EPIDEMIOLOGY

Metabolic maturation in the first 2 years of life in resource-constrained settings and its association with postnatal growth

N. Giallourou[1]*, F. Fardus-Reid[1]*, G. Panic[1], K. Veselkov[1], B. J. J. McCormick[2], M. P. Olortegui[3], T. Ahmed[4], E. Mduma[5], P. P. Yori[3,6], M. Mahfuz[4], E. Svensen[5,7], M. M. M. Ahmed[4], J. M. Colston[6,8], M. N. Kosek[3,6]†, J. R. Swann[1,9]

Malnutrition continues to affect the growth and development of millions of children worldwide, and chronic undernutrition has proven to be largely refractory to interventions. Improved understanding of metabolic development in infancy and how it differs in growth-constrained children may provide insights to inform more timely, targeted, and effective interventions. Here, the metabolome of healthy infants was compared to that of growth-constrained infants from three continents over the first 2 years of life to identify metabolic signatures of aging. Predictive models demonstrated that growth-constrained children lag in their metabolic maturity relative to their healthier peers and that metabolic maturity can predict growth 6 months into the future. Our results provide a metabolic framework from which future nutritional programs may be more precisely constructed and evaluated.

INTRODUCTION

Infants who experience stunting are likely to undergo poorer physical and cognitive development throughout childhood, placing them on an early path to lost human potential (*1*). Stunting, a consequence of impaired linear growth, is determined by the World Health Organization (WHO) as a length-for-age Z score (LAZ) of two or more SDs below the WHO median (*2*). The causes of stunting are diverse and include intrauterine growth retardation, restriction in dietary intake, infections, social status, and maternal factors such as age, stature, and nutritional status (*3*). While nutritional interventions, especially in the first 2 years of life, have helped to reduce the number of stunted infants by 50 million since 2000, the progress toward reducing the prevalence of stunting has been notably recalcitrant to interventions (*4*). It has been estimated that even if current nutritional interventions were scaled up to 90% coverage, the prevalence of stunting would only be reduced by 20% (*5*). This lack of effectiveness may be explained by the myriad of interacting contributing factors (*6*) as well as a misalignment between nutritional interventions and the demands of the developing biochemical system under challenging environmental conditions, particularly in the setting of multiple infections.

We and others have measured the metabolic phenotypes, or phenomes, of stunted infants and have shown that a number of metabolic pathways are modulated compared to children who are not stunted (*7*). In a Brazilian infant case-control study, betaine and tryptophan metabolism, as well as microbial-host cometabolism, was found to be altered in stunted infants (*8*). Moreover, in a popu-lation of severely acutely malnourished infants from Zambia, various metabolic derangements were observed with gut pathology associated with environmental enteric dysfunction (*9*). Although these studies were cross-sectional, we have also shown in longitudinal studies that plasma citrulline and tryptophan were associated with linear growth for the 6 months following their assessment in early life in children from Peru and Tanzania (*10*). A more comprehensive understanding of early longitudinal biochemical development and the changing demands of the maturing metabolic system across the first 1000 days of life is critical to improving the impact of nutritional interventions. Previous work has explored the age-dependent evolution of the gut microbial community structure in infants from resource-constrained settings (*11*, *12*). Defined as the microbiota-for-age Z score (MAZ), this metric was used to assess relative gut microbial maturity and its relationship with nutritional intervention responses and provided a target against which novel interventions could be developed.

In this work, we define the development of the infant metabolic system and its relationship with growth faltering in resource-constrained settings by analyzing samples from infants in the Peru, Bangladesh, and Tanzania sites of the MAL-ED birth cohort (*13*–*16*). Defining age-related biochemical maturation provides a normative reference for healthy metabolic age-specific profiles in children from low-income settings and identifies pathways that are perturbed and vulnerable windows to permit for more focused nutritional interventions to optimize growth in early childhood.

RESULTS

Urine samples were collected at 3, 6, 9, 15, and 24 months of age for Peru ($n = 281$ infants, $N = 1057$ samples) and at 3, 6, 9, and 15 months for Bangladesh ($n = 249$ infants, $N = 860$ samples). Infants from Tanzania were sampled at 6, 15, and 24 months of age ($n = 249$ infants, $N = 506$ samples). Plasma samples were collected from infants in Peru and Bangladesh at 7 and 15 months of age (Peru, $n = 230$ infants; Bangladesh, $n = 223$ infants). Anthropometry was recorded every month from birth to 24 months of age (*13*). Infants with an LAZ

[1]Division of Integrative Systems Medicine and Digestive Diseases, Imperial College London, London, UK. [2]Fogarty International Center, Bethesda, MD, USA. [3]Asociación Benéfica PRISMA, Unidad de Investigación Biomedica, Iquitos, Peru. [4]International Center for Diarrhoeal Disease Research, Dhaka, Bangladesh. [5]Haydom Global Health Institute, Haydom, Tanzania. [6]Division of Infectious Diseases and International Health, University of Virginia, Charlottesville, VA, USA. [7]Haukeland University Hospital, Bergen, Norway. [8]Department of International Health, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA. [9]School of Human Development and Health, Faculty of Medicine, University of Southampton, UK.
*These authors contributed equally to this work.
†Corresponding author. Email: mkosek@virginia.edu

score of ≥−0.75 at birth and ≥−1.25 at 24 months of age (Peru, $n = 22$; Bangladesh, $n = 28$; Tanzania, $n = 13$; total infants, $n = 63$) were classified as an internal "healthy" reference group (details provided in table S1). Although their attained height and growth is below the median of populations of well-nourished children in resource-rich environments, this group was considered to be more metabolically optimized, as their LAZ at both birth and 24 months of age represented the highest decile in their context. Infants below these thresholds at either or both time points were categorized as "growth constrained" (Peru, $n = 259$; Bangladesh, $n = 221$; Tanzania, $n = 236$).

Age-associated metabolic variation in healthy and growth-constrained infants

Untargeted ^1H nuclear magnetic resonance (NMR) spectroscopy was used to comprehensively characterize the metabolic profiles of the urine samples [estimated coverage, ~400 metabolites (17)]. To identify age-related biochemical variation in infants of different growth trajectories, projection to latent structures (PLS) models were constructed using the full unannotated urinary ^1H NMR spectral profiles from all sampling points. In these models, all spectral features were included as the descriptor variables and the age of the infant at the time of sampling was used as the response variable. Significant PLS predictive models were built for both healthy and growth-constrained infants from all sites (table S2). Significant covariate adjusted–PLS (CA-PLS) models were also obtained combining the healthy infants from all three cohorts and growth-constrained infants from all cohorts, separately, adjusting for site (annotated CA-PLS coefficient plots for healthy and growth-constrained infants shown in figs. S1 and S2, respectively). Additional covariates were assessed for their ability to affect the urinary metabolic profile using PLS models. This included sex, breast-feeding, diarrhea, WAMI [a socioeconomic status index (18)], acute lower respiratory infections, and objective fever at the time of sampling. However, none of these covariates were found to consistently alter the biochemical profile across the sites (table S3). From the significant site-adjusted models, the discriminatory spectral features covarying with age were identified. In total, 32 metabolites were identified to be significantly associated with age in either healthy or growth-constrained infants (Fig. 1A). Several biochemical pathways were altered in an age-dependent manner consistently across the sites. In both healthy and growth-constrained infants, 12 metabolites were excreted in higher amounts early in life and their excretion decreased with age: this included choline/betaine-related metabolites [betaine, dimethylglycine (DMG), and glycine], tricarboxylic acid (TCA) cycle intermediates (citrate, fumarate, and succinate), amino acids alanine and taurine, sugars (galactose and mannitol), the niacin catabolite N-methylnicotinamide (NMND), and N-acetylglycoprotein (NAG), a biomarker of general inflammation. In contrast, metabolites arising from microbial-host cometabolism were found to increase with infant age indicative of the dynamic maturation of the gut microbiome. This included 3-indoxyl sulfate (3-IS; from tryptophan), 4-cresol sulfate (from tyrosine), phenylacetylglutamine (PAG; from phenylalanine), and 2-hydroxyisobutyrate (2-HIB; protein degradation). 4-Hydroxyphenylacetate (4-HPA; from tyrosine), hippurate (from polyphenols), and 4-hydroxyhippurate (from polyphenols) also increased with age. These metabolites can be derived from the biochemical interactions between the gut microbiota and the host and can also be obtained from dietary sources or the host processing of dietary precursors. Dietary sucrose and creatine were also positively correlated with aging, and

excretion of N-methylnicotinic acid (NMNA; trigonelline), a marker of niacin deficiency (19), also exhibited a strong positive association to infant aging.

A targeted ultraperformance liquid chromatography–mass spectrometry (UPLC-MS) approach was used to measure the concentrations of 24 plasma amino acids and related metabolites in the infants from Peru and Bangladesh at 7 and 15 months of age (table S4 lists amino acids measured). Significant age-related PLS models were obtained from the plasma amino acid signatures from the healthy and growth-constrained infants (model diagnostics in table S2). In total, four plasma metabolites were significantly correlated with age in infants from both Peru and Bangladesh irrespective of growth status (Fig. 1B). Plasma ethanolamine, glutamine, and tryptophan declined with age, while citrulline was positively associated with age.

Phenome-for-age Z score—A metric for biochemical maturity

The existence of age-discriminatory metabolites common across the cohorts indicates that the infant metabolic maturation is consistent across these highly diverse epidemiologic settings. As such, we sought to identify core urinary metabolites that could be used to calculate a "phenome age" (PA) to describe metabolic maturity. Urine was selected over plasma because of the diversity of metabolic changes observed with aging and its ease of collection facilitating the potential implementation of this metric in the field. To calculate the PA of an infant, a PLS model ($Q^2\hat{Y} = 0.46$) was built on the internal healthy reference population ($n = 63$ infants) using the metabolites significantly associated with chronological age in these infants (Fig. 1A). This included betaine, DMG, citrate, fumarate, succinate, alanine, taurine, creatine, galactose, sucrose, NMNA, 3-IS, 4-cresyl sulfate (4-CS), hippurate, PAG, 4-HPA, 4-hydroxyhippurate, 2-HIB, and NAG.

To facilitate the use of metabolic age as a potential diagnostic tool in the field, the metabolites used to calculate PA were further refined. Linear mixed-effects models (to account for repeated measures) were constructed to estimate the mean effect of each metabolite, normalized to mean zero and unit SD, on PA (fig. S3). The variance in PA explained by each biochemical feature was calculated from the type II sum of squares from the models. Betaine, DMG, citrate, succinate, creatine, hippurate, PAG, and 4-CS had the strongest effect on PA (Fig. 1C). In Peru, these eight metabolites explained 98.8% of variance in PA, 99.1% in Bangladesh, and 96.8% in Tanzania. These eight metabolites were subsequently used to recalculate the PA of both healthy and growth-constrained infants. The time-dependent variation in the excretion of these metabolites is shown in Fig. 2 (fig. S4 shows the excretion trajectories by cohort). Reference concentrations of the eight urinary metabolites used for the calculation of PA are presented in tables S5 and S6 for healthy and growth-constrained infants, respectively. Citrate had the largest influence on PA, with a one-SD increase in citrate excretion resulting in a 2.92-month younger PA. Citrate was excreted in relatively high amounts in the healthy children at 3 and 6 months of age before excretion sharply reduced at 9 months. In the growth-constrained children, citrate excretion was significantly lower than healthy children at 6 months, and the age-related decline in citrate excretion was more gradual. Succinate, another TCA cycle intermediate, followed a similar but less pronounced trend. As with citrate, betaine was excreted in high amounts at 3 months of age before sharply reducing at 6 months before plateauing from 9 months onward.
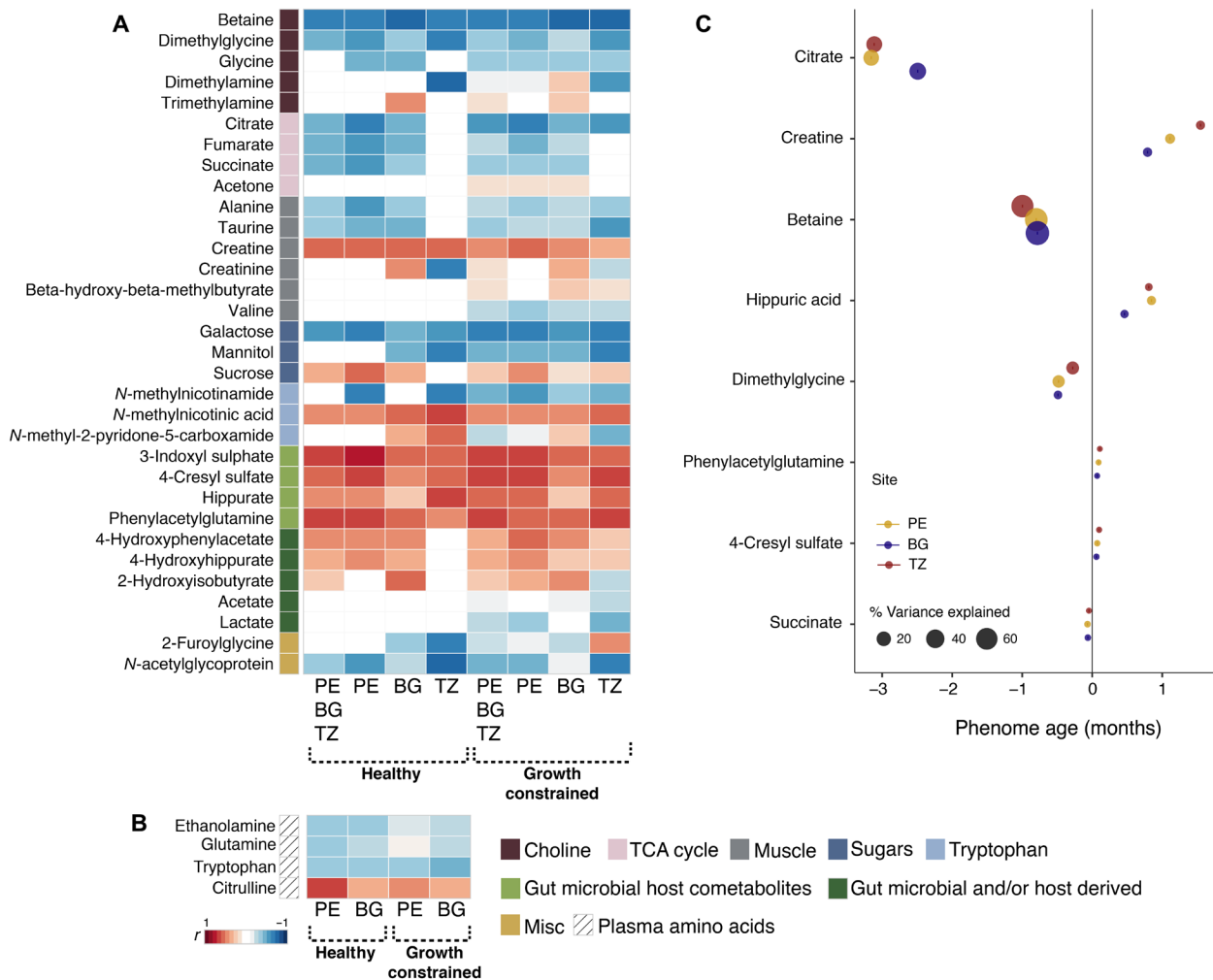
Fig. 1. Biochemical variation associated with aging. (**A**) urinary and (**B**) plasma metabolic profiles of infants from Peru (PE), Bangladesh (BG), and Tanzania (TZ). The heat map presents the correlation coefficient (*r*) obtained from PLS models: Blue colors indicate a negative association with age, and reds represent metabolic shifts positively associated with infant aging. (**C**) Mean effect size of age-discriminatory urinary metabolites on the PA based on children with healthy growth trajectories. The effect size is depicted as the estimated change in PA for each SD change in a metabolite concentration. Color indicates cohort site, and the size of the symbols indicates the percentage PA variance explained.

DMG, a catabolite of betaine, followed a similar, albeit less pronounced, age-related decline in excretion. Conversely, the excretion of creatine and the gut microbial-host cometabolites 4-CS, hippurate, and PAG gradually increased with age.

Following the strategy of Subramanian *et al.* (*11*) who produced microbiome-for-age measures, the PA of each child was compared to the median PA of the healthy reference population at the same chronological age (fig. S5) and converted into a *Z* score [phenome-for-age *Z* score (PAZ); Eq. 2] to standardize values for comparison across the three sites. By definition, the PAZ scores of the children following healthy growth trajectories were approximately constant (mean PAZ, −0.06 ± 0.99) through time, representing stable metabolic maturation (Fig. 3). In contrast, growth-constrained children from the three sites consistently had lower PAZ scores than their healthy growing equivalents (Fig. 3). This indicates that growth-constrained infants consistently lag in their metabolic maturity compared to their healthier peers. For Peru and Bangladesh, the difference in PAZ between the two groups is apparent at 3 months

of age (Mann-Whitney *U*: Peru, $P = 9.67 \times 10^{-6}$; Bangladesh, $P = 9.97 \times 10^{-9}$) and progressively widens until 15 months of age (Peru, $P = 1.68 \times 10^{-10}$; Bangladesh, $P = 3.88 \times 10^{-13}$), where approximately two SDs of difference exist. Urine samples were not available for Tanzanian infants at 3 months of age, but at 6 months, lower PAZ scores were apparent in the growth-constrained individuals compared to the healthy infants ($P = 0.0019$). This difference was exaggerated at 15 months ($P = 0.018$) before narrowing at 24 months of age (Tanzania, $P = 0.019$). By 24 months of age, the gap had also reduced for Peru ($P = 3.79 \times 10^{-9}$).

PAZ as a predictor of linear growth

To investigate whether there was any relationship between PAZ and future linear growth, linear mixed-effects models were constructed, pooling all ages and including random intercepts for child nested in their respective site. Figure 4 shows the mean effect of PAZ on the LAZ between 1 and 6 months after the urine sample, adjusting for the LAZ at the time of the urine sample. The PAZ has a consistently
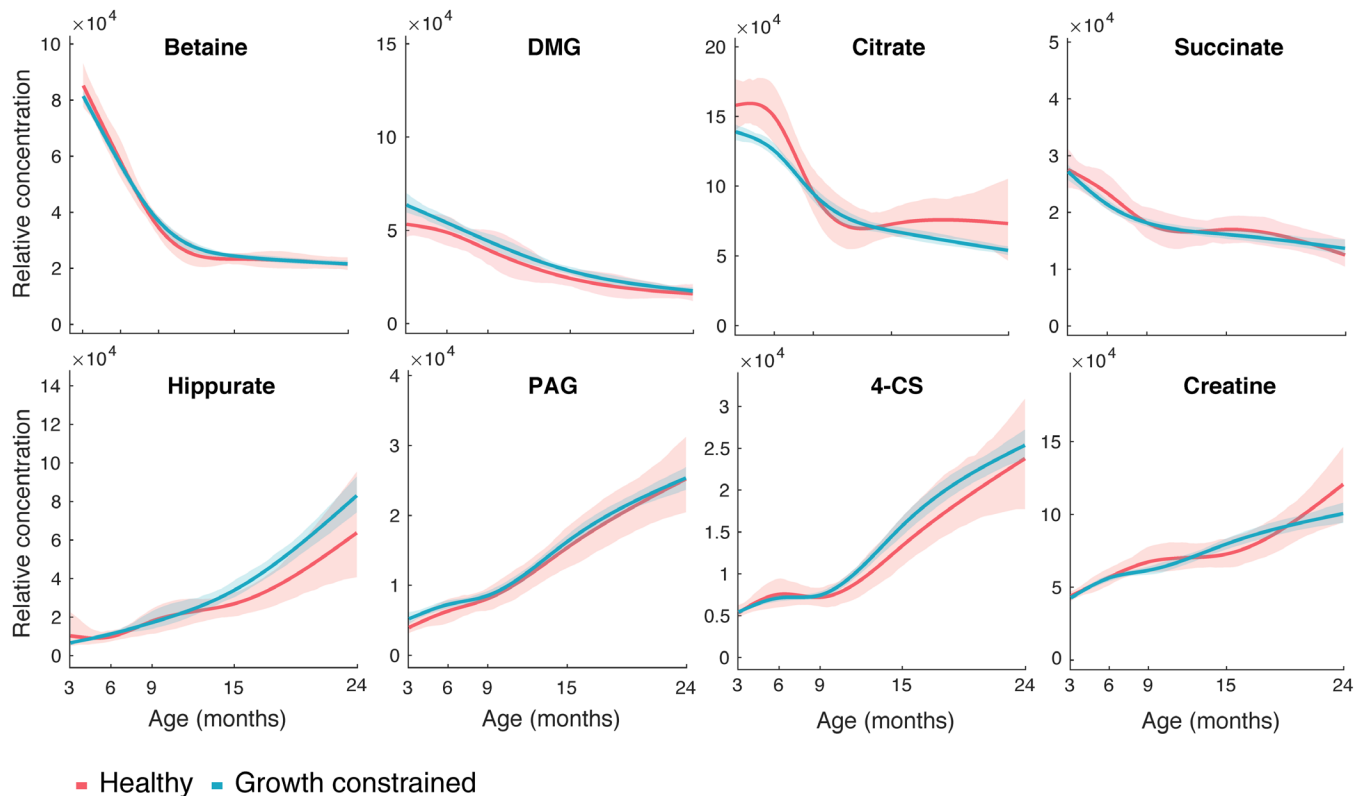
Fig. 2. Time-dependent variation in the eight urinary metabolites used to calculate the PA of the study children. Relative concentrations of metabolites were obtained by measuring the area under selected spectral regions corresponding to betaine, DMG, citrate, succinate, hippurate, PAG, 4-CS, and creatine. Shaded area represents 95% CI.

statistically significant positive association with the LAZ in future months, increasing from 1 to 3 months and then stabilizing at approximately 0.04 SD of LAZ for every increase SD in PAZ by one (the mean effect of a 1-month increase in PAZ is 0.038; 95% confidence interval, 0.017 to 0.058 on the LAZ 3 months later). Similarly, PAZ was associated with markers of wasting (weight-for-length) and underweight (weight-for-age Z score) up to 6 months in the future (see fig. S6).

DISCUSSION

Comprehensive metabolic phenotyping was used to study the maturation of the phenome of infants from resource-constrained settings spanning three continents over the first 2 years of life, a period of time critical for normal human development. An age-dependent metabolic signature was identified in both urine and plasma samples, with a selection of core urinary metabolites changing consistently with age across the cohorts despite different staples and dietary practices. The existence of a common maturation process in children from these settings enabled a model to be constructed to compute the PA of an infant. The urinary metabolites used to predict the PA were refined to the eight most influential age-discriminatory metabolites, increasing its potential for implementation in field settings, via noninvasive sampling. On the basis of these metabolites, early life PAZ was found to be significantly associated with linear and ponderal growth. Furthermore, biochemical immaturity was linked to poor growth attainment, with this immaturity evident as early as 3 months of age persisting until at least the second year of life. Metabolic maturity

evolves over different time frames between individuals and influences the needs of the developing infant. PAZ enables an infant's position along this biochemical maturation continuum to be determined in real time and provides a flexible, sensitive, and noninvasive measure against which the effectiveness of interventions can be measured. These succinct signatures may also enable interventions to be targeted to an individual's specific PA rather than their chronological age, allowing more precise interventions along the spectrum of infancy and early childhood.

Energy requirements of healthy well-nourished children are known to decline with age. Reduction in growth velocity over the first 12 months of life accounts for this change. Growth is an energy-demanding process representing approximately 35% of the total energy requirements during the first 3 months of life, falling to 17.5% at 6 months, and only 3% at 12 months of age (20). The TCA cycle is key for energy production, and the excretion of important TCA cycle intermediates declined with age in these infants. This included citrate and succinate, which were used to define PA, and fumarate, which declined with age in children from Peru and Bangladesh. Complementary to these findings, urinary NMND, a biomarker of nicotinamide and NAD[+] (nicotinamide adenine dinucleotide) (TCA cycle cofactor) availability, also diminished with age in the study children. Consistent with this finding, we have previously observed urinary NMND to be positively associated with future growth in Brazilian children (8). Tryptophan is an essential amino acid that provides 90% of NAD[+] via the de novo synthesis pathway and is also important for growth in infants (10, 21). As with the TCA cycle intermediates and NMND, this circulating amino
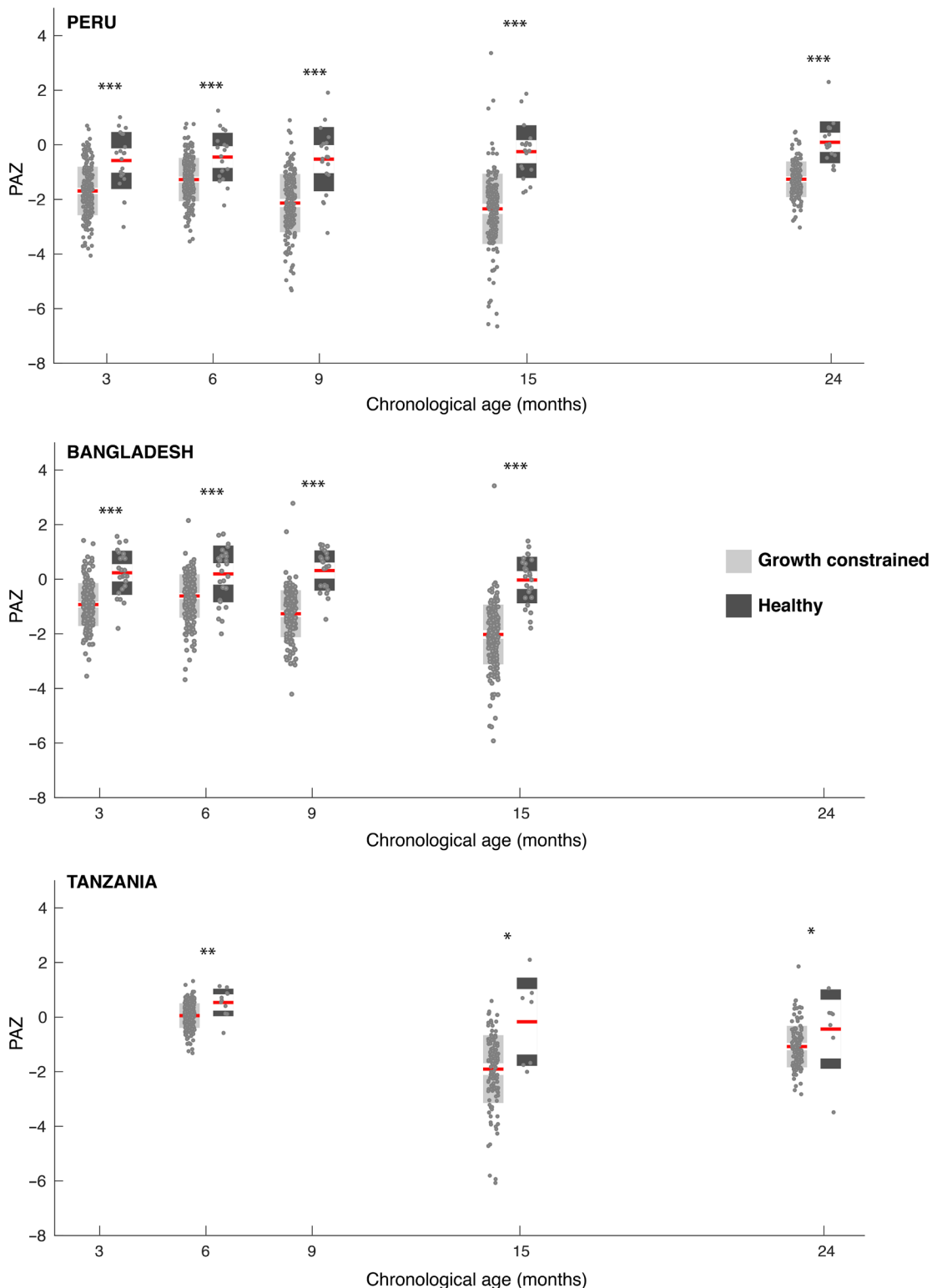
Fig. 3. PAZ of growth-constrained and healthy infants relative to their chronological age across the three sites. The PAZ score of healthy and growth-constrained infants from each site was calculated from eight age-discriminatory urinary metabolites. Significant differences were observed between healthy and growth-constrained children at all sampling points in all cohorts. Mann-Whitney $\underline{U}$ test, $*P < 0.01$, $**P < 0.001$, and $***P < 0.0001$ (Healthy Peru: $N_{3 \text{ months}} = 21$, $N_{6 \text{ months}} = 19$, $N_{9 \text{ months}} = 20$, $N_{15 \text{ months}} = 20$, $N_{24 \text{ months}} = 18$; Growth constrained Peru: $N_{3 \text{ months}} = 220$, $N_{6 \text{ months}} = 214$, $N_{9 \text{ months}} = 197$, $N_{15 \text{ months}} = 183$, $N_{24 \text{ months}} = 141$; Healthy Bangladesh: $N_{3 \text{ months}} = 25$, $N_{6 \text{ months}} = 27$, $N_{9 \text{ months}} = 24$, $N_{15 \text{ months}} = 26$; Growth constrained Bangladesh: $N_{3 \text{ months}} = 196$, $N_{6 \text{ months}} = 197$, $N_{9 \text{ months}} = 181$, $N_{15 \text{ months}} = 184$; Healthy Tanzania: $N_{6 \text{ months}} = 11$, $N_{15 \text{ months}} = 7$, $N_{24 \text{ months}} = 7$; Growth constrained Tanzania: $N_{6 \text{ months}} = 209$, $N_{15 \text{ months}} = 129$, $N_{24 \text{ months}} = 122$).
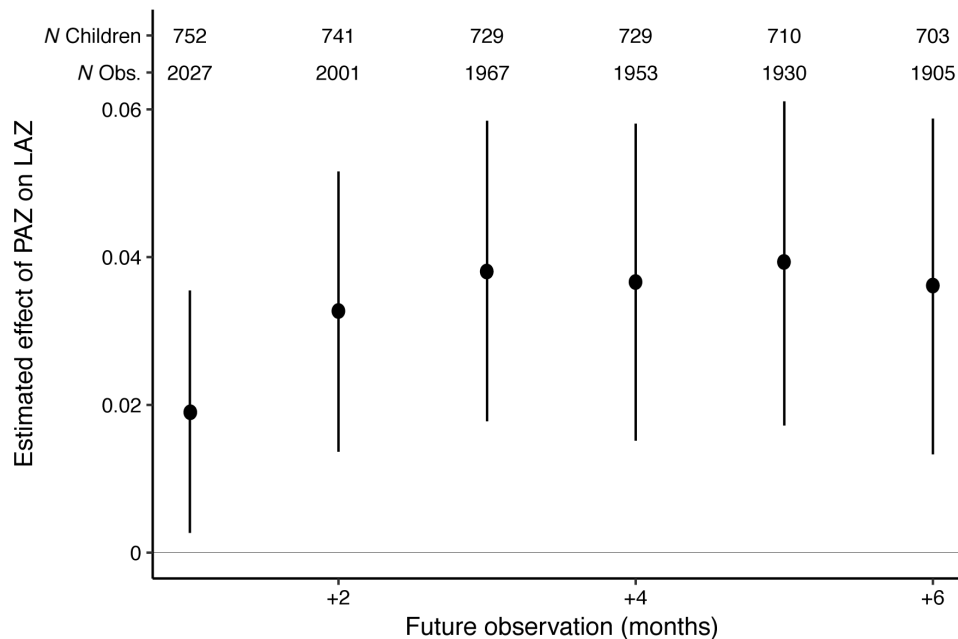
Fig. 4. Mean effect of each additional month of the PAZ on the LAZ 1 to 6 months after the urine sample, adjusting for LAZ at the time of the PAZ estimate and site.

acid was observed to decrease with age. These observations illuminate the critical need for substrate availability during this energetically costly growth phase.

Age-related depletion of plasma glutamine and rises in circulating citrulline and urinary creatine similarly reflect growth and development. Creatine had a substantial positive contribution to PA and most likely reflects increasing muscle mass with age (*22*). Most creatine is stored in skeletal muscle as free creatine or phosphocreatine, which provides a major energy source for the body. Although breast milk contains creatine, most of an infant's creatine is obtained from endogenous synthesis from glycine, arginine, and methionine, and is a major burden on amino acid metabolism (*23*). Similarly, plasma citrulline, a proposed biomarker of functional enterocyte mass, increased with age (*24*). Citrulline is mainly produced by enterocytes in the small intestine and, as it is not incorporated into proteins, reflects small intestinal development. Despite mixed reports regarding age dependence, it seems clear from this large data set that citrulline is age dependent. Conversely, plasma glutamine, a nonessential amino acid abundant in the muscles of humans, was observed to decline with age. Glutamine provides nitrogen for protein synthesis, is a precursor for nucleic acids, and is an important substrate for rapidly proliferating cells of the immune system (*25*) and gastrointestinal tract (*26*). It also acts as a gluconeogenic substrate in the liver and intestine and follows the same trend for age-associated depletion as the TCA cycle intermediate. Glutamine has previously been observed to be lower in the plasma of stunted children from Malawi compared to their nonstunted comparators (*27*).

S-adenosylmethionine (SAMe) is the sole methyl donor for a variety of cellular methylation reactions including DNA and histone methylation, which are major epigenetic events important in the process of metabolic programming. SAMe is generated from the homocysteine-methionine pathway using methyl groups from betaine, demethylating it to DMG. The excretions of betaine and DMG

were important determinants of PA declining sharply with age. This indicates a higher consumption of betaine in the homocysteine-methionine pathway early in life compared to the later sampling points possibly reflecting the higher demands for SAMe during this developmental phase (*28*). Betaine can be obtained directly from the diet or via its precursor choline. Breast milk is a rich source of choline being present at relatively constant amounts (70 to 200 nmol/ml) after the first week of lactation. In this study, infants were breast-fed until 18 months of age, so these changes are unlikely to be driven by the weaning process. The abundance of choline and betaine is high in neonates compared to adults (*29*) and both have been seen to decrease in the sera of Malawian children with age over the first 5 years of life (*30*). Choline is an essential nutrient important for growth and development (*31*) including the linear growth of bone (*32*). The progressive loss of DNA methylation has been implicated in the slowing of bone growth (*33*). The age-associated decline in betaine and DMG, in parallel with the diminishing energy and growth-related metabolites, is consistent with a period of rapid growth in the first 6 months of life followed by a relative slowing of growth 12 to 18 months later. Consistently, we have previously observed betaine and DMG excretion to be positively associated with growth in Brazilian children (*8*), while stunted children from Malawi had lower circulating amounts of choline compared to nonstunted children (*27*). On the basis of these observations, betaine demand is particularly high during the first 6 months of life, diminishing at subsequent time points. This suggests a potential critical window for supplementation that may influence epigenetic programming and subsequent long-term health effects (*34*, *35*).

The microbiota has an important role in digestion and gut health, shaping the bioavailability of certain nutritional components. Several metabolites arising from the combined metabolism of the intestinal microbiota and the host were noted to increase with age, reflecting the functional maturation of the microbiome and an expansion of the nutritional inputs they receive. For example, hippurate and

4-hydroxyhippurate arise from the bacterial breakdown of plant-derived polyphenols (*36*). Hippurate has been previously correlated with *Faecalibacterium prausnitzii* and *Clostridiales* sp., both of which were found to establish in the weaning phase of a healthy Bangladeshi infant cohort (*11*, *37*). 4-CS was another microbial-host cometabolite observed to increase with age. This is produced from the progressive bacterial metabolism of tyrosine to 4-HPA (also found to increase with age) and then 4-cresol, which is absorbed and undergoes hepatic sulfation to form 4-CS. Similarly, phenyl-alanine can be degraded to phenylacetate by the intestinal bacteria before host conjugation with glutamine and excretion as PAG. Both 4-CS and PAG have been positively associated with a number of infections endemic in these settings. As such, increases in their excretion with age may also relate to the increasing burden of infections in these infants (*38*, *39*).

Age-related increases in the excretion of hippurate, PAG, and 4-CS indicate the greater functionality of the microbiome with age and have been found to be positively associated with microbiome diversity (*37*). All these microbial-derived metabolites contributed to the calculation of PA across all cohorts. The assembly of the fecal microbiota in healthy children from Bangladesh has been previously described (*11*). In this work, the authors used compositional data to compute a MAZ to measure the relative maturity of the microbiota. Severely acutely malnourished infants had a lower MAZ than the healthy infants, reflecting an immature microbiota following severe nutritional deficiency (*40*). As PAZ partly reflects the functional maturity of the microbiome and its biochemical cross-talk with the host, it complements the idea of composite maturity provided by MAZ. This maturation of the microbiome and its influence on nutrient flow to the host emphasizes the need to consider the overall supra-organism, with its multidimensional interactions, when studying human development and designing and implementing novel therapeutics with the goal of optimizing early childhood nutrition.

Here, we report metabolic signatures of infant aging and how they diverge between infants on different growth trajectories, across three distinct geographical locations of similar socioeconomic profiles. A PAZ score computed on a simplified battery of urinary metabolites provides an index of the dynamic biochemical maturity of a child at a given age and is effective in predicting future linear growth. Windows of intervention defined by chronological age may eventually be replaced with more dynamic individualized measures of maturity such as the PA.

MATERIALS AND METHODS
Study design
The Etiology, Risk Factors, and Interactions of Enteric Infections and Malnutrition and the Consequences for Child Health (MAL-ED) Study uses a prospective longitudinal design to investigate how undernutrition and repeated enteric infections reinforce one another and affect growth and cognitive development, as well as other health or illness indicators, in infant cohorts across eight countries with high incidence of diarrheal disease and undernutrition (*13*). The detailed study design and methods of data collection have been outlined in the inaugural study manuscript (*13*). Because of the large scope of the study, detailed information on data and sample collection—pertaining to common infant and childhood illnesses, medication usage, administered vaccines, growth measurements, breastfeeding and dietary intake assessments, stool collections for microbiological and gut functional assays and antigen detection, blood collections for micronutrients, serological responses to vaccines, urine collections for micronutrients and gut functional assays, and cognitive testing at various ages—has been separately published, as listed in the introductory MAL-ED study design paper (*13*).

Sample collection
Urine and plasma samples were collected from three cohorts in Peru, Bangladesh, and Tanzania, with anthropometric measures recorded every month from birth (*13*). Urine samples (5-hour collections) were collected at 3, 6, 9, 15, and 24 months for Peru ($n = 273$) and at 3, 6, 9, and 15 for Bangladesh ($n = 249$). The collection time points for Tanzania ($n = 249$) were at 6, 15, and 24 months. Plasma samples were collected from fasted infants in Peru and Bangladesh at 7 and 15 months of age. Samples were stored at −80°C and shipped in dry ice to Imperial College London for analysis. The urine sample set for Peru included 1057 samples, 860 for Bangladesh, and 506 for Tanzania.

^{1}H NMR spectroscopy and UPLC-MS–based metabolic profiling
Urinary metabolic profiles were measured by ^{1}H NMR spectroscopy using the protocols described by Dona *et al.* (*41*) and Beckonert *et al.* (*42*). Briefly, 630 μl of urine sample was combined with 70 μl of phosphate buffer solution (pH 7.4, 100% D_2O) containing 1 mM of the internal standard 3-trimethylsilyl-1-[2,2,3,3-2H4] propionate (TSP). Samples were then vortexed and spun (12,000*g*) for 10 min at 4°C before transfer to 5-mm NMR tubes. A pooled urine quality control sample was prepared by combining 5 μl of each individual sample of the study and was used to monitor variability of the analytical platform. One-dimensional 600-MHz ^{1}H NMR spectra were acquired on a Bruker NMR spectrometer (Bruker BioSpin GmbH, Rheinstetten, Germany), equipped with a SampleJet system and a cooling rack of refrigerated tubes at 6°C. ^{1}H NMR spectra acquisition was achieved using a standard one-dimensional solvent suppression pulse sequence (relaxation delay, 90° pulse, 4-μs delay, 90° pulse, mixing time, 90° pulse, acquire free induction decay). For each sample, 32 transients were collected in 64,000 frequency domain points with a spectral window set to 20 ppm (parts per million). A relaxation delay of 4 s, a mixing time of 10 ms, an acquisition time of 2.73 s, and 0.3-Hz line broadening were used. Spectra were referenced to the TSP resonance at δ 0.0. Spectral phasing and baseline correction were automatically performed using Topspin 3.2 (Bruker Biospin GmbH, Rheinstetten, Germany). The resulting raw NMR spectra were digitized, aligned, and normalized using the Imperial Metabolic Profiling and Chemometrics Toolbox (https://github.com/csmsoftware/IMPaCTS) in MATLAB (version 2018a, MathWorks Inc.). Briefly, after digitization of the spectra, redundant peaks (TSP, H_2O, and urea) were removed and the resulting spectra were manually aligned to reference peaks using recursive segment-wise peak alignment (*43*). The aligned spectra were normalized using probabilistic quotient normalization (*44*). This approach adjusts the metabolite concentrations for differences in sample dilution as a result of differences in liquid and food intakes between infants, offsetting any potential confounding effect from these factors (*45*).

A targeted amino acid and tryptophan metabolite assay was carried out using UPLC-MS for the plasma samples using a Waters Acquity UPLC coupled to a Xevo TQ-XS mass spectrometer following the method published by Gray *et al.* (*46*). For the plasma samples, data were acquired and processed using Waters MassLynx

(version 4.2) followed by multivariate analysis using SIMCA (version 15, Sartorius Stedim Biotech).

Data analysis

PLS regression models were constructed to identify urinary metabolic features associated with infant aging, where the ^{1}H NMR metabolic profiles served as the descriptor matrix and chronological age of the infants was the response variable. Age discriminant features were then identified using in-house databases and the Human Metabolome Database (http://www.hmdb.ca/). The predictive power of each model was calculated using sevenfold cross-validation approach, and model validity (given as P value) was calculated by permutation testing (1000 permutations; significance for $P_{per} \leq 0.05$). All model diagnostics are presented in table S2.

To adjust for potential confounders, CA-PLS models were also constructed using ^{1}H NMR metabolic profiles of healthy and growth-constrained infants (47). Model robustness was assessed using Monte-Carlo cross-validation using a total of 100 models for the 24,706 centered and scaled spectral variables. A P value was calculated for each variable on the basis of 25 bootstrap resamplings of the data in each of the 100 models to estimate the variance and the mean coefficient across the 100 models. Spectral variable importance was assessed with the false discovery rate Q value, with a value of ≤ 0.05.

The relative concentrations of all age-associated metabolites were calculated from the spectral data using trapezoidal numerical integration, and the correlation coefficients (rho) between each of these features to age were extracted from the PLS models and plotted as a heat map to demonstrate positive and negative correlations of metabolites to infant aging.

Estimated reference concentrations for metabolites were calculated from the individual metabolite integrals relative to the TSP integral. The area under the 9-proton TSP signal corresponds to 0.11 mM in each sample. For metabolites with more than one resonance, the resonance in a nonoverlapping spectral region was used. Concentrations of the metabolites were calculated using the following equation (Eq. 1)

$$\frac{M}{S} = \frac{I_M}{I_S} \times \frac{N_S}{N_M}$$

where M corresponds to the unknown concentration of the metabolite, S is the known concentration of the internal standard (TSP), I_M and I_S are the integrals of the metabolite and TSP, and N_S and N_M correspond to the number of protons of the metabolite and TSP. Concentrations of the eight metabolites used for the calculation of PA are expressed as ratios relative to urinary creatinine. It should be noted that creatinine is known to change with age and growth and across different sites.

PAZ calculation

A PA score was calculated to assess biochemical maturity relative to chronological age. An in-sample reference population of infants on a healthy growth trajectory was created based on a LAZ score of ≥ -0.75 at birth and ≥ -1.25 at 24 months of age (Peru, $n = 22$; Bangladesh, $n = 28$; Tanzania, $n = 13$), while children below these thresholds were categorized as "growth constrained" (Peru, $n = 259$; Bangladesh, $n = 221$; Tanzania, $n = 236$). Significant age-associated metabolites were determined in the group of children on the healthy growth trajectory from the three sites combined using PLS regression modeling. Linear regression was used to model the relationship

between the true chronological age of the healthy infants and the calculated scores from the PLS model. Similarly, a PLS regression model was built using the metabolic profiles and the true chronological age of the growth-constrained infants. The scores from the growth-constrained PLS model, together with the linear model predictor variables from the healthy infants, were used to predict the PA of growth-constrained infants. This predicted value therefore reflected the biochemical or PA of all infants—healthy and growth constrained. The predictive capacity of the regression strategy was assessed using a sevenfold cross-validation approach represented as the $Q^2\hat{Y}$ value.

The list of urinary metabolites used in the calculation of PA was refined by calculating the magnitude of influence of each of the biochemical features on the PA. The estimated effect size, as well as the PA variance explained by each metabolite, was calculated using a linear mixed-effects model (accounting for repeated measures of each child at different ages) for each of the three sites separately. Metabolites were ranked on the basis of their contribution to PA (type II sum of squares), and a refined list of urinary metabolic features was taken forward to recalculate the PA of each child.

On the basis of a similar approach used by Subramanian et al. (11) studying microbiome-for-age measures, a predicted PAZ measure was calculated using Eq. 2

$$PAZ = \frac{PA \text{ of child} - \text{Median PA of healthy of same chronological age}}{\text{Standard deviation of healthy of same chronological age}}$$

Models were constructed to examine the association between PAZ and future LAZ. PAZ was pooled by age and site, and linear mixed-effects models were constructed to estimate the LAZ 1 to 6 months after urine samples were taken. The models adjusted for the LAZ at the time of the urine sample (and thereby when the PAZ was calculated) and random intercepts were included for each child nested in their respective site to account for repeated observations. The models yield the average SD change in LAZ for each additional month change in PAZ.

SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at http://advances.sciencemag.org/cgi/content/full/6/15/eaay5969/DC1

View/request a protocol for this paper from *Bio-protocol*.

REFERENCES AND NOTES

1. S. Grantham-McGregor, Y. B. Cheung, S. Cueto, P. Glewwe, L. Richter, B. Strupp; International Child Development Steering Group, Developmental potential in the first 5 years for children in developing countries. *Lancet* **369**, 60–70 (2007).
2. UNICEF/WHO/World Bank, Malnutrition, UNICEF Data (2019); https://www.who.int/news-room/fact-sheets/detail/malnutrition.
3. R. E. Black, C. G. Victora, S. P. Walker, Z. A. Bhutta, P. Christian, M. de Onis, M. Ezzati, S. Grantham-McGregor, J. Katz, R. Martorell, R. Uauy; Maternal and Child Nutrition Study Group, Maternal and child undernutrition and overweight in low-income and middle-income countries. *Lancet* **382**, 427–451 (2013).
4. The Independent Expert Group of the Global Nutrition Report, *Global Nutrition Report: The Burden of Malnutrition* (Development Initiatives Poverty Research Ltd., 2018); https://globalnutritionreport.org/reports/global-nutrition-report-2018/burden-malnutrition/.
5. Z. A. Bhutta, J. K. Das, A. Rizvi, M. F. Gaffey, N. Walker, S. Horton, P. Webb, A. Lartey, R. E. Black; Lancet Nutrition Interventions Review Group; Maternal and Child Nutrition Study Group, Evidence-based interventions for improvement of maternal and child nutrition: What can be done and at what cost? *Lancet* **382**, 452–477 (2013).
6. K. Mennon, S. Puthussery, A. Ravalia, P. Panchal, R. Rana, S. Mistry, P. Tseng, J. Bhandol, D. Mavalankar, *Effectiveness of Nutrition Interventions in Low and Middle Income Countries:*

An Evidence Summary (Indian Institute of Public Health Gandhinagar; Public Health Foundation of India; University of Bedfordshire, UK and BRAC, Bangladesh, 2018), https://www.gov.uk/dfid-research-outputs/effectiveness-of-nutrition-interventions-in-low-and-middle-income-countries-an-evidence-summary.

7. J. Mayneris-Perxachs, J. R. Swann, Metabolic phenotyping of malnutrition during the first 1000 days of life. *Eur. J. Nutr.* **58**, 909–930 (2019).

8. J. Mayneris-Perxachs, A. A. M. Lima, R. L. Guerrant, Á. M. Leite, A. F. Moura, N. L. Lima, A. M. Soares, A. Havt, S. R. Moore, R. Pinkerton, J. R. Swann, Urinary N-methylnicotinamide and β-aminoisobutyric acid predict catch-up growth in undernourished Brazilian children. *Sci. Rep.* **6**, 19780 (2016).

9. M. Farràs, K. Chandwe, J. Mayneris-Perxachs, B. Amadi, J. Louis-Auguste, E. Besa, K. Zyambo, R. Guerrant, P. Kelly, J. R. Swann, Characterizing the metabolic phenotype of intestinal villus blunting in Zambian children with severe acute malnutrition and persistent diarrhea. *PLOS ONE* **13**, e0192092 (2018).

10. M. N. Kosek, E. Mduma, P. S. Kosek, G. O. Lee, E. Svensen, W. K. Y. Pan, M. P. Olortegui, J. H. Bream, C. Patil, C. R. Asayag, G. M. Sanchez, L. E. Caulfield, J. Gratz, P. P. Yori, Plasma tryptophan and the kynurenine–tryptophan ratio are associated with the acquisition of statural growth deficits and oral vaccine underperformance in populations with environmental enteropathy. *Am. J. Trop. Med. Hyg.* **95**, 928–937 (2016).

11. S. Subramanian, S. Huq, T. Yatsunenko, R. Haque, M. Mahfuz, M. A. Alam, A. Benezra, J. DeStefano, M. F. Meier, B. D. Muegge, M. J. Barratt, L. G. VanArendonk, Q. Zhang, M. A. Province, W. A. Petri Jr., T. Ahmed, J. I. Gordon, Persistent gut microbiota immaturity in malnourished Bangladeshi children. *Nature* **510**, 417–421 (2014).

12. L. V. Blanton, M. R. Charbonneau, T. Salih, M. J. Barratt, S. Venkatesh, O. Ilkaveya, S. Subramanian, M. J. Manary, I. Trehan, J. M. Jorgensen, Y.-m. Fan, B. Henrissat, S. A. Leyn, D. A. Rodionov, A. L. Osterman, K. M. Maleta, C. B. Newgard, P. Ashorn, K. G. Dewey, J. I. Gordon, Gut bacteria that prevent growth impairments transmitted by microbiota from malnourished children. *Science* **351**, aad3311 (2016).

13. The MAL-ED Network Investigators, The MAL-ED Study: A multinational and multidisciplinary approach to understand the relationship between enteric pathogens, malnutrition, gut physiology, physical growth, cognitive development, and immune responses in infants and children up to 2 years of age in resource-poor environments. *Clin. Infect. Dis.* **59**, S193–S206 (2014).

14. P. P. Yori, G. Lee, M. P. Olórtegui, C. B. Chávez, J. T. Flores, A. O. Vasquez, R. Burga, S. R. Pinedo, C. R. Asayag, R. E. Black, L. E. Caulfield, M. Kosek, Santa Clara de Nanay: The MAL-ED cohort in Peru. *Clin. Infect. Dis.* **59**, S310–S316 (2014).

15. T. Ahmed, M. Mahfuz, M. M. Islam, D. Mondal, M. I. Hossain, A. S. Ahmed, F. Tofail, S. A. Gaffar, R. Haque, R. L. Guerrant, W. A. Petri, The MAL-ED cohort study in Mirpur, Bangladesh. *Clin. Infect. Dis.* **59**, S280–S286 (2014).

16. E. R. Mduma, J. Gratz, C. Patil, K. Matson, M. Dakay, S. Liu, J. Pascal, L. McQuillin, E. Mighay, E. Hinken, A. Ernst, C. Amour, R. Mvungi, E. Bayyo, Y. Zakaria, S. Kivuyo, E. R. Houpt, E. Svensen, The etiology, risk factors, and interactions of enteric infections and malnutrition and the consequences for child health and development study (MAL-ED): Description of the Tanzanian site. *Clin. Infect. Dis.* **59**, S325–S330 (2014).

17. V. Ghini, D. Quaglio, C. Luchinat, P. Turano, NMR for sample quality assessment in metabolomics. *Nat. Biotechnol.* **52**, 25–34 (2019).

18. S. R. Psaki, J. C. Seidman, M. Miller, M. Gottlieb, Z. A. Bhutta, T. Ahmed, A. S. Ahmed, P. Bessong, S. M. John, G. Kang, M. Kosek, A. Lima, P. Shrestha, E. Svensen, W. Checkley; MAL-ED Network Investigators, Measuring socioeconomic status in multicountry studies: Results from the eight-country MAL-ED study. *Popul. Health Metr.* **12**, 8 (2014).

19. J. S. Sandhu, D. R. Fraser, The metabolic origin of trigonelline in the rat. *Biochem. J.* **200**, 495–500 (1981).

20. FAO/WHO/UNU, Human energy requirements, FAO Food and Nutrition Technical Report Series 1 (Report of a Joint FAO/WHO/UNU Expert Consultation, 2001); http://www.fao.org/3/y5686e/y5686e05.htm.

21. Y. B. Shen, G. Voilqué, J. D. Kim, J. Odle, S. W. Kim, Effects of increasing tryptophan intake on growth and physiological changes in nursery pigs. *J. Anim. Sci.* **90**, 2264–2275 (2012).

22. R. Catherwood, G. Stearns, Creatine and creatinine excretion in infancy. *J. Biol. Chem.* **119**, 201–204 (1937).

23. J. T. Brosnan, R. P. da Silva, M. E. Brosnan, The metabolic burden of creatine synthesis. *Amino Acids* **40**, 1325–1331 (2011).

24. P. Crenn, B. Messing, L. Cynober, Citrulline as a biomarker of intestinal failure due to enterocyte mass reduction. *Clin. Nutr.* **27**, 328–339 (2008).

25. V. Cruzat, M. Macedo Rogero, K. Noel Keane, R. Curi, P. Newsholme, Glutamine: Metabolism and immune function, supplementation and clinical translation. *Nutrients* **10**, E1564 (2018).

26. M.-H. Kim, H. Kim, The roles of glutamine in the intestine and its implication in intestinal diseases. *Int. J. Mol. Sci.* **18**, E1051 (2017).

27. R. D. Semba, M. Shardell, F. A. Sakr Ashour, R. Moaddel, I. Trehan, K. M. Maleta, M. I. Ordiz, K. Kraemer, M. A. Khadeer, L. Ferrucci, M. J. Manary, Child stunting is associated with low circulating essential amino acids. *EBioMedicine* **6**, 246–252 (2016).

28. L. E. McBreairty, R. F. Bertolo, The dynamics of methionine supply and demand during early development. *Appl. Physiol. Nutr. Metab.* **41**, 581–587 (2016).

29. A. M. Wiedeman, S. I. Barr, T. J. Green, Z. Xu, S. M. Innis, D. D. Kitts, Dietary choline intake: Current state of knowledge across the life cycle. *Nutrients* **10**, E1513 (2018).

30. R. D. Semba, P. Zhang, M. Gonzalez-Freire, R. Moaddel, I. Trehan, K. M. Maleta, M. I. Ordiz, L. Ferrucci, M. J. Manary, The association of serum choline with linear growth failure in young children from rural Malawi. *Am. J. Clin. Nutr.* **104**, 191–197 (2016).

31. S. H. Zeisel, K.-A. da Costa, Choline: An essential nutrient for public health. *Nutr. Rev.* **67**, 615–623 (2009).

32. J. Kular, J. C. Tickner, N. J. Pavlos, H. M. Viola, T. Abel, B. S. Lim, X. Yang, H. Chen, R. Cook, L. C. Hool, M. H. Zheng, J. Xu, Choline kinase β mutant mice exhibit reduced phosphocholine, elevated osteoclast activity, and low bone mass. *J. Biol. Chem.* **290**, 1729–1742 (2015).

33. O. Nilsson, R. D. Mitchum Jr., L. Schrier, S. P. Ferns, K. M. Barnes, J. F. Troendle, J. Baron, Growth plate senescence is associated with loss of DNA methylation. *J. Endocrinol.* **186**, 241–249 (2005).

34. C. Junien, Impact of diets and nutrients/drugs on early epigenetic programming. *J. Inherit. Metab. Dis.* **29**, 359–365 (2006).

35. G. C. Burdge, K. A. Lillycrop, Nutrition, epigenetics, and developmental plasticity: Implications for understanding human disease. *Annu. Rev. Nutr.* **30**, 315–339 (2010).

36. H. J. Lees, J. R. Swann, I. D. Wilson, J. K. Nicholson, E. Holmes, Hippurate: The natural history of a mammalian–microbial cometabolite. *J. Proteome Res.* **12**, 1527–1546 (2013).

37. T. Pallister, M. A. Jackson, T. C. Martin, J. Zierer, A. Jennings, R. P. Mohney, A. MacGregor, C. J. Steves, A. Cassidy, T. D. Spector, C. Menni, Hippurate as a metabolomic marker of gut microbiome diversity: Modulation by diet and relationship to metabolic syndrome. *Sci. Rep.* **7**, 13670 (2017).

38. E. Holmes, J. V. Li, T. Athanasiou, H. Ashrafian, J. K. Nicholson, Understanding the role of gut microbiome–host metabolic signal disruption in health and disease. *Trends Microbiol.* **19**, 349–359 (2011).

39. L. A. Bartelt, D. T. Bolick, J. Mayneris-Perxachs, G. L. Kolling, G. L. Medlock, E. I. Zaenker, J. Donowitz, R. V. Thomas-Beckett, A. Rogala, I. M. Carroll, S. M. Singer, J. Papin, J. R. Swann, R. L. Guerrant, Cross-modulation of pathogen-specific pathways enhances malnutrition during enteric co-infection with Giardia lamblia and enteroaggregative Escherichia coli. *PLOS Pathog.* **13**, e1006471 (2017).

40. M. I. Smith, T. Yatsunenko, M. J. Manary, I. Trehan, R. Mkakosya, J. Cheng, A. L. Kau, S. S. Rich, P. Concannon, J. C. Mychaleckyj, J. Liu, E. Houpt, J. V. Li, E. Holmes, J. Nicholson, D. Knights, L. K. Ursell, R. Knight, J. I. Gordon, Gut microbiomes of Malawian twin pairs discordant for kwashiorkor. *Science* **339**, 548–554 (2013).

41. A. C. Dona, B. Jiménez, H. Schäfer, E. Humpfer, M. Spraul, M. R. Lewis, J. T. M. Pearce, E. Holmes, J. C. Lindon, J. K. Nicholson, Precision high-throughput proton NMR spectroscopy of human urine, serum, and plasma for large-scale metabolic phenotyping. *Anal. Chem.* **86**, 9887–9894 (2014).

42. O. Beckonert, H. C. Keun, T. M. D. Ebbels, J. Bundy, E. Holmes, J. C. Lindon, J. K. Nicholson, Metabolic profiling, metabolomic and metabonomic procedures for NMR spectroscopy of urine, plasma, serum and tissue extracts. *Nat. Protoc.* **2**, 2692–2703 (2007).

43. K. A. Veselkov, J. C. Lindon, T. M. D. Ebbels, D. Crockford, V. V. Volynkin, E. Holmes, D. B. Davies, J. K. Nicholson, Recursive segment-wise peak alignment of biological ^{1}H NMR spectra for improved metabolic biomarker recovery. *Anal. Chem.* **81**, 56–66 (2009).

44. F. Dieterle, A. Ross, G. Schlotterbeck, H. Senn, Probabilistic quotient normalization as robust method to account for dilution of complex biological mixtures. Application in ^{1}H NMR metabonomics. *Anal. Chem.* **78**, 4281–4290 (2006).

45. A. K. Rosen Vollmar, N. J. W. Rattray, Y. Cai, Á. J. Santos-Neto, N. C. Deziel, A. M. Z. Jukic, C. H. Johnson, Normalizing untargeted periconceptional urinary metabolomics data: A comparison of approaches. *Metabolites* **9**, E198 (2019).

46. N. Gray, R. Zia, A. King, V. C. Patel, J. Wendon, M. J. W. McPhail, M. Coen, R. S. Plumb, I. D. Wilson, J. K. Nicholson, High-speed quantitative UPLC-MS analysis of multiple amines in human plasma and serum via precolumn derivatization with 6-aminoquinolyl-N-hydroxysuccinimidyl carbamate: Application to acetaminophen-induced liver failure. *Anal. Chem.* **89**, 2478–2487 (2017).

47. J. M. Posma, I. Garcia-Perez, T. M. D. Ebbels, J. C. Lindon, J. Stamler, P. Elliott, E. Holmes, J. K. Nicholson, Optimized phenotypic biomarker discovery and confounder elimination via covariate-adjusted projection to latent structures from metabolic spectroscopy data. *J. Proteome Res.* **17**, 1586–1595 (2018).

ScienceAdvances

Metabolic maturation in the first 2 years of life in resource-constrained settings and its association with postnatal growth

N. GiallourouF. Fardus-ReidG. PanicK. VeselkovB. J. J. McCormickM. P. OlorteguiT. AhmedE. MdumaP. P. YoriM. MahfuzE. SvensenM. M. M. AhmedJ. M. ColstonM. N. KosekJ. R. Swann

View the article online
https://www.science.org/doi/10.1126/sciadv.aay5969
Permissions
https://www.science.org/help/reprints-and-permissions

Use of this article is subject to the Terms of service