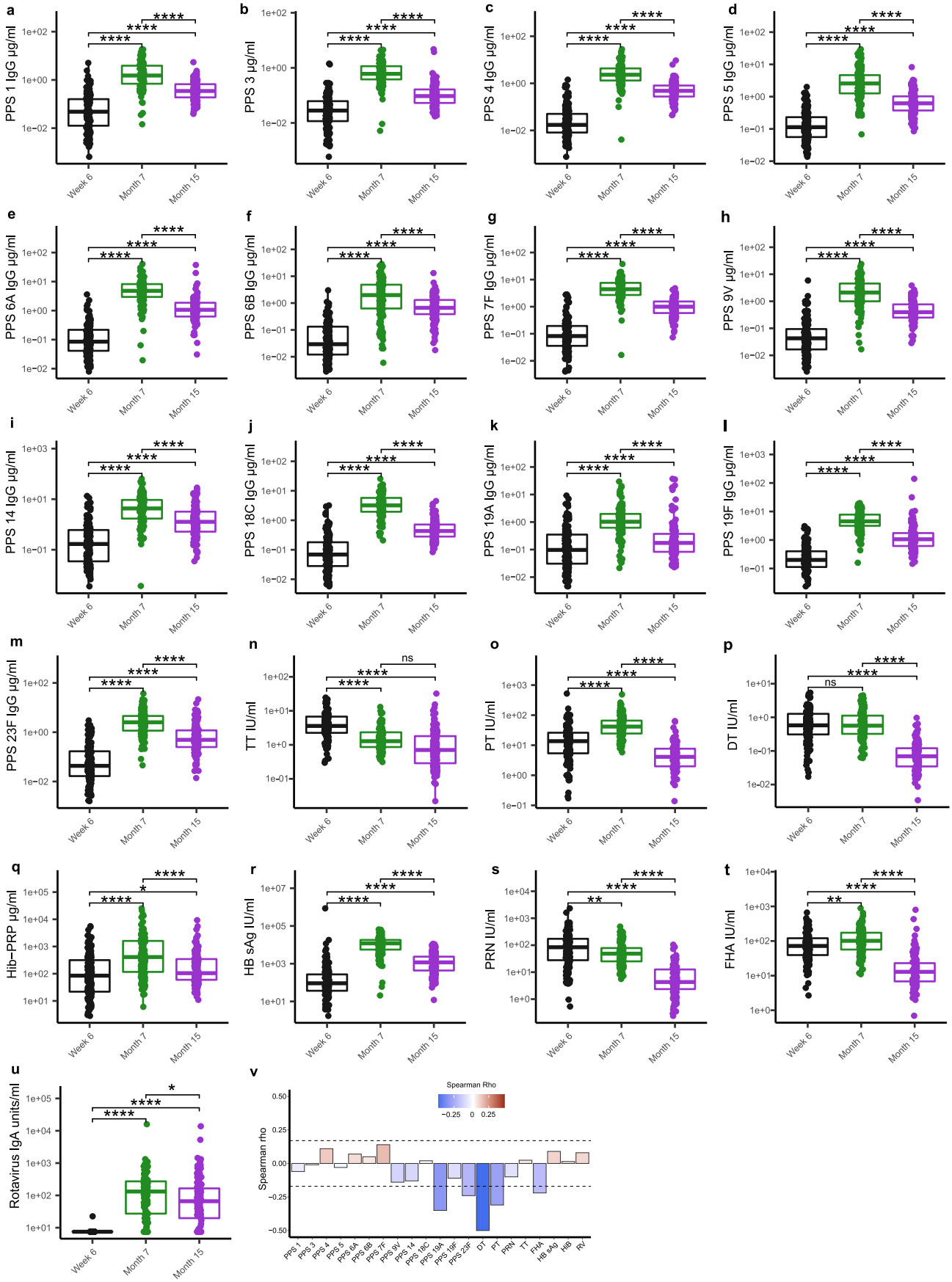
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41586-025-08796-4.

Correspondence and requests for materials should be addressed to David J. Lynn.

Peer review information *Nature* thanks Martin Blaser and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

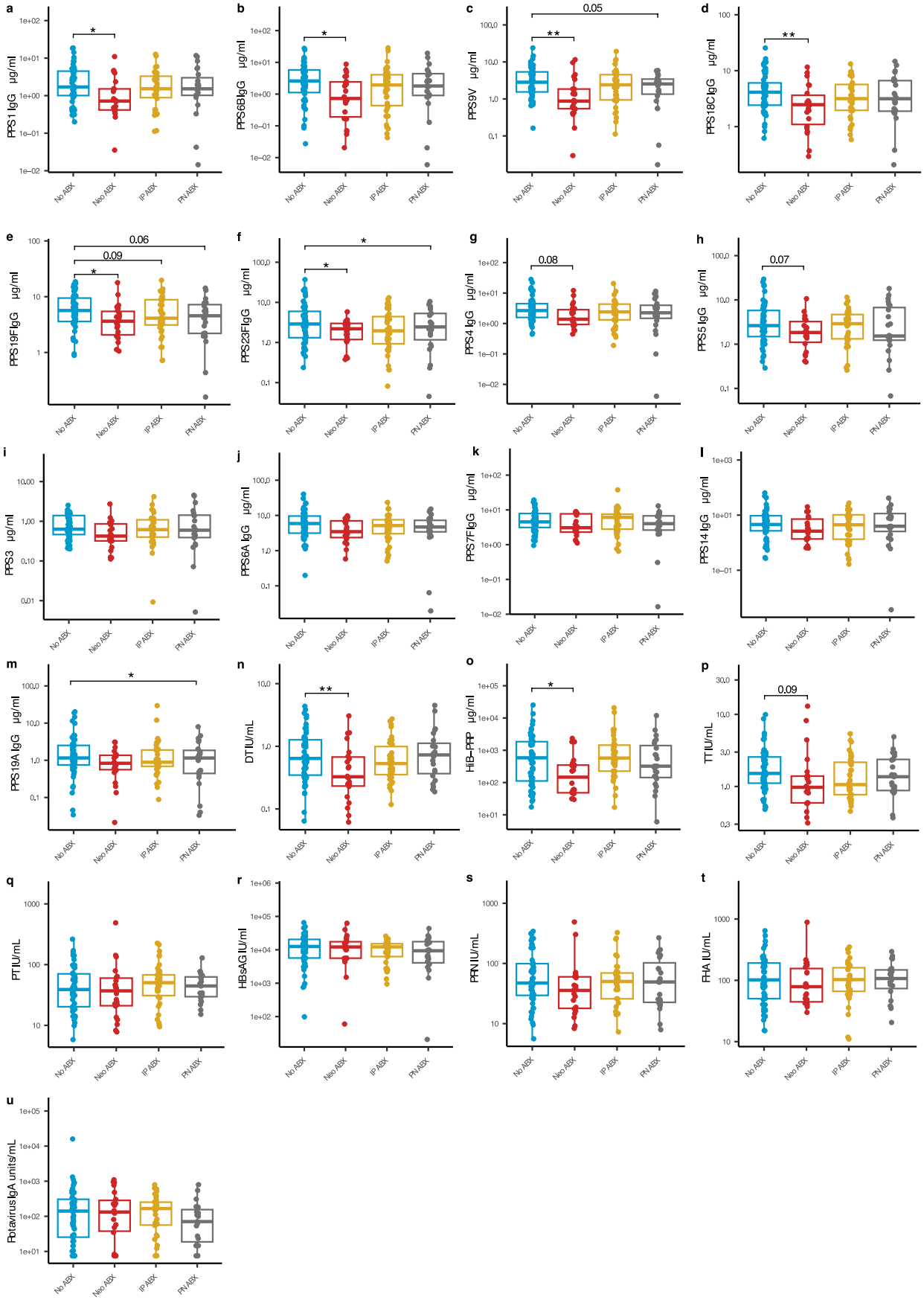


Extended Data Fig. 1 | Study cohort flow diagram. Flow of participant recruitment and progress through the Antibiotics and Immune Responses (AIR) study.



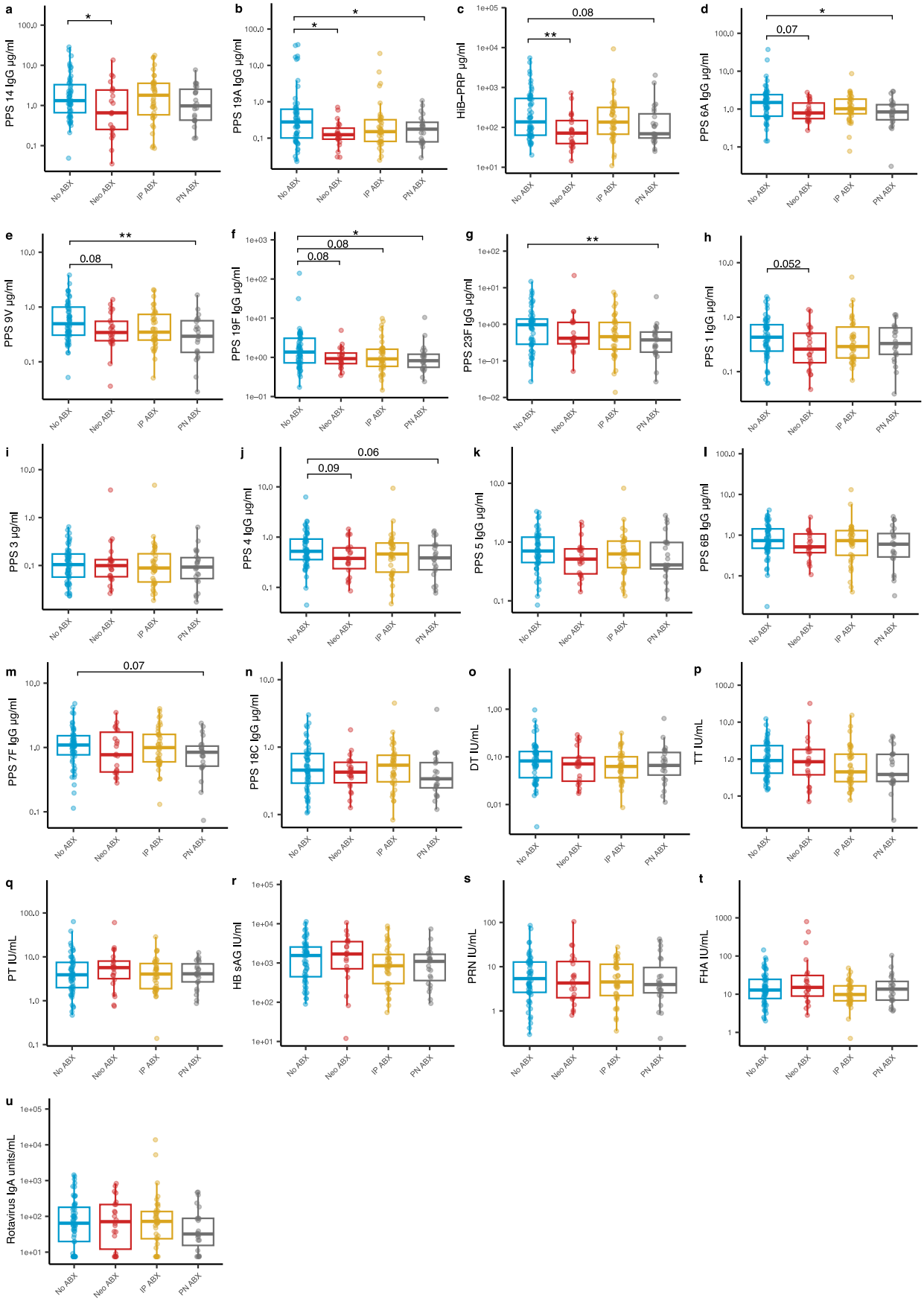
Extended Data Fig. 2 | See next page for caption.

Extended Data Fig. 2 | Antibody titres for each of the measured vaccine antigens at baseline and at 7 and 15 months. (**a**-**m**) IgG titres against the 13 capsular pneumococcal polysaccharide serotypes (PPS) in the PCV13 vaccine in sera collected from infants at baseline, 7 months and 15 months (n = 139 infants per timepoint; only infants with matched samples across 3 timepoints included). (**n**-**t**) IgG titres against (**n**) tetanus toxoid (TT), (**o**) pertussis toxoid (PT), (**p**) diphtheria toxoid (DT), (**q**) *Haemophilus influenzae* type b (Hib) polyribosylribitol phosphate (PRP), (**r**) hepatitis B surface antigen (HBsAg), (**s**) pertactin (PRN), and (**t**) filamentous haemagglutinin (FHA) (all antigens in the 6-in-1 Infanrix Hexa® vaccine) in sera. (**u**) anti-rotavirus vaccine IgA titres in sera (n = 139). (**v**) Spearman correlations between baseline antibody titres and titres at 7 months (n = 139). Data in **a**-**u** are represented as boxplots with the box denoting the 25th and 75th percentiles, the whiskers the 5th and 95th percentiles, the middle bar is the median. Statistical significance was assessed in **a**-**u** using a generalised linear model adjusting for sex and baseline antibody titres. Data shown in **v** were determined using the *cor.test* function in *R*. * unadjusted $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, **** $P < 0.0001$. Exact P values are provided in the Source Data File. All P values are two-sided.



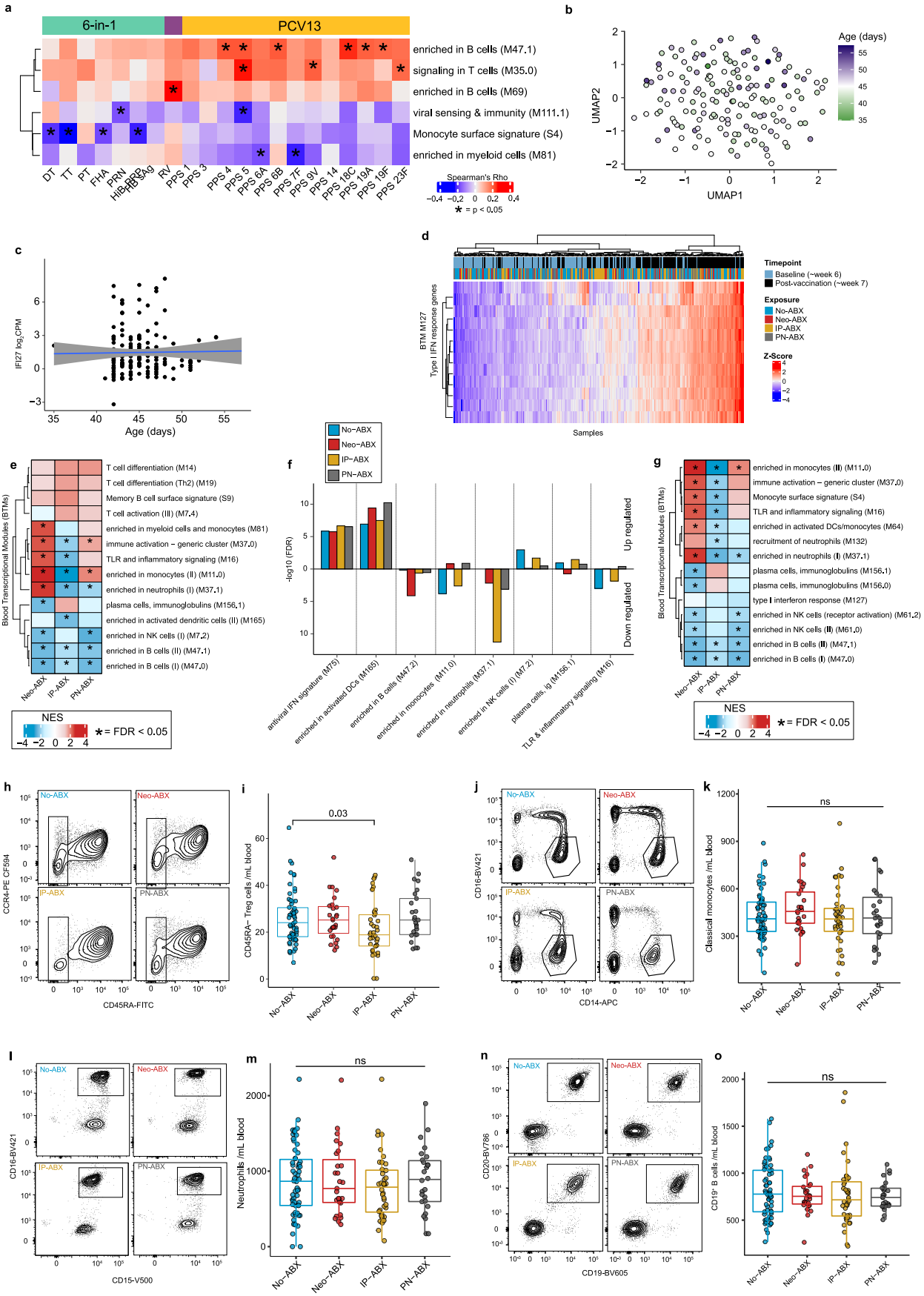
Extended Data Fig. 3 | See next page for caption.

Extended Data Fig. 3 | Antibody titres for each of the measured vaccine antigens at 7 months of life subdivided by antibiotic exposure group. (**a**-**f**) IgG titres in sera at 7 months against 6 of the 13 capsular pneumococcal polysaccharides (PPS1, 6B, 9 V, 18 C, 19 F and 23 F) in the PCV13 vaccine that were significantly lower in Neo-ABX infants. (**g**-**h**) IgG titres against PPS4 and PPS5, which trended lower in Neo-ABX infants at 7 months. (**i**-**m**) IgG titres against other capsular pneumococcal polysaccharides that were not lower in Neo-ABX infants at 7 months. (**n**-**t**) IgG titres against (**n**) DT, (**o**) Hib-PRP, (**p**) TT, (**q**) PT, (**r**) HB sAG, (**s**) PRN and (**t**) FHA in sera collected from infants at 7 months. (**u**) anti-rotavirus vaccine IgA titres in sera collected from infants at 7 months. (**a**-**u**) No-ABX (n = 64 infants), Neo-ABX (n = 27 infants), IP-ABX (n = 43 infants), PN-ABX (n = 26 infants). Data in **a**-**u** are represented as boxplots with the box denoting the 25th and 75th percentiles, the whiskers the 5th and 95th percentiles, the middle bar is the median. Statistical significance was assessed in **a**-**u** using a generalised linear model (Gaussian distribution) adjusting for sex and baseline antibody titres. *unadjusted $P < 0.05$, ** $P < 0.01$. Exact P values are provided in the Source Data File. All P values are two-sided.



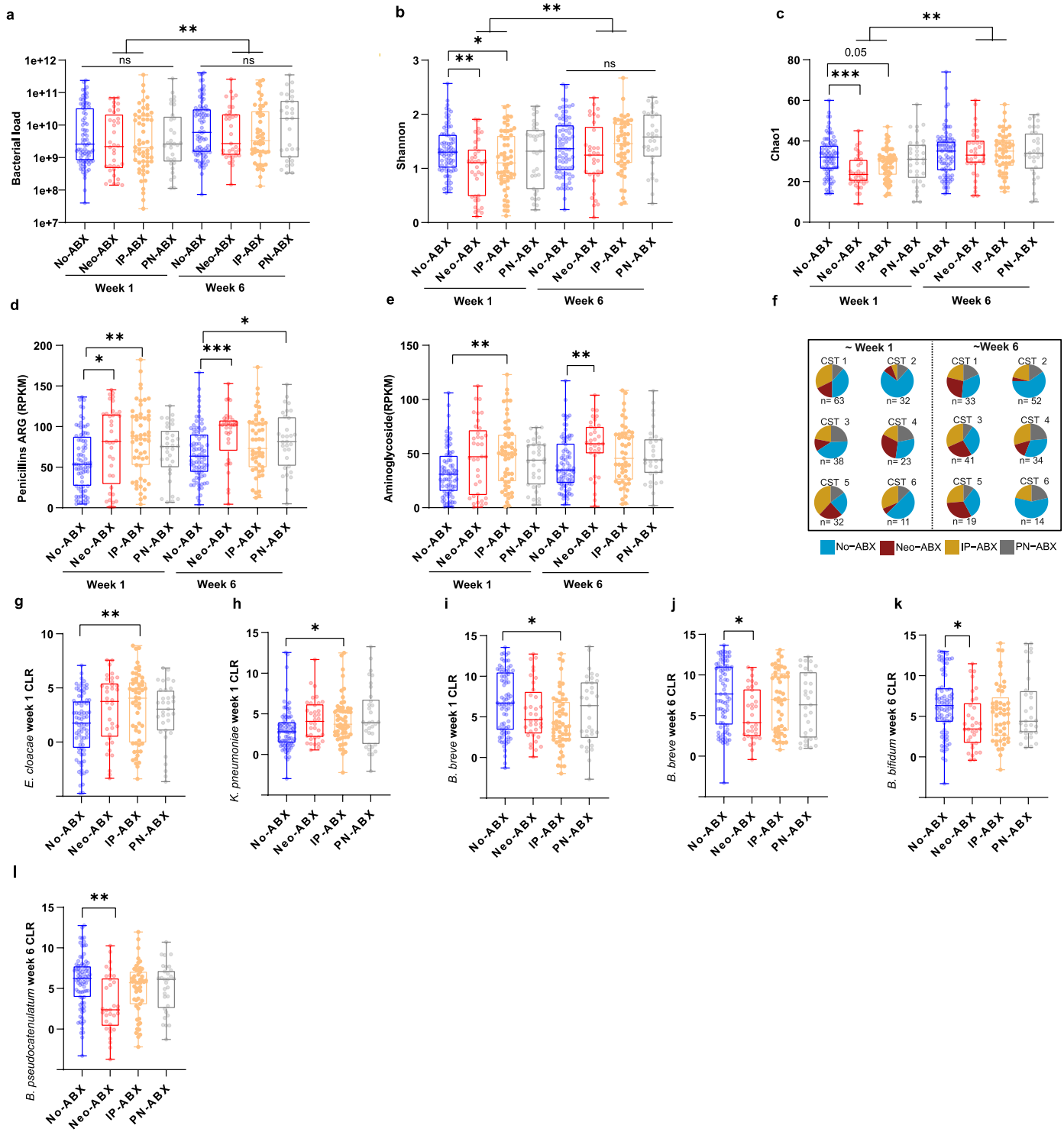
Extended Data Fig. 4 | See next page for caption.

Extended Data Fig. 4 | Antibody titres for each of the measured antigens at 15 months of life subdivided by antibiotic exposure group. (**a**-**u**) IgG titres in sera at 15 months against (**a**) PPS14, (**b**) PPS19A, (**c**) Hib-PRP, (**d**) PPS6A, (**e**) PPS9V, (**f**) PPS19F, (**g**) PPS23F (all significantly lower in Neo-ABX and/or PN-ABX infants), (**h**) PPS1, (**i**) PPS3, (**j**) PPS4, (**k**) PPS5, (**l**) PPS6B, (**m**) PPS7F, (**n**) PPS18C, (**o**) DT, (**p**) TT, (**q**) PT, (**r**) HB sAG, (**s**) PRN, (**t**) FHA. (**u**) anti-rotavirus vaccine IgA titres in sera collected from infants at 15 months**. a-u** No-ABX (n = 65 infants), Neo-ABX (n = 26 infants), IP-ABX (n = 42 infants), PN-ABX (n = 27 infants). Data in **a-u** are represented as boxplots with the box denoting the 25th and 75th percentiles, the whiskers the 5th and 95th percentiles, the middle bar is the median. Statistical significance was assessed in **a-u** using a generalised linear model adjusting for sex and baseline antibody titres. * unadjusted $P < 0.05$, ** $P < 0.01$. Exact P values are provided in the Source Data File. All P values are two-sided.

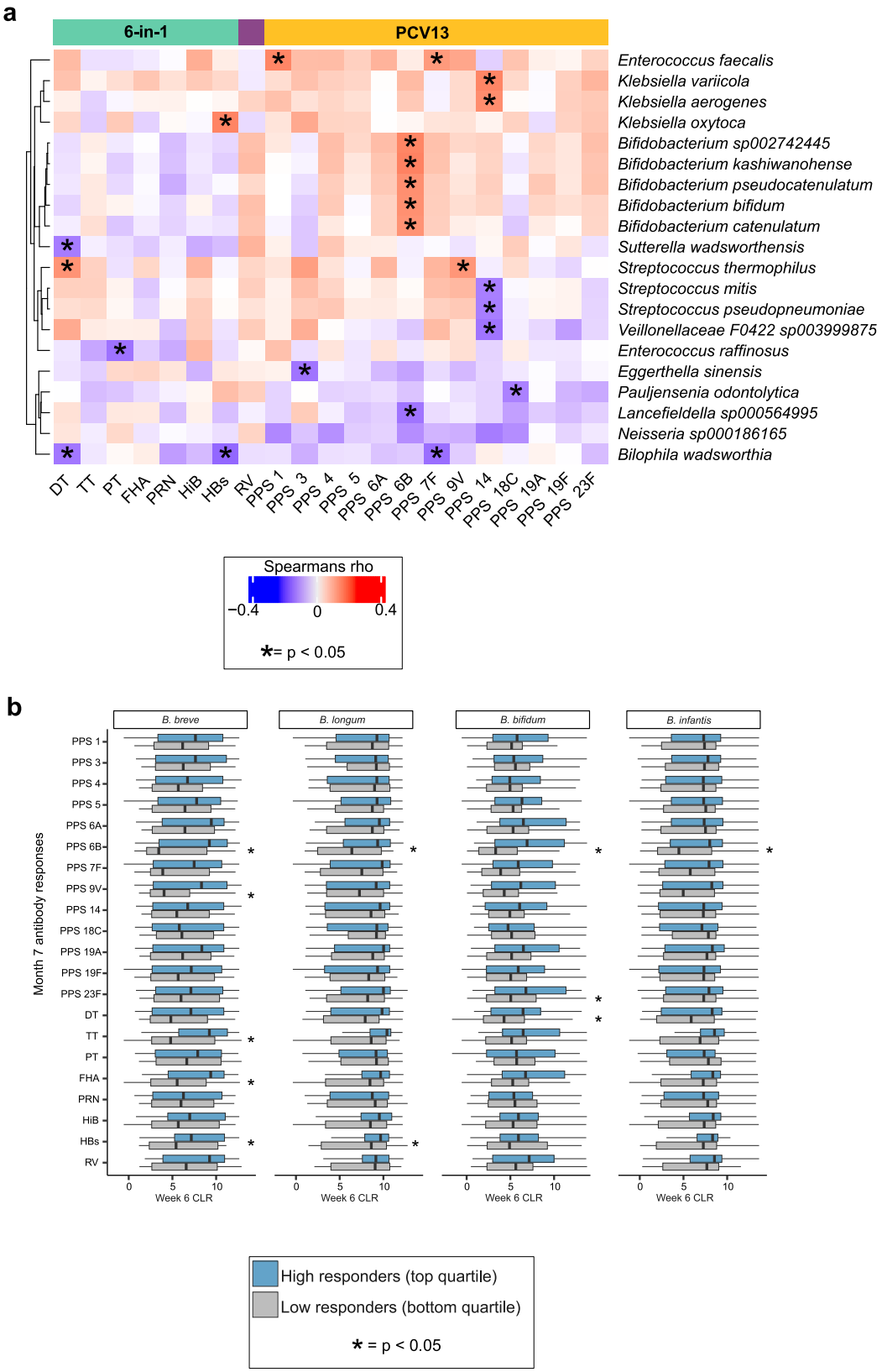


Extended Data Fig. 5 | See next page for caption.

Extended Data Fig. 5 | Whole-blood gene expression responses in antibiotic exposed and unexposed infants. RNAseq analysis was performed on blood collected from infants at week 6 and 7 (n = 329 infant blood samples). (**a**) Heatmap representing the correlations (Spearman's Rho) between the activity of selected Blood Transcriptional Modules (BTMs) at week 6 of life and antibody responses at 7 months of life (n = 139 infants). (**b**) UMAP projection of baseline (pre-vaccination) whole-blood gene expression data, adjusted for sex and batch using SVAseq. Data points coloured by infant age at sample collection (n = 159 infants). (**c**) Normalised expression of one of the DEGs identified post-vaccination, *IFI27*, showing that expression is not correlated with infant age (n = 159 infants). (**d**) Heatmap showing the normalised expression of type I interferon (IFN) response genes (annotated in BTM M127) pre- and post-vaccination (n = 329 infant blood samples). (**e**) Heatmap showing the Normalised Enrichment Scores (NES) in Neo-, IP- and PN-ABX infants for selected BTMs post-vaccination (week 7), relative to No-ABX infants (n = 147 infants). (**f**) Gene set enrichment analysis (GSEA) of selected BTMs comparing their expression in No-, Neo-, IP-, and PN-ABX infants at week 7 to No-ABX infants at week 6 (n = 294 samples). (**g**) Heatmap, as per panel e, showing the NES for selected BTMs post-vaccination in Neo-ABX (n = 27), IP-ABX (n = 44) and PN-ABX (n = 28) infants compared to No-ABX infants (n = 71). (**h-o**) Flow cytometry analysis was performed on fresh peripheral blood collected from No-ABX (n = 63 infants), Neo-ABX (n = 27 infants), IP-ABX (n = 40 infants), and PN-ABX (n = 26 infants) infants at week 7. Representative gates and counts/mL of blood for (**h-i**) $CD45RA^-$ Tregs, (**j-k**) $CD16^-CD14^+$ classical monocytes, (**l-m**) $CD16^+CD15^+$ neutrophils, and (**n-o**) $CD19^+$ B cells. Data in **i, k, m, o** are represented as boxplots with the box denoting the 25th and 75th percentiles, the whiskers the 5th and 95th percentiles, the middle bar is the median. Statistical significance was assessed in **a** using the *cor.test* function in *R*, in **e-g** using the *fgsea* package in *R* and in **i, k, m, o** using two-tailed Wilcoxon signed-rank tests. *P < 0.05, ns = not significant. Exact FDR values in **e** and **g** can be found in Supplementary Tables 5 and 7. All P values are two-sided.

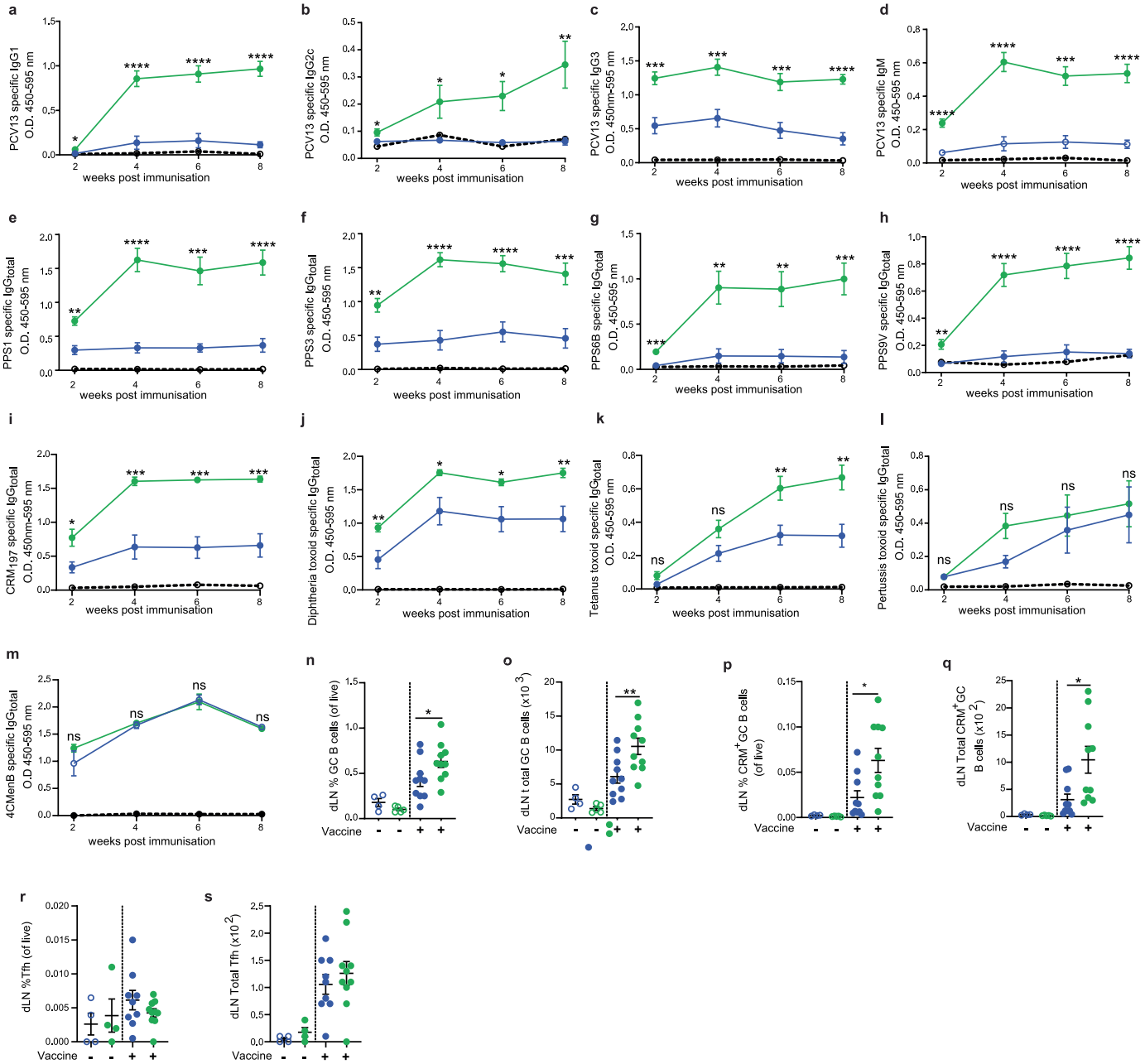


Extended Data Fig. 6 | Extended analyses of the gut microbiota at week 1 and 6 of life in infants recruited to the AIR study. (**a**) Bacterial load (16S rRNA gene copies) in stool samples collected from No-, Neo-, IP-, and PN-ABX infants at week 1 (n = 205 infants) and 6 of life (n = 190 infants). (**b**) Shannon diversity (**c**) Chao1 richness. (**d-e**) The abundance of penicillin and aminoglycoside antimicrobial resistance genes (ARGs) in No-, Neo-, IP-, and PN-ABX infants at week 1 and 6 of life. RPKM = reads per kb per million mapped reads. (**f**) The proportion of No-, Neo-, IP-, and PN-ABX infants at week 1 and 6 of life among the six CSTs. (n = 396 stool samples in **b-f**). (**g-l**) The centered-log ratio (CLR) abundance of selected differentially abundant taxa, see Supplementary Table 11 for a complete list. (**g-i** n = 203, **j-l** n = 193 stool samples). Data in **a-e**, **g-l** are represented as boxplots with the box denoting the 25th and 75th percentiles, the whiskers the 5th and 95th percentiles, the middle bar is the median. Statistical significance was assessed in **a-e** using two-tailed Wilcoxon signed-rank tests and in **g-l** using the *LinDA* method in the MicrobiomeStat R package. All statistical tests were two-sided. * unadjusted P < 0.05, **P < 0.01, ***P < 0.001, ns = not significant. Exact P values are provided in the Source Data File.

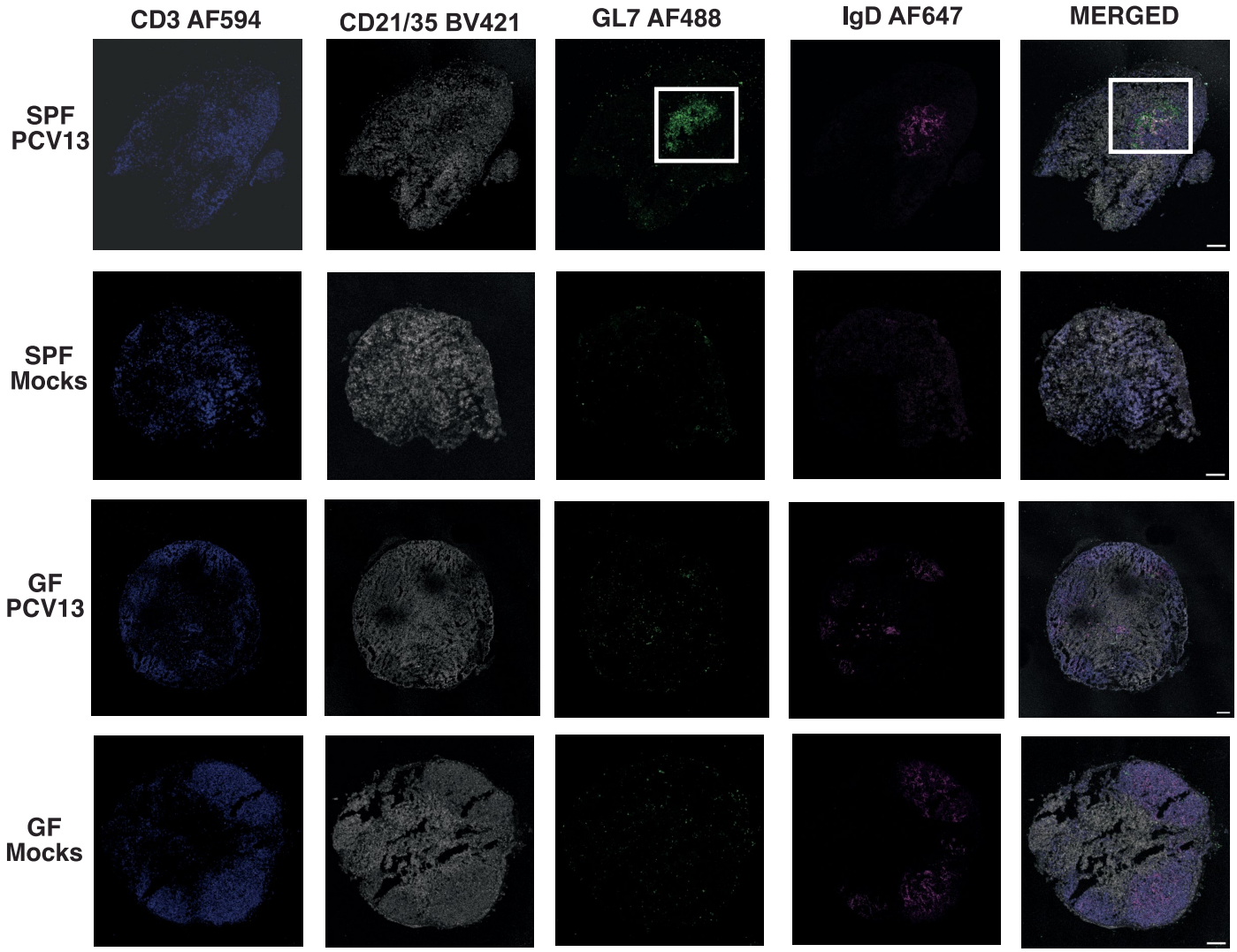


Extended Data Fig. 7 | See next page for caption.

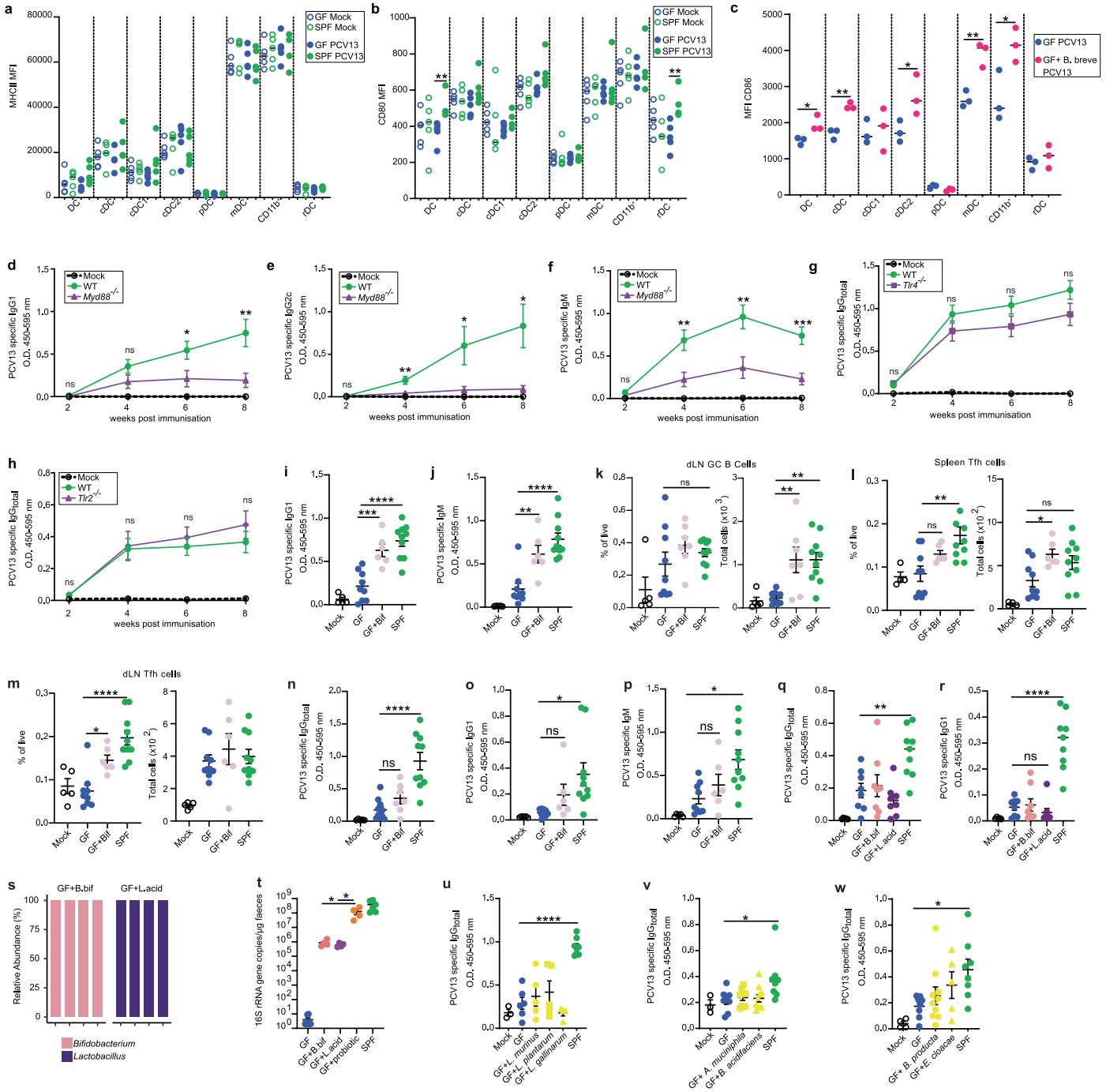
Extended Data Fig. 7 | Additional analyses assessing the relationship between the composition of the gut microbiota and vaccine response. (**a**) Heatmap representing the correlations between the normalised abundance of top 10 most positively correlated, and top 10 most negatively correlated taxa in the infant faecal microbiota at week 1 and antibody responses to vaccination (adjusted for sex and baseline titres) at month 7 (n = 134 infants with matched data across the 3 timepoints). Exact correlation coefficients are provided in Supplementary Table 13. (**b**) The centred log ratio (CLR) normalised abundance of the *Bifidobacterium* species that were most depleted in Neo-ABX infants at week 6 of life comparing high (75%ile, n = 40) vs. low (25%ile, n = 40) vaccine responders for each of the antigens assessed. Data in **a** are represented as a heatmap where the blue-red gradient represents negative and positive Spearman correlations, respectively. Data in **b** are represented as boxplots with the box denoting the 25th and 75th percentiles, the whiskers the 5th and 95th percentiles, the middle bar is the median. Statistical significance was assessed in **a** using the cor.test function in R and in **b** using the *LinDA* method in the MicrobiomeStat R package. All statistical tests were two-sided. * unadjusted $P < 0.05$.



Extended Data Fig. 8 | Germ-free mice have impaired serum antibody responses to the PCV13 and Infanrix Hexa vaccines, but not the 4CMenB meningococcal B vaccine. (**a**) PCV13-specific IgG1, (**b**) IgG2$_c$, (**c**) IgG3, (**d**) IgM antibodies were assessed by ELISA in serum collected from mock (PBS, n = 3 mice) and PCV13 vaccinated GF (n = 9 mice) and SPF (n = 10 mice) mice that were vaccinated at day 28 of life with two doses of PCV13 vaccine i.p., two weeks apart. IgG$_{total}$ responses against (**e**) PPS1, (**f**) PPS3, (**g**) PPS6B, (**h**) PPS9V were also assessed in these mice. (**i**) IgG$_{total}$ responses against CRM$_{197}$ were assessed by ELISA in serum collected from mock (PBS, n = 3 mice) and PCV13 vaccinated GF (n = 8 mice) and SPF (n = 8 mice) mice. (**j**) Diphtheria toxoid, (**k**) tetanus toxoid, and (**l**) pertussis toxoid IgG$_{total}$ responses were assessed by ELISA in serum collected from mock vaccinated (n = 2 mice) mice and GF (n = 8 mice) and SPF (n = 8 mice) mice vaccinated i.p. with two doses of Infanrix Hexa vaccine at day 28 of life. (**m**) 4CMenB-specific IgG$_{total}$ responses were assessed by ELISA in serum collected from mock (n = 8 mice) vaccinated mice and GF (n = 7 mice) and SPF (n = 10 mice) mice vaccinated i.p. with 4CMenB (meningococcal B) vaccine at day 28 of life. The frequency (as % of live) and total number of (**n-o**) GC B cells, (**p-q**) CRM$_{197}^+$ GC B cells and (**r-s**) T follicular helper (Tfh) cells in the dLN of mock (n = 9 mice), GF (n = 10 mice) and SPF (n = 10 mice) mice at 2 weeks after the 2nd dose of PCV13 was administered. Data in **a-s** are represented as mean ± SEM. Data in **a-i, n-s** are representative of two independent experiments. Data in **j-l** and **m** are from two different single independent experiments. Statistical significance in **a-m** was assessed using two-way ANOVA with Tukey's post-test analysis for multiple comparisons and in **n-s** using two-tailed Student's t-tests. *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.001, ns = not significant. Exact P values are provided in the Source Data File.



Extended Data Fig. 9 | Germ-free mice have impaired germinal centre formation following vaccination with the PCV13 vaccine. Confocal microscopy of dLNs collected from mock (PBS) and PCV13 vaccinated GF and SPF mice that were vaccinated at day 28 of life with one 3 µg dose of PCV13 vaccine i.p. Data shown are at two weeks post-immunisation. CD3, CD21/35, GL7 and IgD staining are visualised separately and then merged. A white box has been placed around the GL7 staining, which visible only in the vaccinated SPF mice. Data are from a single experiment. The scale bar represents 100 µm.



Extended Data Fig. 10 | See next page for caption.

Extended Data Fig. 10 | Extended data on immune responses to PCV13 in germ-free, *Myd88*$^{-/-}$, *Tlr4*$^{-/-}$ and *Tlr2*$^{-/-}$ mice. (**a**) MHCII and (**b**) CD80 MFI on indicated myeloid cell populations in the dLN of GF and SPF mice 24 h following PCV13/mock vaccination (n = 5 mice per group). (**c**) CD86 MFI on indicated myeloid cell populations in the dLN of GF and GF mice colonised with *B. breve* from birth, 24 h following PCV13/mock vaccination (n = 3 mice per group). PCV13-specific (**d**) IgG1, (**e**) IgG2c, (**f**) IgM antibodies as assessed by ELISA in the serum of mock (n = 5 mice) and PCV13 vaccinated *Myd88*$^{-/-}$ (n = 10 mice) and littermate wildtype (WT, n = 10 mice) mice. PCV13-specific IgG_{total} antibodies as assessed by ELISA in the serum of mock and PCV13 vaccinated (**g**) *Tlr4*$^{-/-}$ (n = 10 mice) and (**h**) *Tlr2*$^{-/-}$ (n = 11 mice) and littermate WT mice (n = 10, and 9, respectively). Mice in **d-h** were vaccinated at day 28 of life with two 3 µg doses of PCV13 vaccine i.p., two weeks apart. PCV13-specific (**i**) IgG1 and (**j**) IgM in the serum of mock vaccinated (n = 5 mice) and PCV13 vaccinated (two doses i.p., two weeks apart) SPF (n = 10 mice), GF (n = 9 mice), and GF mice born to dams colonised with a consortium of *Bifidobacterium* species in pregnancy and administered these strains again at day 7 and 14 post-birth (GF+Bif, n = 6 mice). Data shown are at 2 weeks post-boost. (**k**) The frequency (as % of live) and total number of GC B cells in the dLN of mice in (**i**) at 2 weeks post-boost. The frequency (as % of live) and total number of Tfh cells in the (**l**) spleen and (**m**) dLN of mice in (**i**) at 2 weeks post-boost. PCV13-specific (**n**) IgG_{total}, (**o**) IgG1 and (**p**) IgM antibodies as assessed by ELISA in serum of mice in (**i**) collected two weeks after PBS (mock) or PCV13 vaccination (one 3 µg dose i.p.). PCV13-specific (**q**) IgG_{total} and (**r**) IgG1 antibodies as assessed by ELISA at 2 weeks post-boost in the serum of mock (n = 5 mice), SPF (n = 9 mice), GF and GF mice colonised at day 21 with either the *B. bifidum* (GF + B.bif) or *L. acidophilus* (GF + L.acid) strains isolated from the Infloran probiotic (n = 8 mice per group). (**s**) 16S rRNA gene sequencing of faecal samples (n = 4 mice per group) was used to confirm colonisation with the Infloran strains. (**t**) Bacterial load was assessed using 16S rRNA gene qPCR in the indicated mice (n = 8 GF, n = 4 colonised and n = 6 SPF mice). PCV13-specific IgG_{total} antibodies were assessed by ELISA in the serum of mock and PCV13 vaccinated SPF, GF or GF mice colonised with (**u**) *L. murinus* (n = 5 mice), *L. plantarum* (n = 5 mice), *E. gallinarum* (n = 4 mice), (**v**) *A. muciniphila* (n = 10 mice), *B. acidifaciens* (n = 10 mice) or (**w**) *B. producta* (n = 10 mice), *E. cloacae* (n = 5 mice). Mice in **u-w** were vaccinated at day 28 of life with two 3 µg doses of PCV13 vaccine i.p., two weeks apart. Data shown are at 2 weeks post-boost. Data in **a-r**, **t-w** are represented as mean ± SEM. Data in **d-f, g, i-m** are each representative of two independent experiments. Data in **a-b**, **c**, **h**, **n-p**, **q-t, u, v, w** are all from independent single experiments. Statistical significance was assessed in **a-c** using two-tailed Student's t-test without adjustment for multiple comparisons, in **d-h** using two-way ANOVA with Tukey's post-test analysis for multiple comparisons and in **i-r**, **t-w** using one-way ANOVA with Dunnett's post-test analysis for multiple comparisons. *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.001, ns = not significant. Exact P values are provided in the Source Data File. All P values are two-sided.

Extended Data Table 1 | Antibiotics and Immune Responses (AIR) study demographics

Vaccinated cohort (n=191)		Antibiotic exposed			Non exposed
		Neo-ABX (n=32)	IP-ABX (n=49)	PN-ABX (n=30)	No-ABX (n=80)
Mother	Age at delivery (Mean years, SD)	31 (5)	32 (4)	33 (4)	32 (4)
	BMI (mean kg/m^2, SD)	24 (3)	24 (3)	23 (3)	24 (3)
	Pertussis vaccine in pregnancy (n, %)	31 (97%)	47 (96%)	30 (100%)	73 (94%)
Infant	Birthweight (mean kg, SD)	3.5 (0.5)	3.3 (0.4)	3.3 (0.4)	3.2 (0.5)
	Gestation at delivery (mean weeks, SD)	39 (1)	39 (1)	39 (1)	39 (1)
	Sex (n, % male)	17 (53%)	30 (61%)	15 (50%)	41/80 (51%)
	Apgar at 5 min (mean, SD)	8.6 (0.8)	8.9 (0.5)	8.9 (0.5)	8.9 (0.4)
	Breastfed (n, %)	32 (100%)	49 (100%)	30 (100%)	79 (99%)
	Received any formula (n, %)	5 (16%)	7 (14%)	5 (15%)	15 (19%)
	Probiotics during study (n, %)	6 (19%)	7 (14%)	5 (15%)	3 (2.5%)

The AIR study was a prospective observational study which divided participants in a priori defined antibiotic exposure groups. See Methods for further details on study design.

nature portfolio

Corresponding author(s): David J. Lynn

Last updated by author(s): Jan 27, 2025

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	REDcap V14.5.28
Data analysis	No custom code was developed as part of this manuscript. R scripts and processed data for reproducing the analyses can be accessed at https://bitbucket.org/lynnlab/air. Other 3rd party code used: Flow cytometry analysis: FlowJo™ v10.6.1; RNAseq analysis: FastQC v0.11.4; MultiQC v1.8; Trimmomatic version 0.38; HISAT2 v2.1.0; FeatureCounts v.5.0-p2; R v4.2.0; R packages: EdgeR v3.38, svaseq v3.46; fgsea v1.22.0, GSVA v1.44.1; Hmisc v4.7-0. Metagenomic Data Analysis: FastQC v0.11.4; MultiQC v1.8; Cutadapt v1.18; Trimmomatic v0.38; Kraken2 v2.1.1; metaSPAdes v3.15.2; MetaBAT2 v2.15; bowtie2 v2.5.0; GTDB-Tk v2.1.0; PhyloSeq v1.28; DirichletMultinomial v1.38.0; HUMAnN v3.0; SAMtools v1.17 Antimicrobial resistance gene analysis: prodigal v2.6.3; cd-hit v4.8.1. Differential abundance analysis: LinDA as implemented in the R library Microbiome v1.18 Causal mediation analysis: mediation R package v4.5.0. 16S rRNA gene sequence analysis: QIIME2 (release 2023.2); DADA2; ggplot2 v3.3.6 in R version 4.2.0 Mouse data analyses: Prism v10 Confocal: Fiji in ImageJ

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

Data that support the findings of this study have been deposited deposited in the Gene Expression Omnibus (GEO) under accession code "GSE198276" (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE198276) and in the Sequence Read Archive under bioProject accession code PRJNA807448 (https://www.ncbi.nlm.nih.gov/bioproject/?term=PRJNA807448). All other data supporting the findings of this study are available within the paper and its supplementary information files or are available via the Lynn Laboratory BitBucket repository (https://bitbucket.org/lynnlab/air). Source data for figures 1-5, extended data figures 2-8 and extended data figure 10 are provided with the paper.

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	Sex was reported by the infant participant's parent/guardian and confirmed via medical records. The expression of the XIST gene in the RNASeq data was used as a way to ensure the correct sex was assigned.
Reporting on race, ethnicity, or other socially relevant groupings	No data on race, ethnicity or other socially relevant groupings is reported.
Population characteristics	Male and female healthy, term, vaginally-born infants were recruited for this study shortly after birth and were followed for the first 15 months of life. Exclusion criteria were <37 weeks' gestation, maternal Body Mass Index (BMI) of >30 (at first trimester antenatal visit), maternal sepsis (defined by laboratory confirmed bacterial infection in blood cultures or CSF) during pregnancy, infant delivered by caesarean section, confirmed sepsis in infant (defined by laboratory confirmed bacterial infection in blood cultures or CSF), a known or suspected disorder of the immune system that would prevent an immune response to the vaccines, such as participant with congenital or acquired immunodeficiency or those receiving systemic immunosuppressive therapy, infant had suspected or confirmed HIV, major congenital abnormality or serious illness, maternal or infant participation in a clinical study that may interfere with participation in this study, anything that would place the individual at increased risk or preclude the individual's full compliance with, or completion of, the study. For immunogenicity analyses, only participants who received at least one vaccine at study visit 1 were included.
Recruitment	Infants were recruited at the Women's and Children's Hospital (WCH) Adelaide, Australia between April 2017 and March 2021 as part of the Antibiotics and Immune Responses (AIR study, a prospective, observational clinical study (ACTRN12617000856314). Potential participants were approached by study research nurses in postnatal wards and invited to participate if the mother and their infant met the study eligibility criteria. Study information had also been previously provided to women delivering at the Women's and Children's Hospital through pamphlets in antenatal care packs. Following written informed consent, stool collection packs were provided to families and medical case notes were reviewed to confirm eligibility criteria and antibiotic exposure status. Ethics approval for the study was obtained from Human Research Ethics Committee of the Women's and Children's Health Network (approval number HREC/17/WCHN/19) and the authors affirm that they have complied with all relevant ethical regulations.
Ethics oversight	Ethics approval for the study was obtained from Human Research Ethics Committee of the Women's and Children's Health Network (approval number HREC/17/WCHN/19) and the authors affirm that they have complied with all relevant ethical regulations.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Human study - We designed and powered the study to describe the impact of antibiotic exposure on the primary outcomes of proportion of infants achieving seroprotective PCV antibody levels at 7 months (following 3 doses of 13vPCV vaccination at 2, 4 and 6 months). Using an assumption that the proportion achieving PCV antibody levels of ≥0.35 ug/ml to common PCV serotypes would be at least 88% in the control

(non-exposed) group, we require 89 individuals per analysis group (exposed vs not exposed) to detect a 20% or greater difference (alpha 0.05) with 90% power. Infants were consented into this study in the first week of life, with the aim of including approximately 200 vaccinated infants (vaccinated cohort) at approximately 6 weeks of age (Study Visit 1). At this visit, we classified participants as either antibiotics exposed, or unexposed. We designed the study to enrol and vaccinate 200 infants with the aim of achieving approximately 100 infants enrolled and vaccinated with no prior antibiotic exposure by 6 weeks of age, and approximately 100 infants enrolled and vaccinated with either direct or maternal intrapartum antibiotic exposure prior to vaccination at 6 weeks of age. This target sample size of 200 allowed for withdrawals and/ or insufficient blood collections for approximately 10%, while ensuring we were adequately powered to detect the difference between exposed and not exposed groups in achieving PCV antibody levels of ≥0.35 ug/ml to common PCV serotypes.

Animal study - Sample size was determined based upon prior experiments evaluating the role of the microbiota in vaccine responses previously undertaken by our team (See Lynn et al. Cell Host & Microbe, 2018, DOI:https://doi.org/10.1016/j.chom.2018.04.009)

Data exclusions

In Figure 4c-d metagenomics data from 13 infants who were withdrawn from the study by study visit 1 were excluded from further analysis. In the analysis shown in Extended Data Fig. 2, only infants with matched samples across 3 timepoints were included in the analysis. No other data that passed QC were excluded from analysis

Replication

The human AIR study was not replicated. Technical replicates were performed on infant samples were sample volume limitations permitted. Replication of data from animal model studies is reported in the figure legends. Replication attempts were successful where animal experiments were performed without issue (e.g. no contamination of germ-free mice). Any experiments where the germ-free status of control mice was breached were excluded and are not included in the manuscript.

Randomization

The clinical study was a prospective, observational study therefore no randomisation was performed. In the animal experiments mice were randomised to treatment groups.

Blinding

In the clinical study, investigators were blinded to antibiotic group exposure allocation during sample collection, processing and analysis. Unblinding occured when all exposure, outcome and covariate data has been entered, second checked for accuracy and the database has been cleaned for spurious or discrepant data (with data queries resolved). Investigators were not blinded to the treatment groups in the animal experiments.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☐	☒ Animals and other organisms
☐	☒ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☐	☒ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used

Infant flow cytometry analysis: Panel 1 was a pan-leukocyte panel allowing for identification neutrophils, monocytes, dendritic cells, eosinophils, basophils, T cells, B cells, and NK cells. Cells were stained with a cocktail of anti-CD3 FITC (HIT3a, 1:20, BD Biosciences), anti-CD11c PE (B-ly6, 1:20, BD Biosciences), anti-HLA-DR APC-H7 (G46-6, 1:40, BD Biosciences), anti-CD15 V500 (HI98, 1:80, BD Biosciences), anti-CD45 BUV395 (HI30, 1:80, BD Biosciences), anti-CD16 BV421 (3G8, 1:143, BD Biosciences), anti-CD14 AF647 (Mop-9, 1:143, BD Biosciences), anti-CD123 BV711 (9F5, 1:143, BD Biosciences), anti-CD56 PE-Cy7 (5.1H11, 1:40, Biolegend), anti-CD19 BV605 (HIB19, 1:80, Biolegend), anti-CD20 BV786 (2H7, 1:80, Biolegend), anti-FcεRIα PerCP Cy5.5 (AER-37, 1:80, Biolegend), and anti-Siglec8 PE Dazzle (7C9, 1:143, Biolegend). Panel 2 enabled the characterisation of ~20 functional subsets of B and T cells including naïve, central memory and effector memory CD4+ and CD8+ T cells; regulatory T cells; and naïve, memory and plasma B cells and used a cocktail of anti-CD19 PE (HIB19, 1:10, BD Biosciences), anti-CCR7/CD197 AF647 (150503, 1:10, BD Biosciences), anti-CD127 BV421 (A019D5, 1:40, Biolegend), anti-CD38 BV510 (HIT2, 1:40, BD Biosciences), anti-HLA-DR APC H7 (G46-6, 1:40, BD Biosciences), anti-CD3 BUV395 (UCHT1, 1:40, BD Biosciences), anti-CD4 BV605 (RPA-T4, 1:80, BD Biosciences), anti-CD8 Cy7 (RPA-T8, 1:80, BD Biosciences), anti-IgD PerCP Cy5.5 (IA6-2, 1:80, BD Biosciences), anti-CD27 BV711 (L128, 1:143, BD Biosciences) and anti-CD25 PE Dazzle594 (M-A51, 1:80, Biolegend), anti-CD20 BV786 (2H7, 1:80, Biolegend), anti-CD45RA FITC (HI100, 1:143, Biolegend). Panel 3 was designed to assess specific T helper subsets, including Th1, Th2, Th9, Th17, Th22 and NKT cells using anti-CD3 BUV395 (UCHT1, 1:50, BD Biosciences), anti-CD4 BV605 (RPA-T4, 1:100, BD Biosciences), anti-CD8 PE Cy7 (RPA-T8, 1:100, BD Biosciences), anti-CXCR3 BV421 (1C6/CXCR3, 1:100, BD Biosciences), anti-CCR10 APC (1B5, 1:100, BD Biosciences), anti-CCCR6 BV711 (11A9, 1:50, BD Biosciences), anti-CCR4 PECF594 (1G1, 1:50, BD Biosciences), anti-CD45RA FITC (HI100, 1:100, Biolegend) and CD1d-Tet PBS57 PE (1:5,000, NIH)

Murine flow cytometry analysis:
GC/Tfh assessment: CD3 (142-C11, 1:300, BD Biosciences), CD4 (RM4, 1:300, BD Biosciences), CD8 (53-6.7, 1:600, BD Biosciences), CD19 (1D3, 1:300, BD Biosciences), CD38 (90, 1:600, Biolegend), CD44 (IM7, 1:600, BD Biosciences), CD62L (MEL-14, 1:600, BD

Biosciences), CD95 (J02, 1:300, BD Biosciences), CD138 (281-2, 1:600, BD Biosciences), B220 (RA3-6B2, 1:600, BD Biosciences), CXCR5 (L138D7, 1:200, Biolegend), GL7 (GL7, 1:600, BD Biosciences), IgD (11-26c.2a, 1:300, Biolegend), IgM (R6-60.2, 1:300, BD Biosciences), PD-1 (RMP1-30, 1:300, BD Biosciences), streptavidin-PECF549 (1:400, Biolegend).

CD86, CD80 and MCHII activation: CD19 (ID3, 1:200, BD Biosciences), CD3 (145-2C11, 1:150. BD Biosciences), CD8 (53-6.7, 1:200, BD Biosciences), NK1.1 (PK136, 1:200, BD Biosciences), Ly6G (1A8, 1:600, BD Biosciences), B220 (RA3-6B2, 1:3000, BD Biosciences), CD11b (M1/70, 1:500, BD Biosciences), CD11c (HL3, 1:400, BD Biosciences), MHCII (M5/114.15.2, 1:150, Miltenyi), CD86 (PO3.3, 1:150, Miltenyi), CD80 (16-10A1, 1:150, Tonbo Biosciences, San Diego, USA).

Immunofluorescence staining: CD16/32 (1:100, BD Biosciences), CD21/35 clone 7E9 (1:200, Biolegend), GL7 (1:100, Biolegend) CD3 clone 17A2 (1:200, Biolegend), IgD clone 11-26c.2a (1:200, Biolegend).

Validation

No specific validation experiments were conducted in house. All antibodies were commercially sourced and validated by the companies. A certificate of analysis and technical data sheet is available for each antibody utilised from the suppliers website.

Animals and other research organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research, and Sex and Gender in Research

Laboratory animals	C57BL/6J SPF and GF mice. Young mice (aged ~3 weeks), dams (aged 8-16 weeks) and pups from birth to 21 days were used as described in the manuscript.
Wild animals	This study did not involve wild animals.
Reporting on sex	Equal numbers of each sex were incorporated into the experimental design for this study.
Field-collected samples	This study did not involve field collected samples.
Ethics oversight	SAHMRI Animal Ethics Committee.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Clinical data

Policy information about clinical studies

All manuscripts should comply with the ICMJE guidelines for publication of clinical research and a completed CONSORT checklist must be included with all submissions.

Clinical trial registration	ACTRN12617000856314.
Study protocol	This study was an observational study and not a randomised controlled trial, as such the study protocol was not published in advance.
Data collection	Infants were recruited at the Women's and Children's Hospital (WCH) Adelaide, Australia between April 2017 and March 2021.
Outcomes	The primary objective of the AIR study was to determine whether antibody responses to the PCV13 vaccine at 7 months of age are different between antibiotic exposed and unexposed groups. Specifically comparing the following: 1) Percentage of participants achieving a seroprotective antibody response (greater than or equal to 0.35 micrograms/ml) against pneumococcal polysaccharide serotypes in the PCV13 vaccine at approximately 7 months of age. 2) Geometric mean concentrations in micrograms/ml of anti-PPS IgG as measured by serum assay at approximately 7 months of age. As secondary outcomes, we also compared PCV13 vaccine responses at 15 months of age, as well as antibody responses to the Infanrix Hexa® 6-in-1 vaccine (6-in-1, DTPa-HepB-IPV-Hib), and the Rotarix® oral rotavirus vaccine (ORV) in antibody exposed and unexposed infants. Multivariable linear regression adjusting for baseline pre-vaccination titres and sex was used to test for an association between antibiotic exposure and IgG GMCs for each of the vaccine antigens assessed. Multivariable logistic regression was used to test for an association between antibiotic exposure and the proportion of infants achieving a seroprotective response. The seroprotective threshold for each vaccine antigen is listed in Table S1. Statistical significance was defined was $P < 0.05$.

Plants

Seed stocks: N/A

Novel plant genotypes: N/A

Authentication: N/A

Flow Cytometry

Plots

Confirm that:

☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☒ All plots are contour plots with outliers or pseudocolor plots.

☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation

For human flow cytometry analysis, 450 µl of infant blood was transferred from a BD Vacutainer K3 EDTA tube into a sterile 10 mL tube within 3 hours of collection and red blood cells were lysed with addition of 4.5 mL of 1X BD Pharm Lyse for 15 min at room temperature. Cells were then pelleted by centrifugation at 200 x g for 5 min, supernatant aspirated and cell pellet resuspended in 10 mL flow cytometry buffer and centrifuged again. Cells were then resuspended in 150 µl flow cytometry buffer and divided equally into 3 tubes for staining. Fc receptors were first blocked Human TruStain FcX (Biolegend, San Diego, USA) in Brilliant Stain Buffer (BD) for 10 min on ice. Three panels were designed for same day flow cytometry analysis (see Supplementary Table 3 for complete list of cell subsets identified on each panel). After each antibody cocktail was added, samples were incubated in the dark on ice for 30 min before being washed and resuspended in 300 µl of flow cytometry buffer. Panel 1 was added to a BD Trucount FACS tube (BD), while panels 2 and 3 were added to round bottomed 5 mL FACS tubes alongside 10 µl liquid counting beads (BD). Counts were generated as previously described 49. Dead cells were stained by addition of DNA binding dye DAPI prior to running on a flow cytometer and were excluded from analysis.

Mouse flow cytometry analysis:
Single cell suspensions of murine spleen or lymph nodes were prepared by mechanical dissociation and filtration through a 70 µm nylon filter (Merck Millipore, Burlington, USA). Cells were stained with Fixable Viability Stain 780 (BD Biosciences) and murine FC blocking antibody Clone 2.4G2 (1:100, BD Biosciences) in PBS for 15 minutes. Cells were washed before staining with fluorochrome-conjugated antibodies in FACS buffer (PBS, 0.1% BSA, 2 mM EDTA) for staining with antibody cocktails from 1 of 2 panels (detailed in methods section). Surface antibody cocktails were added directly after staining for 20 minutes at 4C. Data were acquired on BD FACS Symphony and analyzed using FlowJo v10.

Instrument: FACSymphony™ (BD) flow cytometer

Software: FlowJo™ v10.6.1

Cell population abundance: BD Trucount FACS tubes were used to enumerate cells.

Gating strategy

Gating strategy for the human pan-leukocyte panel. Beads from each Trucount tube were recorded from FSCxSSC and then PEhi. For all analyses, cells were gated by singlet discrimination, followed by live, CD45+ and leukocytes. Leukocytes were split into SSC high and low populations. Eosinophils and neutrophils were gated on SSC low and high populations, respectively. T cells were gated on SSC low CD3hi and CD3- cells were further defined as CD19+CD20+ B cells. The CD19- population was gated on to define CD16-CD56+ NK cells. CD56- cells were further gated on CD16 and CD14 expression to identify monocyte populations. Remaining CD14-CD16- cells were further gated on HLA-DR and CD123 expression to identify CD123+HLA-DR- Basophils. HLA-DR+ cells were gated on CD11c and CD123 in order to identify various DC populations. CD3+ T cells were gated further on their CD56 expression to identify CD56+CD3+ NKT-like cells.

Gating strategy for the human B and T cell panel. For all analyses, cells were gated by singlet discrimination, live and for leukocytes using SSCxFSC. CD3+ T cells were gated on CD4+ and CD8+ expression. CD4+ T cells were then defined as CD45RA-HLA-DR+ CD4+ T cells, CCR7+CD45RA+ Naïve CD4+ T cells, CCR7+CD45RA- CD4+ central memory T cells, CCR7-CD45RA- CD4 effector memory T cells. CD4+ T cells were also gated on CD127 and CD25 expression to identify CD127-CD25+ Tregs, which were further gated for CD45RA and HLA-DR expression to identify CD45RA-HLADR+ Tregs. CD8+ T cells were then defined as CD45RA-HLADR+ CD8+ T cells, CCR7+CD45RA+ naïve CD8+ T cells (naïve T cells), CCR7+CD45RA- CD8+ central memory T cells

(Tcm), and CCR7-CD45RA- CD8 effector memory T cells (Tem). The CD3- population from b was gated further on CD20 and CD19 so that CD20-CD19- cells could be gated out and CD19+ B cells could then be gated for. These were further divided into a CD38 and CD20 gate to identify CD38+CD20- plasma cells and an IgD and CD27 gate to identify IgD-CD27+ Memory B cells and IgD+ naïve B cells.

Gating strategy for the extended human B and T cell panel. Cells were gated by singlet discrimination, live cells and for lymphocytes using SSCxFSC. NKT cells were identified using a tetramer specific for CD1d. CD3+ T cells were gated on CD4+ and CD8+ expression. CD4+ T cells were then defined as CCR4+ CCR3+ Th1 cells and CCR4+ CD4 T cells. This CCR4+ gate was used to define CCR4+CCR6- Th2 cells and CCR4+CCR6+ cells. The CCR4+CCR6+ cells were further gated to identify CCR10 +CCR6+ Th22 cells and CCR10-CCR6+ Th17 cells. Th17 cells could be further distinguished by identifying the CXCR3- population of Th17 cells. CD45RA- T cells were identified using the CD4+ T cell gate. This CD45RA- T cell gate was used to identify CCR4-CXCR3+ Th1 memory cells (mTh1) and a CCR4+ memory population that was further gated to identify CCR4 +CCR6- Th2 memory cells (mTh2). The CD8+ T cell gate was further gated on the CD45RA- population to identify CD8+ memory T cells (mCD8). This population was further gated to identify CCR10+CCR4- CD8+ mTc1 cells. Finally, the mCD8 population was also gated to identify a CCR6+CCR10- population.

Gating strategy for the mouse GC B cells panel. For all analyses, lymphocytes were gated by singlet discrimination, followed by live discrimination. Live cells were gated on B220 and CD19 to identify CD19+ B220+ B cells. These cells were gated on IgD and IgM to identify IgD- B cells. IgD- B cells were gated on GL7 and Fas to identify GL7+Fas+ germinal centre B cells. GC B cells were gated on CRM where biotinylated CRM, (Fina Biosolutions, Rockville, USA) was detected using streptavidin-PECF549 antibody to identify CRM+GC B cells.

Gating strategy for the mouse T follicular helper (Tfh) panel. For all analyses, lymphocytes were gated by singlet discrimination, followed by live discrimination. Live cells were gated on CD3 and CD19 to identify CD3+ cells. CD3+ cells were gated on CD4 and CD8 to identify CD4+ T cells. CD4+ T cells were gated on CD44 and CD62 to identify CD44hiCD62L T cells. CD44hiCD62L T cells were gated on CXCR5 and PD1 to identify CXCR5+PD-1+ Tfh cells.

Gating strategy for the mouse cell activation panel. For all analyses, lymphocytes were gated by singlet discrimination, followed by live discrimination. Live cells were gated on CD19 and CD3 and B220 to gate out B cells. Cells were gated on CD8 to gate out CD8+ T cells. Remaining cells were gated on CD3 to gate out CD4+ T cells. Non- B and T cells were gated on Ly6G to gate out neutrophils and NK1.1 to gate out NK cells. Remaining cells were gated on CD11b and CD11c to identify CD11b +CD11c- Monocytes and CD11c+ DCs. DCs were gated on B220 and MHCII to identify MHCIIloB220+ pDCs and MHCII+B220- cDCs. cDCs were gated on CD8 and CD11b to identify cDC1 (MHCII+B220-CD11b-CD8a+) and cDC2 (MHCII+B220-CD11b +CD8a-) cDC populations. DCs were also gated on CD11c and MHCII to identify (CD11chiMHCIIlo) rDCs and (CD11chiMHCIIhi mDCs).

☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.