MAJOR ARTICLE

Association between maternal breastmilk microbiota composition and rotavirus vaccine response in African, Asian, and European infants: a prospective cohort study

Jonathan Mandolo[1,2,3]*, Edward P. K. Parker[4]*, Christina Bronowski[5], Kulandaipalayam Natarajan C. Sindhu[6], Alistair C. Darby[7], Nigel A. Cunliffe[5,8,9], Gagandeep Kang[6]†, Miren Iturriza-Gómara[10]†, Arox W. Kamng'ona[1,3]†, Khuzwayo C. Jere[1,5,8,11]†, on behalf of the RoVI study group[§]

[1] Virology Research Group, Malawi–Liverpool–Wellcome Trust Clinical Research Programme, Blantyre, 312225, Malawi; [2] Department of Clinical Sciences, Liverpool School of Tropical Medicine, Liverpool, L3 5QA, UK; [3] Department of Biomedical Sciences, School of Life Sciences and Allied Health Professions, Kamuzu University of Health Sciences, Blantyre, 312225, Malawi. [4] Department of Clinical Research, London School of Hygiene and Tropical Medicine,

---

* These authors contributed equally

† These authors jointly supervised this work

Corresponding authors: Khuzwayo C. Jere: Department of Infection Biology and Microbiomes, Institute of Infection, Veterinary and Ecological Sciences, University of Liverpool, Liverpool, L69 7BE, UK, Email: khuzwayo.jere@liverpool.ac.uk

Alternate corresponding authors: Edward P. K. Parker: Department of Clinical Research, London School of Hygiene and Tropical Medicine, London, WC1E 7HT, UK, Email: edward.parker@lshtm.ac.uk

[§] Other members of the RoVI study group: Sudhir Babji, Blossom Benny, Noelia Carmona-Vicente, Nedson Chasweka, End Chinyama, Queen Dube, Sidhartha Giri, Nicholas C. Grassly, Annai Gunasekaran, Deborah Howarth, Sushil Immanuel, Jenna Lowe, Ira Praharaj, Bakthavatsalam Sandya Rani, Sophia Silas, Vivek Kumar Srinivasan, Mark Turner, Srinivasan Venugopal, Valsan Philip Verghese.

London, WC1E 7HT, UK; 5 Department of Clinical Infection, Microbiology and Immunology, Institute of Infection, Veterinary and Ecological Sciences, University of Liverpool, Liverpool, L69 7BE, UK; 6 Wellcome Trust Research Laboratory, Division of Gastrointestinal Sciences, Christian Medical College, Vellore, Tamil Nadu, India; 7 Department of Infection Biology and Microbiomes, Institute of Infection, Veterinary and Ecological Sciences, University of Liverpool, Liverpool, L69 7BE, UK; 8 National Institute for Health and Care Research Health Protection Research Unit in Gastrointestinal Infections, University of Liverpool, L69 7BE, UK; 9 National Institute for Health and Care Research Global Health Research Group on Gastrointestinal Infections, University of Liverpool, L69 7BE, UK; 10 Centre for Vaccine Innovation and Access, Program for Appropriate Technology in Health (PATH), Geneva, 1218, Switzerland; 11 Department of Medical Laboratory Sciences, School of Life Sciences and Allied Health Professions, Kamuzu University of Health Sciences, Blantyre, 312225, Malawi

Background. Maternal breastmilk is a source of pre- and pro-biotics that impact neonatal gut microbiota colonisation. Since oral rotavirus vaccines (ORVs) are administered at a time when infants are often breastfed, breastmilk microbiota composition may have a direct or indirect influence on vaccine take and immunogenicity.

Methods. Using standardised methods across sites, we compared breastmilk microbiota composition in relation to geographic location and ORV response in cohorts prospectively followed up from birth to 18 weeks of age in India ($n = 307$), Malawi ($n = 119$), and the UK ($n = 60$).

Results. Breastmilk microbiota diversity was higher in India and Malawi than the UK across three longitudinal samples spanning weeks of life 1 to 13. Dominant taxa such as *Streptococcus* and *Staphylococcus* were consistent across cohorts; however, significant geographic differences were observed in the prevalence and abundance of common and rare genera throughout follow-up. No consistent associations were identified between breastmilk microbiota composition and ORV outcomes including seroconversion, post-dose 1 vaccine shedding, and/or post-vaccination rotavirus-specific IgA level.

Conclusions. Our findings suggest that breastmilk microbiota composition may not be a key factor in shaping trends in ORV response within or between countries.

Keywords. breastmilk; rotavirus; microbiota; immunogenicity

INTRODUCTION

Breastmilk is a key source of nutrients and contains immunoglobulins, growth hormones, and oligosaccharides that are critical to infant gut homeostasis and immune development [1]. The *Bifidobacteriaceae, Pseudomonadaceae, Streptococcaceae, Enterococcaceae* and *Staphylococcaceae* families have consistently been identified as core constituents of the

breastmilk microbiota [2–4]. These and other breastmilk taxa may act as a source of commensal bacteria for the developing gut microbiota [5]. Geographic region, delivery mode, and maternal health are among the factors associated with breastmilk microbiota composition [2,3,6–9].

More than 100 countries have incorporated oral rotavirus vaccine (ORV) into their national immunization programs [10]. Malawi and the UK introduced the live-attenuated G1P[8] Rotarix vaccine into their national immunisation programs in 2012 and 2013, respectively [11,12]. India introduced a live-attenuated G9P[11] vaccine into its immunisation program in 2016 [13]. These vaccines have reduced the burden of rotavirus, although in India and Malawi the estimated mortality burden due to rotavirus remained significant as of 2016 (9.2 and 31.2 per 100,000, respectively, compared to 0.1 per 100,000 in England) [14]. ORV immunogenicity and efficacy is significantly reduced in low- and middle-income countries (LMICs) compared with high-income countries [15]. Given that LMICs account for approximately 95% of rotavirus deaths worldwide [16], the public health burden associated with impaired ORV response is considerable.

Several mechanisms may contribute to the impaired performance of ORV in LMICs. In Malawi and India, we reported infant gut microbiota diversity to be negatively correlated with ORV response [17,18]. Maternal rotavirus-specific IgG and IgA antibodies in breastmilk and serum were also negatively correlated with ORV response [17,19], although similar correlations were absent in the UK [17]. Other factors which may impact ORV response include histo-blood group antigen status, environmental enteric dysfunction (EED), and pre-vaccination rotavirus exposure [20].

We hypothesised that breastmilk microbiota composition may be associated with ORV response, either by directly interacting with the vaccine viruses or indirectly via the developing infant gut microbiota. We tested this hypothesis using standardised methods across cohorts in Malawi, India, and the UK [17,18].

MATERIALS AND METHODS

Study cohort

This is a follow-up to the Rotavirus Vaccine Immunogenicity (RoVI) study – a multi-site cohort study exploring the impact of maternal antibodies, microbiota development, and EED on ORV response (CTRI/2015/11/006354). The study design, sample handling, lab assays, and primary outcomes of the study have been described previously [17,18]. Briefly, pregnant women were recruited across sites in Blantyre (Malawi), Vellore (India), and Liverpool (UK). Infants received routine immunisations including two doses of Rotarix according to the national immunisation schedule at each study site (weeks of life 6 and 10 in India and Malawi; weeks of life 8 and 12 in the UK). Rotavirus-specific IgA (RV-IgA) was measured in infant blood samples collected pre-

and 4 weeks post-vaccination. Rotavirus shedding was measured in six longitudinal stool samples per infant, including 1 week after each ORV dose. Breastmilk samples were collected in week of life 1 and 1 week after each ORV dose (**Figure 1A**).

Sample processing

Breastmilk samples were collected in sterile sample pots by participants and transferred to the site-specific laboratory by courier within 24 hours (refrigerated throughout) in the UK or within 4 hours in India and Malawi. Upon receipt, samples were kept at 4°C for a maximum of 8 hours until processing and subsequently stored in 2 ml aliquots in SuperLock tubes (Starlab) at −70 °C for a maximum of 2 weeks prior to DNA extraction. DNA was extracted from 1 ml of breastmilk but otherwise followed the stool-specific protocol previously described [17]. A negative extraction control was included in each extraction batch. DNA extracts from Malawi and India were shipped on dry ice to the University of Liverpool for library preparation and sequencing.

Microbiota sequencing

Breastmilk microbiota composition was determined by sequencing the 16S rRNA gene V3–V4 region. Amplicon generation, library preparation, and sequencing were performed as previously described for stool [17], but with 15 cycles (vs 10) for the initial amplicon PCR and 20 cycles (vs 15) for subsequent indexing PCR to ensure robust amplification from the low-biomass samples. We sequenced amplicons for 1,301 separate breastmilk samples (894 from India, 275 from Malawi, and 132 from the UK) across 6 Illumina HiSeq2500 lanes (v2 chemistry with 600 cycles in rapid run mode). Samples from each participant were processed on the same plate. Sequencing was batched by geographic location according to sample availability. Each PCR plate included: a no-template PCR control; a breastmilk control sample provided by a mother in the UK who was not enrolled in the study; DNA from a mock community (Zymo Research D6306); and a pool of extraction controls corresponding to the samples contained on each plate. Due to shipment challenges, extraction controls corresponding to 141/243 (58%) of samples from Malawi were included in the extraction pools. To better define the amplicon profile of extraction controls, we sequenced a further 49 pools containing 1–5 controls from extraction batches performed in India or the UK. Final libraries contained up to four 96-well PCR plates (384 amplicons). Breastmilk DNA samples were amplified on separate plates to stool samples, though we allowed mixing of stool and breastmilk PCR plates in a given library. To validate the robustness of the sequencing protocol, 90 breastmilk DNA samples (30 per cohort, all collected in week of life 1) were transferred to Imperial College London and sequenced according to the methods above with minor modifications, as previously described [17].

Bioinformatic processing

Adapters were trimmed from raw sequences using cutadapt version 1.18 [21]. We merged, filtered, and denoised the amplicon sequences using the DADA2 pipeline in QIIME2 (version

2018.11) [22]. Forward and reverse reads were truncated to 270 bp and 200 bp, respectively. Taxonomic assignment was performed via the *dada2* package (version 1.14.1) using the RDP naïve Bayesian classifier trained on the Silva rRNA database (version 132). Additional data management and filtering steps were performed using the *phyloseq* package in R (version 1.38.0). Ribosomal sequence variants (RSVs) were retained if they were 390–440 bp in length (given an amplicon length distribution after primer trimming with peaks at 400–410 bp and 420–430 bp), assigned as bacterial, detectable at ≥0.1% abundance in at least one sample, and passed frequency-based contamination filtering using the *decontam* package in R (version 3.6.1) [23]. Nanodrop readings (ng/μl) were used to define concentration of the input template.

Given the additional amplification involved in library preparation for breastmilk samples, reads were frequently detected in extraction controls (n = 56 individual or pooled controls with >10,000 reads after the filtering steps above). Several additional filtering steps were therefore included. First, we retained RSVs if they were detectable at ≥0.1% abundance in ≥1% of breastmilk samples from at least one country. Second, we applied prevalence-based filtering using the *decontam* package with a p value threshold of 0.05 to exclude RSVs that were more common in extraction controls. Finally, we removed samples if their mean Bray–Curtis distance (based on either weighted or unweighted metrics) from breastmilk extraction controls was smaller than their mean distance from other breastmilk samples collected from the same country (**Supplementary Figure 1**).

Outcomes

We compared breastmilk microbiota composition by country and ORV response. Our primary indicator of ORV response was seroconversion status – defined as a 4-fold increase in RV-IgA concentration or detection of antibodies at ≥20 IU/ml in previously seronegative infants. Secondary outcomes included post-vaccination RV-IgA concentration and rotavirus shedding 1 week after the first dose of ORV (as an indicator of vaccine virus take). Shedding was detected via real-time PCR targeting the Rotarix *NSP2* gene [24]. We also performed an exploratory analysis of alpha and beta diversity to identify demographic and clinical factors associated with breastmilk composition.

Statistical analysis

Analyses were performed in the programming language R following the statistical pipeline previously described for stool samples with minor modifications [17]. Alpha and beta diversity were calculated at a rarefaction depth of 15,000 sequences per sample. We performed cross-sectional analyses of alpha diversity via analysis of variance (ANOVA), logistic regression (binary ORV outcomes), Pearson's r with two-sided hypothesis testing (log-transformed RV-IgA), and linear regression (exploratory covariates). Pearson's r was also used to compare alpha diversity in paired breastmilk and infant stool samples, with data for the latter obtained as

previously described [17]. We assessed beta diversity using permutational multivariate ANOVA (PERMANOVA) with 999 permutations based on genus-level unweighted Bray–Curtis distances. For binary outcomes, discriminant genera and RSVs were identified via two-sided Fisher's exact test (differences in prevalence) and Aldex2 (two-sided Wilcoxon rank-sum test of centred log-ratio transformed sequence counts), with taxa classified as discriminant if they had a p value of <0.05 based on either method after Benjamini–Hochberg false discovery rate (FDR) adjustment. Aldex2 was used to identify taxa correlated with log-transformed RV-IgA (FDR-adjusted p value of <0.05 based on two-sided Spearman's rank test). Taxa were included if they were detected with a prevalence of >5% in at least one of the groups being compared. We supplemented cross-sectional analyses with longitudinal mixed-effects models of Shannon index and taxon abundances (zero-inflated negative binomial models of genus-level sequence counts), including week of life as a covariate and study ID as a random effect. Genera were included in longitudinal models if they were present in 20% of samples in a given country.

We applied Random Forests to predict country and ORV outcome based on genus or RSV relative abundances. For each analysis, we performed 20 iterations of 5-fold cross-validation. For binary outcomes, we standardised the baseline accuracy at 50% by fitting each iteration on a random subset of 50 samples per group (or the number of samples in the minority group if this was <50). Models were excluded if there were <10 samples in the minority group. For regression models, accuracy was quantified using linear regression to determine the out-of-bag R^2 values for predicted vs observed RV-IgA values.

For positive controls and technical replicates, we used linear regression (alpha diversity and common genera abundances) and PERMANOVA (beta diversity) to quantify the proportion of variance explained by sample ID.

Raw sequence data for this study have been deposited in the European Nucleotide Archive (accession code PRJEB38948). Processed data and analysis code are available on Github (https://github.com/eparker12/RoVI).

Ethics approval

The study was approved by the Institutional Review Board at the Christian Medical College in Vellore (IRB No. 9472/24.06.2015), the College of Medicine Research and Ethics Committee in Blantyre (P.01/16/1853), and the North West—Liverpool Central Research Ethics Committee in Liverpool (15/NW/0924).

RESULTS

Study cohort

Overall, 664 mother–infant pairs (395 in India, 187 in Malawi, and 82 in the UK) were enrolled in the study and the primary endpoint (measurement of seroconversion or dose 1 shedding) was reached for 484 (307 in India, 119 in Malawi, and 60 in the UK). Baseline characteristics, wild-type rotavirus infection status, EED biomarker levels, and infant stool microbiota composition have previously been compared by country and ORV outcome [17]. Exclusive breastfeeding was reported by 265/307 (86%) mothers in India, 108/119 (91%) in Malawi, and 26/60 (43%) in the UK, with partial breastfeeding reported by a further 32/307 (10%) in India, 11/119 (9%) in Malawi, and 20/60 (33%) in the UK. Exclusive breastfeeding was positively correlated with ORV seroconversion and post-vaccination infant RV-IgA levels in India but not in other cohorts. Breastmilk RV-IgA levels were negatively correlated with infant RV-IgA levels in India and Malawi [17].

ORV shedding and immunogenicity

As previously reported [17], seroconversion was observed in 27/51 (53%) infants in the UK, 85/305 (28%) in India, and 24/103 (23%) in Malawi. Rotavirus shedding 1 week after the first dose of ORV was detected in 55/60 (92%) infants in the UK, 82/305 (27%) in India, and 56/101 (55%) in Malawi. Geometric mean concentrations (GMCs) of RV-IgA (IU/ml) after vaccination were 27 (17–45) in the UK, 20 (95% CI 16–25) in India, and 9 (6–12) in Malawi.

Indian infants were characterised by high rates of neonatal rotavirus infection, defined as detection of wild-type rotavirus shedding in week 1 of life or baseline seropositivity (pre-vaccination RV-IgA ≥20 IU/ml). This was observed in 166/304 (55%) infants in India, 10/90 (11%) in Malawi, and 2/54 (4%) in the UK. Given the potential impact of neonatal infection on ORV response [17], we report results for the Indian cohort overall and stratified by neonatal infection status below.

Geographic differences in breastmilk microbiota composition

Of 1,301 breastmilk samples sequenced, 1,124 yielded high-quality microbiota profiles (≥15,000 sequences after quality filtering; $95,075 \pm 113,894$ [mean±s.d.] sequences per sample). Microbiota profiles of positive controls and technical replicates were consistent across sequencing runs and facilities (**Supplementary Figure 2**).

There were marked differences in breastmilk microbiota composition between cohorts. Microbiota diversity was significantly lower in UK than both other cohorts at all timepoints. Diversity was similar in India and Malawi at week of life 1, but higher in India than Malawi at weeks of life 7 and 11 (**Figure 1B**). Samples clustered by individual (PERMANOVA R^2 = 49%, p = 0.001), with country accounting for 7–9% of variation depending on age (**Figure 1C**).

Although 350 genera were detected overall, a small proportion were dominant in each cohort (**Figure 1C** and **Supplementary Figure 3**). Among dominant genera, *Streptococcus* was depleted in India compared with both other cohorts, while *Acinetobacter* and *Corynebacterium* were enriched (FDR $p<0.05$ based on longitudinal models; **Supplementary Figure 4**). *Staphylococcus* followed a parallel trajectory in each cohort, peaking in week of life 1, and was less abundant in Malawi than both other cohorts. *Bifidobacterium* was less abundant in the UK, reflecting the pattern previously reported for stool samples [17], while *Gemella* was enriched in this cohort (**Figure 1D**). Several of these discrepancies in dominant genera were also evident in cross-sectional analyses of prevalence and/or abundance (**Supplementary Table 1**).

Additional discriminant taxa detected by longitudinal and cross-sectional models are reported in **Supplementary Figure 4** and **Supplementary Table 1**. Based on longitudinal models, 17 genera were enriched in India compared with both other cohorts, including nine Proteobacteria (e.g. *Aeromonas* and *Alishewenalla*), three Firmicutes (e.g. *Enterococcus* and *Aerococcus*) and five Actinobacteria (e.g. *Dermacoccus*). Nine genera were enriched in Malawi compared with both other cohorts, including the Bacteroidetes genus *Prevotalla* 9 alongside eight Firmicutes (e.g. *Faecalibacterium* and *Lachnospiraceae*). Three genera – *Gemella*, *Haemophilus*, and *Enterobacter* – were enriched in the UK compared with the other cohorts.

Random Forests discriminated samples by country with high accuracy (median cross-validation accuracies of 85–95%; baseline accuracy 50%; **Supplementary Figure 5**). Genera underlying model accuracy (based on mean importance scores) were consistent with the discriminant taxa described above (**Supplementary Table 1**).

We also assessed alpha and beta diversity of breastmilk samples in relation to individual-level variables measured in each cohort (**Figure 2**). With the exception of infant serum α1 acid glycoprotein level (a marker of systemic inflammation), which was modestly associated with beta diversity in Malawian samples (R^2 3.7%), no covariates were significantly associated with breastmilk microbiota composition. Alpha diversity in paired breastmilk and infant stool samples was not significantly correlated at week of life 1 or at the time of either ORV dose in any cohort (**Supplementary Figure 6**).

Breastmilk microbiota composition versus ORV response

Based on longitudinal models of Shannon index, we observed no significant differences in microbiota diversity according to seroconversion status in any cohort (**Figure 3A**). This was also the case for cross-sectional analyses, with the exception of comparisons in Malawi at week 7 of life (the week after the first dose of ORV), wherein Shannon index was negatively correlated with seroconversion. Beta diversity analyses based on genus-level unweighted Bray–Curtis distances did not reveal any significant association between breastmilk microbiota composition and seroconversion status (**Figure 3B**). Likewise, Random Forest models based on genus or

RSV abundances failed to accurately predict seroconversion (**Figure 3C**), and no discriminant taxa were identified based on cross-sectional analyses of prevalence or abundance after FDR correction. Longitudinal models of common genera (≥20%) revealed frequent age-associated changes in taxon abundance but only one significant association with seroconversion (a negative correlation between *Alloprevotella* abundance and seroconversion in Malawi; **Supplementary Table 2**).

Cross-sectional analyses of secondary ORV endpoints, including post-vaccination RV-IgA concentration (Supplementary Figure 7) and dose 1 ORV shedding (Supplementary Figure 8) mirrored those for seroconversion, revealing no consistent associations. In the UK, beta diversity at week of life 9 was modestly associated with post-vaccination RV-IgA (R^2 6%, p = 0.005), but this was not the case for any other time point or in other cohorts. Very few discriminant genera were identified with respect to these secondary outcomes based on longitudinal models of genus abundance (Supplementary Table 2).

DISCUSSION

Breastmilk is a key source of pre- and pro-biotics that shape infant gut microbiota configuration and immune development. We documented significant differences in breastmilk microbiota composition between Malawi, India, and the UK. However, no consistent differences in breastmilk microbiota composition were observed with respect to ORV response.

Despite the geographic differences in microbiota composition, several genera were dominant across the three cohorts. Together, *Streptococcus, Staphylococcus, Acinetobacter, Bifidobacterium, Veillonella, Gemella, Corynebacterium* and *Pseudomonas* formed approximately 75% of the breastmilk microbiota – consistent with the dominant taxa reported elsewhere [2–4,25]. The relative abundances of these dominant genera changed over time, with *Staphylococcus* declining in abundance while *Streptococcus* and *Veillonella* increased. This is similar to the trajectory in breastmilk microbiota composition reported in Kenya [26]. The infant salivary microbiota is known to be colonised by *Streptococcus* [27,28], such that the continued dominance of *Streptococcus* in maternal breastmilk may partly reflect breastmilk–saliva interplay. Skin-associated genera including *Staphylococcus* and *Corynebacterium* were also among the dominant genera in maternal breastmilk, consistent with previous findings [29].

Breastmilk microbiota diversity was higher in Malawi and India than the UK. This contrasts with discrepancies we reported in stool microbiota diversity, which was higher in Malawi than both India and the UK at week 1 of life but converged over the ensuing 6–8 weeks [17]. In a previous cross-sectional study spanning 11 sites, breastmilk microbiota diversity was highest in rural Ethiopia and lowest in Ghana, with intermediate levels across other sites in Africa, Europe, North America, and South America [8]. *Streptococcus* was more abundant in Malawi and the UK than in India, while *Bifidobacterium* was depleted in the UK compared with both other cohorts.

Prior studies have also highlighted geographically distinct abundance profiles including depletion of *Bifidobacterium* in European compared with African samples [8]. Together, these studies highlight the significant regional variation that occurs in breastmilk microbiota diversity and composition. To delineate overarching global trends (e.g. urban vs rural, high-income vs LMIC), future studies integrating representative data from multiple countries, such as the present, are warranted.

We did not observe consistent associations between breastmilk microbiota composition and ORV response. At the time of the first ORV dose, breastmilk microbiota diversity in Malawi was negatively correlated with ORV seroconversion – a correlation that was also apparent among infant stool samples in this cohort [17]. However, while we reported consistent correlations between diversity and seroconversion among Indian and Malawian infants' stool samples, there was no consistent discrepancy across cohorts in breastmilk. Moreover, we did not observe a significant correlation between the microbiota diversity of paired breastmilk and infant stool samples, suggesting that the negative correlations between microbiota diversity and ORV seroconversion in Malawi are not causally related. A previous study in India documented higher *Enterobacter/Klebsiella* abundance in breastmilk and infant stool samples of infants with symptomatic rotavirus disease compared to those with asymptomatic or no infection [30]. However, no significant discrepancies were observed between neonates with asymptomatic infection and those lacking infection, which is consistent with the lack of association reported here in relation to attenuated viral exposure via ORV.

To our knowledge, this is the first study to explore the link between breastmilk microbiota composition and ORV response. Our study is strengthened by the use of standardised methods across cohorts, including multiple indicators of ORV response. Nonetheless, several limitations of the present study should be considered. Breastmilk samples were collected one week after each ORV dose, so may offer an imperfect proxy for microbiota composition at the time of vaccination. Owing to recruitment challenges in Malawi [17], we fell short of the target sample size in this cohort ($n = 119$ rather than 150), potentially undermining our ability to detect relevant associations in this cohort. Because of their low biomass, breastmilk samples were subjected to extra rounds of amplification to attain adequate material for sequencing, leading to amplification from extraction controls. We accounted for this via stringent filtering of contaminants. Nonetheless, the potential contribution of contamination and site-specific batch effects to the observed trends cannot be discounted.

Our findings suggest that breastmilk microbiota composition may not be a key factor shaping ORV response within or between countries. Other components of human milk would be a valuable focus of future investigation. Human milk oligosaccharides such as lacto-*N*-tetraose have previously been linked with symptomatic rotavirus infection in Indian neonates, possibly via an effect on neonatal G10P[11] rotavirus infectivity [30]. Future studies of the breastmilk metabolome may help discern whether similar factors influence the immunogenicity and efficacy of ORV.

Supplementary materials

Supplementary Figures 1–8 and Supplementary Tables 1–2 are attached.

FIGURE LEGENDS

Figure 1. Geographic differences in breastmilk microbiota composition. (A) Sample collection strategy. (B) Analysis of alpha diversity, based on genus-level Shannon index. Cross-sectional comparisons were performed using ANOVA with post-hoc Tukey tests. Longitudinal comparisons were performed using mixed-effects regressions with false discovery rate correction of pairwise comparisons. (C) Proportion of variation in microbiota composition associated with country, calculated via PERMANOVA using genus-level unweighted Bray–Curtis distances. (D) Longitudinal plot of mean genus abundances. Genera are included if present with a mean relative abundance of ≥5% in at least one country at one or more timepoints. (E) Longitudinal relative abundance plots for major genera by country. Lines show local weighted regression (loess) fits with 95% confidence intervals. IND, India; MLW, Malawi; ns, not significant; †, +2 weeks samples collected at weeks of life 7 and 11 in the UK due to later vaccination schedule; *$p <$ 0.05; **$p = 0.001$; ***$p < 0.0005$.
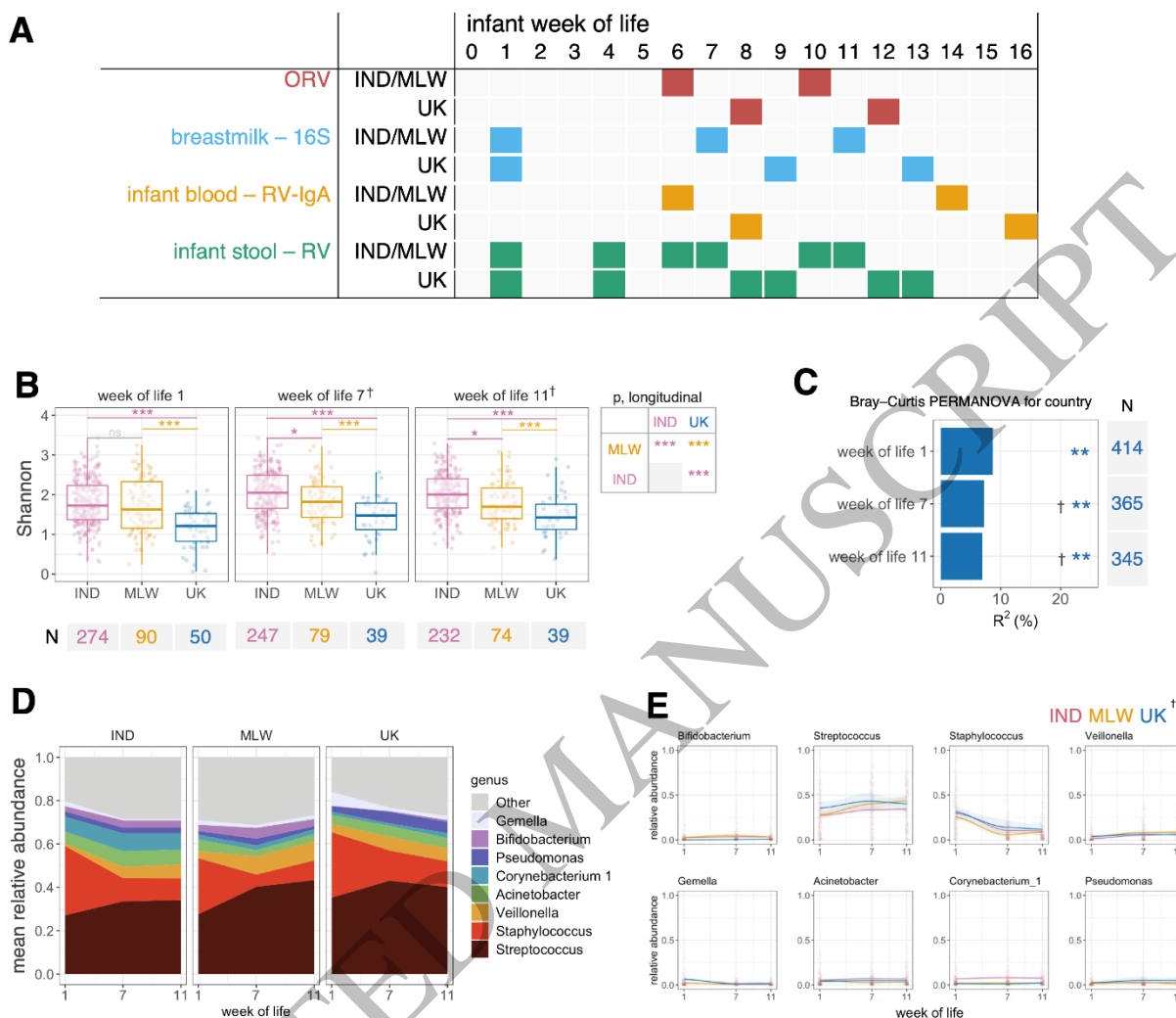
Figure 2. Cofactors associated with breastmilk microbiota composition. Samples collected 1 week after the first dose of oral rotavirus vaccine were included (week of life 7 in India and Malawi; week of life 9 in the UK). The left panel, presenting data for Indian samples (n = 247), contains the full list of exploratory variables (with the exception of HIV exposure status, which was also assessed for Malawi). For analyses of samples from Malawi and the UK (right panels; n = 79 and 39, respectively), variables were excluded if they were not measured or exhibited limited variability (n<10 in either comparison group). PERMANOVA was performed using genus-level unweighted Bray–Curtis distances. Shannon index was calculated at genus level and assessed as an outcome variable via linear regression. A1AT, α1-antitrypsin; α1AG, α1 acid glycoprotein; IND, India; MLW, Malawi; MPO, myeloperoxidase; ORV, oral rotavirus vaccine; RV, rotavirus; * FDR p <0.05.
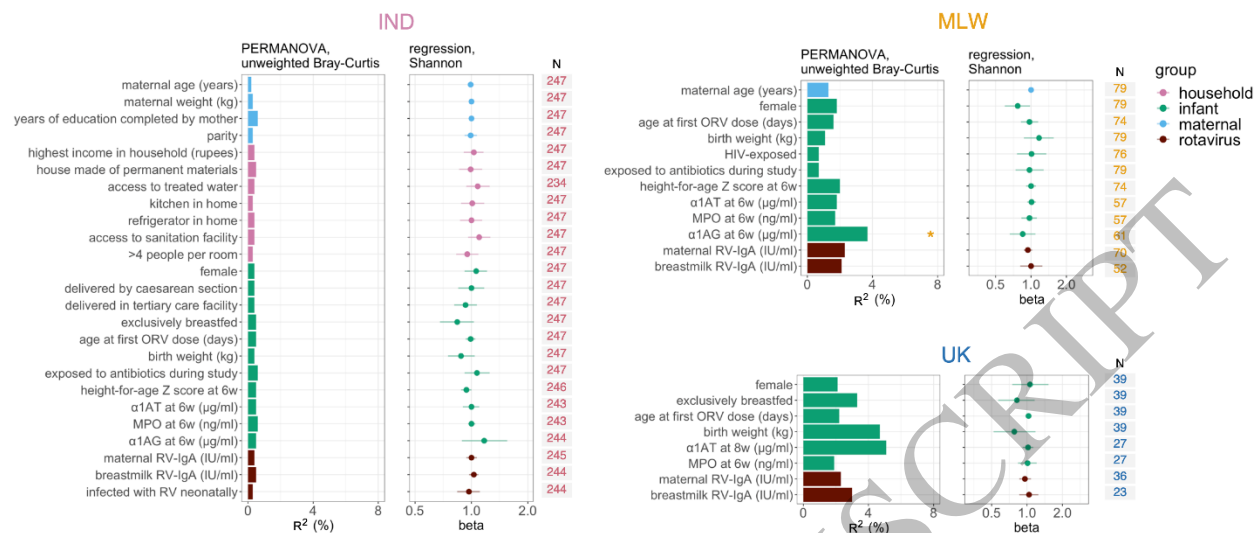
Figure 3. Association between breastmilk microbiota composition and oral rotavirus vaccine seroconversion. (A) Analysis of alpha diversity, based on genus-level Shannon index. Cross-sectional comparisons were performed using logistic regression. Longitudinal comparisons were performed using mixed-effects models. (B) Proportion of variation in microbiota composition associated with seroconversion, calculated via PERMANOVA using genus-level unweighted Bray–Curtis distances. (C) Cross-validation accuracy of Random Forests for prediction of seroconversion. Median out-of-bag accuracy (proportion correctly assigned) and interquartile range across 20 iterations of 5-fold cross-validation are displayed. Each iteration included an equal number of responders and non-responders (50 per group where possible, or else the number in the minority group if this was <50). Taxa were classified as discriminant if they had an FDR-adjusted p value of <0.05 based on either two-sided Fisher's exact test (differences in prevalence) or Aldex2 with two-sided Wilcoxon rank-sum test (differences in abundance). CV, cross-validation; IND, India; MLW, Malawi; neo+, infected with rotavirus neonatally (defined by detection of rotavirus shedding in week of life 1 or baseline seropositivity); neo−, uninfected with rotavirus neonatally; ns, not significant; RF, Random Forests; RSV, ribosomal sequence variant; *p < 0.05.

13

NOTES

Author contributions

Conceptualisation, M.I.G., A.C.D., E.P.K.P., K.C.J., A.W.K., and G.K.; Methodology, J.M., C.B., E.P.K.P., A.C.D., M.I.G., and G.K.; Software, E.P.K.P.; Validation, E.P.K.P.; Formal Analysis, J.M. and E.P.K.P.; Investigation, J.M., C.B., E.P.K.P., A.W.K., and K.C.J.; Data Curation, C.B., E.P.K.P., and J.M.; Writing – Original Draft, J.M. and E.P.K.P.; Writing – Review & Editing, A.C.D., N.A.C., M.I.G, A.W.K. and K.C.J.; Visualisation, E.P.K.P and J.M.; Supervision, N.A.K. and K.J.; Project Administration, C.B., M.I.G., K.N.S., K.C.J., and G.K.; Funding Acquisition, M.I.G., and G.K. All authors read and approved the manuscript.

Conference presentation

Part of this work was presented at the 14th International dsRNA Virus Symposium 2022, Banff, Alberta, Canada, 10th October - 14th October 2022 and the 1st Kamuzu University of Health (KUHeS) Research Dissemination Conference, Blantyre, Malawi, 24th - 25th November 2022

Declaration of Interests

References

1. Bode L. Human milk oligosaccharides: Every baby needs a sugar mama. Glycobiology. **2012**; 22(9):1147–1162.
2. Kumar H, Toit E du, Kulkarni A, et al. Distinct patterns in human milk microbiota and fatty acid profiles across specific geographic locations. Front Microbiol. **2016**; 7(OCT).
3. Li SW, Watanabe K, Hsu CC, et al. Bacterial composition and diversity in breast milk samples from mothers living in Taiwan and Mainland China. Front Microbiol. **2017**; 8(MAY):1–15.
4. Ding M, Qi C, Yang Z, et al. Geographical location specific composition of cultured microbiota and: Lactobacillus occurrence in human breast milk in China. Food Funct [Internet]. **2019**; 10(2):554–564. Available from: https://doi.org/10.1039/C8FO02182A
5. Martín V, Maldonado-Barragán A, Moles L, et al. Sharing of bacterial strains between breast milk and infant feces. J Hum Lact. **2012**; 28(1):36–44.
6. Cabrera-Rubio R, Collado MC, Laitinen K, Salminen S, Isolauri E, Mira A. The human milk microbiome changes over lactation and is shaped by maternal weight and mode of delivery. Am J Clin Nutr. **2012**; 96(3):544–551.
7. Ding M, Qi C, Yang Z, et al. Composition of gut microbiota and its influence on the immunogenicity of oral rotavirus vaccines. Food Funct. Elsevier Ltd; **2019**; 36(24):554–564.
8. Lackey KA, Williams JE, Meehan CL, et al. What's normal? Microbiomes in human milk and infant feces are related to each other but vary geographically: The inspire study. Front Nutr. **2019**; 6(April).
9. Khodayar-Pardo P, Mira-Pascual L, Collado MC, Martínez-Costa C. Impact of lactation stage, gestational age and mode of delivery on breast milk microbiota. J Perinatol. **2014**; 34:599–605.
10. Hallowell BD, Tate J, Parashar U. An overview of rotavirus vaccination programs in developing countries. Expert Rev Vaccines [Internet]. Taylor & Francis; **2020**; 19(6):529–537. Available from: https://doi.org/10.1080/14760584.2020.1775079
11. Bar-Zeev N, King C, Phiri T, et al. Impact of monovalent rotavirus vaccine on diarrhoea-associated post-neonatal infant mortality in rural communities in Malawi: a population-based birth cohort study. Lancet Glob Heal. **2018**; 6(9):e1036–e1044.

12.   Thomas SL, Walker JL, Fenty J, et al. Impact of the national rotavirus vaccination programme on acute gastroenteritis in England and associated costs averted. Vaccine [Internet]. The Authors; **2017**; 35(4):680–686. Available from: http://dx.doi.org/10.1016/j.vaccine.2016.11.057

13.   Nair NP, Reddy SN, Giri S, et al. Rotavirus vaccine impact assessment surveillance in India: Protocol and methods. BMJ Open. **2019**; 9(4):1–8.

14.   Troeger C, Khalil IA, Rao PC, et al. Rotavirus Vaccination and the Global Burden of Rotavirus Diarrhea among Children Younger Than 5 Years. JAMA Pediatr. **2018**; 172(10):958–965.

15.   Walker CLF, Black RE. Rotavirus vaccine and diarrhea mortality: quantifying regional variation in effect size. (Special Issue: Technical inputs, enhancements and applications of the Lives Saved Tool (LiST).). BMC Public Health. **2011**; 11(Suppl. 3):5–11.

16.   Tate JE, Burton AH, Boschi-Pinto C, Steele AD, Duque J, Parashar UD. 2008 estimate of worldwide rotavirus-associated mortality in children younger than 5 years before the introduction of universal rotavirus vaccination programmes: a systematic review and meta-analysis. Lancet Infect Dis [Internet]. Elsevier; **2012**; 12(2):136–141. Available from: https://doi.org/10.1016/S1473-3099(11)70253-5

17.   Parker EPK, Bronowski C, Sindhu KNC, et al. Impact of maternal antibodies and microbiota development on the immunogenicity of oral rotavirus vaccine in African, Indian, and European infants. Nat Commun. Springer US; **2021**; 12(1).

18.   Sindhu KNC, Cunliffe N, Peak M, et al. Impact of maternal antibodies and infant gut microbiota on the immunogenicity of rotavirus vaccines in African, Indian and European infants: Protocol for a prospective cohort study. BMJ Open. **2017**; 7(3):1–10.

19.   Kazimbaya KM, Chisenga CC, Simuyandi M, et al. In-vitro inhibitory effect of maternal breastmilk components on rotavirus vaccine replication and association with infant seroconversion to live oral rotavirus vaccine. PLoS One [Internet]. **2020**; 15(11 November):1–13. Available from: http://dx.doi.org/10.1371/journal.pone.0240714

20.   Parker EPK, Ramani S, Lopman BA, et al. Causes of impaired oral vaccine efficacy in developing countries. Future Microbiol. **2018**; 13(1):97–118.

21.   Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnetjournal. **2011**; 17(1):10–12.

22.   Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. DADA2: High resolution sample inference from Illumina amplicon data. Nat Methods. **2016**; 13(7):581–583.

23.   Davis NM, Proctor DiM, Holmes SP, Relman DA, Callahan BJ. Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data. Microbiome. Microbiome; **2018**; 6(1):1–14.

24.   Gautam R, Mijatovic-Rustempasic S, Esona MD, Tam KI, Quaye O, Bowen MD. One-step multiplex real-time RT-PCR assay for detecting and genotyping wild-type group A rotavirus strains and vaccine strains (Rotarix® and RotaTeq®) in stool samples. PeerJ. **2016**; 2016(1).

25.   Hunt KM, Foster JA, Forney LJ, et al. Characterization of the diversity and temporal stability of bacterial communities in human milk. PLoS One. **2011**; 6(6):1–8.

26.   Maqsood R, Skidmore PT, Holland LA, et al. Dynamic Changes in Breast Milk Microbiome in the Early Postpartum Period of Kenyan Women Living with HIV Are Influenced by Antibiotics but Not Antiretrovirals. Microbiol Spectr. American Society for Microbiology; **2022**; 10(2):1–10.

27.   Aas JA, Bruce J. P, Lauren N. S, Dewh FE, Aas JA. Defining the Normal Bacterial Flora of the Oral Cavity. J Clin Microbiol. **2005**; 43(11):5721–5732.

28. Davé V, Street K, Francis S, et al. Bacterial microbiome of breast milk and child saliva from low-income Mexican-American women and children. Pediatr Res. **2016**; 79(6):846–854.

29. Hunt KM, Foster JA, Forney LJ, et al. Characterization of the diversity and temporal stability of bacterial communities in human milk. PLoS One. **2011**; 6(6):1–8.

30. Ramani S, Stewart CJ, Laucirica DR, et al. Human milk oligosaccharides, milk microbiome and infant gut microbiome modulate neonatal rotavirus infection. Nat Commun [Internet]. Springer US; **2018**; 9(1):1–12. Available from: http://dx.doi.org/10.1038/s41467-018-07476-4