Food & Function

ROYAL SOCIETY OF CHEMISTRY

PAPER

View Article Online
View Journal

Check for updates

Cite this: DOI: 10.1039/d5fo02791h

Enhanced dietary monitoring using fecal genomics for childhood malnutrition interventions

Ammara Aqeel,[a] Najeeha T. Iqbal,[b] Sanam Iram Soomro,[c] Sheraz Ahmed,[c] Teresa K. McDonald,[a] Olivia Osborne,[a] Sharon Jiang,[a] Nolan Ives,[a] Kazi Ahsan,[d,e] Fayaz Umrani,[c] Michael J. Barratt,[d,e] Jeffrey I. Gordon,[d,e] Syed Asad Ali[c] and Lawrence A. David *[a,f]

Received 29th June 2025,
Accepted 4th March 2026

DOI: 10.1039/d5fo02791h

rsc.li/food-function

Ready-to-use therapeutic and supplementary foods (RUTF/RUSF) are a primary treatment for childhood malnutrition in Community Management of Acute Malnutrition (CMAM) programs. However, measuring intervention compliance is labor intensive and unreliable. We applied FoodSeq, a fecal genomic dietary assessment biomarker, in malnourished infants (3–15 months) from Matiari, Pakistan. FoodSeq identified a significant increase in the abundance of DNA from chickpea, a primary RUSF ingredient, during RUSF administration and captured region-specific complementary feeding practices, including age-inappropriate feeding practices such as wide-spread tea consumption. Our findings highlight the potential of dietary genomics as an empirical and scalable tool for compliance monitoring and dietary analysis in community-based malnutrition programs.

Open Access Article. Published on 18 March 2026. Downloaded on 4/8/2026 5:33:14 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

Introduction

Childhood malnutrition affected over 194 million children in 2022 and accounted for 45% of deaths in children under five globally.[1] Driven by a combination of socio-economic, environmental, and biologic factors,[2–6] malnutrition disproportionately impacts low- and middle-income countries (LMICs), leading to long-term health, social, and economic consequences.[1,7] Global treatment regimens utilize Community-Based Management of Acute Malnutrition (CMAM), deploying community health workers (CHWs) to screen and enroll children with malnutrition in an outpatient treatment regimen.[8,9] For severe acute malnutrition (SAM); weight-for-height *z*-score (WHZ) < −3 standard deviations (SD) from median WHO growth standards or mid-upper arm circumference (MUAC) < 115 mm,[10] a weekly prescription of weight-adjusted ready-to-use therapeutic food (RUTF) and systematic antibiotic treatment is provided.[8,9,11] For moderate acute malnutrition (MAM); WHZ between −2 and −3 SD or a MUAC between 115–125 mm (ref. 10) ready-to-use supplemental food (RUSF) is provided and an antibiotic regimen is prescribed for underlying concomitant infection as needed.[12]

RUTF/SFs are formulated as a shelf-stable, single-dose foil package of an energy-nutrient dense paste eaten directly from the packet.[13] Although adherence to RUTF/SF therapy is crucial for adequate weight gain,[14] both socio-economic and dietary factors can impact treatment efficacy. For example, RUTF might be shared among household members or sold for supplementary income,[15,16] or age-inappropriate foods might be used for complementary feeding which can impair weight gain.[17] However, monitoring RUTF compliance involves regular CHW home-visits and/or collection of empty RUTF packets,[8,9,14,18,19] requiring extensive training and resources.[20] Novel approaches for objective monitoring of both RUTF/SF adherence and potential impacts of demographic, socioeconomic and/or dietary factors on compliance are needed. Dietary genomics, specifically FoodSeq, offer an objective, cost-effective alternative, characterizing dietary intake from residual food DNA in human stool[21] and have recently also been applied to verify dietary improvements in a U.S.-based grocery provisioning intervention among children with obesity.[22]

Here, we extend FoodSeq application (*trnL-P6* chloroplast marker for plants[21] and *12SV5* mitochondrial marker for vertebrates[23–25]) to acute malnutrition through a subset of an infant cohort from Matiari, Pakistan,[18] assessing its utility in a resource-limited, community-based public health context. A total of 150 stool samples were included from 60 infants with

[a]*Department of Molecular Genetics and Microbiology, Duke University School of Medicine, Durham, NC 27710, USA. E-mail: lawrence.david@duke.edu*
[b]*Department of Biological and Biomedical Sciences and Department of Pediatrics and Child Health, Aga Khan University Hospital, Karachi, Pakistan*
[c]*Department of Community Health Sciences, Aga Khan University, Karachi, Pakistan*
[d]*Edison Family Center for Genome Sciences and Systems Biology, Washington University School of Medicine, St. Louis, MO 63110, USA*
[e]*Newman Center for Gut Microbiome and Nutrition Research, Washington University School of Medicine, St. Louis, MO 63110, USA*
[f]*Duke Microbiome Center, Duke University School of Medicine, Durham, NC 27710, USA*

Open Access Article. Published on 18 March 2026. Downloaded on 4/8/2026 5:33:14 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

Fig. 1 Study design and FoodSeq-derived characterization of infant dietary composition across development. (a) A cohort of 350 malnourished infants was followed longitudinally in the SEEM study through ages 3–6 months to 15 months.[18] From this cohort, our study obtained stool samples from a subset of 30 infants with moderate acute malnutrition (MAM; weight-for-height *z*-scores (WHZ) ≤ −2) at 3–6 months of age when caregivers received an educational intervention covering WHO recommended breast feeding and complementary feeding practice and at approximately 9 months (Row 1) and another subset of 30 malnourished infants at approximately 10 months of age when caregivers were provided a locally-produced, chickpea-based RUSF (*Acha Mum*) for children and subsequently at approximately 14 months of age (Row 2). Thirty stool samples from healthy infants (WHZ > 0, HAZ > −1.0) between the ages of 12.1–15.2 months were also included in the analysis (Row 3). (b) Bar plot showing the presence counts of top 25 food (plant and vertebrate combined) taxa detected in participant stool samples across all timepoints except >12 months (healthy controls, see Fig. S3c for top 25 taxa in healthy controls). (c) Principal Component Analysis (PCA) of plant dietary composition with biplot indicating relative contributions of top 10 dietary components. Points colored by timepoint (bottom of figure) and compared along PC1 values using linear mixed effect models with gender and PCR batch as fixed effects and participant ID as a random effect. Significant differences as marked.

MAM (3.1–15.2 months of age) and 30 healthy infants (12.1–15.2 months of age)[18] (Fig. 1a). In this study, caregivers were provided breast-feeding and complementary feeding education at home for four weeks after enrollment (infants 3–6 months-of-age) (Fig. 1a). A locally produced RUSF *Acha Mum*, made from chickpea, edible oil, dried milk, and fortified

This journal is © The Royal Society of Chemistry 2026

Open Access Article. Published on 18 March 2026. Downloaded on 4/8/2026 5:33:14 PM. This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

with a vitamin and mineral mix,[18,19] was administered to wasted children of weaning age (mean age of 9.89 ± 0.99 months) for an average duration of 57.8 ± 6.3 days (Fig. 1a).

Materials and methods

Study design and dietary intervention

This study analyzed a subset of 90 participants from the Study of Environmental Enteropathy (SEEM; NCT03588013; 365 malnourished with WHZ ≤ −2, 51 healthy with WHZ > 0, height-for-age *Z*-score [HAZ] > −1.0), a prospective longitudinal study conducted in rural Matiari, Pakistan (March 2016-March 2019)[18] (Fig. 1a; [Created in BioRender. Aqeel, A. (2022) https://BioRender.com/f56s058Wti]). SEEM followed infants from birth to 24 months through a multi-national collaboration between Aga Khan University, Washington University in St. Louis, University of Virginia, and Cincinnati Children's Hospital Medical Center, with funding from the Bill and Melinda Gates Foundation.[18] Our subset comprised three groups: 30 malnourished infants with paired samples at 3–6 and 9 months, 30 malnourished infants with paired samples at 10–14 months, and 30 healthy controls sampled after 12 months of age (exact infant ages corresponded with age at the time of stool collection). This group structure reflected sample accessibility at the time of design of the present study and statistical analyses accounted for this structure as appropriate. Caregivers of malnourished infants received complementary feeding education at enrollment (3–6 months), with anthropometric measurements recorded monthly.[18] Infants maintaining WHZ < −2 at 9 months were provided a two-month intervention of one sachet (100 g) per day of *Acha Mum* at no cost, a locally-produced ready-to-use supplementary food (RUSF; ingredients: chickpea, edible oil, dried milk, sugar, vitamins [A, B1, B2, B3, B5, B6, B9, B12, C, D, E, K, H], zinc, folate, iodine, calcium *etc*. emulsifier).[18] Study staff monitored compliance through weekly home visits, with additional visits as needed.[18] A previous Phase 1 study demonstrated 90.5% intervention adherence (assessed as sachets consumed/provided) monitored through weekly home visits by study staff[18,26] and CHW home-visit RUSF consumption data similarly obtained was evaluated for the cohort in this study. The RUSF intervention was administered for 57.8 ± 6.3 days beginning at 9.89 ± 0.99 months.[27] Non-responders (WHZ < −2 post-intervention) without evidence of celiac disease or other identifiable growth failure etiologies underwent esophagogastroduodenoscopy (EGD) for investigation of underlying pathophysiology.[18]

Stool collection, processing and DNA extraction

Fecal samples were collected by CHWs at 3–6, 9, 10, 14, and ≥12 months (approximately 1 g per participant per timepoint) within 10 minutes of production and transferred to liquid nitrogen pre-charged dry shippers without additives or preservatives (Taylor Wharton, CX-100) for transfer to and storage at −80 °C at Aga Khan University (AKU) research facilities prior to shipping on dry ice to Washington University in St. Louis (WUSTL), USA. At WUSTL, frozen fecal samples underwent pulverization in liquid nitrogen using a sterile mortar and pestle. A 50 mg aliquot of pulverized material was transferred to a 2 mL screw-cap vial. A solution consisting of 500 µL of phenol : chloroform : isoamyl alcohol (25 : 24 : 1), 210 µL of 20% SDS and 500 µL of buffer (200 mM NaCl, 200 mM Trizma base, 20 mM EDTA) containing 500 µL of 0.1 mm diameter zirconia/silica beads was added to each sample, followed by bead beating (Mini-Beadbeater-8; Biospec) to extract DNA. DNA was purified (Qiaquick columns, Qiagen), eluted in 70 µL of Tris-EDTA (TE) buffer, and quantified (Quant-iT dsDNA broad range kit; Invitrogen).

FoodSeq: residual dietary DNA amplification, sequencing and taxonomic assignment

Residual dietary plant DNA and animal DNA in stool samples was assessed using FoodSeq, as previously described.[21,25] Briefly, FoodSeq libraries were generated using a two-step PCR protocol. Primary amplification was conducted using locus-specific primers (trnL *g-h* for plants and 12SV5F/12SV5R76 for animals) with Illumina overhang adapter sequences, a human blocking primer (DeBarba14 HomoB77 for 12SV5), and SYBR Green for qPCR. PCR reactions were performed in a total volume of 10 µl using either AccuStart II PCR SuperMix (12SV5)[25] or previously established conditions (trnL).[21] Cycling conditions included an initial denaturation at 94 °C for 3 minutes, followed by 35 cycles of denaturation, annealing, and extension. Each PCR batch included positive and negative controls, and batches were repeated if controls failed.[25] Secondary PCR was conducted in a 50 µl reaction volume using KAPA HiFi polymerase to add Illumina adapters and 8 bp dual indices for sample multiplexing.[25] The final libraries were cleaned, pooled separately for trnL and 12SV5, and sequenced on independent Illumina runs.[25]

Curated reference databases were constructed by compiling a list of edible plant and animal taxa from global food surveys and reference sources.[25] Sequences containing *trnL-P6* or *12SV5* regions were retrieved from GenBank and RefSeq, filtered for primer binding sites with a ≤20% mismatch tolerance, and trimmed to the target amplicon regions.[25] Identical sequences from different accessions were de-duplicated, while intra-taxon variability and conserved sequences across taxa were retained.[25] Demultiplexed reads were processed using BBDuk (adapter trimming), Cutadapt (primer filtering and additional trimming), and DADA2 (denoising, merging, and amplicon sequence variant (ASV) inference).[25] Reads were quality-filtered using an expected error threshold of 2 and truncated at the first base with a quality score ≤2.[25] For trnL, ASVs were assigned using DADA2's assignSpecies function, with exact sequence matching to the custom trnL reference database.[25] In cases of ambiguous matches, taxa were assigned to the last common ancestor (*e.g.*, a sequence matching both wheat and rye was classified at the Poaceae family level).[25] For 12SV5, ASVs were assigned using DADA2's assignTaxonomy and agglommerated at the lowest common taxonomic assignment, due to reduced performance of exact sequence matching

Open Access Article. Published on 18 March 2026. Downloaded on 4/8/2026 5:33:14 PM. This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

caused by PCR polymerase-induced mismatches.[25] Suspected contaminants were identified using decontam and removed based on DNA quantitation data.[25] ASV count tables, taxonomic assignments, and metadata were structured using phyloseq (version 1.38.0).[25]

Statistical analyses

Statistical Programming Language *R* (V. 4.5.0) was used to analyze and visualize data. Taxonomic read counts were converted to relative abundances and normalized through center-log-ratio transformation to preserve overall data structure followed by removal of taxa with fewer than 5 sequence reads, removal of taxa with no assignment at the superkingdom level in trnL and removal of human ASVs from the 12SV5 data. Plant dietary compositionality was analyzed using Principal Component Analysis (PCA, 'prcomp' function, parameters: centered, not scaled). Plant FoodSeq richness (pFR) was calculated as the number of unique taxa with at least one read count in each sample as previously described.[21] The Dietary Diversity Score (DDS) was derived from 24 h dietary recall data.[28] The data collection process, as well as the complete methodology for standardization and homogenization has been previously described.[28] The reported food items were categorized into 19 predefined food groups following a standardized food classification system.[28] Food groups considered included: cereals and their products; pulses, legumes, and their products; nuts, seeds, and their products; leafy vegetables and their products; vegetables and their products; starchy roots and tubers; flesh meat and their products; fish, seafood, and their products; organ meats; fruits and their products; beverages; milk and their products; sweets and their products; spices, condiments, and seasoning foods; fats and oils; non-food items; supplements; and eggs and their products. Although standard DDS calculations incorporate up to ten food groups, for this study we classified food in greater detail for improved comparison with genomic dietary data resolution. Each participant's dietary intake was assessed to determine whether at least one item from each food group was consumed within the recall period. A binary scoring system was applied, where the presence of a food group was assigned a score of '1', while its absence was assigned a score of '0'. The DDS was calculated as the sum of all consumed food groups:

$$\mathrm{DDS} = \sum_{i=1}^{n} \mathrm{Score}\, i$$

where Score represents the presence (1) or absence (0) of the *i*th food group, and *n* is the total number of food groups considered (19 in this study).

Presence-absence concordance between FoodSeq and time-point-matched 24 hour recall data was evaluated for three major dietary taxa through counts of each taxon detected in either measure across all 150 samples. Detections were categorized as present in both methods, FoodSeq only, recall only, or neither, and overall agreement was calculated as the sum of samples classified as either both-present or both-absent.

CHW home visit derived compliance reporting was used to calculate (1) total sachets consumed across all visits/total sachets provided across all visits. (2) Sachets consumed/provided for the closest home visit matching to stool sampling date (home visits occurring after stool sample date were chosen in case of ties due to visit reporting RUSF consumption over past seven days and stool sample capturing approximately last 72 h of residual food DNA). (3) Visit-matched values from (2) normalized by number of days between home visit date and stool collection.

Where applicable, analyses explicitly accounted for temporal mismatch between dietary exposure and stool collection by aligning the closest adherence and recall date to stool sample collection dates and by including an adherence analysis with compliance data normalized by the interval between visit date and sample collection. For comparative testing, Wilcoxon rank-sum test was used between unpaired samples and Wilcoxon signed-rank test for paired samples. Relationships spanning all timepoints were assessed using linear mixed effects models with participant ID as a random effect, and PCR amplification batch, age and gender, wealth, food insecurity, wealth, and family size as fixed effects. Age and gender were not added as covariates for analyses involving WHZ, WAZ, and WHZ as these metrics already incorporate age and gender adjustments. Given modest subgroup sample sizes, wealth quintiles were collapsed into three categories (below middle, middle, above middle) and food insecurity quartiles were dichotomized (none *vs.* insecure) to ensure adequate cell counts and improve model stability in multivariable analyses.

Results

Dietary DNA was successfully amplified in 149/150 samples, with the only non-amplifying sample corresponding to a 3–6 month old malnourished infant expected to be exclusively breastfed per standard recommendations.[17] Low levels of food DNA detected in other infants of the same age were consistent with previous reports of inappropriate early complementary feeding in this region.[17] Overall, 89 unique plant and 16 vertebrate taxa were identified (Fig. 1b). Wheat was detected across >50% of samples; banana was the most frequent fruit (31.6%); peas the most frequent legume (35%), and water buffalo the most common animal (35.8%) (Fig. 1b) (the abundance of detected reads per target DNA is provided in SI 1). PCA analysis identified the following taxa as top drivers of variation in dietary composition among participants: wheat; tea; banana; cowpea family; chickpea; carrot, parsley, cumin family; pea; wild rice; pepper family; coriander, and cilantro (Fig. 1c and Table S1).

Chickpea DNA spike aligns with RUSF administration

FoodSeq detected a spike in chickpea DNA (*Cicer arietinum*; primary plant ingredient of *Acha Mum*) center-log-ratio normalized relative abundances (hereafter "CLR value") (median = 5.27, IQR [7.75], Fig. 2a) in stool samples collected during *Acha*

This journal is © The Royal Society of Chemistry 2026

Open Access Article. Published on 18 March 2026. Downloaded on 4/8/2026 5:33:14 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

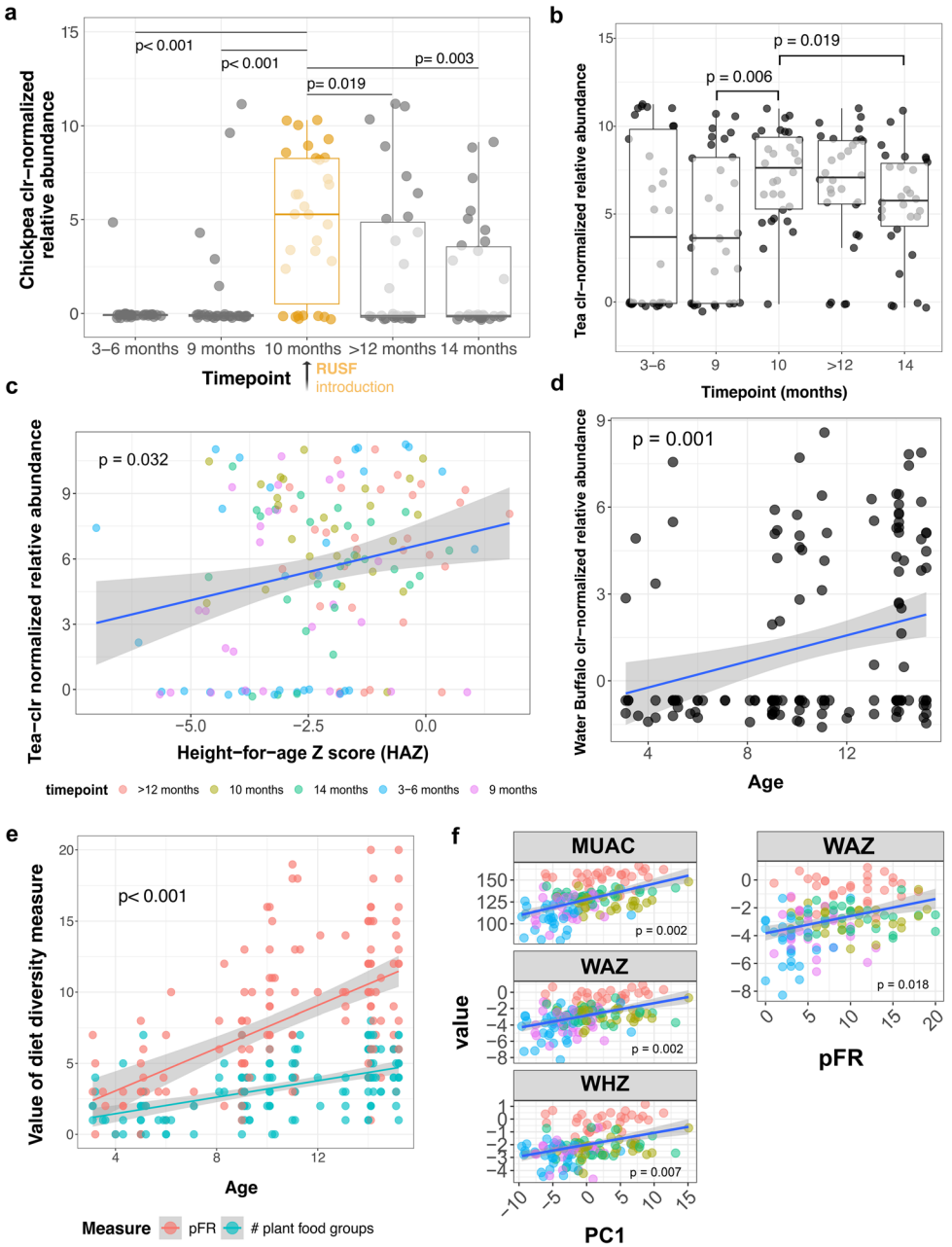

Fig. 2 Increase in chickpea DNA level aligns with RUSF administration, FoodSeq captures dietary signals relevant to nutritional development and infant health. (a) CLR-normalized relative abundances of chickpea (*Cicer arietinum*) DNA amplified from stool samples compared across all timepoints. The 10-month timepoint when RUSF was administered is highlighted in orange. The >12 months samples correspond to healthy infants. Box plots show median and interquartile ranges, while individual dots represent samples from each subject. Wilcoxon signed-rank test for 3–6 months *vs.* 9 months and 10 months *vs.* 14 months and Wilcoxon rank-sum test for all other timepoint comparisons were conducted. Significance results of statistical testing are marked. (b) CLR-normalized relative abundances of tea (*Camellia sinensis*) DNA amplified from stool samples compared across all timepoints. Box plots show median and interquartile ranges, while individual dots represent samples from each subject. Wilcoxon signed-rank test for 3–6 months *vs.* 9 months and 10 months *vs.* 14 months and Wilcoxon rank-sum test for all other timepoint comparisons were conducted. Significance results of statistical testing are marked. (c) Relationship between CLR-normalized tea (*Camellia sinensis*) relative abundance and height-for-age *Z*-score (HAZ) across all timepoints. Each point represents an individual stool sample, colored by timepoint (p = 0.032, linear mixed-effects model: participant ID as a random effect; PCR batch as a fixed effect). Shaded area represents the 95% confidence interval. (d) CLR-normalized relative abundance values of water buffalo (*Bubalus bubalis*) DNA detected in stool samples compared across age. Individual dots represent samples from each subject. Linear mixed effects model (pcr batch, gender and socioeconomics as fixed effects and participant ID as random effect). Significance results as marked. (e) Age-dependent increase in dietary diversity measured by traditional dietary diversity score (DDS, salmon) and plant FoodSeq Richness (pFR, cyan). Both metrics show significant positive correlation with age (p < 0.0001; linear mixed effects models). Shaded areas represent 95% confidence intervals. (f) Relationships between the first principal component of FoodSeq dietary composition (PC1; left panels) or plant FoodSeq richness (pFR; right panels) and anthropometric measures including mid-upper arm circumference (MUAC), weight-for-age *Z*-score (WAZ), and weight-for-height *Z*-score (WHZ). Each point represents an individual sample, colored by timepoint (legend same as for panel c). Significance values marked (linear mixed-effects model).

Open Access Article. Published on 18 March 2026. Downloaded on 4/8/2026 5:33:14 PM. This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

Mum administration at 10 months (mean infant age = 10.56 months) compared to pre-administration time points (median = −0.07, IQR [0.06] at 3–6 months, $p < 0.001$ and median = −0.13, IQR [0.10], $p < 0.001$ at 9 months; Wilcoxon rank-sum test), in healthy controls (median = −0.13, IQR [5.06], $p = 0.019$; Wilcoxon rank-sum test), and post-administration in MAM samples at 14 months (median = −0.12, IQR [3.75], $p = 0.003$; Wilcoxon signed-rank test) (Fig. 2a). To account for cohort-level differences, we further compared values between the 9–10 month timepoints through a linear model adjusting for PCR batch, gender, wealth index, food insecurity, and family size which demonstrated the same Chickpea CLR value spike at 10 months ($\beta = 4.08$, $p = 6.9 \times 10^{-5}$) indicating that this increase was not attributable to cohort-level heterogeneity. A similar spike during RUSF administration was not observed in any of the other major dietary drivers (Fig. S1).

Comparison of chickpea CLR-values with CHW-home visit derived compliance metrics showed positive but non-significant trends, likely impacted by the low sample size and lag between compliance visits and stool sample collection (only 1/30 sample had 100% compliance reported on the same date as the stool sample but no chickpea detected). Overall, 22/30 samples were true positives for chickpea and 1/30 a true negative, yielding a sensitivity and sensitivity of 0.76 and 1 respectively.

Chickpea CLR-values at 10 months were also not associated with DDS, pFR, major dietary patterns (plant PCs) (Fig. S2b) and did not partition by wealth index, food insecurity, or family size (Fig. S2c), suggesting RUSF compliance was not affected by local socio-economic factors or dietary practices.

Capture of dietary relationships relevant to infant health and intervention efficacy

Notably, tea (*Camellia sinensis*), an age-inappropriate food per WHO infant feeding guidelines,[29] was present in >50% of samples, detected as early as 3–6 months (consistent with caregiver reports), and did not partition by socioeconomic variables (wealth and food insecurity). Tea CLR-values showed a marked increase from 9 to 10 months of age (median 3.63 [IQR 8.31] to 7.63 [IQR 4.09], $p = 0.006$; Wilcoxon rank-sum test) which remained consistent to cohort-heterogeneity adjustment through a linear model controlling for PCR batch, gender, wealth index, food insecurity, and family size ($\beta = 2.93$, $p = 0.0043$). This increase was followed by a decline by 14 months (median 5.76 [IQR 3.59], $p = 0.019$; Wilcoxon signed-rank test) (Fig. 2b).

While tea was also modestly associated with higher HAZ scores ($\beta = 0.041$, $p = 0.029$, linear mixed effects model, PCR batch and socioeconomics as fixed effects and participant ID as random effect) (Fig. 2c), having no food insecurity was the greatest contributor to higher HAZ ($\beta = 0.834$, $p = 0.021$, linear mixed effects model, PCR batch and socioeconomics as fixed effects and participant ID as random effect). Tea CLR-values were not related to chickpea values at the 10 month timepoint.

Water buffalo, the most abundant animal taxon (Fig. 1b) and a common source of milk in South Asia,[30] displayed increasing CLR values with age ($\beta = 0.234$, $p = 0.001$, linear mixed effects model adjusted for PCR batch, gender and socioeconomics as fixed effects and participant ID as random effect) (Fig. 2d), reflecting diluted water buffalo milk for complementary feeding reported in 24 hour recalls.

Evaluation of water buffalo (milk proxy) and tea co-consumption demonstrated a modest but significant relationship ($\beta = 0.28$, $p = 0.012$; linear mixed-effects model controlling for age, gender, PCR batch, and socioeconomics as fixed effects with participant ID as a random effect) (Fig. S2d). However, tea and milk DNA were not consistently co-detected across samples, with tea frequently detected in the absence of detectable water buffalo DNA.

Although FoodSeq data (from stool samples) was temporally well-separated from 24 h recalls (median = 10 days, IQR = 17), broad concordance estimates for selected major taxa detected in FoodSeq (wheat as most frequent and staple food, water buffalo as most frequent animal, and tea as notable anomaly) demonstrated highest agreement for wheat (72.7%), with complete alignment of detections in both measures; moderate for water buffalo (60.7%), with more frequent reporting in recalls than detection in stool; and lowest for tea (59.3%), detected much more frequently by FoodSeq even when not reported (Table S2).

Given the established expectations of complementary feeding and dietary expansion with infant maturation,[29] we compared both a traditionally calculated Dietary Diversity Score (DDS) and plant FoodSeq Richness (pFR), the number of unique identified plant taxa detected per sample,[21] against age to evaluate FoodSeq's ability to capture expected dietary relationships. Both metrics were positively correlated with age ($\beta = 0.27$, $p < 0.0001$ and $\beta = 0.69$, $p < 0.0001$ respectively) (Fig. 2e) but not with each other, suggesting capture of distinct aspects of dietary diversity. pFR further demonstrated a positive relationship with WAZ ($\beta = 0.066$, $p = 0.018$ Bonferroni-Hochberg corrected; linear mixed effects model with participant ID as random effect; robust to socioeconomic adjustment) (Fig. 2f) while DDS associated with WAZ ($\beta = 0.15$, $p = 0.011$; robust to socioeconomic control) but not with WHZ (data not shown).

Consistent with prior findings from a global infant cohort, the primary axis of dietary variation (PC1) suggested capture of total plant material from complementary feeding, with dietary loadings pointing almost exclusively in the same direction[31] (Fig. 1c). In line with this interpretation, PC1 scores also increased with age ($p < 0.001$; linear mixed effects model: gender and participant ID adjusted) (Fig. 1c) with age exerting a stronger influence on PC1 than health status (Fig. S2a). PC1 was also positively associated with WHZ, WAZ, and MUAC ($p < 0.01$ for all, Bonferroni-Hochberg corrected, linear mixed effect model with participant ID as random effect, age and gender included for MUAC) with these relationships robust to socioeconomic adjustment (Fig. 2f).

This journal is © The Royal Society of Chemistry 2026

Open Access Article. Published on 18 March 2026. Downloaded on 4/8/2026 5:33:14 PM. This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

Discussion

Although free RUT/SF provision through CMAM programs is now the standard of care for acute malnutrition,[32] socioeconomic barriers continue to influence consumption and adherence behaviours, necessitating intensive compliance monitoring. Although chickpea DNA was detectable across all timepoints in the study since chickpea is a common part of local diet,[33] FoodSeq captured a distinct and statistically significant increase during RUSF administration, demonstrating sensitivity to relative changes in intake rather than simple presence-absence detection. This finding highlights the potential of stool genomics to simultaneously capture geographic and cultural dietary patterns through the universal language of DNA, augmenting adherence tracking in CMAM and revealing contextual barriers to treatment.

In our cohort, the most frequently detected taxa were staple crops in Pakistan and ingredients of common complementary foods: wheat in roti (flatbread) and daliya (porridge), rice and derivatives, and banana,[17] consistent with 24 h recall reports. The detection of water buffalo, South Asia's primary cattle species, widely used for milk and meat,[30] underscores the ability to taxonomically resolve regionally-specific sources of common foods. The consistent increase in water buffalo DNA with age further suggests weaning of infants from breast milk to cattle dairy, supported by increased rates of dairy milk consumption reported across age in the 24 h recalls. The widespread presence of tea DNA suggested a more extensive consumption during complementary feeding than previously reported (only one study documented use in >50% of infants), contrary to WHO guidelines.[17] Notably, however, tea detection was not associated with RUSF intake, WHZ change, or socioeconomic variables. The modest relationship between tea and HAZ likely also reflects an age-related variation (height increase with age) in tea provision rather than a causal effect. Similarly, socioeconomic or dietary factors did not directly impact RUSF intake, and chickpea-based RUSF was well-accepted for outpatient care, though measures to monitor tea provision in future interventions might be beneficial given the high prevalence detected in both FoodSeq and 24 h recalls across all timepoints as tea tannins can inhibit appetite and iron absorption.[17]

As expected, concordance between FoodSeq and 24 h reports showed high agreement for staple foods while highlighting discrepancies for items prone to low-volume intake or under-reporting such as tea. Frequent detection of tea DNA despite absence of reporting aligns with cultural feeding practices[17] and highlights the value of genomic monitoring for capturing subtle or unrecognized dietary exposures. Moderate concordance for water buffalo likely reflects reduced detectability of DNA from milk given the comparative ease of digestion *vs.* plant foods. These findings support FoodSeq's ability to complement traditional dietary assessment by identifying both reported and unreported exposures relevant to child nutrition.[17]

Furthermore, although tea is typically prepared with milk in this setting, tea and water buffalo DNA demonstrated only partial co-detection in stool samples. This is consistent with known biological and technical factors affecting DNA recovery such as the greater stability of plant structures in the digestive tract as well as leaves (which tea is brewed from) containing high density of chloroplasts which are the target of the *trnL-P6* marker, increasing amplifiable DNA yield. In comparison, DNA in milk, detected *via* the mitochondrial *12SV5* marker, can be more susceptible to degradation and may be present at lower effective concentrations. Because FoodSeq is sensitive to relative amplifiable residual DNA biomass in stool, one-to-one concordance is not expected.

While the observed increase in genomic dietary diversity with age supports the ability to detect expected complementary feeding expansion with age,[29] associations of pFR and PC1 with measures such as WHZ, WAZ, and MUAC suggest broad potential to capture nutritional information relevant to infant health through biologically relevant variation in diet rather than simply capturing household-level socioeconomic differences.

Interestingly, though *Acha Mum* contains dried-milk, the pattern of milk-producing taxa abundances was not comparable to chickpea (Fig. 2d and Fig. S3a). This discrepancy may stem from DNA degradation during milk drying, low milk content in RUSF, or masking of RUSF-specific milk signal from milk in complementary feeding. These findings indicate a need to optimize detection methods for animal-derived RUT/SF components. Given FoodSeq's potential to reflect portion size for certain plant taxa,[21] studies should quantify the relationship between RUT/SF intake and residual DNA levels, and assess how formulations influence detectability across contexts. Currently, the interpretation of taxon-specific concordance between FoodSeq and 24 h recalls is also constrained due to the temporal mismatch between the methods in this retrospective analysis. Furthermore, due to the partially longitudinal sub-cohorts and the resulting limited sample size, long-term within-infant dietary trajectories could not be assessed across the full 3–14 month window and compliance analysis results did not fully untangle the contribution of adherence variability, methodological limitations or biological differences in DNA digestion to false negatives for chickpea detection. Prospective studies with better synchronization between dietary assessment methods and compliance tracking as well as fully longitudinal sampling across all key developmental stages could further resolve these relationships and clarify the evolution of dietary DNA signals across growth and intervention stages. Such future work could also integrate genomic dietary assessment with traditional clinical and nutritional measures in efforts to disentangle underlying drivers of malnutrition. Though next-generation sequencing may be cost-prohibitive in low-resource settings, the use of ethanol-based preservation and room-temperature shipping,[22] along with targeted PCR assays and emerging portable PCR tools,[34] could lower costs and accelerate results. Lastly, stool collection faces barriers related to cultural sensitivities, logistics, and

Open Access Article. Published on 18 March 2026. Downloaded on 4/8/2026 5:33:14 PM. This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

compliance; solutions may include community engagement and integration into existing healthcare workflows.

Conclusions

Overall, our study demonstrates the promise of FoodSeq as a novel tool for monitoring RUSF compliance in the CMAM framework, with potential extension to other standardized therapeutic foods such as RUTF. Future trials could leverage FoodSeq to track adherence to other nutritional interventions or detect inappropriate feeding practices that compromise treatment. Standardizing ingredient detection, *e.g.*, through marker food or synthetic DNA spike-ins, may enhance specificity and enable more precise quantification. As the method matures, FoodSeq insights could inform broader efforts to combat undernutrition, evaluate novel therapeutic foods, and shape child feeding policies.

Author contributions

A. A., N. T. I., S. A. A., and L. A. D. conceptualized the study. A. A., N. T. I., S. I. S., S. A., F. U., K. A., M. J. B. performed data curation. A. A., S. I. S., S. A. conducted formal analysis. N. T. I., S. A. A., and L. A. D. provided funding for the study. A. A., T. K. M., O. O., N. I., N. T. I., S. A., F. U. performed experimental investigation. A. A., T. K. M., L. A. D., S. I. S. formulated methodology. A. A., S. J., and M. J. B. administered the project. L. A. D., N. T. I., A. A., M. J. B. and J. I. G. provided resources. A. A. and T. K. M. performed bioinformatic processing. L. A. D., N. T. I., and S. A. A. supervised the project. A. A. created data visualizations. A. A. and S. I. S. wrote the original manuscript draft. A. A., L. A. D., N. T. I., S. A., K. A., M. J. B. reviewed and edited the final manuscript.

Conflicts of interest

Lawrence A. David consults for City of Hope on dietary assessment technology and is a member of the Global Grants for Gut Health Colloquium (Nature Portfolio/Yakult). There are no other conflicts of interest to declare.

Data availability

Deidentified analysis-ready data, data dictionaries, code notebooks, and clinical metadata for reproducing manuscript results from processed trnL-P6 and 12SV5 data have been provided through the Figshare repository (https://doi.org/10.6084/m9.figshare.c.7900001). Human 12SV5 sequences in the processed data were replaced with a matched synthetic DNA sequence which does not correspond to any real individual and is provided solely to enable replication of data analysis without compromising participant privacy. Raw trnL and 12SV5 amplicon sequencing data have been deposited in the NCBI Sequence Read Archive (SRA) under accession number PRJNA1428497. Human-derived ASVs were again anonymized prior to deposition using the privacy-preserving filtering algorithm described at https://github.com/LAD-LAB/mb-pipeline/tree/main/reference/anonymizer, in accordance with participant confidentiality requirements. 24 h recall data is included for select taxa used in the concordance analysis; further data can be provided upon request.

Supplementary figures and tables are provided. See Supplementary information (SI) DOI: https://doi.org/10.1039/d5fo02791h.

Acknowledgements

This study was funded by the Bill and Melinda Gates Foundation (OPP1144149 and OPP1138727), the National Institutes of Health (R01-DK116187 and R01-DK128611) Schmidt Futures, Burroughs Wellcome Fund Pathogenesis of Infectious Disease Award, the Chan Zuckerberg Initiative and the Gerber Foundation. This study was further supported by the High-performance computing facility (Duke Compute Cluster) provided by the Duke Office of Information Technology and Duke Research Computing. Matiari data management team and field staff for study support; families and their children for study participation; Marty Meier for technical assistance in purifying DNA from fecal samples; Ashish Subramanian for development of Human DNA anonymization approach for raw *trnL-P6* and *12SV5* sequencing data; Michelle Kirtley and Anna Bauer for editorial guidance; Tonya Snipes, Lisa Alston-Latta, Joey Biondo, and Christopher Sutton for lab space maintenance and cleaning.

References

1 WHO. Fact sheets - Malnutrition. [cited 2022 Jul 13]. Available from: https://www.who.int/news-room/fact-sheets/detail/malnutrition.
2 F. Birhanu, K. Yitbarek, F. T. Bobo, E. Atlantis and M. Woldie, Undernutrition in children under five associated with wealth-related inequality in 24 low- and middle-income countries from 2017 to 2022, *Sci. Rep.*, 2024, **14**(1), 3326, DOI: 10.1038/s41598-024-53280-0.
3 D. Gallegos, Effects of Food and Nutrition Insecurity on Global Health, *N. Engl. J. Med.*, 2025, **392**(7), 686–697, DOI: 10.1056/NEJMra2406458.
4 A. R. Patlán-Hernández, H. C. Stobaugh, O. Cumming, A. Angioletti, D. Pantchova, J. Lapègue, *et al.*, Water, sanitation and hygiene interventions and the prevention and treatment of childhood acute malnutrition: A systematic review, *Matern. Child Nutr.*, 2022, **18**(1), e13257, DOI: 10.1111/mcn.13257.
5 S. Subramanian, S. Huq, T. Yatsunenko, R. Haque, M. Mahfuz, M. A. Alam, *et al.*, Persistent gut microbiota

 This journal is © The Royal Society of Chemistry 2026

Open Access Article. Published on 18 March 2026. Downloaded on 4/8/2026 5:33:14 PM. This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

immaturity in malnourished Bangladeshi children, *Nature*, 2014, **510**(7505), 417–421, DOI: 10.1038/nature13421.

6 S. Syed, A. Ali and C. Duggan, Environmental Enteric Dysfunction in Children, *J. Pediatr. Gastroenterol. Nutr.*, 2016, **63**(1), 6–14, DOI: 10.1097/MPG.0000000000001147.

7 V. J. B. Martins, T. M. M. Toledo Florêncio, L. P. Grillo, M. d. C. Franco, P. A. Martins, A. P. G. Clemente, *et al.*, Long-Lasting Effects of Undernutrition, *Int. J. Environ. Res. Public Health*, 2011, **8**(6), 1817–1846, DOI: 10.3390/ijerph8061817. PubMed PMID: 21776204; PubMed Central PMCID: PMC3137999.

8 G. Teshale, A. Debie, E. Dellie and T. Gebremedhin, Evaluation of the outpatient therapeutic program for severe acute malnourished children aged 6–59 months implementation in Dehana District, Northern Ethiopia: a mixed-methods evaluation, *BMC Pediatr.*, 2022, **22**(1), 374, DOI: 10.1186/s12887-022-03417-9.

9 I. Hussain, A. Habib, S. Ariff, G. N. Khan, A. Rizvi, S. Channar, *et al.*, Effectiveness of management of severe acute malnutrition (SAM) through community health workers as compared to a traditional facility-based model: a cluster randomized controlled trial, *Eur. J. Nutr.*, 2021, **60**(7), 3853–3860, DOI: 10.1007/s00394-021-02550-y.

10 L. Lenters, K. Wazny and Z. A. Bhutta, Management of Severe and Moderate Acute Malnutrition in Children, in *Reproductive, Maternal, Newborn, and Child Health: Disease Control Priorities*, ed. R. E. Black, R. Laxminarayan, M. Temmerman and N. Walker, The International Bank for Reconstruction and Development/The World Bank, Washington (DC), 3rd edn, 2016, vol. 2. [cited 2025 Feb 21] PubMed PMID: 27227221. Available from: https://www.ncbi.nlm.nih.gov/books/NBK361900/.

11 S. T. Kangas, C. Salpéteur, V. Nikièma, C. Ritz, H. Friis, A. Briend, *et al.*, Predictors of time to recovery and non-response during outpatient treatment of severe acute malnutrition, *PLoS One*, 2022, **17**(5), e0267538, DOI: 10.1371/journal.pone.0267538.

12 L. M. Lenters, K. Wazny, P. Webb, T. Ahmed and Z. A. Bhutta, Treatment of severe and moderate acute malnutrition in low- and middle-income settings: a systematic review, meta-analysis and Delphi process, *BMC Public Health*, 2013, **13**(3), S23, DOI: 10.1186/1471-2458-13-S3-S23.

13 A wonder 'food' for the world's children | UNICEF. [cited 2024 Feb 28]. Available from: https://www.unicef.org/nutrition/RUTF.

14 R. C. Purwestri, V. Scherbaum, D. A. Inayati, N. N. Wirawan, J. Suryantan, M. A. Bloem, *et al.*, Impact of Daily versus Weekly Supply of Locally Produced Ready-to-Use Food on Growth of Moderately Wasted Children on Nias Island, Indonesia, *Int. Sch. Res. Notices*, 2013, **2013**(1), 412145, DOI: 10.5402/2013/412145.

15 B. Beggs, M. Bustos, L. J. Brubacher, M. Little, L. Lau and W. Dodd, Facilitators and barriers to implementing complex community-based interventions for addressing acute malnutrition in low- and lower-middle income countries: A scoping review, *Nutr. Health*, 2024, **30**(3), 447–462, DOI: 10.1177/02601060241253327.

16 E. Tadesse, Y. Berhane, A. Hjern, P. Olsson and E. C. Ekström, Perceptions of usage and unintended consequences of provision of ready-to-use therapeutic food for management of severe acute child malnutrition. A qualitative study in Southern Ethiopia, *Health Policy Plan.*, 2015, **30**(10), 1334–1341, DOI: 10.1093/heapol/czv003.

17 L. Manikam, A. Sharmila, A. Dharmaratnam, E. C. Alexander, J. Y. Kuah, A. Prasad, *et al.*, Systematic review of infant and young child complementary feeding practices in South Asian families: the Pakistan perspective, *Public Health Nutr.*, 2018, **21**(4), 655–668, DOI: 10.1017/S1368980017002956. PubMed PMID: 29151370; PubMed Central PMCID: PMC5851056.

18 N. T. Iqbal, S. Syed, K. Sadiq, M. N. Khan, J. Iqbal, J. Z. Ma, *et al.*, Study of Environmental Enteropathy and Malnutrition (SEEM) in Pakistan: protocols for biopsy based biomarker discovery and validation, *BMC Pediatr.*, 2019, **19**(1), 1, DOI: 10.1186/s12887-019-1564-x.

19 A. Sarfraz, S. Ahmed, S. Muhammad, N. Rehman, S. I. Soomro, K. Qureshi, *et al.*, Standard RUTF vs. locally-made RUSF for acutely malnourished children: A quasi-experimental comparison of the impact on growth and compliance in a rural community of Pakistan, *PLoS One*, 2023, **18**(7), e0287962, DOI: 10.1371/journal.pone.0287962. PubMed PMID: 37437065.

20 R. Micha, J. Coates, C. Leclercq, U. R. Charrondiere and D. Mozaffarian, Global Dietary Surveillance: Data Gaps and Challenges, *Food Nutr. Bull.*, 2018, **39**(2), 175–205, DOI: 10.1177/0379572117752986. PubMed PMID: 29478333.

21 B. L. Petrone, A. Aqeel, S. Jiang, H. K. Durand, E. P. Dallow, J. R. McCann, *et al.*, Diversity of plant DNA in stool is linked to dietary quality, age, and household income, *Proc. Natl. Acad. Sci. U. S. A.*, 2023, **120**(27), e2304441120, DOI: 10.1073/pnas.2304441120.

22 A. Aqeel, M. C. Kay, J. Zeng, B. L. Petrone, C. Yang, T. Truong, *et al.*, Grocery intervention and DNA-based assessment to improve diet quality in pediatric obesity: a pilot randomized controlled study, *Obesity*, 2025, **33**(2), 331–345, DOI: 10.1002/oby.24205.

23 J. Schneider, E. Mas-Carrió, C. Jan, C. Miquel, P. Taberlet, K. Michaud, *et al.*, Comprehensive coverage of human last meal components revealed by a forensic DNA metabarcoding approach, *Sci. Rep.*, 2021, **11**(1), 8876, DOI: 10.1038/s41598-021-88418-x.

24 T. Riaz, W. Shehzad, A. Viari, F. Pompanon, P. Taberlet and E. Coissac, ecoPrimers: inference of new DNA barcode markers from whole genome sequence analysis, *Nucleic Acids Res.*, 2011, **39**(21), e145, DOI: 10.1093/nar/gkr732.

25 B. L. Petrone, A. Bartlett, S. Jiang, A. Korenek, S. Vintila, C. Tenekjian, *et al.*, A pilot study of metaproteomics and DNA metabarcoding as tools to assess dietary intake in humans, *Food Funct.*, 2025, **16**(1), 282–296, DOI: 10.1039/D4FO02656J.

Open Access Article. Published on 18 March 2026. Downloaded on 4/8/2026 5:33:14 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

26 S. Syed, S. Yeruva, J. Herrmann, A. Sailer, K. Sadiq, N. Iqbal, *et al.*, Environmental Enteropathy in Undernourished Pakistani Children: Clinical and Histomorphometric Analyses, *Am. J. Trop. Med. Hyg.*, 2018, **98**(6), 1577–1584, DOI: 10.4269/ajtmh.17-0306.

27 N. T. Iqbal, R. Y. Chen, N. W. Griffin, M. C. Hibberd, A. Khalid, K. Sadiq, *et al.*, A shared group of bacterial taxa in the duodenal microbiota of undernourished Pakistani children with environmental enteric dysfunction, *mSphere*, 2024, **9**(6), e00196–e00124, DOI: 10.1128/msphere.00196-24.

28 S. I. Soomro, Z. Jamil, N. Memon, S. Ahmed, F. Umrani, I. A. Choudhri, *et al.*, Nutrient dataset development via FAO/INFOODS approach for infant nutritional survey in rural Matiari, Pakistan, *J. Food Compos. Anal.*, 2024, **133**, 106471, DOI: 10.1016/j.jfca.2024.106471.

29 WHO Guideline for complementary feeding of infants and young children 6–23 months of age. [cited 2025 Jan 22]. Available from:https://www.who.int/publications/i/item/9789240081864.

30 G. K. Deb, T. N. Nahar, P. G. Duran and G. A. Presicce, Safe and Sustainable Traditional Production: The Water Buffalo in Asia, *Front. Environ. Sci.*, 2016, **4**, 1–7, DOI: 10.3389/fenvs.2016.00038.

31 T. McDonald, A. Aqeel, B. Neubert, A. Bauer, S. Jiang, O. Osborne, *et al.* Dietary plant diversity predicts early life microbiome maturation. *medRxiv*, 2025 [cited 2025 May 5]. p. 2025.02.28.25323117. Available from: https://www.medrxiv.org/content/10.1101/2025.02.28.25323117v1. DOI: 10.1101/2025.02.28.25323117.

32 WHO, WFP, SCN, UNICEF. Joint Statement on Community-Based Management of Severe Acute Malnutrition. 2007 [cited 2024 Nov 25]. Available from:https://www.unicef.org/media/96981/file/Statement-WHO-WFP-SCN-and-UNICEF-on-Community-Based-Management-of-SAM.

33 F. Javed, S. Jabeen, M. K. Sharif, I. Pasha, A. Riaz, M. F. Manzoor, *et al.*, Development and storage stability of chickpea, mung bean, and peanut–based ready–to–use therapeutic food to tackle protein–energy malnutrition, *Food Sci. Nutr.*, 2021, **9**(9), 5131–5138, DOI: 10.1002/fsn3.2479. PubMed PMID: 34532022; PubMed Central PMCID: PMC8441456.

34 Bento Lab. [cited 2024 Dec 2]. Bento Lab | A Portable PCR Workstation. Available from:https://bento.bio/.

This journal is © The Royal Society of Chemistry 2026