Metabolomics

Metabolic derangements identified through untargeted metabolomics in a cross-sectional study of Nigerian children with severe acute malnutrition
--Manuscript Draft--

Manuscript Number:	MEBO-D-16-00240R1
Full Title:	Metabolic derangements identified through untargeted metabolomics in a cross-sectional study of Nigerian children with severe acute malnutrition
Article Type:	Original Article
Keywords:	severe acute malnutrition, metabolomics, Nigerian children, gut microbiota, LC-MS
Corresponding Author:	Stephen J. Allen, Ph.D Liverpool School of Tropical Medicine Liverpool, UNITED KINGDOM
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	Liverpool School of Tropical Medicine
Corresponding Author's Secondary Institution:	
First Author:	Amy McMillan, Ph.D.
First Author Secondary Information:	
Order of Authors:	Amy McMillan, Ph.D.
	Adebola E Orimadegun, MBBS, M.Sc.
	Mark W Sumarah, Ph.D
	Justin Renaud, Ph.D
	Magdalena Muc da Encarnacao, Ph.D
	Gregory B. Gloor, Ph.D
	Olusegun O Akinyinka, Ph.D
	Gregor Reid, Ph.D
	Stephen J. Allen, Ph.D
Order of Authors Secondary Information:	
Funding Information:	Yakult	Dr. Stephen J. Allen

| Abstract: | Introduction: Severe acute malnutrition (SAM) is a major cause of child mortality worldwide, however the pathogenesis of SAM remains poorly understood. Recent studies have uncovered an altered gut microbiota composition in children with SAM, suggesting a role for microbes in the pathogenesis of malnutrition.
Objectives: To elucidate the metabolic consequences of SAM and whether these changes are associated with changes in gut microbiota composition.
Methods: We applied an untargeted multi-platform metabolomics approach (gas chromatography-mass spectrometry (GC-MS) and liquid chromatography-mass spectrometry (LC-MS)) to stool and plasma samples from 47 Nigerian children with SAM and 11 control children. The composition of the stool microbiota was assessed by 16S rRNA gene sequencing.
Results: The plasma metabolome discriminated children with SAM from controls, while no significant differences were observed in the microbial or small molecule composition of stool. The abundance of 585 features in plasma were significantly altered in malnourished children (Wilcoxon test, FDR corrected P < 0.1), representing approximately 15% of the metabolome. Consistent with previous studies, children with SAM exhibited a marked reduction in amino acids/dipeptides and phospholipids, and an increase in acylcarnitines. We also identified numerous metabolic perturbations |

which have not been reported previously, including increased disaccharides, truncated fibrinopeptides, angiotensin I, dihydroxybutyrate, lactate, and heme, and decreased bioactive lipids belonging to the eicosanoid and docosanoid family.

Conclusion: Our findings provide a deeper understanding of the metabolic consequences of malnutrition. Further research is required to determine if specific metabolites may guide improved management, and/or act as novel biomarkers for assessing response to treatment.

Dear Royston Goodacre,
Editor-in-Chief
Metabolomics

We are very grateful for the valuable comments raised by the reviewers. Please see our responses below; we trust that these are appropriate. I have also attached a revised version of the manuscript with the changes highlighted.

Please note that, in response to comments by reviewer 3, we have moved details of methodology to the main manuscript. This has increased the length of the manuscript but we trust that this is acceptable.

Please do let me know if there are any remaining issues.

Sincerely, Stephen Allen

Reviewer #1: This study is about metabolic derangements identified through untargeted metabolomics in a cross-sectional study of Nigerian children with severe acute malnutrition. It is an interesting and relevant study because studies investigating the pathogenesis of SAM are lacking. To this respect the study presented in the manuscript has some qualities, however as the authors point it out it has several limitations. Only the plasma samples gave some results which were able to discriminate between the control but this discrimination is very weak as the total variance is around 12%. The lack of difference in the stool samples could be due to the fact that the authors did not take into account the demographic characteristics in their model, these can be confounding variables.

Response: We acknowledge in the discussion that there are many unaccounted for sources of variation in the plasma metabolome along with a number of other limitations. However, despite the small sample size and moderate percent variation explained, we were able to identify a number of novel metabolites significantly associated with SAM. We believe our findings have the potential to improve our understanding of malnutrition once validated in other cohorts, and are therefore worthy of publication.

We have explored residence as a potential confounder and could not find any plasma metabolites which differed significantly between rural, urban and peri-urban children (Kruskal-Wallis test, FDR corrected P>0.1) ,indicating residence does not significantly affect the metabolome in this population. These findings held true when SAM/Control samples were analyzed individually or together. We have added statements to the manuscript clarifying this.

Reviewer #1: On the "Study design and sample collection" part the sample collection is not clear. In which period the samples were collected? Were all the samples collected in the same period? Here the samples collection is not standardized as it should be. The period collection may have an impact also on the results of a metabolomic study.

Response: Dates of patient recruitment and sample collection are noted in Supplementary Table 1. All samples were collected between June 25th and Sept 21st 2012. A statement has been added to the manuscript clarifying this.

Reviewer #1: Instead of PCA, the PLS is more appropriate for the discrimination in metabolomic analyses because of its ability to model highly collinear spectral data with considerable amounts of both random and systematic noise.

Response: We prefer to use unsupervised methods when possible due to risks of over-fitting with supervised models. However, to further validate the separation between groups we have included supplementary PLS-DA models for both plasma and stool. The plasma model had moderate predictability and was significant (R2=0.98,Q2=0.69,CV-ANOVA p=1.43E-05, 5 PCs), while the stool model had low predictability and was not

significant (R2=0.46,Q2=0.08,CV-ANOVA p=0.085,1PC), further validating conclusions from unsupervised analyses. We have added a statement to the results section including these findings.

Reviewer #1: Line 43-51 p9 should be in the material and method.
Response: This section has been moved to the materials and methods

Reviewer #1: A limited number of volunteers were included in the control group.
Response: We acknowledge this is a limitation of the study. These were the maximum number of controls we were able to recruit during the study duration.

Reviewer #1: Line 36-43 p10 correct P > 0.1 to P < 0.1.
Response: These findings were not significant and therefore "P >0.1" is the appropriate notation.

Reviewer #1: Line 48 p17 to correct Bartz et al (Bartz et al. 2014) to Bartz et al. (2014).
Response: This has been corrected

Reviewer #1: Line 26 p18 to correct Zhang et al (Zhang et al. 2013) to Zhang et al (2013)
Response: This has been corrected

Reviewer #3: The paper describes a comparison of plasma and stool samples from children with severe acute malnutrition and a normal control group. LC-MS, GC-MS as well as rRNA gene sequencing and calprotectin and lactoferrin analyses were conducted. The paper is very well written; clear, concise and the interpretation is as such sensible. Unfortunately the paper has one severe issue and that is the design of the study itself. The control group is small (n=11) even if you would think that recruiting controls should not have been difficult. This made it impossible to control possible confounding factors such as urban/rural, ethnicity, age, formula fed or not etc. The reported demographics data unfortunately suggest that the SAM and control groups are not well matched. The Ethnicity but especially the Residence differences is a major concern. It would be easy to imagine that the difference in Residence could influence things such as diet (as is seem with the formula/no formula numbers) that would certainly cause a difference in the plasma metabolome. You must try to address this issue before it is possible to trust that the markers you find are markers of SAM and not something else. One suggestion could be to try to subset the SAM group to create matched groups. It is to the authors' credit that they discuss honestly and detailed the issues. The reservations one should have about making conclusions based on the reported study are thus clear, but should be investigated more in detail.
Response: We acknowledge the small number of controls is a major limitation of the study. Unfortunately, these were the maximum number of controls we could recruit during the study duration. In terms of the potential effects of residence on the metabolome, we found no evidence that this is a confounding variable. There were no plasma metabolites which differed significantly between rural, urban and peri-urban (Kruskal Wallace test, FDR corrected P>0.1) indicating residence does not significantly affect the metabolome in this population. These findings held true when SAM/Control samples were analyzed individually or together. There was also no affect of formula feeding on the metabolome. We have added statements to the manuscript presenting these findings.

Major issues

Reviewer #3: Fig 1A: This is a bad choice of plot to represent the data. The stacking of the bars make it impossible to see the difference between the taxa. I would suggest that a heatmap might be a better way to represent the data. Using clustering you might also be able to visualize sub-groups in the heatmap should such be present in addition to observing if the two groups are perhaps more similar intra-group than inter-group.

Response: Thank you for your suggestion. We have tried this approach, however due to the large number of rare taxa we find the heatmap to be less informative compared to barplots (ie most of the plot is blank space representing rare taxa). One could show only the most abundant taxa, however the clustering in this case would be misleading due to the compositional nature of 16S data. The PCoA plot was used to display clustering of samples as an alternative to heatmap clustering.

Reviewer #3: Fig 3: In the caption phospholipids are mentioned as significantly different but they are not marked as such in the text.
Response: The phospholipids findings are mentioned in pg 12, line 7-9 of the manuscript.

Reviewer #3: Pre-processing description should be in the main manuscript.
Response: These details were moved to the supplemental methods due to manuscript word limit recommendations. We have restored them to the main manuscript for the revised version assuming the editorial board deems this acceptable.

Reviewer #3: Pre-processing: Peakwidth settings missing. For grouping minsamp and bw parameters is missing. The bw parameter is one of the most critical.
Response: The "bw" parameter was included in the supplementary methods ("allowable retention time deviation was 5 seconds"). "minfrac" was used in place of "minsamp" to select the proportion of samples in which a feature must be found to be retained in the dataset. These details have been moved to the main manuscript for the revised version.

Reviewer #3: I suggest using CAMERA also. This will give you an estimation of the number of compounds instead of features.
Response: We have included the CAMERA output in the Table S5.

Reviewer #3: Using CAMERA would also help your identification. For example the compound you have identified as "Tetrahydro-2,5-furan-diacetic acid" looks like an isotope of your feature ID = 562 which is in turn the [M-NH3+H]+ fragment of your feature 640 which is Tryptophan, I am pretty sure. I looked as this manually in excel plotting the intensities against each other to see correlations but CAMERA would do that for you. So you might get fewer false positive identifications.
Response: Thank you for catching this error. Feature 162 has been re-annotated as Tryptophan.

Reviewer #3: For the dipeptides it would be better in figures and tables to be explicit that you don't know if you have Glu-Val or Val-Glu for example.
Response: We are able to determine the specific order of dipeptides due to differences in MS/MS spectras. As a result of amide bond cleavage during fragmentation, the C-terminal amino acid will remain intact in the MS/MS spectra while the N-terminal amino acid will appear as M-H2O, allowing one to determine the C and N-term amino acids in each peptide.

Reviewer #3: p=12, l=4: you mention Leucine yet your tables indicate that you cannot measure Leu independant of Ile.
Response: "Leucine" has been replaced with "leucine/isoleucine" in all instances in the manuscript.

Reviewer #3: The results says the age was from 6-59 months. Yet table 1 indicates 6-48. The real range should be in the results. Not the inclusion criteria.
Response: This has been corrected.

Reviewer #3: You state that GC-MS was able to discriminate the disaccharides that LC-MS found to be statistically significant. Yet they are not significant in GC-MS. Please comment on how this could be.
Response: The precise species of disaccharides differ between individuals, and therefore only when all disaccharides are analyzed together are they significant. As different disaccharides were resolved by our GC-MS method, they did not reach significance, however sucrose trended towards being increased in SAM (FDR p=0.2). Lower sensitivity with GC-MS compared to LC-MS may also be a factor as many of the disaccharides were close to the limit of detection when analyzed by GC-MS.

Minor issues:

Reviewer #3: p=7,l=32: "directly injected" to me indicates no LC. But an LC was used.
Response: This has been corrected.

Reviewer #3: p=7,l=27: u --> μ
Response: This has been corrected.

Reviewer #3: "The authors declare that they have not conflict of interests": not --> no?
Response: This has been corrected.

Reviewer #3: space before degree sign or not is not consistent.
Response: This has been corrected.

Reviewer #3: LC-MS analysis (gradient, voltage etc) should be in the main manuscript.
Response: This has been corrected.

Reviewer #3: Fig 2: For the loadings I would suggest that points instead of arrows might be easier to visualize.
Response: We believe arrows make the plots more intuitive as they indicate in which direction the features are pulling the samples on the corresponding scoreplot, and are visually distinct from the scoreplot which uses points.

Reviewer #3: In my opinion Figure 1 should be moved to the supplementary since all the results are negative.
Response: Although there were no differences between groups, the composition of the microbiota of this population remains under sampled (Reid et al, 2014, Microbiome 2:12), and therefore may be of interest to readers.

Reviewer #3: Fig 1B: What is the point of using 3D? To me you cannot see where the points are in the 3 dimensions and there is no effect anyway. I would assume PC1 vs PC2 would suffice to show the point.
Response: This is simply the way plots are constructed in Qiime, a widely used software for analyzing 16S data.

Reviewer #3: The Pearson reference is a little off: "Pearson a D" --> "Pearson AD"
Response: This has been corrected.

Reviewer #3: What is the purpose of the ROC curves? What is the reason to predict SAM from metabolites when a diagnosis is not difficult? Please state this in the manuscript. If you want to make a prediction model I suggest building a multivariate model.
Response: The purpose of the ROC curves was to identify metabolites which might be most useful as predictors of mortality/recovery in future studies. A statement has been added to the manuscript clarifying their purpose.

: Pre-processing: I am concerned about the ppm tolerance of 1 ppm which is very optimistic. Even for orbitrap; the tails of peaks need to be within that too. I suggestion checking for split peaks and running with ppm= 10 - 20 and check for differences. This is of course instrument dependent but in my hands ppm=1 was never a good choice.

Response: Our initial analyses had set the ppm higher as you suggest (5, 2.5 ppm), however we consistently received warnings indicating the ppm tolerance was too high and therefore settled upon 1 ppm. We did not observe any instances of peak splitting with these settings.

Reviewer #3: Pre-processing: "grouping included features present in at least 25% of samples". The settings refer to the percentage in each group. If you didn't group the data in XCMS the statement is correct but could be more explicit to say "grouping included features present in at least 25% of all the samples".

Response: Samples were not grouped into SAM/Control for xcms pre-processing. We have added statement to clarify this as you suggest.

Reviewer #3: p=11, l=26: The separation in Fig 2. is far from clear. There are far fewer controls than SAM samples so it is difficult to see if they really occupy a different space. Their spread does seem to be smaller which I guess is to be expected. I would be less absolute in the statements regarding the PCA analysis. A supervised analysis could be considered instead. PLS-DA or similar perhaps.

Response: We prefer to use unsupervised methods when possible due to risks of over-fitting with supervised models. However, to further validate the separation between groups we have included supplementary PLS-DA models for both plasma and stool. The plasma model had moderate predictability and was significant ($R2=0.98,Q2=0.69$,CV-ANOVA $p=1.43E-05$, 5 PCs), while the stool model had low predictability and was not significant ($R2=0.46,Q2=0.08$,CV-ANOVE $p=0.085$,1PC), further validating conclusions from unsupervised analyses. We have added a statement to the results section including these findings.

Reviewer #3: But you already have the univariate analysis so I assume the PCAs are just for illustrative purposes? If so it might be more informative to do a heatmap of the significant features

Response: Thank you for the suggestion. We have experimented with heatmaps as alternatives to PCA in the past and concluded the following: The purpose of the PCA in this case is to demonstrate the grouping of samples based on the entire detected metabolome, with the later stripcharts used to display the variation within specific compounds of interest. In comparison to these plots, we feel a heatmap is less appropriate as it does not display the entire dataset in an informative manor (ie control/SAM do not cluster well together as would be expected based on components 1 and 2 of the PCAs), nor does it allow easy interpretation of the spread of the data for individual metabolites.

Reviewer #3: "2C6 disaccharides" needs to be defined. I finally got the notation but I have never seen it before and it seems neither has google.

Response: We have changed the notation to "Dihexoses" to reflect more conventionally used nomenclature.

1

Metabolic derangements identified through untargeted metabolomics in a cross-sectional study of Nigerian children with severe acute malnutrition

ABSTRACT

*Introduction***:** Severe acute malnutrition (SAM) is a major cause of child mortality worldwide, however the pathogenesis of SAM remains poorly understood. Recent studies have uncovered an altered gut microbiota composition in children with SAM, suggesting a role for microbes in the pathogenesis of malnutrition.

Objectives: To elucidate the metabolic consequences of SAM and whether these changes are associated with changes in gut microbiota composition.

Methods: We applied an untargeted multi-platform metabolomics approach (gas chromatography-mass spectrometry (GC-MS) and liquid chromatography-mass spectrometry (LC-MS)) to stool and plasma samples from 47 Nigerian children with SAM and 11 control children. The composition of the stool microbiota was assessed by 16S rRNA gene sequencing.

Results: The plasma metabolome discriminated children with SAM from controls, while no significant differences were observed in the microbial or small molecule composition of stool. The abundance of 585 features in plasma were significantly altered in malnourished children (Wilcoxon test, FDR corrected $P < 0.1$), representing approximately 15% of the metabolome. Consistent with previous studies, children with SAM exhibited a marked reduction in amino acids/dipeptides and phospholipids, and an increase in acylcarnitines. We also identified numerous metabolic perturbations which have not been reported previously, including increased disaccharides, truncated fibrinopeptides, angiotensin I, dihydroxybutyrate, lactate, and heme, and decreased bioactive lipids belonging to the eicosanoid and docosanoid family.

Conclusion: Our findings provide a deeper understanding of the metabolic consequences of malnutrition. Further research is required to determine if specific metabolites may guide improved management, and/or act as novel biomarkers for assessing response to treatment.

1 Introduction

Severe acute malnutrition (SAM) is a leading cause of global child mortality, is associated with growth faltering, and results in impaired cognitive development (Kar et al. 2008). A significant proportion of deaths due to numerous illnesses including malaria, measles and diarrheal diseases have also been attributed to malnutrition (Caulfield et al. 2004). SAM is most frequently defined by a weight-for-length/height Z-score (WHZ) < -3 and/or presence of oedema (Who 2009). Middle upper arm circumference (MUAC) < 11.5 cm is also commonly used in children aged 6-– months (Who 2009). Despite its prevalence, the metabolic changes that occur during malnutrition are poorly understood. A better understanding of these metabolic derangements may shed light on pathophysiology and/or adaptation to malnutrition, and inform improved clinical management.

Multiple studies have demonstrated amino acid depletion in children with SAM (Ghisolfi et al. 1978; Bartz et al. 2014), but untargeted studies pertaining to other metabolites are limited. Bartz et al 2014 applied a targeted and untargeted approach to identify biomarkers associated with recovery and mortality in Ugandan children with SAM (Bartz et al. 2014). The concentrations of a number of metabolites were significantly altered upon rehabilitation, including amino acids, acylcarnitines and leptin. However, non-malnourished controls were not included, and the untargeted analysis was limited to gas chromatography-mass spectrometry (GC-MS) which covers only a small portion of the metabolome.

Malnutrition in juvenile pigs has also revealed alterations in the metabolome, including changes in amino acids, choline metabolites, and products of microbial-mammalian co-metabolism (Jiang et al. 2015), suggesting the gut microbiota may be altered by malnutrition.

Differences in gut microbiota composition of children with SAM compared to healthy co-twins and unrelated controls have been described in Malawian and Bangladeshi children (Smith et al. 2013; Subramanian et al. 2014). Additionally, specific taxa capable of causing enteropathy when combined with a low protein diet have been identified in mice (Brown et al. 2015). Enteropathy resulting in malabsorption and increased intestinal permeability is associated with SAM in humans, as indicated by intestinal biopsies and dual sugar permeability tests (Denno et al. 2014). This phenomenon may be a more severe form of environmental enteric dysfunction (EED), which occurs almost universally in people exposed to poor hygiene and sanitation. The etiology of enteropathy in SAM is not well understood, but may involve repeated exposure to fecal enteropathogens combined with a low protein diet (Hashimoto et al. 2012; Brown et al. 2015).

The aim of this study was to identify conserved changes in the metabolome of children with SAM compared to controls in order to improve our understanding of the pathophysiology of malnutrition. Given the association with enteropathy and recent studies linking changes in the microbiota to SAM (Smith et al. 2013; Subramanian et al. 2014), we also measured biomarkers of intestinal inflammation and profiled the gut microbiota of these children.

2 Materials and methods

Study design and sample collection

We aimed to undertake a pragmatic study of unselected children admitted to a health facility in a region with endemic malnutrition to maximise the relevance of our findings to clinical practice. We recruited 58 children aged 6-48 months: 47 with SAM and 11 well-nourished hospital controls. Demographic and clinical characteristics at recruitment are shown in Table 1. Children aged 6 – 48 months with SAM (WHZ <-3 or MUAC <11.5 cms and/or nutritional oedema) admitted to the Federal Medical Centre, Gusau, Zamfara State, Nigeria between June 25th and September 21st, 2012 were invited to participate in the study. The hospital serves both urban and rural populations and recruitment occurred during the rainy season when malnutrition is most common. Written information about the study and a verbal explanation in the appropriate local language was provided to parents/carers and signed or thumb-printed consent secured. Clinical management was performed according to usual practice based on WHO guidelines including those for the management of SAM (WHO 1999). Well-nourished children (MUAC >12.5 cms or WHZ score ≥ -1 and no nutritional oedema) were recruited from the paediatric ward or outpatient clinics as controls during the same time period (**Supplementary Table 1**). HIV positive individuals were excluded from the study. In both groups, baseline demographic and clinical information was collected on standard forms by one of five clinicians trained in the research methods (**Supplementary Table 1**). The first available stool sample was collected and its consistency and the presence of visible blood and mucus were recorded. At the time of clinical sampling, approximately 2.5 mL of venous blood was collected into EDTA tubes for the purposes of the study. Ethical approval was provided by the Joint Ethical Review Committees of the University of Ibadan / University College Hospital, Ibadan, Nigeria.

Metabolite extraction from stool

Approximately 250 mg of wet stool was lyophilized overnight. After drying, 40 mg was weighed into microcentrifuge tubes and extracted with 8:2 methanol:H_2O to a final concentration of 40 mg/mL. Samples were then vortexed for 30 sec, followed by centrifugation for 15 min at 10 000 rpm. 50 µL of this supernatant was dried in a speedvac for GC-MS analysis and the remaining stored at -80 °C for LC-MS.

Metabolite extraction from plasma

Metabolites were extracted from plasma according to the methods of Dunn et al, 2011 (Dunn et al. 2011). Briefly, plasma samples were thawed on ice for 30 min. Once thawed, 805 µL of 8:2 methanol:H_2O was added to 230 µL of plasma to make a 4.5 fold dilution. Samples were vortexed for 15 sec and centrifuged at 15 000 rpm for 15 min to pellet precipitated proteins. 370 µL of supernatant was then transferred to separate vials and dried down for GC-MS and LC-MS using a speedvac with no heat.

Untargeted GC-MS analyses

For stool, samples were derivatized with 40 µL of 2% methoxyamine-HCl in pyridine (MOX) incubated at 50 °C for 90 min, followed by 20 µL of N- Methyl-N-(trimethylsilyl) trifluoroacetamide (MSTFA) for 30 min at 50 °C. Samples were then transferred to micro-inserts before analysis by GC-MS. For plasma, 50 µL of MOX was added to dried samples and incubated for 90 min at 50 °C. 50 µL of MSTFA was then added and incubated for 30 min at 50 °C. Samples were then transferred to 1.5 ml microcentrifuge tubes and centrifuged at

15 000 rpm for 5 min to pellet debris. This supernatant was then transferred to HPLC vials

containing inserts for analysis.

One µL of sample was injected into an Agilent 7890A GC, 5975 inert MSD with triple

axis detector. Samples were injected using pulsed splitless mode using a 30 m DB5-MS

column with 10 m duraguard, diameter 0.35mm, thickness 0.25 µm (J&W Scientific, Folsom,

USA). Helium was used as the carrier gas at a constant flow rate of 1 mL/min. Oven

temperature was held at 70 °C for 5 min then increased at a rate of 5 °C/min to 300 °C and

held for 10 min. Solvent delay was set to 7 min, and total run time was 61 min. Masses

between 25 *m/z* and 600 *m/z* were selected by the detector. All samples were run in random

order and a single sample was run with every batch as a quality control to ensure machine

consistency.

Chromatogram files were deconvoluted and converted to ELU format using AMDIS

Mass Spectrometry software (Stein 1999) with the sensitivity set to low, resolution to medium,

and support threshold to high. Chromatograms were aligned using Spectconnect

(http://spectconnect.mit.edu) (Styczynski et al. 2007) with the support threshold set to low.

The integrated signal (IS) matrix output was used for all further analysis. Zeros were replaced

with two thirds the minimum detected value on a per metabolite basis (Timotej 2011),

followed by a log base 2 transformation. All further analyses were performed using these log-

transformed values.

Metabolites were initially identified by comparison to the NIST 11 standard reference

database (http://www.nist.gov/srd/nist1a.cfm). Identities of metabolites of interest were then

confirmed by authentic standards if available.

Untargeted LC-MS analyses

For stool, 2.5 µL of 1 µg/mL $^{13}C_6$ phenylalanine internal standard (Cambridge Isotopes, Tewksbury, USA) was added to 47.5 µL of extracted stool. For plasma, dried samples were reconstituted in 85.5 µL of ddH$_2$O. 4.5 µL of 1 ug/mL $^{13}C_6$ phenylalanine in ddH$_2$O was then - added to each vial as an internal standard. Samples were vortexed for 15 sec, then transferred to microinserts and injected into an Agilent 1290 Infinity HPLC coupled to a Q-Exactive Orbitrap mass spectrometer (Thermo Fisher Scientific, Waltham, USA) with a HESI (heated electrospray ionization) source. For HPLC, 2 µL of each sample was injected into a ZORBAX Eclipse plus C18 2.1 x 50mm x 1.8 micron column. Mobile phase (A) consisted of 0.1% formic acid in water and mobile phase (B) consisted of 0.1% formic acid in acetonitrile. The initial composition of 0% (B) was held constant for 30 s and increased to 100% over 3.0 min. For stool, mobile phase B was held at 100% for 1 minute and returned to 0% over 30s for a total run time of 5 min. For plasma, mobile phase B was held at 100% for 2 minutes and returned to 0% over 30s for a total run time of 6 min.

Full MS scanning between the ranges of *m/z* 50-750 was performed on all samples in both positive and negative mode at 140 000 resolution. The HESI source was operated under the following conditions: nitrogen flow of 30 and 8 arbitrary units for the sheath and auxiliary gas respectively, probe temperature and capillary temperature of 450 °C and 250 °C respectively and spray voltage of 3.9 kV and 3.5 kV in positive and negative mode respectively. The automatic gain control (AGC) target and maximum injection time were 1e6 and 500 ms respectively. For molecular characterization, every tenth sample was also analyzed with a data dependent MS/MS method where a 35 000 resolution full MS scan identified the top 12 signals above a 1e5 threshold which were subsequently selected at a 1.2 *m/z* isolation window for MS/MS. Normalized collision energy for MS/MS was 28, resolution

17 500, AGC target 1E5 and maximum injection time was 60ms. Blanks of pure methanol were run between every sample to limit carryover, and a single sample was run multiple times with every batch to account for any machine inconsistency. After data acquisition Thermo .RAW files were converted to .MZML format and centroided using ProteoWizard (Kessner et al. 2008). Files were then imported into R using the XCMS package (Patti et al. 2012) for chromatogram alignment and deconvolution. Features were detected with the "xcmsSet" function using the "centWave" method and a ppm tolerance of 1. Prefilter was set to 3-5000, noise 1E5, and signal to noise threshold was set to 5. Due to a lower overall noise and signal in negative mode, noise was set to 1E3 for this mode. Retention time correction was conducted using the "obiwarp" method, grouping included features present in at least 25% of all samples, allowable retention time deviation was 5 seconds, and *m/z* width set to 0.015. Areas of features below the signal to noise threshold in the data were integrated using the "fillPeaks" function with default settings. Any remaining zeros in the data were then replaced with two-thirds the minimum value on a per mass basis (Timotej 2011) before log base 2 transformation. The log-transformed mass list was then exported as a single .txt file and used for all further analyses. All further analyses were carried out in R unless otherwise specified. Positive and negative mode data were treated as two independent datasets for all analyses.

Metabolites were putatively identified based on accurate mass and LC-MS/MS fragmentation patterns (**Supplemental Figure 1**). Predictions were made mainly by *de novo* compound identification from in depth investigation of individual MS/MS spectra, utilizing the METLIN database (http://metlin.scripps.edu) (Smith et al. 2005), Human Metabolome Database (www.hmdb.ca) (Wishart et al. 2007), and CFM-ID (http://cfmid.wishartlab.com)

(Allen et al. 2014) whenever possible to aid in identification. Metabolites of interest were confirmed by authentic standards when available based on accurate mass, retention time and MS/MS spectra (**Supplemental Table 2**).

Statistical analysis of metabolome data

Principal Component Analysis (PCA) was conducted in R using the "FactoMineR" package with pareto scaling. Components 1 and 2, representing approximately 15 % and 9% of the variation respectively, could not be explained by any of the metadata collected (**Supplemental Figure 2**). We therefore selected the lowest components which maximized separation between SAM and controls (components 3 and 4) for qualitative analysis of the metabolome. Partial Least Squares Discriminant Analysis (PLS-DA) models were constructed in SIMCA 13.0.3.0 (Umetrics) using pareto scaling and 7-fold cross-validation. The number of components was selected using the autofit function which selects the maximum number of significant components according to cross-validation rules. Significant models were defined as those with cross-validated analysis of variance (CV-ANOVA) P values less than 0.05. Metabolites that differed between SAM and controls were determined independently of PCA analysis using unpaired Wilcoxon tests with Benjamini-Hochberg (False Discovery Rate (FDR)) corrections to account for multiple hypothesis testing (Benjamini and Hochberg 1995). Metabolites with a corrected $P < 0.1$ and average fold change > 2 in either group (SAM or control) were selected for further investigation. Fold changes were calculated using the geometric mean to limit inflation of fold change values due to outliers. Boxplots and stripcharts were constructed in R using the ggplot2 package. Odds ratios of metabolites to identify SAM from controls were calculated from conditional logistic regressions performed on

all metabolites using the glm function in R with 10 000 iterations and a binomial distribution. Receiver operating characteristic (ROC) curves and forest plots were built in R using the pROC and forestplot packages respectively. In the case of Des ADS-FPA, the ^{13}C isotope was used for statistical analyses and plotting due to a large singly-charged interfering peak with mass similar to the ^{12}C Des ADS-FPA. All raw data files were manually inspected to ensure the data was accurately represented by the ^{13}C peak. The effect of potential confounding variables was investigated using Spearman's correlations for continuous variables and Wilcoxon or Kruskal Wallace tests for discrete variable. Tests with FDR corrected P-values less than 0.1 were considered significant.

Calprotectin and lactoferrin stool ELISAs

Calprotectin was measured using the IDK® Calprotectin ELISA Kit along with the IDK Extract® stool extraction kit (Immundiagnostik, Bensheim, Germany). Lactoferrin was measured using the Lactoferrin Scan™ ELISA kit (Techlab®, Blacksburg, USA). All protocols were followed as per the manufacturer's instructions.

Microbiome profiling

DNA was extracted from stool samples using the PowerSoil-htp 96 Well Soil DNA isolation kit from MoBio (Carlsbad, CA) according to the manufacturer's protocol, with modifications as outlined by the Earth Microbiome Project (version 4_13). Approximately 250 mg of fecal sample was used for the extractions. Samples were sequenced by amplifying the V4 hypervariable region of the 16S rRNA gene. Sample amplification for sequencing was carried out using the forward primer

(ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNN(8)GTGCCAGCMGCCGCGGTAA)

and the reverse primer

(CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCTNNNN(8)GGACTACHVGGGTWT

CTAAT) where nnnn indicates four randomly incorporated nucleotides, and (8) was a sample-

specific nucleotide barcode where the barcodes differed by an edit distance of at least 4. The

5' end is the adapter sequence for the Illumina MiSeq sequencer and the sequences

following the barcode are complementary to the V4 rRNA gene region. Amplification was

carried out in 42 µL with each primer present at 3.2 pmol/µL, 20 µL GoTaq hot start colorless

master mix (Promega) and 2 µL extracted DNA. The PCR protocol was as follows: initial

activation step at 95 °C for 2 minutes and 25 cycles of 1 minute 95 °C, 1 minute 50 °C and 1

minute 72 °C.

All subsequent work was carried out at the London Regional Genomics Centre (LRGC,

lrgc.ca, London, Ontario, Canada). Briefly, PCR products were quantified with a Qubit 2.0

Flourometer and the high sensitivity dsDNA specific fluorescent probes (Life Technologies).

Samples were mixed at equimolar concentrations and purified with the QIAquick PCR

Purification kit (QIAGEN). Samples were paired-end sequenced on an Illumina Mi-Seq with

the 600 cycle version 3 reagents with 2x220 cycles.

Resulting reads were extracted and de-multiplexed using modifications of in-house

Perl and UNIX-shell scripts with operational taxonomic units (OTUs) clustered at 97%

identity, similar to our reported protocol (Gloor et al. 2010). Automated taxonomic

assignments were carried out by comparison to the SILVA database (http://www.arb-

silva.de/). **Supplemental Table 1** displays the nucleotide barcodes and their corresponding

samples. To control for background contaminating sequences, a no-template control was

also sequenced. Barplots were constructed with R (r-project.org) using proportional values. Rare OTUs found at less than 1% abundance in all samples were grouped in the remainder for barplots only. All other analyses were conducted with all OTUs. Principal Coordinate Analysis (PCoA) plots were constructed in Qiime (Caporaso et al. 2010) (qiime.org) using weighted UniFrac distances.

To avoid inappropriate statistical inferences made from compositional data, centred log-ratios (clr), a method previously described by Aitchison (Aitchison 1982), and adapted to microbiome data was used with unpaired Wilcoxon tests for comparisons of OTU level data (Fernandes et al. 2013; Fernandes et al. 2014). The Benjamini Hochberg (FDR) method was used to control for multiple testing with a significance threshold of 0.1. All statistical analysis, unless otherwise indicated, was carried out using R (r-project.org).

3 Results

3.1 Study population

The demographics of the study population is displayed in Table 1. Amongst the 58 children recruited, 47 were diagnosed with SAM (WHZ < -3 and/or MUAC < 11.5 and/or oedema), of which 26 (55.3%) were diagnosed with kwashiorkor. Apathy was present in 34 (72.3%), anorexia in 42 (89.4%), thin hair in 42 (89.4%), glossitis in 27 (57.4%), one or more Bitot spots in 9 (19.1%), dermatitis in 22 (46.8%), stomatitis in 27 (57.4%) and 18 (38.3%) had oral ulceration. Abdominal distension was present in 11 (23.4%), hepatomegaly in 28 (59.6%) and splenomegaly in 26 (55.3%). Loose or watery stools (Bristol Stool Form Scale 5-7) occurred in 36 (76.6%) of the SAM cases compared with 4 (36.4%) controls (Table 1).

None of the children had bloody stools. The distribution of ethnicity and area of residence differed significantly between cases and controls.

3.2 The fecal microbiota of Nigerian children

To determine if the composition of the gut microbiota differed between Nigerian children with SAM and controls, we profiled the microbiota of stool samples by amplifying the V4 region of the 16S rRNA gene. Two samples did not have enough material for DNA extraction, leaving 45 SAM and 11 controls remaining. The fecal microbiota was dominated by *Enterobacteraciae, Bifidobacterium, Enterococcus, Pediococcus, Lactobacillus* and *Streptococcus* species (Fig. 1A). There were no differences in alpha or beta diversity between SAM and controls (Fig. 1B,C), nor were there any individual taxa that differed significantly between groups (Wilcoxon test, FDR corrected $P > 0.1$, Supplementary Table 3). Although there was a trend for lower bacterial diversity in children with SAM compared to controls, the difference was not significant, even when age was taken into account (Fig. 1C, ANCOVA $P > 0.05$).

3.3 Fecal metabolome and inflammatory markers do not distinguish SAM from controls

To obtain a global view of nutrient deficiency in SAM and gain insight into pathophysiology, we performed a comprehensive untargeted analysis of metabolites in stool from all 58 children using two different methods; gas chromatography-mass spectrometry (GC-MS) and liquid chromatography-mass spectrometry (LC-MS). Surprisingly, both methods yielded no metabolites significantly affected by SAM (Wilcoxon test, FDR corrected $P > 0.1$, Supplementary Fig. 3, Supplementary Table 4). There was no separation of groups based on

qualitative PCA analysis, and PLS-DA models had low predictability and were not significant (Positive ESI model R2=0.46, Q2=0.08, CV-ANOVA p=0.085,1 component) (Supplementary Fig. 4). There were also no significant differences in the inflammatory proteins calprotectin or lactoferrin, as measured by ELISA (Supplementary Fig. 5).

3.4 Children with SAM have a distinct plasma metabolome

In contrast to stool results, the abundance of approximately 15% of LC-MS features (defined as a unique *m/z* and retention time) detected in plasma were significantly altered by SAM (Wilcoxon test, FDR corrected *P* < 0.1, Supplementary Table 5). Moderate separation between groups was observed along the 3rd and 4th component of principal component analysis (PCA) plots built from LC-MS data (Fig. 2; Supplementary Fig. 2 shows principal components 1 and 2), demonstrating the plasma metabolome of children with SAM is distinct from non-malnourished children. To further validate these findings, supervised PLS-DA models were constructed from positive ESI data (Supplementary Fig. 6). The PLS-DA model showed moderate predictability and was significant (R2=0.98, Q2=0.69, CV-ANOVA p=1.43E-05, 5 components), indicating that the plasma metabolome of children with SAM was unique from healthy controls. The vast majority of differential features were detected exclusively by LC-MS, with only three metabolites differing significantly by GC-MS (valine, leucine/isoleucine and aspartic acid). There were no metabolites that differed between children with non-oedematous malnutrition (marasmus, N = 21) or oedematous malnutrition (kwashiorkor, N = 26), and no significant effect of sex, age, formula feeding, or residence was observed (data not shown).

3.5 Oxylipins, phospholipids and amino acids are depleted in the plasma of children with SAM

Significant features in LC-MS data with greater than 2-fold change between SAM and controls were selected for further identification. As expected, free amino acids and dipeptides were lower in children with SAM including glutamine, arginine, tyrosine, leucine/isoleucine, valine, and the tryptophan metabolite kynurenine (Fig. 3A). A number of ether-linked single chain phospholipids belonging to the phosphatidylcholine (PC) and phosphatidylethanolamine (PE) families were also significantly decreased (Fig. 3B). Additionally, there was a marked reduction in a number of oxylipins belonging to the eicosanoid and docosanoid family. In depth investigation of fragmentation patterns of these lipids revealed that each feature (represented as a single boxplot pair in Fig. 3C) contained at least two different oxylipin species, differing only in the location of the hydroxyl group(s). The precise species contributing to the differences between SAM and controls could therefore not be determined. Arachidonic acid (AA), a precursor to the eicosanoids was also significantly lower in children with SAM. While median abundance of docosahexaenoic acid (DHA, the precursor to the docosanoids), was lower in children with SAM, the difference was not significant (FDR corrected $P = 0.2$) (Fig. 3C).

3.6 Metabolites elevated in plasma of children with SAM

Unlike metabolites negatively associated with SAM, the classes of metabolites elevated in malnourished children were more diverse, and included sugars, peptides, lipids, short chain fatty acids and porphyrins among others. Of interest, we detected a dihexose (a disaccharide composed of two six carbon hexoses) by LC-MS in both positive and negative

ionization modes that was significantly elevated in children with SAM (Wilcoxon test, FDR corrected $P < 0.1$, Fig. 4). Authentic standards of different dihexoses confirmed that our LC-MS measurements represented the total dihexoses in plasma. Analysis of standards by GC-MS, which is able to resolve dihexoses, and LC-MS/MS determined that sucrose was the most abundant dihexose in plasma, with smaller amounts of lactose and maltose also present.

Two forms of truncated fibrinopeptide A (FPA) missing 2, or 3 N-terminal amino acids were also detected in significantly greater amounts in children with SAM. Specifically, des-AD (lacking N-terminal alanine and aspartate) and des-ADS FPA (lacking N-terminal alanine, aspartate and serine) were both increased approximately 6-fold (Fig. 4). Intact FPA was outside the mass range of our initial analysis, and therefore we conducted a second analysis with a larger mass range to include the intact FPA peptide. Surprisingly, the abundance of intact FPA was not significantly different between SAM and controls (Supplementary Fig. 7), indicating that the truncated forms are not simply a degradation product of increased intact FPA, but have some other unknown origin.

Other metabolites of interest elevated in SAM included several acylcarnitines, the peptide hormone angiotensin I (Ang I), heme, lactate, oleoyl ethanolamide, 2,4 and 3,4-dihyroxybutyrate, an uncharacterized sphingoid base, a hydroxyvitamin D3 derivative and several other unknown compounds (Fig. 4, Supplementary Table 5).

3.7 Assessment of metabolites to discriminate SAM from controls

To measure the strength of the association between metabolites and SAM, we calculated odds ratios (OR) based on conditional logistic regressions of all metabolites

elevated in children with SAM (Supplementary Table 6). Dihexoses and lactate were among

the metabolites with the highest OR, ranging from 2.25-2.20 respectively (Fig. 5A). ROC

curves identified decanoylcarnitine, dihexoses, an uncharacterized sphingoid base,

angiotensin I, and heme as the metabolites that maximized the sensitivity and specificity for

SAM, as shown by the area under the curve (AUC), which ranged from 0.83- 0.81 for these

compounds (Fig. 5B, Supplementary Table 6).

4 Discussion

We report the first LC-MS-based untargeted metabolomic study of stool and plasma

from children with SAM. In contrast to the stool microbiota and metabolome, which did not

discriminate SAM from controls, approximately 15% of the plasma metabolome, equating to

585 features, were significantly altered in malnutrition. The explanation for the lack of

differences in stool are likely multifaceted, but may include the large effect of diet and inter-

individual variation which cannot be accounted for with small cross-sectional studies. It is also

worth noting that feces represent the net result of nutrient consumption, digestion and

absorption. Malabsorption has been widely reported in SAM (Kvissberg et al. 2016), and

therefore it is possible that these children both consume and absorb fewer nutrients

compared with non-malnourished controls, resulting in little net difference in stool nutrient

composition.

The stool microbiota data mirrored the metabolome results as neither the composition

nor diversity were significantly altered by malnutrition, even when children were matched for

age. This is in contrast to previous studies that have observed differences in the relative

abundance of specific taxa and diversity for age in twin pairs and unrelated controls (Smith et

al. 2013; Subramanian et al. 2014). The relatively small number of controls in our study limited our ability to model the relationship between microbial diversity and age, and therefore it is possible that a larger sample size may have revealed significant differences. Our cohort also tended to be older than the children included in the study by Subramanian et al (Subramanian et al. 2014), which could also partially explain the lack of differences in the microbiota. As the microbiota has been shown to stabilize by around two years of age (Subramanian et al. 2014), and a large proportion (41%) of the children in our study were two years or older, it is possible that the window where microbiota stunting is most apparent was not captured by our study. However, none of the plasma metabolites affected by SAM were of clear microbial origin, providing further evidence that the microbiota may not be playing a major role in this particular cohort.

Among the plasma metabolites discriminating SAM from controls, the total dihexoses had one of the strongest positive associations, as determined by ROC and OR analyses, and were detected in both positive and negative ionization modes. Regardless of their structure and content in the diet, disaccharides are not readily absorbed, and therefore must pass through the intercellular space of the intestinal mucosa to reach systemic circulation. For this reason, sucrose, cellobiose, lactulose and other disaccharides have been used as indicators of intestinal permeability (Pearson et al. 1982; Sutherland et al. 1994). Malabsorption and increased intestinal permeability are associated with SAM and EED, as measured by dual sugar permeability tests such as the lactulose/mannitol test (Denno et al. 2014; Kvissberg et al. 2016). Although these tests are fairly reliable, the requirement for fasting prior to administration raises ethical concerns for children with SAM, and urine must be collected

over several hours for accurate results. A non-administered permeability test, for which the total dihexoses are an attractive candidate, would therefore be highly valuable.

Although elevated dihexoses are suggestive of enteropathy in the SAM group, we did not observe any difference in the inflammatory markers calprotectin or lactoferrin in stool. These proteins are used as indicators of inflammation in other pathologies such as inflammatory bowel disease (Smith and Gaya 2012), but have not been evaluated as biomarkers of enteropathy in SAM. The weak correlation between lactoferrin and calprotectin (Spearman's R = 0.33, P =0.01) limits our ability to make conclusions as to the intestinal inflammatory status of these children. Future studies evaluating the association of non-administered disaccharides with enteropathy are warranted to determine if they might also be useful in identifying EED.

The majority of metabolites discriminating SAM from controls in plasma were consistent with nutritional depletion due to malnutrition. As expected, a number of amino acids/dipeptides were reduced, indicative of protein deficiency. Oxylipins were also significantly decreased, and to the best of our knowledge, have not been examined in children with SAM previously. Importantly, each oxylipin feature contained at least two different species, and therefore the number and diversity of oxylipins altered by SAM may be vastly under-represented by our study. Oxylipins are bioactive lipids formed by oxidation of long chain polyunsaturated fatty acids (LCPUFA), with the most well studied being the AA-derived eicosanoids (Gabbs et al. 2015). These bioactive lipids perform a wide array of functions, including tissue repair, blood clotting, and regulation of the immune system. Oxylipins within the same family can have similar or opposing effects (i.e. pro- or anti-inflammatory), and therefore we cannot determine the precise biological consequence of

reduced oxylipins during SAM (Gabbs et al. 2015). However, as children with SAM exhibit an impaired immune response to a variety of pathogens (Schaible and Kaufmann 2007), a lack of oxylipin mediators may be a contributing factor and warrants further investigation.

Previous studies have confirmed that children with SAM are deficient in LCPUFA (Holman et al. 1981; Leichsenring et al. 1995), including the oxylipin precursors AA and DHA. A large proportion of LCPUFA are stored as acyl-linked phospholipids, which could not detectable by our method, and therefore a comparison of the total LCPUFA in our cohort was not possible. Nevertheless, the significant reduction in free AA, non-significant trend for lower free DHA, and significant decreases in multiple phospholipids and oxylipins are suggestive of LCPUFA deficiency in our cohort.

Despite previous evidence of LCPUFA deficiency in SAM, the levels of LCPUFA in ready-to-use therapeutic food (RUTF) are low (Jones et al. 2015). Recently, a RUTF formulation supplemented with fish oil has been developed, which translated into increased LCPUFA in recipient children (Jones et al. 2015). More studies are required to determine if the incorporation of fish oil into refeeding programs restores oxylipin levels as well and whether this leads to any clinical benefit.

Apart from dietary deficiencies, other factors may also influence LCPUFA and oxylipin levels. Children with SAM have decreased desaturase activity, resulting in decreased synthesis of AA and DHA from their precursors (Koletzko et al. 1986; Smit et al. 2004). Additionally, a large proportion of dietary fatty acids are lost to beta-oxidation, a process that is induced during starvation (Nagao et al. 1993; Cunnane and Anderson 1997). Interestingly, a number of even-chain acylcarnitines, the bi-products of beta oxidation, were significantly elevated in children with SAM in our study. This suggests beta-oxidation of lipids may

contribute in part to LCPUFA deficiency in SAM. Bartz et al. 2014 found that even-chain

acylcarnitines, including the C10, C8 and C2 species identified in our study, decreased

significantly upon nutritional intervention in children with SAM from Uganda. Acylcarnitines

may therefore be useful not only as biomarkers of a malnourished state, but also as

measures of treatment efficacy.

The truncated FPAs were also metabolites of interest as they are unique peptides of

unknown origin and were elevated approximately 6-fold in SAM. The des-ADS form in

particular has not been reported previously. FPA is a 16 amino acid peptide produced upon

cleavage of the fibrinogen alpha chain by thrombin during the coagulation cascade. Thus it is

an indicator of thrombosis, and is elevated in plasma during a number of inflammatory

conditions including Crohn's disease, gastric cancer, and coronary thrombosis (Eisenberg et

al. 1985; Edwards et al. 1987; Ebert et al. 2006). As the intact peptide was not elevated in

SAM, these fragments are not simply a result of increased inflammation and thrombosis, but

have another unknown origin. One hypothesis is that they may arise from increased protease

activity as proposed by Zhang et al. 2013 who reported elevated des-A FPA in gastric cancer

with lymph node metastasis. Protease activity is increased during starvation to supply the

body with additional amino acids (Medina et al. 1991), and therefore it is possible that these

truncated forms are a consequence of this up-regulation.

Lactate was also elevated in the plasma of malnourished children, and was one the

metabolites with the highest OR for SAM. Lactate is an endproduct of cellular respiration

under anaerobic conditions, and is elevated in blood due to a number of etiologies, including

anemia, sepsis, trauma, and malignancy (Reddy et al. 2015). In the context of malnutrition,

elevated lactate may result from insufficient oxygen supply due to the anemia or increased

infections associated with SAM. Interestingly, several studies have shown blood lactate to be a good predictor of mortality due to a variety of illnesses (Krishna et al.; Shapiro et al. 2005; Mikkelsen et al. 2009), including a study of Tanzanian children with any febrile illness (Mtove et al. 2011). Further research should evaluate lactate as a predictor of mortality in SAM.

Finally, we observed a significant increase in free iron (III) heme in the plasma of malnourished children. Heme is normally bound to hemeproteins, but is released under conditions of oxidative stress and as a result of hemolysis. Free heme causes tissue damage, systemic inflammation (Balla et al. 1991; Wagener et al. 2001), and exacerbates sepsis and malaria in animal models (Seixas et al. 2009; Larsen et al. 2010). In humans, malaria severity correlates with the levels of free heme, and sepsis-related mortality is associated with decreased levels of the heme binding protein hemopexin (Larsen et al. 2010; Dalko et al. 2015). These diseases frequently affect children with SAM in the developing world, and malaria in particular is significantly associated with malnutrition (Berkley et al. 2009; Khogali et al. 2011). Our findings indicate elevated heme during SAM may contribute to increased severity of these infections, and therefore treatments targeting free heme may be beneficial.

Our study has several limitations. Firstly, reliable information regarding other infections and pathologies was not available due to the limited local diagnostic facilities, so their possible impact on the microbiota and metabolome remains unknown. Children with malnutrition were also more likely to originate from rural settings, and therefore we cannot distinguish the role of environment from that of malnutrition with absolute certainty. However, we could not identify any metabolites that differed significantly between children from urban, peri-urban or rural settings, indicating residence does significantly impact the plasma metabolome in this cohort. Also, although we identified clear differences in the plasma

metabolome of children according to nutritional status, this was not the only source of variation. We speculate that other unaccounted for sources of variation may include time of sample collection as it relates to circadian rhythms (Dallmann et al. 2012) and feeding time, diet and other environmental exposures, and/or the proximity of the plasma sampling to the plasma-red blood cell interface. The samples in this study were taken at random time points, with no requirements for fasting or feeding prior to collection. This may explain why we did not identify some of the classical signs of fasting, such as ketones elevated in the SAM group. However, our pragmatic collection method emphasized metabolites robust to external factors such as time of day and feeding. This strengthens the applicability of our findings to a clinical setting, where controlling for such variables may not be feasible.

In conclusion, we have demonstrated the plasma metabolome discriminates children with SAM from controls and identified a number of previously unreported metabolic derangements in malnutrition, providing new insight into disease mechanisms and management. Future studies are needed to monitor these metabolites during intervention to identify those most correlated with mortality and/or recovery. Validation of such biomarkers may enable better identification of children at highest risk of poor outcomes, and could guide improved management for a leading cause of childhood mortality worldwide.

Compliance with ethical standards

Conflict of Interest

The authors declare that they have no conflict of interests in relation to the work described.

Ethical approval

This study was approved by the Joint Ethical Review Committees of the University of Ibadan / University College Hospital, Ibadan, Nigeria, and was conducted in accordance with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent

Human stool and peripheral blood samples were collected after informed consent from all parents of children enrolled in the study.

References

Aitchison J (1982) The Statistical Analysis of Compositional Data. J R Stat Soc Ser B Methodol 44:139–177. doi: 10.2307/2345821

Allen F, Pon A, Wilson M, et al (2014) CFM-ID: A web server for annotation, spectrum prediction and metabolite identification from tandem mass spectra. Nucleic Acids Res 42:94–99. doi: 10.1093/nar/gku436

Balla G, Vercellotti GM, Muller-Eberhard U, et al (1991) Exposure of endothelial cells to free heme potentiates damage mediated by granulocytes and toxic oxygen species. Lab Invest 64:648–55.

Bartz S, Mody A, Hornik C, et al (2014) Severe acute malnutrition in childhood: hormonal and metabolic status at presentation, response to treatment, and predictors of mortality. J Clin Endocrinol Metab 99:2128–37. doi: 10.1210/jc.2013-4018

Benjamini Y, Hochberg Y (1995) Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J. R. Stat. Soc. Ser. B 57:289 – 300.

Berkley JA, Bejon P, Mwangi T, et al (2009) HIV infection, malnutrition, and invasive bacterial infection among children with severe malaria. Clin Infect Dis 49:336–343. doi: 10.1086/600299

Brown EM, Wlodarska M, Willing BP, et al (2015) Diet and specific microbial exposure trigger features of environmental enteropathy in a novel murine model. Nat Commun 6:7806. doi: 10.1038/ncomms8806

Caporaso JG, Kuczynski J, Stombaugh J, et al (2010) QIIME allows analysis of high-throughput community sequencing data. Nat Methods 7:335–6. doi: 10.1038/nmeth.f.303

Caulfield LE, de Onis M, Blössner M, Black RE (2004) Undernutrition as an underlying cause of child deaths associated with diarrhea, pneumonia, malaria, and measles. Am J Clin Nutr 80:193–8.

Cunnane SC, Anderson MJ (1997) The majority of dietary linoleate in growing rats is beta-oxidized or stored in visceral fat. J Nutr 127:146–52.

Dalko E, Das B, Herbert F, et al (2015) Multifaceted Role of Heme during Severe Plasmodium falciparum Infections in India. Infect Immun 83:3793–9. doi:

10.1128/IAI.00531-15

Dallmann R, Viola AU, Tarokh L, et al (2012) The human circadian metabolome. Proc Natl Acad Sci U S A 109:2625–9. doi: 10.1073/pnas.1114410109

Denno DM, VanBuskirk K, Nelson ZC, et al (2014) Use of the lactulose to mannitol ratio to evaluate childhood environmental enteric dysfunction: a systematic review. Clin Infect Dis 59 Suppl 4:S213–9. doi: 10.1093/cid/ciu541

Dunn WB, Broadhurst D, Begley P, et al (2011) Procedures for large-scale metabolic profiling of serum and plasma using gas chromatography and liquid chromatography coupled to mass spectrometry. Nat Protoc 6:1060–1083. doi: 10.1038/nprot.2011.335

Ebert MP a, Niemeyer D, Deininger SO, et al (2006) Identification and confirmation of increased fibrinopeptide a serum protein levels in gastric cancer sera by magnet bead assisted MALDI-TOF mass spectrometry. J Proteome Res 5:2152–2158. doi: 10.1021/pr060011c

Edwards RL, Levine JB, Green R, et al (1987) Activation of blood coagulation in Crohn's disease. Increased plasma fibrinopeptide A levels and enhanced generation of monocyte tissue factor activity. Gastroenterology 92:329–37.

Eisenberg PR, Sherman L a., Schectman K, et al (1985) Fibrinopeptide A: a marker of acute coronary thrombosis. Circulation 71:912–918. doi: 10.1161/01.CIR.71.5.912

Fernandes AD, Macklaim JM, Linn TG, et al (2013) ANOVA-like differential expression (ALDEx) analysis for mixed population RNA-Seq. PLoS One 8:e67019. doi: 10.1371/journal.pone.0067019

Fernandes AD, Reid JN, Macklaim JM, et al (2014) Unifying the analysis of high-throughput sequencing datasets: characterizing RNA-seq, 16S rRNA gene sequencing and selective growth experiments by compositional data analysis. Microbiome 2:15. doi: 10.1186/2049-2618-2-15

Gabbs M, Leng S, Devassy JG, et al (2015) Advances in Our Understanding of Oxylipins Derived from Dietary PUFAs. Adv Nutr 6:513–40. doi: 10.3945/an.114.007732

Ghisolfi J, Charlet P, Ser N, et al (1978) Plasma free amino acids in normal children and in patients with proteinocaloric malnutrition: fasting and infection. Pediatr Res

12:912–7.

Gloor GB, Hummelen R, Macklaim JM, et al (2010) Microbiome profiling by illumina sequencing of combinatorial sequence-tagged PCR products. PLoS One 5:e15406. doi: 10.1371/journal.pone.0015406

Hashimoto T, Perlot T, Rehman A, et al (2012) ACE2 links amino acid malnutrition to microbial ecology and intestinal inflammation. Nature 487:477–483. doi: 10.1038/nature11228

Holman RT, Johnson SB, Mercuri O, et al (1981) Essential fatty acid deficiency in malnourished children. Am J Clin Nutr 34:1534–9.

Jiang P, Stanstrup J, Thymann T, et al (2015) Progressive Changes in the Plasma Metabolome during Malnutrition in Juvenile Pigs. J Proteome Res. doi: 10.1021/acs.jproteome.5b00782

Jones KD, Ali R, Khasira M a, et al (2015) Ready-to-use therapeutic food with elevated n-3 polyunsaturated fatty acid content, with or without fish oil, to treat severe acute malnutrition: a randomized controlled trial. BMC Med 13:93. doi: 10.1186/s12916-015-0315-6

Kar BR, Rao SL, Chandramouli BA (2008) Cognitive development in children with chronic protein energy malnutrition. Behav Brain Funct 4:31.

Kessner D, Chambers M, Burke R, et al (2008) ProteoWizard: open source software for rapid proteomics tools development. Bioinformatics 24:2534–6. doi: 10.1093/bioinformatics/btn323

Khogali M, Zachariah R, Keiluhu A, et al (2011) Detection of malaria in relation to fever and grade of malnutrition among malnourished children in Ethiopia. Public Heal Action 1:16–18. doi: 10.5588/pha.11.0008

Koletzko B, Abiodun PO, Laryea MD, Bremer HJ (1986) Fatty acid composition of plasma lipids in Nigerian children with protein-energy malnutrition. Eur J Pediatr 145:109–15.

Krishna S, Waller DW, ter Kuile F, et al Lactic acidosis and hypoglycaemia in children with severe malaria: pathophysiological and prognostic significance. Trans R Soc Trop Med Hyg 88:67–73.

Kvissberg MA, Dalvi PS, Kerac M, et al (2016) Carbohydrate malabsorption in acutely

malnourished children and infants: a systematic review. Nutr Rev 74:48–58. doi: 10.1093/nutrit/nuv058

Larsen R, Gozzelino R, Jeney V, et al (2010) A central role for free heme in the pathogenesis of severe sepsis. Sci Transl Med 2:51ra71. doi: 10.1126/scitranslmed.3001118

Leichsenring M, Sütterlin N, Less S, et al (1995) Polyunsaturated fatty acids in erythrocyte and plasma lipids of children with severe protein-energy malnutrition. Acta Paediatr 84:516–20.

Medina R, Wing SS, Haas A, Goldberg AL (1991) Activation of the ubiquitin-ATP-dependent proteolytic system in skeletal muscle during fasting and denervation atrophy. Biomed Biochim Acta 50:347–56.

Mikkelsen ME, Miltiades AN, Gaieski DF, et al (2009) Serum lactate is associated with mortality in severe sepsis independent of organ failure and shock. Crit Care Med 37:1670–7. doi: 10.1097/CCM.0b013e31819fcf68

Mtove G, Nadjm B, Hendriksen ICE, et al (2011) Point-of-care measurement of blood lactate in children admitted with febrile illness to an African District Hospital. Clin Infect Dis 53:548–54. doi: 10.1093/cid/cir471

Nagao M, Parimoo B, Tanaka K (1993) Developmental, nutritional, and hormonal regulation of tissue-specific expression of the genes encoding various acyl-CoA dehydrogenases and alpha-subunit of electron transfer flavoprotein in rat. J Biol Chem 268:24114–24.

Patti GJ, Tautenhahn R, Siuzdak G (2012) Meta-analysis of untargeted metabolomic data from multiple profiling experiments. Nat Protoc 7:508–16. doi: 10.1038/nprot.2011.454

Pearson AD, Eastham EJ, Laker MF, et al (1982) Intestinal permeability in children with Crohn's disease and coeliac disease. Br Med J (Clin Res Ed) 285:20–21. doi: 10.1136/bmj.285.6334.20

Reddy AJ, Lam SW, Bauer SR, Guzman JA (2015) Lactic acidosis: Clinical implications and management strategies. Cleve Clin J Med 82:615–24. doi: 10.3949/ccjm.82a.14098

Schaible UE, Kaufmann SHE (2007) Malnutrition and Infection: Complex Mechanisms

and Global Impacts. PLoS Med 4:e115. doi: 10.1371/journal.pmed.0040115

Seixas E, Gozzelino R, Chora A, et al (2009) Heme oxygenase-1 affords protection against noncerebral forms of severe malaria. Proc Natl Acad Sci U S A 106:15837–42. doi: 10.1073/pnas.0903419106

Shapiro NI, Howell MD, Talmor D, et al (2005) Serum lactate as a predictor of mortality in emergency department patients with infection. Ann Emerg Med 45:524–8. doi: 10.1016/j.annemergmed.2004.12.006

Smit EN, Muskiet FAJ, Boersma ER (2004) The possible role of essential fatty acids in the pathophysiology of malnutrition: A review. Prostaglandins Leukot. Essent. Fat. Acids 71:241–250.

Smith CA, O'Maille G, Want EJ, et al (2005) METLIN: a metabolite mass spectral database. Ther Drug Monit 27:747–751. doi: 10.1097/01.ftd.0000179845.53213.39

Smith L a., Gaya DR (2012) Utility of faecal calprotectin analysis in adult inflammatory bowel disease. World J Gastroenterol 18:6782–6789. doi: 10.3748/wjg.v18.i46.6782

Smith MI, Yatsunenko T, Manary MJ, et al (2013) Gut microbiomes of Malawian twin pairs discordant for kwashiorkor. Science 339:548–54. doi: 10.1126/science.1229000

Stein SE (1999) An integrated method for spectrum extraction and compound identification from gas chromatography/mass spectrometry data. J Am Soc Mass Spectrom 10:770–781. doi: 10.1016/S1044-0305(99)00047-1

Styczynski MP, Moxley JF, Tong L V, et al (2007) Systematic identification of conserved metabolites in GC/MS data for metabolomics and biomarker discovery. Anal Chem 79:966–73. doi: 10.1021/ac0614846

Subramanian S, Huq S, Yatsunenko T, et al (2014) Persistent gut microbiota immaturity in malnourished Bangladeshi children. Nature 509:417–21. doi: 10.1038/nature13421

Sutherland LR, Verhoef M, Wallace JL, et al (1994) A simple, non-invasive marker of gastric damage: Sucrose permeability. Lancet 343:998–1000. doi: 10.1016/S0140-6736(94)90125-2

Timotej V (2011) A comparison of parameters below the limit of detection in

geochemical analyses by substitution methods. RMZ Mater Geoenvironment 58:393–404.

Wagener FA, Eggert A, Boerman OC, et al (2001) Heme is a potent inducer of inflammation in mice and is counteracted by heme oxygenase. Blood 98:1802–1811. doi: 10.1182/blood.V98.6.1802

Who (2009) WHO child growth standards and the identification of severe acute malnutrition in infants and children. WHO Libr 1–12.

Wishart DS, Tzur D, Knox C, et al (2007) HMDB: the Human Metabolome Database. Nucleic Acids Res 35:D521–6. doi: 10.1093/nar/gkl923

Zhang MH, Xu XH, Wang Y, et al (2013) A Prognostic Biomarker for Gastric Cancer With Lymph Node Metastases. Anat Rec Integr Anat Evol Biol 296:590–594. doi: Doi 10.1002/Ar.22642

Figure legends

Fig. 1: The stool microbiota does not discriminate SAM from controls. **(A)**. Stool microbiota profiled using the V4 region of the 16S rRNA gene. Each bar in the barplot represents a single sample from a single child and each color a different bacterial taxa (OTU). OTUs present at less than one percent in every sample were placed in the remainder displayed in black. Samples are ordered by their WHZ displayed in the heatmap above. **(B)**. Principal Coordinate Analysis (PCoA) plot of microbiota profiles built from weighted UniFrac distances. Each point represents a single sample from a single child. Positions of points display dissimilarities in the microbiota, with points further from one another being more dissimilar. (**C**). Shannon's diversity of the stool microbiota in children with SAM compared with controls. The boxes represent the 25th and 75th quartiles, and the line displays the median value within each group. Points extending beyond the lines are outliers defined as values greater or less than 1.5 times the interquartile range.

Fig. 2. Principal component analysis (PCA) plots of plasma metabolome from children with SAM and controls as determined by LC-MS using positive (top) or negative (bottom) electrospray ionization (ESI). Only components 3 and 4 are shown. The plots on the left display individual samples (scores). Each point represents a single sample from a single child. Positions of points display dissimilarities in the metabolome, with points furthest from one another being most dissimilar. Plots on the right display individual LC-MS features (loadings).

Each ray represents a single LC-MS feature, with those significantly elevated in SAM or Controls highlighted in black or red respectively (Wilcoxon test, FDR corrected P < 0.1).

Fig. 3. Free amino acids/dipeptides **(A)**, phospholipids **(B),** and oxylipins **(C)** significantly decreased in children with SAM compared to controls (> 2 fold change, Wilcoxon test, FDR corrected P < 0.1, with the exception of DHA which was not significant). Panels **(A)** and **(B)** were detected by positive ESI LC-MS and **(C)** by negative ESI LC-MS. Each point represents a single sample from a single child. The boxes represent the 25th and 75th quartiles, and the line displays the median value within each group. Points extending beyond the lines are outliers defined as values greater or less than 1.5 times the interquartile range. PC: Phosphatidylcholine, PE: Phosphatidylethanolamine. See **Supplementary Table 5** for phospholipid identities. (*) Metabolite ID confirmed by authentic standards.

Fig. 4. Metabolites significantly elevated at least 2-fold in the plasma of children with SAM compared to controls (Wilcoxon test, FDR corrected P < 0.1). (Left) Metabolites detected using positive ESI LC-MS. (Right) Metabolites detected using negative ESI LC-MS. Each point represents a single sample from a single child. The boxes represent the 25th and 75th quartiles, and the line displays the median value within each group. Points extending beyond the lines are outliers defined as values greater or less than 1.5 times the interquartile range. FPA:

Fibrinopeptide A, HTD: 3-Hydroxy-cis-5-tetradecenoyl. (*) Metabolite ID confirmed by authentic standards.

Fig. 5. Evaluation of biomarkers to identify SAM from controls. **(A)** Odds Ratios as determined by conditional logistic regressions of all validated metabolites positively associated with SAM. Bars represent 95% confidence intervals. **(B)** Receiver Operator Characteristics (ROC) curves. Metabolites with the highest area under the curve (AUC) are shown.

Metabolic derangements identified through untargeted metabolomics in a cross-sectional study of Nigerian children with severe acute malnutrition

ABSTRACT

*Introduction***:** Severe acute malnutrition (SAM) is a major cause of child mortality worldwide, however the pathogenesis of SAM remains poorly understood. Recent studies have uncovered an altered gut microbiota composition in children with SAM, suggesting a role for microbes in the pathogenesis of malnutrition.

Objectives: To elucidate the metabolic consequences of SAM and whether these changes are associated with changes in gut microbiota composition.

Methods: We applied an untargeted multi-platform metabolomics approach (gas chromatography-mass spectrometry (GC-MS) and liquid chromatography-mass spectrometry (LC-MS)) to stool and plasma samples from 47 Nigerian children with SAM and 11 control children. The composition of the stool microbiota was assessed by 16S rRNA gene sequencing.

Results: The plasma metabolome discriminated children with SAM from controls, while no significant differences were observed in the microbial or small molecule composition of stool. The abundance of 585 features in plasma were significantly altered in malnourished children (Wilcoxon test, FDR corrected $P < 0.1$), representing approximately 15% of the metabolome. Consistent with previous studies, children with SAM exhibited a marked reduction in amino acids/dipeptides and phospholipids, and an increase in acylcarnitines. We also identified numerous metabolic perturbations which have not been reported previously, including increased disaccharides, truncated fibrinopeptides, angiotensin I, dihydroxybutyrate, lactate, and heme, and decreased bioactive lipids belonging to the eicosanoid and docosanoid family.

Conclusion: Our findings provide a deeper understanding of the metabolic consequences of malnutrition. Further research is required to determine if specific metabolites may guide improved management, and/or act as novel biomarkers for assessing response to treatment.

1 Introduction

Severe acute malnutrition (SAM) is a leading cause of global child mortality, is associated with growth faltering, and results in impaired cognitive development (Kar et al. 2008). A significant proportion of deaths due to numerous illnesses including malaria, measles and diarrheal diseases have also been attributed to malnutrition (Caulfield et al. 2004). SAM is most frequently defined by a weight-for-length/height Z-score (WHZ) < -3 and/or presence of oedema (Who 2009). Middle upper arm circumference (MUAC) < 11.5 cm is also commonly used in children aged 6-– months (Who 2009). Despite its prevalence, the metabolic changes that occur during malnutrition are poorly understood. A better understanding of these metabolic derangements may shed light on pathophysiology and/or adaptation to malnutrition, and inform improved clinical management.

Multiple studies have demonstrated amino acid depletion in children with SAM (Ghisolfi et al. 1978; Bartz et al. 2014), but untargeted studies pertaining to other metabolites are limited. Bartz et al 2014 applied a targeted and untargeted approach to identify biomarkers associated with recovery and mortality in Ugandan children with SAM (Bartz et al. 2014). The concentrations of a number of metabolites were significantly altered upon rehabilitation, including amino acids, acylcarnitines and leptin. However, non-malnourished controls were not included, and the untargeted analysis was limited to gas chromatography-mass spectrometry (GC-MS) which covers only a small portion of the metabolome.

Malnutrition in juvenile pigs has also revealed alterations in the metabolome, including

changes in amino acids, choline metabolites, and products of microbial-mammalian co-

metabolism (Jiang et al. 2015), suggesting the gut microbiota may be altered by malnutrition.

Differences in gut microbiota composition of children with SAM compared to healthy

co-twins and unrelated controls have been described in Malawian and Bangladeshi children

(Smith et al. 2013; Subramanian et al. 2014). Additionally, specific taxa capable of causing

enteropathy when combined with a low protein diet have been identified in mice (Brown et al.

2015). Enteropathy resulting in malabsorption and increased intestinal permeability is

associated with SAM in humans, as indicated by intestinal biopsies and dual sugar

permeability tests (Denno et al. 2014). This phenomenon may be a more severe form of

environmental enteric dysfunction (EED), which occurs almost universally in people exposed

to poor hygiene and sanitation. The etiology of enteropathy in SAM is not well understood,

but may involve repeated exposure to fecal enteropathogens combined with a low protein diet

(Hashimoto et al. 2012; Brown et al. 2015).

The aim of this study was to identify conserved changes in the metabolome of children

with SAM compared to controls in order to improve our understanding of the pathophysiology

of malnutrition. Given the association with enteropathy and recent studies linking changes in

the microbiota to SAM (Smith et al. 2013; Subramanian et al. 2014), we also measured

biomarkers of intestinal inflammation and profiled the gut microbiota of these children.

2 Materials and methods

Study design and sample collection

We aimed to undertake a pragmatic study of unselected children admitted to a health facility in a region with endemic malnutrition to maximise the relevance of our findings to clinical practice. We recruited 58 children aged 6-48 months: 47 with SAM and 11 well-nourished hospital controls. Demographic and clinical characteristics at recruitment are shown in Table 1. Children aged 6 – 48 months with SAM (WHZ <-3 or MUAC <11.5 cms and/or nutritional oedema) admitted to the Federal Medical Centre, Gusau, Zamfara State, Nigeria between June 25th and September 21st, 2012 were invited to participate in the study. The hospital serves both urban and rural populations and recruitment occurred during the rainy season when malnutrition is most common. Written information about the study and a verbal explanation in the appropriate local language was provided to parents/carers and signed or thumb-printed consent secured. Clinical management was performed according to usual practice based on WHO guidelines including those for the management of SAM (WHO 1999). Well-nourished children (MUAC >12.5 cms or WHZ score ≥ -1 and no nutritional oedema) were recruited from the paediatric ward or outpatient clinics as controls during the same time period (**Supplementary Table 1**). HIV positive individuals were excluded from the study. In both groups, baseline demographic and clinical information was collected on standard forms by one of five clinicians trained in the research methods (**Supplementary Table 1**). The first available stool sample was collected and its consistency and the presence of visible blood and mucus were recorded. At the time of clinical sampling, approximately 2.5 mL of venous blood was collected into EDTA tubes for the purposes of the study. Ethical approval was provided by the Joint Ethical Review Committees of the University of Ibadan / University College Hospital, Ibadan, Nigeria.

Metabolite extraction from stool

Approximately 250 mg of wet stool was lyophilized overnight. After drying, 40 mg was weighed into microcentrifuge tubes and extracted with 8:2 methanol:H_2O to a final concentration of 40 mg/mL. Samples were then vortexed for 30 sec, followed by centrifugation for 15 min at 10 000 rpm. 50 μL of this supernatant was dried in a speedvac for GC-MS analysis and the remaining stored at -80 °C for LC-MS.

Metabolite extraction from plasma

Metabolites were extracted from plasma according to the methods of Dunn et al, 2011 (Dunn et al. 2011). Briefly, plasma samples were thawed on ice for 30 min. Once thawed, 805 μL of 8:2 methanol:H_2O was added to 230 μL of plasma to make a 4.5 fold dilution. Samples were vortexed for 15 sec and centrifuged at 15 000 rpm for 15 min to pellet precipitated proteins. 370 μL of supernatant was then transferred to separate vials and dried down for GC-MS and LC-MS using a speedvac with no heat.

Untargeted GC-MS analyses

For stool, samples were derivatized with 40 μL of 2% methoxyamine-HCl in pyridine (MOX) incubated at 50 °C for 90 min, followed by 20 μL of N- Methyl-N-(trimethylsilyl) trifluoroacetamide (MSTFA) for 30 min at 50 °C. Samples were then transferred to micro-inserts before analysis by GC-MS. For plasma, 50 μL of MOX was added to dried samples and incubated for 90 min at 50 °C. 50 μL of MSTFA was then added and incubated for 30 min at 50 °C. Samples were then transferred to 1.5 ml microcentrifuge tubes and centrifuged at

15 000 rpm for 5 min to pellet debris. This supernatant was then transferred to HPLC vials containing inserts for analysis.

One µL of sample was injected into an Agilent 7890A GC, 5975 inert MSD with triple axis detector. Samples were injected using pulsed splitless mode using a 30 m DB5-MS column with 10 m duraguard, diameter 0.35mm, thickness 0.25 µm (J&W Scientific, Folsom, USA). Helium was used as the carrier gas at a constant flow rate of 1 mL/min. Oven temperature was held at 70 °C for 5 min then increased at a rate of 5 °C/min to 300 °C and held for 10 min. Solvent delay was set to 7 min, and total run time was 61 min. Masses between 25 m/z and 600 m/z were selected by the detector. All samples were run in random order and a single sample was run with every batch as a quality control to ensure machine consistency.

Chromatogram files were deconvoluted and converted to ELU format using AMDIS Mass Spectrometry software (Stein 1999) with the sensitivity set to low, resolution to medium, and support threshold to high. Chromatograms were aligned using Spectconnect (http://spectconnect.mit.edu) (Styczynski et al. 2007) with the support threshold set to low. The integrated signal (IS) matrix output was used for all further analysis. Zeros were replaced with two thirds the minimum detected value on a per metabolite basis (Timotej 2011), followed by a log base 2 transformation. All further analyses were performed using these log-transformed values.

Metabolites were initially identified by comparison to the NIST 11 standard reference database (http://www.nist.gov/srd/nist1a.cfm). Identities of metabolites of interest were then confirmed by authentic standards if available.

Untargeted LC-MS analyses

For stool, 2.5 µL of 1 µg/mL $^{13}C_6$ phenylalanine internal standard (Cambridge Isotopes, Tewksbury, USA) was added to 47.5 µL of extracted stool. For plasma, dried samples were reconstituted in 85.5 µL of ddH$_2$O. 4.5 µL of 1 ug/mL $^{13}C_6$ phenylalanine in ddH$_2$O was then - added to each vial as an internal standard. Samples were vortexed for 15 sec, then transferred to microinserts and injected into an Agilent 1290 Infinity HPLC coupled to a Q-Exactive Orbitrap mass spectrometer (Thermo Fisher Scientific, Waltham, USA) with a HESI (heated electrospray ionization) source. For HPLC, 2 µL of each sample was injected into a ZORBAX Eclipse plus C18 2.1 x 50mm x 1.8 micron column. Mobile phase (A) consisted of 0.1% formic acid in water and mobile phase (B) consisted of 0.1% formic acid in acetonitrile. The initial composition of 0% (B) was held constant for 30 s and increased to 100% over 3.0 min. For stool, mobile phase B was held at 100% for 1 minute and returned to 0% over 30s for a total run time of 5 min. For plasma, mobile phase B was held at 100% for 2 minutes and returned to 0% over 30s for a total run time of 6 min.

Full MS scanning between the ranges of *m/z* 50-750 was performed on all samples in both positive and negative mode at 140 000 resolution. The HESI source was operated under the following conditions: nitrogen flow of 30 and 8 arbitrary units for the sheath and auxiliary gas respectively, probe temperature and capillary temperature of 450 °C and 250 °C respectively and spray voltage of 3.9 kV and 3.5 kV in positive and negative mode respectively. The automatic gain control (AGC) target and maximum injection time were 1e6 and 500 ms respectively. For molecular characterization, every tenth sample was also analyzed with a data dependent MS/MS method where a 35 000 resolution full MS scan identified the top 12 signals above a 1e5 threshold which were subsequently selected at a 1.2 *m/z* isolation window for MS/MS. Normalized collision energy for MS/MS was 28, resolution

17 500, AGC target 1E5 and maximum injection time was 60ms. Blanks of pure methanol were run between every sample to limit carryover, and a single sample was run multiple times with every batch to account for any machine inconsistency. After data acquisition Thermo .RAW files were converted to .MZML format and centroided using ProteoWizard (Kessner et al. 2008). Files were then imported into R using the XCMS package (Patti et al. 2012) for chromatogram alignment and deconvolution. Features were detected with the "xcmsSet" function using the "centWave" method and a ppm tolerance of 1. Prefilter was set to 3-5000, noise 1E5, and signal to noise threshold was set to 5. Due to a lower overall noise and signal in negative mode, noise was set to 1E3 for this mode. Retention time correction was conducted using the "obiwarp" method, grouping included features present in at least 25% of all samples, allowable retention time deviation was 5 seconds, and m/z width set to 0.015. Areas of features below the signal to noise threshold in the data were integrated using the "fillPeaks" function with default settings. Any remaining zeros in the data were then replaced with two-thirds the minimum value on a per mass basis (Timotej 2011) before log base 2 transformation. The log-transformed mass list was then exported as a single .txt file and used for all further analyses. All further analyses were carried out in R unless otherwise specified. Positive and negative mode data were treated as two independent datasets for all analyses.

Metabolites were putatively identified based on accurate mass and LC-MS/MS fragmentation patterns (**Supplemental Figure 1**). Predictions were made mainly by *de novo* compound identification from in depth investigation of individual MS/MS spectra, utilizing the METLIN database (http://metlin.scripps.edu) (Smith et al. 2005), Human Metabolome Database (www.hmdb.ca) (Wishart et al. 2007), and CFM-ID (http://cfmid.wishartlab.com)

(Allen et al. 2014) whenever possible to aid in identification. Metabolites of interest were confirmed by authentic standards when available based on accurate mass, retention time and MS/MS spectra (**Supplemental Table 2**).

Statistical analysis of metabolome data

Principal Component Analysis (PCA) was conducted in R using the "FactoMineR" package with pareto scaling. Components 1 and 2, representing approximately 15 % and 9% of the variation respectively, could not be explained by any of the metadata collected (**Supplemental Figure 2**). We therefore selected the lowest components which maximized separation between SAM and controls (components 3 and 4) for qualitative analysis of the metabolome. Partial Least Squares Discriminant Analysis (PLS-DA) models were constructed in SIMCA 13.0.3.0 (Umetrics) using pareto scaling and 7-fold cross-validation. The number of components was selected using the autofit function which selects the maximum number of significant components according to cross-validation rules. Significant models were defined as those with cross-validated analysis of variance (CV-ANOVA) P values less than 0.05. Metabolites that differed between SAM and controls were determined independently of PCA analysis using unpaired Wilcoxon tests with Benjamini-Hochberg (False Discovery Rate (FDR)) corrections to account for multiple hypothesis testing (Benjamini and Hochberg 1995). Metabolites with a corrected $P < 0.1$ and average fold change > 2 in either group (SAM or control) were selected for further investigation. Fold changes were calculated using the geometric mean to limit inflation of fold change values due to outliers. Boxplots and stripcharts were constructed in R using the ggplot2 package. Odds ratios of metabolites to identify SAM from controls were calculated from conditional logistic regressions performed on

all metabolites using the glm function in R with 10 000 iterations and a binomial distribution. Receiver operating characteristic (ROC) curves and forest plots were built in R using the pROC and forestplot packages respectively. In the case of Des ADS-FPA, the ^{13}C isotope was used for statistical analyses and plotting due to a large singly-charged interfering peak with mass similar to the ^{12}C Des ADS-FPA. All raw data files were manually inspected to ensure the data was accurately represented by the ^{13}C peak. The effect of potential confounding variables was investigated using Spearman's correlations for continuous variables and Wilcoxon or Kruskal Wallace tests for discrete variable. Tests with FDR corrected *P*-values less than 0.1 were considered significant.

Calprotectin and lactoferrin stool ELISAs

Calprotectin was measured using the IDK® Calprotectin ELISA Kit along with the IDK Extract® stool extraction kit (Immundiagnostik, Bensheim, Germany). Lactoferrin was measured using the Lactoferrin Scan™ ELISA kit (Techlab®, Blacksburg, USA). All protocols were followed as per the manufacturer's instructions.

Microbiome profiling

DNA was extracted from stool samples using the PowerSoil-htp 96 Well Soil DNA isolation kit from MoBio (Carlsbad, CA) according to the manufacturer's protocol, with modifications as outlined by the Earth Microbiome Project (version 4_13). Approximately 250 mg of fecal sample was used for the extractions. Samples were sequenced by amplifying the V4 hypervariable region of the 16S rRNA gene. Sample amplification for sequencing was carried out using the forward primer

(ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNN(8)GTGCCAGCMGCCGCGGTAA)

and the reverse primer

(CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCTNNNN(8)GGACTACHVGGGTWT
CTAAT) where nnnn indicates four randomly incorporated nucleotides, and (8) was a sample-specific nucleotide barcode where the barcodes differed by an edit distance of at least 4. The 5' end is the adapter sequence for the Illumina MiSeq sequencer and the sequences following the barcode are complementary to the V4 rRNA gene region. Amplification was carried out in 42 µL with each primer present at 3.2 pmol/µL, 20 µL GoTaq hot start colorless master mix (Promega) and 2 µL extracted DNA. The PCR protocol was as follows: initial activation step at 95 °C for 2 minutes and 25 cycles of 1 minute 95 °C, 1 minute 50 °C and 1 minute 72 °C.

All subsequent work was carried out at the London Regional Genomics Centre (LRGC, lrgc.ca, London, Ontario, Canada). Briefly, PCR products were quantified with a Qubit 2.0 Flourometer and the high sensitivity dsDNA specific fluorescent probes (Life Technologies). Samples were mixed at equimolar concentrations and purified with the QIAquick PCR Purification kit (QIAGEN). Samples were paired-end sequenced on an Illumina Mi-Seq with the 600 cycle version 3 reagents with 2x220 cycles.

Resulting reads were extracted and de-multiplexed using modifications of in-house Perl and UNIX-shell scripts with operational taxonomic units (OTUs) clustered at 97% identity, similar to our reported protocol (Gloor et al. 2010). Automated taxonomic assignments were carried out by comparison to the SILVA database (http://www.arb-silva.de/). **Supplemental Table 1** displays the nucleotide barcodes and their corresponding samples. To control for background contaminating sequences, a no-template control was

also sequenced. Barplots were constructed with R (r-project.org) using proportional values. Rare OTUs found at less than 1% abundance in all samples were grouped in the remainder for barplots only. All other analyses were conducted with all OTUs. Principal Coordinate Analysis (PCoA) plots were constructed in Qiime (Caporaso et al. 2010) (qiime.org) using weighted UniFrac distances.

To avoid inappropriate statistical inferences made from compositional data, centred log-ratios (clr), a method previously described by Aitchison (Aitchison 1982), and adapted to microbiome data was used with unpaired Wilcoxon tests for comparisons of OTU level data (Fernandes et al. 2013; Fernandes et al. 2014). The Benjamini Hochberg (FDR) method was used to control for multiple testing with a significance threshold of 0.1. All statistical analysis, unless otherwise indicated, was carried out using R (r-project.org).

3 Results

3.1 Study population

The demographics of the study population is displayed in Table 1. Amongst the 58 children recruited, 47 were diagnosed with SAM (WHZ < -3 and/or MUAC < 11.5 and/or oedema), of which 26 (55.3%) were diagnosed with kwashiorkor. Apathy was present in 34 (72.3%), anorexia in 42 (89.4%), thin hair in 42 (89.4%), glossitis in 27 (57.4%), one or more Bitot spots in 9 (19.1%), dermatitis in 22 (46.8%), stomatitis in 27 (57.4%) and 18 (38.3%) had oral ulceration. Abdominal distension was present in 11 (23.4%), hepatomegaly in 28 (59.6%) and splenomegaly in 26 (55.3%). Loose or watery stools (Bristol Stool Form Scale 5-7) occurred in 36 (76.6%) of the SAM cases compared with 4 (36.4%) controls (Table 1).

None of the children had bloody stools. The distribution of ethnicity and area of residence differed significantly between cases and controls.

3.2 The fecal microbiota of Nigerian children

To determine if the composition of the gut microbiota differed between Nigerian children with SAM and controls, we profiled the microbiota of stool samples by amplifying the V4 region of the 16S rRNA gene. Two samples did not have enough material for DNA extraction, leaving 45 SAM and 11 controls remaining. The fecal microbiota was dominated by *Enterobacteraciae, Bifidobacterium, Enterococcus, Pediococcus, Lactobacillus* and *Streptococcus* species (Fig. 1A). There were no differences in alpha or beta diversity between SAM and controls (Fig. 1B,C), nor were there any individual taxa that differed significantly between groups (Wilcoxon test, FDR corrected $P > 0.1$, Supplementary Table 3). Although there was a trend for lower bacterial diversity in children with SAM compared to controls, the difference was not significant, even when age was taken into account (Fig. 1C, ANCOVA $P > 0.05$).

3.3 Fecal metabolome and inflammatory markers do not distinguish SAM from controls

To obtain a global view of nutrient deficiency in SAM and gain insight into pathophysiology, we performed a comprehensive untargeted analysis of metabolites in stool from all 58 children using two different methods; gas chromatography-mass spectrometry (GC-MS) and liquid chromatography-mass spectrometry (LC-MS). Surprisingly, both methods yielded no metabolites significantly affected by SAM (Wilcoxon test, FDR corrected $P > 0.1$, Supplementary Fig. 3, Supplementary Table 4). There was no separation of groups based on

qualitative PCA analysis, and PLS-DA models had low predictability and were not significant (Positive ESI model R2=0.46, Q2=0.08, CV-ANOVA p=0.085,1 component) (Supplementary Fig. 4). There were also no significant differences in the inflammatory proteins calprotectin or lactoferrin, as measured by ELISA (Supplementary Fig. 5).

3.4 Children with SAM have a distinct plasma metabolome

In contrast to stool results, the abundance of approximately 15% of LC-MS features (defined as a unique m/z and retention time) detected in plasma were significantly altered by SAM (Wilcoxon test, FDR corrected $P < 0.1$, Supplementary Table 5). Moderate separation between groups was observed along the 3rd and 4th component of principal component analysis (PCA) plots built from LC-MS data (Fig. 2; Supplementary Fig. 2 shows principal components 1 and 2), demonstrating the plasma metabolome of children with SAM is distinct from non-malnourished children. To further validate these findings, supervised PLS-DA models were constructed from positive ESI data (Supplementary Fig. 6). The PLS-DA model showed moderate predictability and was significant (R2=0.98, Q2=0.69, CV-ANOVA p=1.43E-05, 5 components), indicating that the plasma metabolome of children with SAM was unique from healthy controls. The vast majority of differential features were detected exclusively by LC-MS, with only three metabolites differing significantly by GC-MS (valine, leucine/isoleucine and aspartic acid). There were no metabolites that differed between children with non-oedematous malnutrition (marasmus, N = 21) or oedematous malnutrition (kwashiorkor, N = 26), and no significant effect of sex, age, formula feeding, or residence was observed (data not shown).

3.5 Oxylipins, phospholipids and amino acids are depleted in the plasma of children with SAM

Significant features in LC-MS data with greater than 2-fold change between SAM and controls were selected for further identification. As expected, free amino acids and dipeptides were lower in children with SAM including glutamine, arginine, tyrosine, leucine/isoleucine, valine, and the tryptophan metabolite kynurenine (Fig. 3A). A number of ether-linked single chain phospholipids belonging to the phosphatidylcholine (PC) and phosphatidylethanolamine (PE) families were also significantly decreased (Fig. 3B). Additionally, there was a marked reduction in a number of oxylipins belonging to the eicosanoid and docosanoid family. In depth investigation of fragmentation patterns of these lipids revealed that each feature (represented as a single boxplot pair in Fig. 3C) contained at least two different oxylipin species, differing only in the location of the hydroxyl group(s). The precise species contributing to the differences between SAM and controls could therefore not be determined. Arachidonic acid (AA), a precursor to the eicosanoids was also significantly lower in children with SAM. While median abundance of docosahexaenoic acid (DHA, the precursor to the docosanoids), was lower in children with SAM, the difference was not significant (FDR corrected $P = 0.2$) (Fig. 3C).

3.6 Metabolites elevated in plasma of children with SAM

Unlike metabolites negatively associated with SAM, the classes of metabolites elevated in malnourished children were more diverse, and included sugars, peptides, lipids, short chain fatty acids and porphyrins among others. Of interest, we detected a dihexose (a disaccharide composed of two six carbon hexoses) by LC-MS in both positive and negative

ionization modes that was significantly elevated in children with SAM (Wilcoxon test, FDR corrected $P < 0.1$, Fig. 4). Authentic standards of different dihexoses confirmed that our LC-MS measurements represented the total dihexoses in plasma. Analysis of standards by GC-MS, which is able to resolve dihexoses, and LC-MS/MS determined that sucrose was the most abundant dihexose in plasma, with smaller amounts of lactose and maltose also present.

Two forms of truncated fibrinopeptide A (FPA) missing 2, or 3 N-terminal amino acids were also detected in significantly greater amounts in children with SAM. Specifically, des-AD (lacking N-terminal alanine and aspartate) and des-ADS FPA (lacking N-terminal alanine, aspartate and serine) were both increased approximately 6-fold (Fig. 4). Intact FPA was outside the mass range of our initial analysis, and therefore we conducted a second analysis with a larger mass range to include the intact FPA peptide. Surprisingly, the abundance of intact FPA was not significantly different between SAM and controls (Supplementary Fig. 7), indicating that the truncated forms are not simply a degradation product of increased intact FPA, but have some other unknown origin.

Other metabolites of interest elevated in SAM included several acylcarnitines, the peptide hormone angiotensin I (Ang I), heme, lactate, oleoyl ethanolamide, 2,4 and 3,4-dihyroxybutyrate, an uncharacterized sphingoid base, a hydroxyvitamin D3 derivative and several other unknown compounds (Fig. 4, Supplementary Table 5).

3.7 Assessment of metabolites to discriminate SAM from controls

To measure the strength of the association between metabolites and SAM, we calculated odds ratios (OR) based on conditional logistic regressions of all metabolites

elevated in children with SAM (Supplementary Table 6). Dihexoses and lactate were among

the metabolites with the highest OR, ranging from 2.25-2.20 respectively (Fig. 5A). ROC

curves identified decanoylcarnitine, dihexoses, an uncharacterized sphingoid base,

angiotensin I, and heme as the metabolites that maximized the sensitivity and specificity for

SAM, as shown by the area under the curve (AUC), which ranged from 0.83- 0.81 for these

compounds (Fig. 5B, Supplementary Table 6).

4 Discussion

We report the first LC-MS-based untargeted metabolomic study of stool and plasma

from children with SAM. In contrast to the stool microbiota and metabolome, which did not

discriminate SAM from controls, approximately 15% of the plasma metabolome, equating to

585 features, were significantly altered in malnutrition. The explanation for the lack of

differences in stool are likely multifaceted, but may include the large effect of diet and inter-

individual variation which cannot be accounted for with small cross-sectional studies. It is also

worth noting that feces represent the net result of nutrient consumption, digestion and

absorption. Malabsorption has been widely reported in SAM (Kvissberg et al. 2016), and

therefore it is possible that these children both consume and absorb fewer nutrients

compared with non-malnourished controls, resulting in little net difference in stool nutrient

composition.

The stool microbiota data mirrored the metabolome results as neither the composition

nor diversity were significantly altered by malnutrition, even when children were matched for

age. This is in contrast to previous studies that have observed differences in the relative

abundance of specific taxa and diversity for age in twin pairs and unrelated controls (Smith et

al. 2013; Subramanian et al. 2014). The relatively small number of controls in our study limited our ability to model the relationship between microbial diversity and age, and therefore it is possible that a larger sample size may have revealed significant differences. Our cohort also tended to be older than the children included in the study by Subramanian et al (Subramanian et al. 2014), which could also partially explain the lack of differences in the microbiota. As the microbiota has been shown to stabilize by around two years of age (Subramanian et al. 2014), and a large proportion (41%) of the children in our study were two years or older, it is possible that the window where microbiota stunting is most apparent was not captured by our study. However, none of the plasma metabolites affected by SAM were of clear microbial origin, providing further evidence that the microbiota may not be playing a major role in this particular cohort.

Among the plasma metabolites discriminating SAM from controls, the total dihexoses had one of the strongest positive associations, as determined by ROC and OR analyses, and were detected in both positive and negative ionization modes. Regardless of their structure and content in the diet, disaccharides are not readily absorbed, and therefore must pass through the intercellular space of the intestinal mucosa to reach systemic circulation. For this reason, sucrose, cellobiose, lactulose and other disaccharides have been used as indicators of intestinal permeability (Pearson et al. 1982; Sutherland et al. 1994). Malabsorption and increased intestinal permeability are associated with SAM and EED, as measured by dual sugar permeability tests such as the lactulose/mannitol test (Denno et al. 2014; Kvissberg et al. 2016). Although these tests are fairly reliable, the requirement for fasting prior to administration raises ethical concerns for children with SAM, and urine must be collected

over several hours for accurate results. A non-administered permeability test, for which the total dihexoses are an attractive candidate, would therefore be highly valuable.

Although elevated dihexoses are suggestive of enteropathy in the SAM group, we did not observe any difference in the inflammatory markers calprotectin or lactoferrin in stool. These proteins are used as indicators of inflammation in other pathologies such as inflammatory bowel disease (Smith and Gaya 2012), but have not been evaluated as biomarkers of enteropathy in SAM. The weak correlation between lactoferrin and calprotectin (Spearman's R = 0.33, P =0.01) limits our ability to make conclusions as to the intestinal inflammatory status of these children. Future studies evaluating the association of non-administered disaccharides with enteropathy are warranted to determine if they might also be useful in identifying EED.

The majority of metabolites discriminating SAM from controls in plasma were consistent with nutritional depletion due to malnutrition. As expected, a number of amino acids/dipeptides were reduced, indicative of protein deficiency. Oxylipins were also significantly decreased, and to the best of our knowledge, have not been examined in children with SAM previously. Importantly, each oxylipin feature contained at least two different species, and therefore the number and diversity of oxylipins altered by SAM may be vastly under-represented by our study. Oxylipins are bioactive lipids formed by oxidation of long chain polyunsaturated fatty acids (LCPUFA), with the most well studied being the AA-derived eicosanoids (Gabbs et al. 2015). These bioactive lipids perform a wide array of functions, including tissue repair, blood clotting, and regulation of the immune system. Oxylipins within the same family can have similar or opposing effects (i.e. pro- or anti-inflammatory), and therefore we cannot determine the precise biological consequence of

reduced oxylipins during SAM (Gabbs et al. 2015). However, as children with SAM exhibit an impaired immune response to a variety of pathogens (Schaible and Kaufmann 2007), a lack of oxylipin mediators may be a contributing factor and warrants further investigation.

Previous studies have confirmed that children with SAM are deficient in LCPUFA (Holman et al. 1981; Leichsenring et al. 1995), including the oxylipin precursors AA and DHA. A large proportion of LCPUFA are stored as acyl-linked phospholipids, which could not detectable by our method, and therefore a comparison of the total LCPUFA in our cohort was not possible. Nevertheless, the significant reduction in free AA, non-significant trend for lower free DHA, and significant decreases in multiple phospholipids and oxylipins are suggestive of LCPUFA deficiency in our cohort.

Despite previous evidence of LCPUFA deficiency in SAM, the levels of LCPUFA in ready-to-use therapeutic food (RUTF) are low (Jones et al. 2015). Recently, a RUTF formulation supplemented with fish oil has been developed, which translated into increased LCPUFA in recipient children (Jones et al. 2015). More studies are required to determine if the incorporation of fish oil into refeeding programs restores oxylipin levels as well and whether this leads to any clinical benefit.

Apart from dietary deficiencies, other factors may also influence LCPUFA and oxylipin levels. Children with SAM have decreased desaturase activity, resulting in decreased synthesis of AA and DHA from their precursors (Koletzko et al. 1986; Smit et al. 2004). Additionally, a large proportion of dietary fatty acids are lost to beta-oxidation, a process that is induced during starvation (Nagao et al. 1993; Cunnane and Anderson 1997). Interestingly, a number of even-chain acylcarnitines, the bi-products of beta oxidation, were significantly elevated in children with SAM in our study. This suggests beta-oxidation of lipids may

contribute in part to LCPUFA deficiency in SAM. Bartz et al. 2014 found that even-chain

acylcarnitines, including the C10, C8 and C2 species identified in our study, decreased

significantly upon nutritional intervention in children with SAM from Uganda. Acylcarnitines

may therefore be useful not only as biomarkers of a malnourished state, but also as

measures of treatment efficacy.

The truncated FPAs were also metabolites of interest as they are unique peptides of

unknown origin and were elevated approximately 6-fold in SAM. The des-ADS form in

particular has not been reported previously. FPA is a 16 amino acid peptide produced upon

cleavage of the fibrinogen alpha chain by thrombin during the coagulation cascade. Thus it is

an indicator of thrombosis, and is elevated in plasma during a number of inflammatory

conditions including Crohn's disease, gastric cancer, and coronary thrombosis (Eisenberg et

al. 1985; Edwards et al. 1987; Ebert et al. 2006). As the intact peptide was not elevated in

SAM, these fragments are not simply a result of increased inflammation and thrombosis, but

have another unknown origin. One hypothesis is that they may arise from increased protease

activity as proposed by Zhang et al. 2013 who reported elevated des-A FPA in gastric cancer

with lymph node metastasis. Protease activity is increased during starvation to supply the

body with additional amino acids (Medina et al. 1991), and therefore it is possible that these

truncated forms are a consequence of this up-regulation.

Lactate was also elevated in the plasma of malnourished children, and was one the

metabolites with the highest OR for SAM. Lactate is an endproduct of cellular respiration

under anaerobic conditions, and is elevated in blood due to a number of etiologies, including

anemia, sepsis, trauma, and malignancy (Reddy et al. 2015). In the context of malnutrition,

elevated lactate may result from insufficient oxygen supply due to the anemia or increased

infections associated with SAM. Interestingly, several studies have shown blood lactate to be a good predictor of mortality due to a variety of illnesses (Krishna et al.; Shapiro et al. 2005; Mikkelsen et al. 2009), including a study of Tanzanian children with any febrile illness (Mtove et al. 2011). Further research should evaluate lactate as a predictor of mortality in SAM.

Finally, we observed a significant increase in free iron (III) heme in the plasma of malnourished children. Heme is normally bound to hemeproteins, but is released under conditions of oxidative stress and as a result of hemolysis. Free heme causes tissue damage, systemic inflammation (Balla et al. 1991; Wagener et al. 2001), and exacerbates sepsis and malaria in animal models (Seixas et al. 2009; Larsen et al. 2010). In humans, malaria severity correlates with the levels of free heme, and sepsis-related mortality is associated with decreased levels of the heme binding protein hemopexin (Larsen et al. 2010; Dalko et al. 2015). These diseases frequently affect children with SAM in the developing world, and malaria in particular is significantly associated with malnutrition (Berkley et al. 2009; Khogali et al. 2011). Our findings indicate elevated heme during SAM may contribute to increased severity of these infections, and therefore treatments targeting free heme may be beneficial.

Our study has several limitations. Firstly, reliable information regarding other infections and pathologies was not available due to the limited local diagnostic facilities, so their possible impact on the microbiota and metabolome remains unknown. Children with malnutrition were also more likely to originate from rural settings, and therefore we cannot distinguish the role of environment from that of malnutrition with absolute certainty. However, we could not identify any metabolites that differed significantly between children from urban, peri-urban or rural settings, indicating residence does significantly impact the plasma metabolome in this cohort. Also, although we identified clear differences in the plasma

metabolome of children according to nutritional status, this was not the only source of variation. We speculate that other unaccounted for sources of variation may include time of sample collection as it relates to circadian rhythms (Dallmann et al. 2012) and feeding time, diet and other environmental exposures, and/or the proximity of the plasma sampling to the plasma-red blood cell interface. The samples in this study were taken at random time points, with no requirements for fasting or feeding prior to collection. This may explain why we did not identify some of the classical signs of fasting, such as ketones elevated in the SAM group. However, our pragmatic collection method emphasized metabolites robust to external factors such as time of day and feeding. This strengthens the applicability of our findings to a clinical setting, where controlling for such variables may not be feasible.

In conclusion, we have demonstrated the plasma metabolome discriminates children with SAM from controls and identified a number of previously unreported metabolic derangements in malnutrition, providing new insight into disease mechanisms and management. Future studies are needed to monitor these metabolites during intervention to identify those most correlated with mortality and/or recovery. Validation of such biomarkers may enable better identification of children at highest risk of poor outcomes, and could guide improved management for a leading cause of childhood mortality worldwide.

Aknowledgments

Compliance with ethical standards

Conflict of Interest

The authors declare that they have <mark>no</mark> conflict of interests in relation to the work described.

Ethical approval

This study was approved by the Joint Ethical Review Committees of the University of Ibadan / University College Hospital, Ibadan, Nigeria, and was conducted in accordance with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent

Human stool and peripheral blood samples were collected after informed consent from all parents of children enrolled in the study.

References

Aitchison J (1982) The Statistical Analysis of Compositional Data. J R Stat Soc Ser B
Methodol 44:139–177. doi: 10.2307/2345821

Allen F, Pon A, Wilson M, et al (2014) CFM-ID: A web server for annotation, spectrum
prediction and metabolite identification from tandem mass spectra. Nucleic Acids
Res 42:94–99. doi: 10.1093/nar/gku436

Balla G, Vercellotti GM, Muller-Eberhard U, et al (1991) Exposure of endothelial cells to
free heme potentiates damage mediated by granulocytes and toxic oxygen species.
Lab Invest 64:648–55.

Bartz S, Mody A, Hornik C, et al (2014) Severe acute malnutrition in childhood:
hormonal and metabolic status at presentation, response to treatment, and
predictors of mortality. J Clin Endocrinol Metab 99:2128–37. doi: 10.1210/jc.2013-
4018

Benjamini Y, Hochberg Y (1995) Controlling the False Discovery Rate: A Practical and
Powerful Approach to Multiple Testing. J. R. Stat. Soc. Ser. B 57:289 – 300.

Berkley JA, Bejon P, Mwangi T, et al (2009) HIV infection, malnutrition, and invasive
bacterial infection among children with severe malaria. Clin Infect Dis 49:336–343.
doi: 10.1086/600299

Brown EM, Wlodarska M, Willing BP, et al (2015) Diet and specific microbial exposure
trigger features of environmental enteropathy in a novel murine model. Nat
Commun 6:7806. doi: 10.1038/ncomms8806

Caporaso JG, Kuczynski J, Stombaugh J, et al (2010) QIIME allows analysis of high-
throughput community sequencing data. Nat Methods 7:335–6. doi:
10.1038/nmeth.f.303

Caulfield LE, de Onis M, Blössner M, Black RE (2004) Undernutrition as an underlying
cause of child deaths associated with diarrhea, pneumonia, malaria, and measles.
Am J Clin Nutr 80:193–8.

Cunnane SC, Anderson MJ (1997) The majority of dietary linoleate in growing rats is
beta-oxidized or stored in visceral fat. J Nutr 127:146–52.

Dalko E, Das B, Herbert F, et al (2015) Multifaceted Role of Heme during Severe
Plasmodium falciparum Infections in India. Infect Immun 83:3793–9. doi:

10.1128/IAI.00531-15

Dallmann R, Viola AU, Tarokh L, et al (2012) The human circadian metabolome. Proc Natl Acad Sci U S A 109:2625–9. doi: 10.1073/pnas.1114410109

Denno DM, VanBuskirk K, Nelson ZC, et al (2014) Use of the lactulose to mannitol ratio to evaluate childhood environmental enteric dysfunction: a systematic review. Clin Infect Dis 59 Suppl 4:S213–9. doi: 10.1093/cid/ciu541

Dunn WB, Broadhurst D, Begley P, et al (2011) Procedures for large-scale metabolic profiling of serum and plasma using gas chromatography and liquid chromatography coupled to mass spectrometry. Nat Protoc 6:1060–1083. doi: 10.1038/nprot.2011.335

Ebert MP a, Niemeyer D, Deininger SO, et al (2006) Identification and confirmation of increased fibrinopeptide a serum protein levels in gastric cancer sera by magnet bead assisted MALDI-TOF mass spectrometry. J Proteome Res 5:2152–2158. doi: 10.1021/pr060011c

Edwards RL, Levine JB, Green R, et al (1987) Activation of blood coagulation in Crohn's disease. Increased plasma fibrinopeptide A levels and enhanced generation of monocyte tissue factor activity. Gastroenterology 92:329–37.

Eisenberg PR, Sherman L a., Schectman K, et al (1985) Fibrinopeptide A: a marker of acute coronary thrombosis. Circulation 71:912–918. doi: 10.1161/01.CIR.71.5.912

Fernandes AD, Macklaim JM, Linn TG, et al (2013) ANOVA-like differential expression (ALDEx) analysis for mixed population RNA-Seq. PLoS One 8:e67019. doi: 10.1371/journal.pone.0067019

Fernandes AD, Reid JN, Macklaim JM, et al (2014) Unifying the analysis of high-throughput sequencing datasets: characterizing RNA-seq, 16S rRNA gene sequencing and selective growth experiments by compositional data analysis. Microbiome 2:15. doi: 10.1186/2049-2618-2-15

Gabbs M, Leng S, Devassy JG, et al (2015) Advances in Our Understanding of Oxylipins Derived from Dietary PUFAs. Adv Nutr 6:513–40. doi: 10.3945/an.114.007732

Ghisolfi J, Charlet P, Ser N, et al (1978) Plasma free amino acids in normal children and in patients with proteinocaloric malnutrition: fasting and infection. Pediatr Res

12:912–7.

Gloor GB, Hummelen R, Macklaim JM, et al (2010) Microbiome profiling by illumina sequencing of combinatorial sequence-tagged PCR products. PLoS One 5:e15406. doi: 10.1371/journal.pone.0015406

Hashimoto T, Perlot T, Rehman A, et al (2012) ACE2 links amino acid malnutrition to microbial ecology and intestinal inflammation. Nature 487:477–483. doi: 10.1038/nature11228

Holman RT, Johnson SB, Mercuri O, et al (1981) Essential fatty acid deficiency in malnourished children. Am J Clin Nutr 34:1534–9.

Jiang P, Stanstrup J, Thymann T, et al (2015) Progressive Changes in the Plasma Metabolome during Malnutrition in Juvenile Pigs. J Proteome Res. doi: 10.1021/acs.jproteome.5b00782

Jones KD, Ali R, Khasira M a, et al (2015) Ready-to-use therapeutic food with elevated n-3 polyunsaturated fatty acid content, with or without fish oil, to treat severe acute malnutrition: a randomized controlled trial. BMC Med 13:93. doi: 10.1186/s12916-015-0315-6

Kar BR, Rao SL, Chandramouli BA (2008) Cognitive development in children with chronic protein energy malnutrition. Behav Brain Funct 4:31.

Kessner D, Chambers M, Burke R, et al (2008) ProteoWizard: open source software for rapid proteomics tools development. Bioinformatics 24:2534–6. doi: 10.1093/bioinformatics/btn323

Khogali M, Zachariah R, Keiluhu A, et al (2011) Detection of malaria in relation to fever and grade of malnutrition among malnourished children in Ethiopia. Public Heal Action 1:16–18. doi: 10.5588/pha.11.0008

Koletzko B, Abiodun PO, Laryea MD, Bremer HJ (1986) Fatty acid composition of plasma lipids in Nigerian children with protein-energy malnutrition. Eur J Pediatr 145:109–15.

Krishna S, Waller DW, ter Kuile F, et al Lactic acidosis and hypoglycaemia in children with severe malaria: pathophysiological and prognostic significance. Trans R Soc Trop Med Hyg 88:67–73.

Kvissberg MA, Dalvi PS, Kerac M, et al (2016) Carbohydrate malabsorption in acutely

malnourished children and infants: a systematic review. Nutr Rev 74:48–58. doi: 10.1093/nutrit/nuv058

Larsen R, Gozzelino R, Jeney V, et al (2010) A central role for free heme in the pathogenesis of severe sepsis. Sci Transl Med 2:51ra71. doi: 10.1126/scitranslmed.3001118

Leichsenring M, Sütterlin N, Less S, et al (1995) Polyunsaturated fatty acids in erythrocyte and plasma lipids of children with severe protein-energy malnutrition. Acta Paediatr 84:516–20.

Medina R, Wing SS, Haas A, Goldberg AL (1991) Activation of the ubiquitin-ATP-dependent proteolytic system in skeletal muscle during fasting and denervation atrophy. Biomed Biochim Acta 50:347–56.

Mikkelsen ME, Miltiades AN, Gaieski DF, et al (2009) Serum lactate is associated with mortality in severe sepsis independent of organ failure and shock. Crit Care Med 37:1670–7. doi: 10.1097/CCM.0b013e31819fcf68

Mtove G, Nadjm B, Hendriksen ICE, et al (2011) Point-of-care measurement of blood lactate in children admitted with febrile illness to an African District Hospital. Clin Infect Dis 53:548–54. doi: 10.1093/cid/cir471

Nagao M, Parimoo B, Tanaka K (1993) Developmental, nutritional, and hormonal regulation of tissue-specific expression of the genes encoding various acyl-CoA dehydrogenases and alpha-subunit of electron transfer flavoprotein in rat. J Biol Chem 268:24114–24.

Patti GJ, Tautenhahn R, Siuzdak G (2012) Meta-analysis of untargeted metabolomic data from multiple profiling experiments. Nat Protoc 7:508–16. doi: 10.1038/nprot.2011.454

Pearson AD, Eastham EJ, Laker MF, et al (1982) Intestinal permeability in children with Crohn's disease and coeliac disease. Br Med J (Clin Res Ed) 285:20–21. doi: 10.1136/bmj.285.6334.20

Reddy AJ, Lam SW, Bauer SR, Guzman JA (2015) Lactic acidosis: Clinical implications and management strategies. Cleve Clin J Med 82:615–24. doi: 10.3949/ccjm.82a.14098

Schaible UE, Kaufmann SHE (2007) Malnutrition and Infection: Complex Mechanisms

and Global Impacts. PLoS Med 4:e115. doi: 10.1371/journal.pmed.0040115

Seixas E, Gozzelino R, Chora A, et al (2009) Heme oxygenase-1 affords protection against noncerebral forms of severe malaria. Proc Natl Acad Sci U S A 106:15837–42. doi: 10.1073/pnas.0903419106

Shapiro NI, Howell MD, Talmor D, et al (2005) Serum lactate as a predictor of mortality in emergency department patients with infection. Ann Emerg Med 45:524–8. doi: 10.1016/j.annemergmed.2004.12.006

Smit EN, Muskiet FAJ, Boersma ER (2004) The possible role of essential fatty acids in the pathophysiology of malnutrition: A review. Prostaglandins Leukot. Essent. Fat. Acids 71:241–250.

Smith CA, O'Maille G, Want EJ, et al (2005) METLIN: a metabolite mass spectral database. Ther Drug Monit 27:747–751. doi: 10.1097/01.ftd.0000179845.53213.39

Smith L a., Gaya DR (2012) Utility of faecal calprotectin analysis in adult inflammatory bowel disease. World J Gastroenterol 18:6782–6789. doi: 10.3748/wjg.v18.i46.6782

Smith MI, Yatsunenko T, Manary MJ, et al (2013) Gut microbiomes of Malawian twin pairs discordant for kwashiorkor. Science 339:548–54. doi: 10.1126/science.1229000

Stein SE (1999) An integrated method for spectrum extraction and compound identification from gas chromatography/mass spectrometry data. J Am Soc Mass Spectrom 10:770–781. doi: 10.1016/S1044-0305(99)00047-1

Styczynski MP, Moxley JF, Tong L V, et al (2007) Systematic identification of conserved metabolites in GC/MS data for metabolomics and biomarker discovery. Anal Chem 79:966–73. doi: 10.1021/ac0614846

Subramanian S, Huq S, Yatsunenko T, et al (2014) Persistent gut microbiota immaturity in malnourished Bangladeshi children. Nature 509:417–21. doi: 10.1038/nature13421

Sutherland LR, Verhoef M, Wallace JL, et al (1994) A simple, non-invasive marker of gastric damage: Sucrose permeability. Lancet 343:998–1000. doi: 10.1016/S0140-6736(94)90125-2

Timotej V (2011) A comparison of parameters below the limit of detection in

geochemical analyses by substitution methods. RMZ Mater Geoenvironment 58:393–404.

Wagener FA, Eggert A, Boerman OC, et al (2001) Heme is a potent inducer of inflammation in mice and is counteracted by heme oxygenase. Blood 98:1802–1811. doi: 10.1182/blood.V98.6.1802

Who (2009) WHO child growth standards and the identification of severe acute malnutrition in infants and children. WHO Libr 1–12.

Wishart DS, Tzur D, Knox C, et al (2007) HMDB: the Human Metabolome Database. Nucleic Acids Res 35:D521–6. doi: 10.1093/nar/gkl923

Zhang MH, Xu XH, Wang Y, et al (2013) A Prognostic Biomarker for Gastric Cancer With Lymph Node Metastases. Anat Rec Integr Anat Evol Biol 296:590–594. doi: Doi 10.1002/Ar.22642

Figure legends

Fig. 1: The stool microbiota does not discriminate SAM from controls. **(A)**. Stool microbiota profiled using the V4 region of the 16S rRNA gene. Each bar in the barplot represents a single sample from a single child and each color a different bacterial taxa (OTU). OTUs present at less than one percent in every sample were placed in the remainder displayed in black. Samples are ordered by their WHZ displayed in the heatmap above. **(B)**. Principal Coordinate Analysis (PCoA) plot of microbiota profiles built from weighted UniFrac distances. Each point represents a single sample from a single child. Positions of points display dissimilarities in the microbiota, with points further from one another being more dissimilar. (**C**). Shannon's diversity of the stool microbiota in children with SAM compared with controls. The boxes represent the 25th and 75th quartiles, and the line displays the median value within each group. Points extending beyond the lines are outliers defined as values greater or less than 1.5 times the interquartile range.

Fig. 2. Principal component analysis (PCA) plots of plasma metabolome from children with SAM and controls as determined by LC-MS using positive (top) or negative (bottom) electrospray ionization (ESI). Only components 3 and 4 are shown. The plots on the left display individual samples (scores). Each point represents a single sample from a single child. Positions of points display dissimilarities in the metabolome, with points furthest from one another being most dissimilar. Plots on the right display individual LC-MS features (loadings).

Each ray represents a single LC-MS feature, with those significantly elevated in SAM or Controls highlighted in black or red respectively (Wilcoxon test, FDR corrected P < 0.1).

Fig. 3. Free amino acids/dipeptides **(A)**, phospholipids **(B),** and oxylipins **(C)** significantly decreased in children with SAM compared to controls (> 2 fold change, Wilcoxon test, FDR corrected P < 0.1, with the exception of DHA which was not significant). Panels **(A)** and **(B)** were detected by positive ESI LC-MS and **(C)** by negative ESI LC-MS. Each point represents a single sample from a single child. The boxes represent the 25th and 75th quartiles, and the line displays the median value within each group. Points extending beyond the lines are outliers defined as values greater or less than 1.5 times the interquartile range. PC: Phosphatidylcholine, PE: Phosphatidylethanolamine. See **Supplementary Table 5** for phospholipid identities. (*) Metabolite ID confirmed by authentic standards.

Fig. 4. Metabolites significantly elevated at least 2-fold in the plasma of children with SAM compared to controls (Wilcoxon test, FDR corrected P < 0.1). (Left) Metabolites detected using positive ESI LC-MS. (Right) Metabolites detected using negative ESI LC-MS. Each point represents a single sample from a single child. The boxes represent the 25th and 75th quartiles, and the line displays the median value within each group. Points extending beyond the lines are outliers defined as values greater or less than 1.5 times the interquartile range. FPA:

Fibrinopeptide A, HTD: 3-Hydroxy-cis-5-tetradecenoyl. (*) Metabolite ID confirmed by authentic standards.

Fig. 5. Evaluation of biomarkers to identify SAM from controls. **(A)** Odds Ratios as determined by conditional logistic regressions of all validated metabolites positively associated with SAM. Bars represent 95% confidence intervals. **(B)** Receiver Operator Characteristics (ROC) curves. Metabolites with the highest area under the curve (AUC) are shown.

Table 1: Demographic and clinical characteristics of cases and controls.[1]

Variable	Severe acute malnutrition (N=47)	Non-malnourished controls (N=11)	*P* value
Male: No (%)	30 (63.8)	7 (63.6)	1.00
Age (months; median, range)	22.0 (6 to 48)	14.0 (6 to 44)	0.76
Ethnicity: No. (%)			
• Hausa	43 (91.5)	9 (81.8)	
• Fulani	3 (6.4)	0 (0)	
• Ibo	1 (2.1)	0 (0)	0.048
• Yoruba	0 (0)	1 (9.1)	
• Bugaje	0 (0)	1 (9.1)	
Residence: No. (%)			
• Urban	21 (44.7)	7 (63.6)	
• Peri-urban	1 (2.1)	2 (18.2)	0.024
• Rural	25 (53.2)	2 (18.2)	
Weight-for-length/height z score (median; IQR)	-5.08 (-10.74 to -2.32)	-0.61 (-1.64 to 0.73)	<0.001
Mid-upper arm circumference (cms: median; IQR)	10.0 (7.0 to 12.0)	14.0 (13.5 to 18.0)	<0.001
Pedal oedema: No. (%)	21 (44.7)	0 (0)	-
Stool appearance (Bristol Stool Form Scale):			
• 2	0 (0)	3 (27.3)	
• 3	9 (19.1)	4 (36.4)	
• 4	2 (4.3)	0 (0)	0.003
• 5	9 (19.1)	2 (18.2)	
• 6	18 (38.3)	2 (18.2)	
• 7	9 (19.1)	0 (0)	
Ever received formula milk: No. (%)	23 (48.9)	2 (18.2)	0.093
Ever breast fed: No. (%)	47(100)	11(100)	1.00

[1]*P* values were calculated using the Wilcoxon test with the exception of ethnicity, residence, formula and breast feeding, for which the Chi squared or Fisher's exact test was used.

Click here to access/download
Supplementary Material
Supplemental figures.docx

Click here to access/download
Supplementary Material
Table S5 plasma metabolome .xlsx