Plasma proteomics reveals markers of metabolic stress in HIV infected children
with severe acute malnutrition: call for more evidence-based nutritional
intervention strategies

Gerard Bryan Gonzales1,2*, James M Njunge3,4, Bonface M Gichuki3,4, Bijun Wen5,6, Isabel
Potani3, Wieger Voskujl3,7,8, Robert Bandsma3,5,6, James A Berkley3,9

1. Department of Gastroenterology, Faculty of Medicine and Health Sciences, Ghent
University, Ghent, Belgium
2. VIB Inflammation Research Centre, Ghent, Belgium
3. The Childhood Acute Illness & Nutrition (CHAIN) Network, Nairobi, Kenya
4. KEMRI/Wellcome Trust Research Programme, Kilifi, Kenya
5. Centre for Global Child Health, The Hospital for Sick Children, Toronto, Ontario,
Canada
6. Department of Nutritional Sciences, Faculty of Medicine, University of Toronto,
Toronto, Canada
7. Global Child Health Group, Emma Children's Hospital, Amsterdam University
Medical Centres, Amsterdam, the Netherlands
8. Amsterdam Institute for Global Health and Development, Department of Global
Health, Amsterdam University Medical Centres, Amsterdam, The Netherlands
9. Centre for Tropical Medicine & Global Health, Nuffield Department of Medicine,
University of Oxford, Oxford, UK

*Correspondence to:
Gerard Bryan Gonzales, PhD
Email: Gerard.gonzales@ugent.be
Abstract (word count = 250)

HIV infection affects up to 30% of children presenting with severe acute malnutrition (SAM) in Africa and is associated with increased mortality. Children with SAM are treated similarly regardless of the presence of HIV although mechanisms of nutritional recovery in HIV and/or SAM are not well understood. We aimed to test the hypothesis that metabolic and inflammatory perturbations among children with SAM differed between those with HIV (n = 54) and without HIV (n = 113) to better understand their specific nutritional and therapeutic requirements. We performed a secondary analysis of a clinical trial and plasma proteomics data among children with complicated SAM in Kenya and Malawi. Compared to children with SAM without HIV, HIV-infected children had evidence (false discovery rate corrected p<0.05) of metabolic stress, including enriched pathways related to inflammation and lipid metabolism. Moreover, we observed reduced plasma levels of zinc-α-2-glycoprotein, butyrylcholinesterase, and increased levels of complement C2 resembling findings in metabolic syndrome, diabetes and other non-communicable diseases. HIV was also associated (false discovery rate corrected p<0.05) with higher levels of plasma chemokines: MCP1, CCL4, GCSF, IL1b, TNFa, IL-2, 5, 7, 8 and 15, IL12p40, IP-10, and IL-1RA. Considering evidence of biomarkers of metabolic stress in children with HIV and SAM, it is of potential concern that our current treatment strategy for SAM regardless of HIV status involves a high-fat therapeutic diet. The results of this study indicate a need for clinical trials of therapeutic foods that meet the specific metabolic needs of children with HIV and SAM.

Keywords: HIV, SAM, lipid metabolism, molecular nutrition
Introduction

Malnutrition, specifically undernutrition in all its forms, remains a global public health burden that accounts for 45% of all death among children under 5 years old [1]. Despite careful monitoring and adherence to guidelines set by the World Health Organization, although in general, uncomplicated SAM cases treated in the community do well, up to 25% of children with complicated severe acute malnutrition (SAM) treated in a hospital environment do not survive [2-5]. Furthermore, about one in five children treated for complicated SAM and discharged alive, die in the first year after discharge in low-resource settings [6-8]. However, our understanding of the pathophysiology underlying the poor prognosis for these children is surprisingly limited.

Infection with the human immunodeficiency virus (HIV) is a common co-morbidity of SAM in sub-Saharan Africa affecting up to 30% of admissions among SAM cases [9]. HIV-infected or exposed children are significantly more likely to be stunted, wasted, and underweight [10]. They also more often present with other clinical complications and greater susceptibility to infections, thus further complicating their clinical management, which may include providing more aggressive antimicrobial therapy and higher caloric nutritional intervention [11]. Moreover, response to clinical management is also less predictable and less well-understood in HIV-infected children compared to their uninfected counterparts [12]. Although acute opportunistic infections play a key role in the outcome of these children, intestinal pathology including inflammation and malabsorption, and metabolic perturbations may also be present. However, mechanisms driving poor nutritional recovery of children with HIV even when detected co-morbidities are treated remain poorly understood [12].
We hypothesised that inflammatory, metabolic and other pathways which are likely to be involved in the response to infection, survival and nutritional recovery differ between children with SAM with and without HIV. We conducted a secondary analysis of data and biological samples from a randomised clinical trial in Kenya and Malawi [13].

**Materials and Methods**

**Patient recruitment and study design**

This is a secondary analysis of a nested case control study from a randomised controlled trial (NCT02246296), which tested the effect of a lactose-free, low-carbohydrate F75 milk to limit carbohydrate malabsorption, diarrhoea and refeeding syndrome among children hospitalized for complicated SAM at Queen Elizabeth Central Hospital in Blantyre, Malawi, Kilifi County Hospital and Coast General Hospital, Mombasa, Kenya [13]. Children aged 6 months to 13 years were eligible for enrolment into the trial at admission to hospital if they had SAM, defined as: mid-upper arm circumference (MUAC) <11.5cm or weight-for-height Z score <-3 if younger than 5 years of age, BMI Z score <-3 if older than 5 years, or oedematous malnutrition at any age and had medical complications or failing an appetite test, as defined by WHO guidelines [14]. Children were excluded if they had a known allergy to milk products and did not provide consent. Biological samples were obtained before the children received the randomised treatment irrespective of HIV status. Unless an HIV positive status was documented, HIV status was assessed by offering an antibody test at admission. For this analysis, patients that tested positive on an HIV antibody test were considered HIV(+) and children with missing or declined HIV test were excluded.

The nested case-control study was designed to investigate inpatient mortality for which proteomic, cytokine, and chemokine data was generated using plasma samples collected at
admission during enrolment to the trial. To compare the proteomic profiles between HIV affected and non-affected children with SAM, we used data from a nested case-control study to investigate inpatient mortality during the trial including 79 deaths and 88 children that were clinically stabilized and discharged from the hospital matched by site of recruitment. For this analysis, all observations from the proteomics dataset were included and the analysis was designed to help overcome selection bias, as given in detail below.

**Proteomics, cytokine and chemokine analysis**

Untargeted proteomics and targeted cytokines and chemokines analysis of plasma samples were performed following methods described previously [15]. The targeted protein panel included: epidermal growth factor (EGF); eotaxin; granulocyte-colony stimulating factor (GCSF); granulocyte-macrophage colony-stimulating factor (GMCSF); interferon alpha-2 (IFNa2); interferon gamma (IFNg); interleukins 10, 12p40, 12p70, 13, 15, 17A, 1a, 1b, 1RA, 2 to 8; interferon gamma-induced protein 10 (IP10); monocyte chemoattractant protein 1 (MCP1), macrophage inflammatory protein 1 alpha and beta (MIP1a & b); tumour necrosis factor alpha (TNFa) and beta (TNFb); and vascular endothelial growth factor (VEGF).

**Data analysis**

Data analyses were performed using R v3.5 [16]. Analysis of the prevalence of HIV(+), nutritional status and their associations with inpatient mortality utilised the entire trial dataset (N=843). Analysis of categorical data was performed using Fisher’s test and generalised linear models for continuous outcomes. Logistic regression was used to analyse binary outcomes adjusting for age, sex, presence of oedema, and site of recruitment. These associations were also adjusted for MUAC. As a sensitivity analysis to address the possibility of confounding due to HIV maternal antibodies in younger children, a test of interaction
between age above or below 18 months and individual proteins towards HIV status was performed.

The proteomics, cytokines and chemokines analyses were secondary analyses of data collected from a nested case-control study with inpatient mortality as its primary outcome, hence with strong selection bias. The analysis for the association between HIV status and individual proteins was therefore performed using logistic regression analysis with inverse probability weighting (IPW) to correct for selection bias [17-20]. Weights were calculated as suggested by Samuelsen [18] wherein the weight for each observation selected into the nested case-control study was computed as the inverse of the probability of being selected for the nested study from the main clinical trial. The probability of inclusion was therefore calculated as:

\[
p(i) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n)}},
\]

where \(p(i)\) is the probability of inclusion in the nested case-control study and \(x_1, x_2, \ldots, x_n\) are HIV status, sex, age, presence of oedema, mid-upper arm circumference, and site of recruitment of the \(i\)-th observation (child) based on the entire trial population. Inverse probability weight is therefore:

\[
w(i) = \frac{1}{p(i)}
\]

Differences in individual proteins abundances were considered statistically significant when \(p<0.05\) after adjustment for multiple comparisons using Benjamini-Hochberg false discovery rate (FDR) [21].
Multivariate analysis was undertaken in order to determine several proteins that are collectively associated with HIV status, some of which may not be significantly associated to HIV independently. This was performed using a weighted elastic net (EN) model implemented using the “glmnet” package in R [22]. EN is a penalized regression approach that was developed to help overcome problems caused by high dimensional data. It is an integration of two regularized approaches, ridge regression and least absolute shrinkage and selection operator (LASSO), wherein the contribution of each of these models to the final EN model is controlled by the $\alpha$ parameter [22, 23]. The strong penalization imposed by LASSO draws coefficients to zero thereby eliminating non-predictive proteins features, whereas ridge regression addresses potential multi-collinearity problems in high-dimensional data [22, 23].

Weighted EN model generation was performed with HIV status as outcome, protein profile as predictors, and $w$ as observation weights. The penalization parameter lambda, which influences the shrinkage of variable coefficients to zero thus eliminating some non-contributing variables, was determined by estimating the area under the receiver operator curve (ROC) of the population using ten-fold cross validation. Several alpha parameter values were assessed and a final value of 0.85 was taken to achieve a compromise between predictive ability and fewer number of features extracted. The final lambda parameter was based on the value which gave the highest area under the ROC (AUROC) value.

Proteins with significant association with HIV status after correction for false discovery and those extracted by the EN model were then uploaded to The Database for Annotation, Visualization and Integrated Discovery (DAVID) v6.8 Bioinformatics Resource [24] to assess the Gene ontology (GO) enriched pathways of the differentially expressed proteins.
EN model validity was judged based on the AUROC and misclassification error rate. The fitted EN model performance measured as optimism-corrected AUC was validated using bootstrap, following the procedure of Smith et al. [25]. Bootstrapping was performed on 2000 iterations using the “BootValidation” package in R. Protein features extracted at least 80% of all iterations by the bootstrap EN model were then considered to be the most relevant protein biomarkers. These proteins were then fitted on a weighted logistic regression with HIV as outcome.

**Visualisation of significantly enriched GO terms**

Bubble plots were used to visualise the significantly enriched pathways (p<0.05 after adjustment for FDR) obtained from DAVID. The p-values in DAVID were obtained using a modified Fisher’s exact test [26]. The y-axis represents the fold enrichment which indicates the magnitude of the enrichment, as calculated in DAVID. Fold enrichment is defined as:

\[
fold\ enrichment = \frac{(m/n)}{M/N}
\]

where \(m\) is the number of proteins significantly associated with HIV status or proteins extracted by the EN model that belong to a particular pathway, while \(M\) is the total number of proteins belonging to the same pathway. Variable \(n\) is the number of all proteins significantly associated with HIV status or extracted by the EN model and \(N\) is the total number of all proteins in the human background. Therefore, a fold enrichment of 10 indicates that 10% of the proteins significantly associated with HIV status belong to a particular pathway, and 1% of all annotated proteins in the human background belongs to the same pathway [26]. However, the proponents of this metric warns that big fold enrichments could be obtained from a small number of proteins, which could be due to small \(n\) or pathways with fewer members.
The x-axis on the hand represents the enrichment z-score for a particular pathway [27], which is calculated as follows:

\[
  z\text{-score} = \frac{(\text{up} - \text{down})}{\sqrt{\text{count}}},
\]

where \(\text{up}\) is the total number of proteins upregulated, \(\text{down}\) is the total number of proteins downregulated, and \(\text{count}\) is the total number of proteins in the input which belongs to a particular pathway. Variables \(\text{up}\) and \(\text{down}\) were based on the weighted logistic regression for each individual protein. Hence, if 5 proteins belonging to pathway \(x\) were upregulated and 2 were downregulated, the z-score for pathway \(x\) would be: \((5-2)/\sqrt{7} = 1.13\). A positive z-score indicates that the particular pathway is overall upregulated in HIV(+), whereas a negative z-score indicates an overall downregulation [27].

Results

Patient characteristics

Table 1 presents the baseline characteristics of the children in the randomised trial. A total of 843 complicated SAM children were recruited for the randomised trial, of which 179 (22%) patients were HIV(+). Age was higher and MUAC was lower in HIV(+) children than HIV(-) counterparts. Most HIV cases were found in Malawi. Sex and the presence of oedema were not associated with HIV status. Mortality was more than 2 times higher among in HIV(+) compared to HIV(-) (\(p < 0.001\)). Children whose HIV status were unknown had the highest mortality of 34%, which indicates the need for more active HIV screening among children with SAM.

Among HIV(+), 33\% were already under an anti-retroviral treatment (ART) regime: 53/179 (30\%) on highly active antiretroviral therapy (HAART), and 7/179 (4\%) on Nevirapine only. About half of the children (90/179) were naïve for ART whereas HIV treatment status was
unknown for 16% (29/179). Mortality was not significantly different among children on
HAART, ART naïve and children with unknown HIV treatment status (Supplementary Table
1).

HIV is associated with increased inflammatory and immune response, dysregulated lipid
metabolism, and increased proteolysis in children with SAM

Table 2 shows the characteristics for children included in the proteomics study. Fifty four
HIV (+) children were compared to 113 HIV(-) children with SAM. In this sub-population,
age was not significantly associated with HIV, and as in the full dataset. HIV(+) children also
had significantly lower MUAC and higher mortality than HIV(-) children. There were no
significant associations between sex and presence of oedema with HIV status.

A total of 204 circulating proteins were annotated and compared between children with and
without HIV infection. Of these, levels of 42 proteins were found to be significantly
associated with HIV status in the initial univariate analysis (Figure 1A) (Supplementary table
2). Specifically, HIV(+) was associated with higher circulating levels of immunoglobulins,
inflammatory proteins such as calprotectin (S100 calcium binding protein A8 and S100
calcium binding protein A9), complement proteins, and proteins related to host response to
infection (i.e. lipopolysaccharide binding protein, galectin 3 binding protein and CD5
molecule-like protein). Enrichment analysis suggested that HIV(+) children have increased
complement activation and immune response, and inflammatory responses than HIV(-)
children. Neutrophil aggregation and chemokine production appeared to be the pathways most
highly enriched in HIV(+) compared to HIV(-) SAM children. To substantiate these results,
we quantified chemokine and cytokine levels in plasma. As shown, most chemokines had the
tendency to be associated with HIV infection, where elevated plasma concentration of 12
were significantly associated with HIV status in SAM children (Figure 1B), namely: monocyte chemoattractant protein 1 (MCP1), macrophage inflammatory protein 1 beta (MIP1b, CCL4), granulocyte colony-stimulating factor (GCSF), interleukin 1 beta (IL1b), tumour necrosis factor alpha (TNFa), interleukins 2, 5, 7, 8 and 15 (IL2, 5, 7, 8, 15), interleukin 12 subunit beta (IL12p40), interferon gamma-induced protein 10 (IP-10), and interleukin-1 receptor antagonist (IL-1RA).

Out of the 43 differentially expressed proteins, three proteins were found to be negatively associated with HIV status on initial univariate analysis, namely: adiponectin, kininogen-1 and peptidase inhibitor 16. Among HIV(+) children, there were no statistically significant associations with receiving HAART (n = 53) compared to ART naïve (n = 90) children (data not shown), although our study may not be powered to observe significant difference for this comparison. Furthermore, sensitivity analysis to address the possibility of HIV maternal antibodies in younger children, showed no significant interaction of age above or below 18 months and individual proteins plasma levels towards HIV status.

The weighted EN model extracted 73 circulating proteins (Figure 2A) that are associated with HIV status with AUROC = 0.97 [95% CI: 0.95 – 0.99] (Figure 2B) and misclassification error rate of 2.4%. Optimism-adjusted validated AUROC after bootstrapping was 0.90 [95% CI: 0.90 – 0.90], indicating a robust model. Pathway enrichment analysis highlighted that apart from increased immune response, HIV(+) children with SAM have increased proteolysis and lipid mobilisation, specifically increased very low-density lipoprotein assembly, indicating metabolic dysregulation related to cholesterol and triglyceride metabolism among HIV(+) patients (Figure 2D).
After 2000 bootstrap iterations during bootstrap validation, 3 proteins were consistently extracted by the EN model >80% of the time (Figure 2E), namely: butyrylcholinesterase (BChE), complement C2 and zinc-α-2-glycoprotein (ZAG), indicating that these 3 proteins are likely to be the most important features associated with HIV in children with complicated SAM. Weighted logistic regression model of these 3 proteins showed good discrimination of HIV status (AUROC = 0.80 [95% CI: 0.74 – 0.87]) (Figure 2F).

Discussion

In this study, we report plasma proteomic differences associated with HIV status, which indicates that HIV imposes additional metabolic and inflammatory insults among HIV(+) children with SAM. Our results show that pathways involved in inflammatory response, complement cascade activation and lipid metabolism dysregulation are associated with HIV status. Levels of several plasma chemokines were also found to be higher in HIV(+) among children with SAM. Greater inflammatory responses in these children could be related to the higher inpatient mortality of HIV(+) compared to HIV(-) children with SAM.

An earlier metabolomics study in Uganda reported reduced serum levels of adiponectin and leptin, whereas serum triglycerides, ketones and even-chain acylcarnitines were higher in HIV(+) children with SAM indicating perturbed lipid metabolism [28]. Our current study therefore concurs with this finding, as we also found reduced plasma levels of adiponectin in HIV(+) SAM children compared to HIV(-) SAM children, along with upregulation of pathways involved in lipid transport and metabolism, specifically very low-density lipoprotein assembly.
Using optimism-adjusted bootstrap validation of the EN model, we found 3 proteins: complement c2, BChE and ZAG robustly distinguished HIV(+) from HIV(-) in children with SAM, demonstrating the ability of multivariate analysis techniques, such as EN, to uncover underlying relationships between protein markers which would be difficult to identify when analysed individually. The activation of the complement system during HIV infection has been previously discussed at length, which is associated with the increased cellular invasion of HIV in cells [29-31]. On the other hand, in a recent study in China, low circulating BChE was found to be highly associated with HIV severity, was predictive of mortality in adults, and was proposed as a plausible strategy for severity classification among adults with HIV [32]. BChE is also reported to be reduced in protein-energy malnutrition, stress and inflammation [33]. In animal studies, BChE deficiency was found to strongly affect fat metabolism and promotes hepatic lipid accumulation [34]. Serum BChE levels have been found to have a significant negative correlation with serum total cholesterol and serum low-density-lipoprotein cholesterol among people with diabetes mellitus [35]. Furthermore, reduced plasma levels of ZAG, an adipokine with protective effect against insulin resistance, was previously reported to be implicated in dyslipidaemia in HIV(+) adults under ART treatment [36]. Reduced circulating levels of ZAG has also been found among adults with clinically diagnosed metabolic syndrome, based on guidelines of the United States National Cholesterol Education Program (NCEP) Expert Panel Adult Treatment Panel (ATP) III criteria [37]. Serum ZAG levels have been reported lower among adults with impaired glucose tolerance and type 2 diabetes mellitus [38]. Taken together, our results therefore suggest that children with both HIV and SAM children manifest hallmarks of metabolic stress similar to those occurring in metabolic syndrome and other non-communicable diseases (NCD).
The results of this study is the first proteomics investigation on the interaction between HIV and SAM, which together with the previously published metabolomics study [28], provides more evidence on the increased metabolic stress and altered metabolic response among children living with both HIV and SAM. Our results also concur with previous studies that reported elevated metabolic stress among non-malnourished adults living with HIV leading to increased prevalence or risk for metabolic syndrome, cardiovascular diseases, diabetes and other non-communicable diseases [39-45].

However, despite our knowledge that HIV-infected populations have altered metabolic requirements compared to HIV-uninfected counterparts, WHO guidelines for the nutritional management for SAM are globally the same regardless of HIV status, which is summarized in Table 3 [46]. Nutritional management for in-patient children with SAM involves provision of a low-protein, low-fat milk-based food, F75, every few hours. F75 is used during clinical stabilization occurring during the first few days after admission and is not intended for weight gain. Once the children are clinically stabilized and are able to tolerate the solute load, children are transitioned to F100, a higher-calorie, high-fat milk intended to boost weight gain. Upon discharge from in-patient care, children are referred to community based nutritional therapeutic centres where they are provided with ready-to-use therapeutic food (RUTF), a peanut-based calorie-dense diet.

Considering evidence of biomarkers of metabolic syndrome and NCD in HIV(+) children with SAM, it is of potential concern that our current treatment strategy involves a high-fat therapeutic diet. About 50% of much needed calories during the growth catch-up phase are supplied as lipids, which HIV(+) children may not be able to efficiently assimilate. Alterations in lipid metabolism in HIV(+) children with SAM may also mean that the high
amounts of dietary lipids could be deposited as ectopic fat in the liver and muscle, predisposing to insulin resistance, diabetes, cardiovascular problems and other NCDs later in life. Although long-term metabolic follow-up studies could be done for HIV(+) children previously treated for either complicated and uncomplicated SAM, significant barriers are the high mortality rate in earlier studies of HIV(+) children with SAM, cost and difficulty tracing them years later. The results of this study indicate a need for clinical trials modifying the composition of F100 or RUTF to meet the specific metabolic needs of HIV(+) children with SAM. This could initially be done in relatively small groups with outcomes that include measuring metabolic stress.

Several studies on nutritional intervention strategies among HIV-infected adults have been reported. For instance, a study in the USA showed that dietary fat intake, specifically saturated fats, was significantly associated with hypertriglyceridemia among HIV-infected adults (18 – 60 years) [47]. Moreover, in a preclinical model, high saturated fat consumption was found to accelerate immunodeficiency virus disease progression in macaques, specifically increased mortality hazard and circulating levels of pro-inflammatory cytokines, especially IL8 [48], which has been previously reported to be associated with lipodystrophy among HIV patients [49]. In our study, we also found a significant association between high plasma IL8 concentration and HIV in SAM children. Hence, modifying the saturated fat composition of the milk-based F75 and F100 could potentially lower metabolic stress.

Furthermore, the European Society for Parenteral and Enteral Nutrition (ESPEN) gave a grade A recommendation for the use of medium-chain triglyceride (MCT)-based diet on HIV(+) patients with diarrhoea and severe undernutrition in its 2006 ESPEN Guidelines on Enteral Nutrition [50]. Grade A recommendations are given to strategies based on meta-analysis or at
least one randomised control trial. In this case, the recommendation was based on a prospective, randomized double-blind comparative trial on 24 adult patients with HIV and diarrhoea of more than 4-wk duration, fat malabsorption, and loss of 10–20% of ideal body weight [51]. In this study, the authors found improved outcomes from diarrhoea and fat malabsorption from MCT than long-chain triglyceride-based diet among HIV(+) adults.

Lastly, the long-term metabolic effect of nutritional intervention strategies for SAM still remains unresolved. Most specifically, the potential metabolic stress associated with the rapid weight gain during the nutritional rehabilitation phase after SAM and its implications on nutritional outcomes during adulthood demands urgent research attention, especially for HIV(+) children with SAM.

Conclusion
Plasma proteomics reveals that HIV(+) children with SAM manifest hallmarks of metabolic stress similar to those observed in non-communicable diseases. This could be related to the poor nutritional recovery and high mortality of HIV(+) children with SAM despite clinical and nutritional intervention. The results of this study indicate a need for clinical trials modifying the composition of F100 or RUTF to meet the specific metabolic needs of HIV(+) children with SAM during rehabilitation phase. This could initially be done in relatively small groups with outcomes that include measuring metabolic stress.

Acknowledgements
The original clinical trial was supported by the Thrasher Fund (grant number: 9403), awarded to JAB. GBG is a postdoctoral fellow of the Research Foundation – Flanders (FWO). The Ghent University –VLIR-UOS Global Minds Fund supported the travel of GBG to Kenya.
JAB and JMN are currently supported by the Bill & Melinda Gates Foundation within the Childhood Acute Illness and Nutrition (CHAIN) Network (grant OPP1131320). JAB is currently supported by the MRC/DfID/Wellcome Trust Joint Global Health Trials scheme (grant MR/M007367/1). The funders had no role in the design, conduct, analysis, or writing of this manuscript.

Study approvals

The secondary analyses of the trial were approved by the Kenyan national ethics committee, KEMRI- SERU (KEMRI/RES/7/3/1). The trial was registered at clinicaltrials.gov (NCT02246296).

Author contributions

GBG, JMN and JAB designed the study and data analysis. JMN and BG performed the proteomics analysis. GBG performed data analysis. BW, RB, WV, IP and JAB were involved in the clinical aspects of the study. GBG wrote the initial drafts of the manuscript and all authors contributed to editing and improving the manuscript.

References

1. Black, R.E., et al., *Maternal and child undernutrition and overweight in low-income and middle-income countries*. The Lancet, 2013. 382(9890): p. 427-451.
2. Heikens, G.T., et al., *Case management of HIV-infected severely malnourished children: challenges in the area of highest prevalence*. Lancet, 2008. 371(9620): p. 1305-7.
3. Irena, A.H., M. Mwambazi, and V. Mulenga, *Diarrhea is a Major killer of Children with Severe Acute Malnutrition Admitted to Inpatient Set-up in Lusaka, Zambia*. Nutrition Journal, 2011. 10: p. 110.
4. Maitland, K., et al., *Children with Severe Malnutrition: Can Those at Highest Risk of Death Be Identified with the WHO Protocol?* PLOS Medicine, 2006. 3(12): p. e500.
5. Hossain, M., et al., *Efficacy of World Health Organization guideline in facility-based reduction of mortality in severely malnourished children from low and middle income countries: A systematic review and meta-analysis*. J Paediatr Child Health, 2017. 53(5): p. 474-479.
6. Kerac, M., et al., *Follow-Up of Post-Discharge Growth and Mortality after Treatment for Severe Acute Malnutrition (FuSAM Study): A Prospective Cohort Study.* PLOS ONE, 2014. 9(6): p. e96030.

7. Berkley, J.A., et al., *Daily co-trimoxazole prophylaxis to prevent mortality in children with complicated severe acute malnutrition: a multicentre, double-blind, randomised placebo-controlled trial.* Lancet Glob Health, 2016. 4(7): p. e464-73.

8. Nemetchek, B., et al., *Paediatric postdischarge mortality in developing countries: a systematic review.* BMJ Open, 2018. 8(12): p. e023445.

9. Fergusson, P. and A. Tomkins, *HIV prevalence and mortality among children undergoing treatment for severe acute malnutrition in sub-Saharan Africa: a systematic review and meta-analysis.* Trans R Soc Trop Med Hyg, 2009. 103(6): p. 541-8.

10. Ndirangu, M., et al., *Nutritional status of under-five children in HIV-affected households in western Kenya.* Food Nutr Bull, 2011. 32(2): p. 159-67.

11. Rose, A.M., C.S. Hall, and N. Martinez-Alier, *Aetiology and management of malnutrition in HIV-positive children.* Archives of Disease in Childhood, 2014. 99(6): p. 546-551.

12. Trehan, I., et al., *Challenges in the Management of HIV-Infected Malnourished Children in Sub-Saharan Africa.* AIDS Research and Treatment, 2012. 2012: p. 8.

13. Bandsma, R.H.J., et al., *A reduced-carbohydrate and lactose-free formulation for stabilization among hospitalized children with severe acute malnutrition: A double-blind, randomized controlled trial.* PLoS Med, 2019. 16(2): p. e1002747.

14. WHO. Management of severe malnutrition: a manual for physicians and other senior health workers. 1999 May 2019]; Available from: http://apps.who.int/iris/bitstream/10665/41999/1/a57361.pdf.

15. Njunge, J.M., et al., *Biomarkers of post-discharge mortality among children with complicated severe acute malnutrition.* Scientific Reports, 2019. 9(1): p. 5981.

16. R Core Team. *R: A language and environment for statistical computing.* R Foundation for Statistical Computing. 2017; Available from: https://www.R-project.org/.

17. Kim, R.S. and R.C. Kaplan, *Analysis of secondary outcomes in nested case-control study designs.* Statistics in medicine, 2014. 33(24): p. 4215-4226.

18. Samuelsen, S.O., *A Pseudolikelihood Approach to Analysis of Nested Case-Control Studies.* Biometrika, 1997. 84(2): p. 379-394.

19. Pan, Y., et al., *Secondary outcome analysis for data from an outcome-dependent sampling design.* Statistics in Medicine, 2018. 37(15): p. 2321-2337.

20. Jiang, Y., A.J. Scott, and C.J. Wild, *Secondary analysis of case-control data.* Statistics in Medicine, 2006. 25(8): p. 1323-1339.

21. Benjamini, Y. and Y. Hochberg, *Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing.* Journal of the Royal Statistical Society. Series B (Methodological), 1995. 57(1): p. 289-300.

22. Friedman, J.H., T. Hastie, and R. Tibshirani, *Regularization Paths for Generalized Linear Models via Coordinate Descent.* 2010, 2010. 33(1): p. 22.

23. Zou, H. and T. Hastie, *Regularization and variable selection via the elastic net.* Journal of the Royal Statistical Society: Series B (Statistical Methodology), 2005. 67(2): p. 301-320.

24. Huang, D.W., B.T. Sherman, and R.A. Lempicki, *Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources.* Nat. Protocols, 2008. 4(1): p. 44-57.

25. Smith, G.C.S., et al., *Correcting for Optimistic Prediction in Small Data Sets.* American Journal of Epidemiology, 2014. 180(3): p. 318-324.
26. Huang da, W., B.T. Sherman, and R.A. Lempicki, Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc, 2009. 4(1): p. 44-57.

27. Walter, W., F. Sanchez-Cabo, and M. Ricote, GOplot: an R package for visually combining expression data with functional analysis. Bioinformatics, 2015. 31(17): p. 2912-4.

28. Mody, A., et al., Effects of HIV infection on the metabolic and hormonal status of children with severe acute malnutrition. PLoS One, 2014. 9(7): p. e102233.

29. Tacnet-Delorme, P., et al., In Vitro Analysis of Complement-Dependent HIV-1 Cell Infection Using a Model System. The Journal of Immunology, 1999. 162(7): p. 4088.

30. Stoiber, H., et al., The supportive role of complement in HIV pathogenesis. Immunological Reviews, 2001. 180(1): p. 168-176.

31. Stoiber, H., C. Speth, and M.P. Dierich, Role of complement in the control of HIV dynamics and pathogenesis. Vaccine, 2003. 21: p. S77-S82.

32. Xu, L., et al., Butyrylcholinesterase Levels on Admission Predict Severity and 12-Month Mortality in Hospitalized AIDS Patients. Mediators Inflamm, 2018. 2018: p. 5201652.

33. Santarpia, L., et al., Butyrylcholinesterase as a prognostic marker: a review of the literature. J Cachexia Sarcopenia Muscle, 2013. 4(1): p. 31-9.

34. Chen, V.P., et al., Butyrylcholinesterase Deficiency Promotes Adipose Tissue Growth and Hepatic Lipid Accumulation in Male Mice on High-Fat Diet. Endocrinology, 2016. 157(8): p. 3086-95.

35. Sridhar, G., et al., Serum butyrylcholinesterase in type 2 diabetes mellitus: a biochemical and bioinformatics approach. Lipids in Health and Disease, 2005. 4(1): p. 18.

36. Ceperuelo-Mallafre, V., et al., Zinc alpha-2 glycoprotein is implicated in dyslipidaemia in HIV-1-infected patients treated with antiretroviral drugs. HIV Med, 2012. 13(5): p. 297-303.

37. Lei, L., et al., Circulating zinc-α2-glycoprotein levels are low in newly diagnosed patients with metabolic syndrome and correlate with adiponectin. Nutrition & Metabolism, 2017. 14(1): p. 53.

38. Yang, M., et al., Zinc-α2-glycoprotein is associated with insulin resistance in humans and is regulated by hyperglycemia, hyperinsulinemia, or liraglutide administration: cross-sectional and interventional studies in normal subjects, insulin-resistant subjects, and subjects with newly diagnosed diabetes. Diabetes care, 2013. 36(5): p. 1074-1082.

39. Bastard, J.P., et al., Diabetes and dyslipidaemia are associated with oxidative stress independently of inflammation in long-term antiretroviral-treated HIV-infected patients. Diabetes Metab, 2019.

40. Nix, L.M. and P.C. Tien, Metabolic syndrome, diabetes, and cardiovascular risk in HIV. Curr HIV/AIDS Rep, 2014. 11(3): p. 271-8.

41. Sobieszczuk, M.E., et al., Metabolic Syndrome After HIV Acquisition in South African Women. J Acquir Immune Defic Syndr, 2016. 73(4): p. 438-445.

42. Policarpo, S., et al., Cardiovascular risk in HIV-infected individuals: A comparison of three risk prediction algorithms. Rev Port Cardiol, 2019. 38(7): p. 463-470.

43. Lagathu, C., et al., Metabolic complications affecting adipose tissue, lipid and glucose metabolism associated with HIV antiretroviral treatment. Expert Opin Drug Saf, 2019. 18(9): p. 829-840.
44. Sears, S., et al., *Metabolic Syndrome Among People Living with HIV Receiving Medical Care in Southern United States: Prevalence and Risk Factors.* AIDS Behav, 2019.

45. Seth, A. and K.E. Sherman, *Fatty liver disease in persons with HIV infection.* Top Antivir Med, 2019. 27(2): p. 75-82.

46. Trehan, I. and M.J. Manary, *Management of severe acute malnutrition in low-income and middle-income countries.* Arch Dis Child, 2015. 100(3): p. 283-7.

47. Joy, T., et al., *Dietary fat intake and relationship to serum lipid levels in HIV-infected patients with metabolic abnormalities in the HAART era.* AIDS (London, England), 2007. 21(12): p. 1591-1600.

48. Mansfield, K.G., et al., *A diet high in saturated fat and cholesterol accelerates simian immunodeficiency virus disease progression.* J Infect Dis, 2007. 196(8): p. 1202-10.

49. Lindegaard, B., et al., *Adipose tissue expression of IL-18 and HIV-associated lipodystrophy.* Aids, 2004. 18(14): p. 1956-8.

50. Ockenga, J., et al., *ESPEN Guidelines on Enteral Nutrition: Wasting in HIV and other chronic infectious diseases.* Clinical Nutrition, 2006. 25(2): p. 319-329.

51. Wanke, C.A., et al., *A medium chain triglyceride-based diet in patients with HIV and chronic diarrhea reduces diarrhea and malabsorption: A prospective, controlled trial.* Nutrition, 1996. 12(11): p. 766-771.
Figure 1. (A) Volcano plot showing several significantly different (FDR adjusted $p$ value < 0.05) proteins and their log2 HIV(+) versus HIV(-) fold change. Red points represent those significantly higher in plasma of HIV(-), blue points significantly enriched in plasma of HIV(+) and orange points significantly higher than 1.5 folds in HIV(+) compared to HIV(-) SAM children. Vertical lines indicate significance level at $p = 0.05$ and 0.01; horizontal lines indicate more than 1.5 folds enrichment. (B) Log odds plots showing association of chemokine markers analysed using Luminex platform and HIV status. Points indicate log odds ratio for every log increase in plasma protein concentration; bars indicate 95% confidence interval. (C) Gene ontology (GO-terms) enrichment analysis of proteins significantly higher in plasma showing biologically relevant processes heightened among HIV(+) SAM children. X-axis represents z-scores; y-axis, fold enrichment, and size of the spheres represent the number of proteins involved in the particular pathway. Only significantly enriched pathways ($p < 0.05$ after FDR adjustment) are plotted. See main text for explanation of the plots. Pathways enriched are identified in the table.
Figure 2. (A) Elastic net (EN) regularized regression lambda parameter optimization curve, optimal lambda parameter was chosen based on the highest area under the receiver operator curve (AUROC); (B) AUROC (0.97 [95% CI: 0.95 – 0.99]) of the EN model generated using the lambda parameter, alpha parameter was set to 0.75, final model extracted 34 protein features; (C) optimism-adjusted bootstrap validation of the generated EN model, validated AUROC = 0.90 [95% CI: 0.90 – 0.90] using 2000 iterations; (D) Gene ontology (GO-terms) enrichment analysis of proteins extracted by the EN model. X-axis represents z-scores; y-axis, fold enrichment, and size of the spheres represent the number of proteins involved in the particular pathway. Gold circles represent pathways enriched in HIV(+) whereas blue circles are pathways more associated with HIV(-). The grey circle indicate that there are as much proteins in this pathway that are significantly upregulated and downregulated in HIV. Only significantly enriched pathways (p < 0.05 after FDR adjustment) are plotted. See main text for explanation of the plots. Pathways enriched are identified in the table. (E) Log odds ratio plot
of the 3 proteins extracted after bootstrap validation with log odds on the x-axis and bars indicating 95% confidence interval obtained using weighted logistic regression with HIV as outcome variable and the 3 proteins as covariates. Weights used were obtained by inverse probability of treatment weights; (F) Predictive ability of the weighted logistic regression model using the 3 bootstrap validated proteins with HIV as outcome variable, AUROC = 0.80 [95% CI: 0.73 – 0.87]
Table 1. Descriptive characteristics of the study participants

|                     | All      | HIV (+)  | HIV (-)  | Unknown HIV status | p*       |
|---------------------|----------|----------|----------|--------------------|----------|
| n (%)               | 843      | 179 (21%)| 618 (73%)| 46 (5%)            |          |
| Median age in months [IQR] | 16 [10 – 25] | 21 [12 – 31] | 16 [10 – 25] | 10 [8 – 17] | <0.001   |
| % girls (n)         | 45% (359)| 45% (81) | 45% (278)| 56% (26)          | 0.95     |
| Mean MUAC in cm [95% CI] | 11.2 [11.1 - 11.3] | 10.5 [10.36 - 10.7] | 11.4 [11.3 -11.5] | 11.2 [10.9 – 11.5] | <0.001   |
| Mean weight-for-age z-score [95% CI] | -4.01 [-4.11 – -3.92] | -4.51 [-4.72 – -4.31] | -3.92 [-4.03 – -3.80] | -3.56 [-3.94 – -3.92] | <0.001   |
| % mortality (n)     | 15% (127)| 26% (47) | 10% (64) | 34% (16)          | <0.001°  |
| % oedematous (n)    | 31% (264)| 30% (54) | 33% (203)| 15% (7)           | 0.50     |
| Site                |          |          |          |                   |          |
| Coast Provincial General Hospital, Kenya | 39% (329)| 25% (45) | 40% (247)| 80% (37)          | Reference|
| Kilifi County Hospital, Kenya | 22% (187)| 22% (40) | 23% (145)| 4% (2)            | 0.08     |
| Queen Elizabeth Central Hospital, Malawi | 39% (327)| 52% (94) | 36% (226)| 15% (7)           | <0.001   |

*comparison between HIV(+) and HIV(-); °adjusted for age, sex and site of recruitment
|                      | All  | HIV (+) | HIV (-) | p*    |
|----------------------|------|---------|---------|-------|
| n                    | 167  | 54      | 113     |       |
| Median age in months [IQR] | 15 [10 – 26] | 15 [10 – 26] | 15 [10 – 24] | 0.433 |
| n girls (%)          | 76 (45%) | 27 (50%) | 49 (43%) | 0.506 |
| Mean MUAC at admission (cm) [95% CI] | 10.9 [10.7 - 11.1] | 10.2 [9.8 - 10.5] | 11.3 [11.0 - 11.5] | <0.001 |
| % oedematous (n)     | 49 (29%) | 15 (28%) | 34 (30%) | 0.856 |
| % mortality (n)      | 79 (47%) | 36 (67%) | 43 (38%) | <0.001° |

*comparison between HIV(+) and HIV(-), °adjusted for age, sex, site of recruitment, oedema
Table 3. Nutritional management protocol for children with severe acute malnutrition [46]

|                  | Stabilization phase | In-patient Rehabilitation phase | Out-patient Rehabilitation phase |
|------------------|---------------------|----------------------------------|----------------------------------|
|                  | Days 1 - 7          | Weeks 2 - 6                      | Lengths vary depending on site   |
| Complicated SAM  | F75                 | F100                             | RUTF                             |
| Uncomplicated SAM| -                   | -                                | RUTF                             |

**Composition**

|                   | Energy (kcal per 100 mL F75/F100 or 100 g RUTF) | Protein (% total energy) | Fat (% total energy) |
|-------------------|-------------------------------------------------|--------------------------|----------------------|
|                   | 75                                               | 5                        | 32                   |
|                   | 100                                              | 12                       | 53                   |
|                   | 5.2 – 5.5                                         | 10 – 12                  | 45 - 60              |
Supplementary Table 1. Prior use of anti-retroviral treatment among all HIV (+) children included in the clinical trial

| Prior HIV treatment* | HAART | NVP only | ART naïve | Unknown |
|---------------------|-------|----------|-----------|---------|
| n (%)°              | 53 (29%) | 7 (4%) | 90 (50%) | 29 (16%) |
| Mean MUAC at admission (cm) [95% CI] | 10.9 [10.5 - 11.4] | 9.8 [8.9 - 10.6] | 10.5 [10.2 - 10.8] | 9.6 [8.1 - 11.1] |
| Mortality n (%)     | 10 (18%) | 1 (28%) | 25 (28%) | 1 (25%) |
| Odds ratio for mortality [95% CI] | 0.97 [0.82 - 1.12] | 1 | Reference | 1.07 [0.88 – 1.30] |
|                     | p=0.63# |         |           | p=0.45# |

*HAART - highly active antiretroviral therapy; NVP – Nevirapine; ° adjusted for age, sex, site and oedema

Reference 1.07 [0.88 – 1.30] p=0.45#
### Supplementary table 2. Association between individual proteins to HIV status

| Uniprot accession | Protein code | Protein names | log FC* | p°  |
|-------------------|--------------|---------------|--------|-----|
| P05109            | S100A8       | Protein S100-A8 (Calgranulin-A) (Calprotectin L1L subunit) (Cystic fibrosis antigen) (CFAG) (Leukocyte L1 complex light chain) (Migration inhibitory factor-related protein 8) (MRP-8) (p8) (S100 calcium-binding protein A8) (Urinary stone protein band A) | 0.56   | 0.00|
| P06702            | S100A9       | Protein S100-A9 (Calgranulin-B) (Calprotectin L1H subunit) (Leukocyte L1 complex heavy chain) (Migration inhibitory factor-related protein 14) (MRP-14) (p14) (S100 calcium-binding protein A9) | 0.55   | 0.00|
| P01708            |              | Immunoglobulin lambda variable 2-11 (Ig gamma lambda chain V-II region DOT) (Ig lambda chain V-II region BOH) (Ig lambda chain V-II region BUR) (Ig lambda chain V-II region NIG-58) (Ig lambda chain V-II region TRO) (Ig lambda chain V-II region WIN) | 0.59   | 0.01|
| P04275            | VWF          | von Willebrand factor (vWF) [Cleaved into: von Willebrand antigen 2 (von Willebrand antigen II)] | 0.33   | 0.01|
| P01880            | IGHD         | Immunoglobulin heavy constant delta (Ig delta chain C region) (Ig delta chain C region NIG-65) (Ig delta chain C region WAH) | 0.60   | 0.01|
| P23083            | IGHV1-2      | Immunoglobulin heavy variable 1-2 (Ig heavy chain V-I region ND) (Ig heavy chain V-I region V35) | 0.60   | 0.01|
| P01598            |              | Immunoglobulin kappa variable 1-5 (Ig kappa chain V-I region CAR) (Ig kappa chain V-I region EU) (Ig kappa chain V-I region HK102) (Ig kappa chain V-I region Kue) | 0.55   | 0.01|
| P80748            | IGLV3-21     | Immunoglobulin lambda variable 3-21 (Ig lambda chain V-III region LOI) (Ig lambda chain V-V region DEL) (Ig lambda chain V-VII region MOT) | 0.58   | 0.01|
| P13796            | LCP1         | Plastin-2 (L-plastin) (LC64P) (Lymphocyte cytosolic protein 1) (LCP-1) | 0.28   | 0.01|
| P01857            | IGHG1        | Immunoglobulin heavy constant gamma 1 (Ig gamma-1 chain C region) (Ig gamma-1 chain C region EU) (Ig gamma-1 chain C region KOL) (Ig gamma-1 chain C region NIE) | 0.64   | 0.01|
| P05362            | ICAM1        | Intercellular adhesion molecule 1 (ICAM-1) (Major group rhinovirus receptor) (CD antigen CD54) | 0.28   | 0.01|
| P06318            |              | Immunoglobulin lambda variable 6-57 (Ig lambda chain V-VI region AR) (Ig lambda | 0.49   | 0.01|
| Accession | Description                                                | Score | E-value |
|-----------|-------------------------------------------------------------|-------|---------|
| P18428    | LBP Lipopolysaccharide-binding protein (LBP)                | 0.31  | 0.01    |
| P01834    | IGKC Immunoglobulin kappa constant (Ig kappa chain C region) | 0.69  | 0.02    |
| P0CG05    | Immunoglobulin lambda constant 3 (Ig lambda chain C region) | 0.67  | 0.02    |
| P0CG05    | Immunoglobulin lambda constant 2 (Ig lambda chain C region) | 0.67  | 0.02    |
| B9A064    | IGLL5 Immunoglobulin lambda-like polypeptide 5 (G lambda-1) | 0.60  | 0.02    |
| H9KV70    | Deleted.                                                    | 0.27  | 0.02    |
| P01597    | IGKV1-39 Immunoglobulin kappa variable 1-39 (Ig kappa chain V-I region DEE) | 0.47  | 0.02    |
| P01714    | IGLV3-19 Immunoglobulin lambda variable 3-19 (Ig lambda chain V-III region SH) | 0.54  | 0.02    |
| P06331    | IGHV4-34 Immunoglobulin heavy variable 4-34 (Ig heavy chain V-II region ARH-77) | 0.49  | 0.02    |
| P19652    | ORM2 Alpha-1-acid glycoprotein 2 (AGP 2) (Orosomucoid-2) (OMD 2) | 0.20  | 0.02    |
| P01610    | Immunoglobulin kappa variable 1-17 (Ig kappa chain V-I region Gal) (Ig kappa chain V-I region WEA) | 0.52  | 0.02    |
| P01781    | Immunoglobulin heavy variable 3-7 (Ig heavy chain V-III region GAL) (Ig heavy chain V-III region GAR) (Ig heavy chain V-III region JON) | 0.55  | 0.02    |
| P02741    | CRP C-reactive protein [Cleaved into: C-reactive protein(1-205)] | 0.51  | 0.02    |
| P01717    | IGLV3-25 Immunoglobulin lambda variable 3-25 (Ig lambda chain V-IV region Hil) | 0.46  | 0.02    |
| P01719    | Immunoglobulin lambda variable 3-21 (Ig lambda chain V-III region LOI) (Ig lambda chain V-V region DEL) (Ig lambda chain V-VII region MOT) | 0.44  | 0.02    |
| P06681  | C2    | Complement C2 (EC 3.4.21.43) (C3/C5 convertase) [Cleaved into: Complement C2b fragment; Complement C2a fragment] | 0.15  | 0.02  |
| Q15848  | ADIPOQ | Adiponectin (30 kDa adipocyte complement-related protein) (Adipocyte complement-related 30 kDa protein) (ACRP30) (Adipocyte, C1q and collagen domain-containing protein) (Adipose most abundant gene transcript 1 protein) (apM-1) (Gelatin-binding protein) | -0.23 | 0.02  |
| P04208  |       | Immunoglobulin lambda variable 1-47 (Ig lambda chain V-I region HA) (Ig lambda chain V-I region WAH) | 0.53  | 0.02  |
| Q15485  | FCN2  | Ficolin-2 (37 kDa elastin-binding protein) (Collagen/fibrinogen domain-containing protein 2) (EBP-37) (Ficolin-B) (Ficolin-beta) (Hucoin) (L-ficolin) (Serum lectin p35) | 0.26  | 0.03  |
| P18135  |       | Immunoglobulin kappa variable 3-20 (Ig kappa chain V-III region B6) (Ig kappa chain V-III region GOL) (Ig kappa chain V-III region HAH) (Ig kappa chain V-III region HIC) (Ig kappa chain V-III region IARC/BL41) (Ig kappa chain V-III region NG9) (Ig kappa chain V-III region SIE) (Ig kappa chain V-III region Ti) (Ig kappa chain V-III region WOL) | 0.57  | 0.03  |
| P01613  |       | Immunoglobulin kappa variable 1D-33 (Ig kappa chain V-I region AG) (Ig kappa chain V-I region Bi) (Ig kappa chain V-I region Lay) (Ig kappa chain V-I region Ni) (Ig kappa chain V-I region Rei) (Ig kappa chain V-I region Roy) (Ig kappa chain V-I region Scw) (Ig kappa chain V-I region WAT) | 0.45  | 0.04  |
| P01602  | IGKV1-5 | Immunoglobulin kappa variable 1-5 (Ig kappa chain V-I region CAR) (Ig kappa chain V-I region EU) (Ig kappa chain V-I region HK102) (Ig kappa chain V-I region Kue) | 0.55  | 0.04  |
| P04211  | IGLV7-43 | Immunoglobulin lambda variable 7-43 (Ig lambda chain V region 4A) | 0.49  | 0.04  |
| P01042_2 | KNG1  | Kininogen-1 (Alpha-2-thiol proteinase inhibitor) (Fitzgerald factor) (High molecular weight kininogen) (HMWK) (Williams-Fitzgerald-Flaujeac factor) [Cleaved into: Kininogen-1 heavy chain; T-kinin (Ile-Ser-Bradykinin); Bradykinin (Kallidin I); Lysyl-bradykinin (Kallidin II); Kininogen-1 light chain; Low molecular weight growth-promoting factor] | -0.28 | 0.04  |
| H0YGL9  |       | Deleted. | 0.82  | 0.04  |
| A0M8Q6  | IGLC7 | Immunoglobulin lambda constant 7 (Ig | 0.54  | 0.05  |
| P01617 | Immunoglobulin kappa variable 2D-28 (Ig kappa chain V-II region FR) (Ig kappa chain V-II region GM607) (Ig kappa chain V-II region MIL) (Ig kappa chain V-II region TEW) | 0.37 | 0.05 |
| P01620 | Immunoglobulin kappa variable 3-20 (Ig kappa chain V-III region B6) (Ig kappa chain V-III region GOL) (Ig kappa chain V-III region HAH) (Ig kappa chain V-III region HIC) (Ig kappa chain V-III region IARC/BL41) (Ig kappa chain V-III region NG9) (Ig kappa chain V-III region SIE) (Ig kappa chain V-III region Ti) (Ig kappa chain V-III region WOL) | 0.44 | 0.05 |
| P01702 | Immunoglobulin lambda variable 1-51 (Ig lambda chain V-I region BL2) (Ig lambda chain V-I region EPS) (Ig lambda chain V-I region NEW) (Ig lambda chain V-I region NIG-64) | 0.65 | 0.05 |
| Q6UXB8_2 | Peptidase inhibitor 16 (PI-16) (Cysteine-rich secretory protein 9) (CRISP-9) (PSP94-binding protein) (CD antigen CD364) | -0.22 | 0.05 |
| P01876 | Immunoglobulin heavy constant alpha 1 (Ig alpha-1 chain C region) (Ig alpha-1 chain C region BUR) (Ig alpha-1 chain C region TRO) | 0.51 | 0.05 |
| P04196 | Histidine-rich glycoprotein (Histidine-proline-rich glycoprotein) (HPRG) | -0.26 | 0.05 |
| P04433 | Immunoglobulin kappa variable 3-11 (Ig kappa chain V-III region VG) | 0.57 | 0.05 |
| P06310 | Immunoglobulin kappa variable 2-30 (Ig kappa chain V-II region RPMI 6410) | 0.52 | 0.05 |
| G8JLA8 | Deleted. | 0.11 | 0.07 |
| P01034 | Cystatin-C (Cystatin-3) (Gamma-trace) (Neuroendocrine basic polypeptide) (Post-gamma-globulin) | 0.16 | 0.07 |
| P01765 | Immunoglobulin heavy variable 3-23 (Ig heavy chain V-III region LAY) (Ig heavy chain V-III region POM) (Ig heavy chain V-III region TEI) (Ig heavy chain V-III region TIL) (Ig heavy chain V-III region TUR) (Ig heavy chain V-III region VH26) (Ig heavy chain V-III region WAS) (Ig heavy chain V-III region ZAP) | 0.49 | 0.07 |
| P04003 | C4BPA | C4b-binding protein alpha chain (C4bp) (Proline-rich protein) (PRP) | 0.10 | 0.07 |
| P01743 | Immunoglobulin heavy variable 1-46 (Ig heavy chain V-I region DOT) (Ig heavy chain V-I region HG3) (Ig heavy chain V-I region Mot) | 0.45 | 0.07 |
| P01031 | C5 | Complement C5 (C3 and PZP-like alpha-2-macroglobulin domain-containing protein 4) [Cleaved into: Complement C5 beta chain; Complement C5 alpha chain; C5a anaphylatoxin; Complement C5 alpha' chain] | 0.07 | 0.07 |
| P01764 | IGHV3-23 | Immunoglobulin heavy variable 3-23 (Ig heavy chain V-III region LAY) (Ig heavy chain V-III region POM) (Ig heavy chain V-III region TEI) (Ig heavy chain V-III region TIL) (Ig heavy chain V-III region TUR) (Ig heavy chain V-III region VH26) (Ig heavy chain V-III region WAS) (Ig heavy chain V-III region ZAP) | 0.49 | 0.07 |
| P01766 | IGHV3-13 | Immunoglobulin heavy variable 3-13 (Ig heavy chain V-III region BRO) | 0.41 | 0.07 |
| P01767 | IGHV3-53 | Immunoglobulin heavy variable 3-53 (Ig heavy chain V-III region BUT) | 0.31 | 0.07 |
| P02748 | C9 | Complement component C9 [Cleaved into: Complement component C9a; Complement component C9b] | 0.12 | 0.07 |
| Q06033_2 | | Inter-alpha-trypsin inhibitor heavy chain H3 (ITI heavy chain H3) (ITI-HC3) (Inter-alpha-inhibitor heavy chain 3) (Serum-derived hyaluronan-associated protein) (SHAP) | 0.13 | 0.07 |
| F5H6I0 | B2M | Beta-2-microglobulin | 0.14 | 0.07 |
| P01871 | IGHM | Immunoglobulin heavy constant mu (Ig mu chain C region) (Ig mu chain C region BOT) (Ig mu chain C region GAL) (Ig mu chain C region OU) | 0.44 | 0.07 |
| P08571 | CD14 | Monocyte differentiation antigen CD14 (Myeloid cell-specific leucine-rich glycoprotein) (CD antigen CD14) [Cleaved into: Monocyte differentiation antigen CD14, urinary form; Monocyte differentiation antigen CD14, membrane-bound form] | 0.13 | 0.08 |
| O00391_2 | | Sulfhydryl oxidase 1 (hQSOX) (EC 1.8.3.2) (Quiescin Q6) | 0.11 | 0.08 |
| P04206 | | Immunoglobulin kappa variable 3-20 (Ig kappa chain V-III region B6) (Ig kappa chain V-III region GOL) (Ig kappa chain V-III region HAH) (Ig kappa chain V-III region HIC) (Ig kappa chain V-III region IARC/GL4I) (Ig kappa chain V-III region NG9) (Ig kappa chain V-III region SIE) (Ig kappa chain V-III region Ti) (Ig kappa chain V-III region WOL) | 0.34 | 0.08 |
| P07225 | PROS1 | Vitamin K-dependent protein S | 0.04 | 0.08 |
| P29622 | SERPI4 | Kallistatin (Kallikrein inhibitor) (Peptidase inhibitor 4) (PI-4) (Serpin A4) | -0.16 | 0.08 |
| H0Y755 | FCGR3A | Low affinity immunoglobulin gamma Fc | 0.18 | 0.08 |
| Gene | Description                                                                 | Value 1 | Value 2 |
|------|-----------------------------------------------------------------------------|---------|---------|
| P10643 | C7 Complement component C7                                                  | 0.06    | 0.08    |
| P01859 | IGHG2 Immunoglobulin heavy constant gamma 2 (Ig gamma-2 chain C region)     | 0.36    | 0.09    |
|        | (Ig gamma-2 chain C region DOT)                                             |         |         |
|        | (Ig gamma-2 chain C region TIL)                                              |         |         |
|        | (Ig gamma-2 chain C region ZIE)                                              |         |         |
| P01625 | IGHG2 Immunoglobulin heavy constant gamma 3 (HDC) (Heavy chain disease protein) (Ig gamma-3 chain C region) | 0.35    | 0.11    |
| O75636 | FCN3 Ficolin-3 (Collagen/fibrinogen domain-containing lectin 3 p35) (Collagen/fibrinogen domain-containing protein 3) (Hakata antigen) | 0.14    | 0.09    |
| P25311 | AZGP1 Zinc-alpha-2-glycoprotein (Zn-alpha-2-GP) (Zn-alpha-2-glycoprotein)    | -0.14   | 0.09    |
| Q96PD5 | PGLYRP2 N-acetylmuramoyl-L-alanine amidase (EC 3.5.1.28) (Peptidoglycan recognition protein 2) (Peptidoglycan recognition protein long) (PGRP-L) | -0.14   | 0.10    |
| P01019 | AGT Angiotensinogen (Serpin A8) [Cleaved into: Angiotensin-1 (Angiotensin 1-10) (Angiotensin I) (Ang I); Angiotensin-2 (Angiotensin 1-8) (Angiotensin II) (Ang II); Angiotensin-3 (Angiotensin 2-8) (Angiotensin III) (Ang III) (Des-Asp[1]-angiotensin II); Angiotensin-4 (Angiotensin 3-8) (Angiotensin IV) (Ang IV); Angiotensin 1-9; Angiotensin 1-7; Angiotensin 1-5; Angiotensin 1-4] | 0.21    | 0.11    |
| P04040 | CAT Catalase (EC 1.11.1.6)                                                   | 0.18    | 0.11    |
| P01860 | IGHG3 Immunoglobulin heavy constant gamma 3 (HDC) (Heavy chain disease protein) (Ig gamma-3 chain C region) | 0.35    | 0.11    |
| P01611 | IGKV1D-12 Immunoglobulin kappa variable 1D-12 (Ig kappa chain V-I region Wes) | 0.30    | 0.11    |
| P07195 | LDHB L-lactate dehydrogenase B chain (LDH-B) (EC 1.1.1.27) (LDH heart subunit) (LDH-H) (Renal carcinoma antigen NY-REN-46) | 0.14    | 0.11    |
| Q9NZP8 | C1RL Complement C1r subcomponent-like protein (C1r-LP) (C1r-like protein) (EC 3.4.21.-) (C1r-like serine protease analog protein) (CLSPA) | 0.10    | 0.11    |
| P19823 | ITIH2 Inter-alpha-trypsin inhibitor heavy chain H2 (ITI heavy chain H2) (ITI-HC2) (Inter-alpha-inhibitor heavy chain 2) (Inter-alpha-trypsin inhibitor complex component II) (Serum-derived hyaluronan-associated protein) (SHAP) | -0.16   | 0.12    |
| P01624 | IGKV3-15 Immunoglobulin kappa variable 3-15 (Ig kappa chain V-III region CLL) (Ig kappa |

Page 33 of 42
| ID      | Name                                      | Description                                                                 | z-score | p-value |
|---------|-------------------------------------------|------------------------------------------------------------------------------|---------|---------|
| P02750  | LRG1                                      | Leucine-rich alpha-2-glycoprotein (LRG)                                      | 0.13    | 0.13    |
| P20851_2| C4b-binding protein beta chain             |                                                                              | 0.03    | 0.13    |
| P07360  | C8G                                       | Complement component C8 gamma chain                                          | 0.05    | 0.13    |
| F5GY80  | C8B                                       | Complement component C8 beta chain                                          | 0.04    | 0.13    |
| K7ER74  | APOC4-APOC2                               | APOC4-APOC2 readthrough (NMD candidate)                                      | 0.24    | 0.13    |
| P04406_2| Glyceraldehyde-3-phosphate dehydrogenase  | (GAPDH) (EC 1.2.1.12) (Peptidyl-cysteine S-nitrosylase GAPDH) (EC 2.6.99.)     | 0.20    | 0.14    |
| B4E1Z4  | cDNA FLJ55673, highly similar to Complement factor B |                                                                              | 0.03    | 0.15    |
| P01011  | SERP13                                    | Alpha-1-antichymotrypsin (ACT) (Cell growth-inhibiting gene 24/25 protein) (Serpin A3) [Cleaved into: Alpha-1-antichymotrypsin His-Pro-less] | 0.13    | 0.15    |
| D6RA08  | Deleted.                                  |                                                                              | 0.12    | 0.16    |
| P00915  | CA1                                       | Carbonic anhydrase 1 (EC 4.2.1.1) (Carbonic dehydratase I) (CA-I)            | 0.10    | 0.16    |
| P02747  | C1QC                                      | Complement C1q subcomponent subunit C                                        | 0.12    | 0.16    |
| H0Y612  | TRIM33                                    | E3 ubiquitin-protein ligase TRIM33 (Fragment)                                | 0.15    | 0.17    |
| P07359  | GP1BA                                     | Platelet glycoprotein Ib alpha chain (GP-Ib alpha) (GP1b-alpha) (GP1bA) (Glycoprotein Ibalpha) (Antigen CD42b-alpha) (CD antigen CD42b) [Cleaved into: Glycocalcin] | 0.06    | 0.17    |
| P20749  | APOH                                      | Beta-2-glycoprotein 1 (APC inhibitor) (Anticardiolipin cofactor) (Apo-L) (Beta-2-glycoprotein I) (B2GPI) (Beta2) | -0.20   | 0.18    |
| P22792  | CPN2                                      | Carboxypeptidase N subunit 2 (Carboxypeptidase N 83 kDa chain) (Carboxypeptidase N large subunit) (Carboxypeptidase N polypeptide 2) (Carboxypeptidase N regulatory subunit) | 0.04    | 0.19    |
| Q6EMK4  | VASN                                      | Vasorin (Protein slit-like 2)                                                | 0.06    | 0.19    |
| B7ZKJ8  | ITIH4                                     | ITIH4 protein (Inter-alpha-trypsin inhibitor heavy chain H4)                  | 0.11    | 0.19    |
| O14791_3| Apolipoprotein L1 (Apolipoprotein L) (Apo-L) (Apol-L) (Apolipoprotein L-I) (ApoL-I) |                                                                              | 0.08    | 0.19    |
| P02649  | APOE                                      | Apolipoprotein E (Apo-E)                                                     | 0.16    | 0.19    |
| P55056  | APOC4                                     | Apolipoprotein C-IV (Apo-CIV) (ApoC-IV) (Apolipoprotein C4)                  | 0.15    | 0.19    |
| P05546  | SERPIND                                   | Heparin cofactor 2 (Heparin cofactor II) (HC-II) (Protease inhibitor leuserpin-2) (HLS2) (Serpin D1) | -0.22   | 0.19    |
| P06276  | BCHE                                      | Cholinesterase (EC 3.1.1.8) (Acetylcholine)                                 | -0.19   | 0.21    |

The copyright holder for this preprint is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.
| P48740 | MASP1 | Mannan-binding lectin serine protease 1 (EC 3.4.21.-) (Complement factor MASP-3) (Complement-activating component of Ra-reactive factor) (Mannose-binding lectin-associated serine protease 1) (MASP-1) (Mannose-binding protein-associated serine protease) (Ra-reactive factor serine protease p100) (RaRF) (Serine protease 5) [Cleaved into: Mannan-binding lectin serine protease 1 heavy chain; Mannan-binding lectin serine protease 1 light chain] | 0.04 | 0.23 |
| P20742 | PZP | Pregnancy zone protein (C3 and PZP-like alpha-2-macroglobulin domain-containing protein 6) | 0.12 | 0.25 |
| P02766 | TTR | Transthyretin (ATTR) (Prealbumin) (TBPA) | -0.18 | 0.27 |
| P32119 | PRDX2 | Peroxiredoxin-2 (EC 1.11.1.15) (Natural killer cell-enhancing factor B) (NKEF-B) (PRP) (Thiol-specific antioxidant protein) (TSA) (Thioredoxin peroxidase 1) (Thioredoxin-dependent peroxide reductase 1) | 0.02 | 0.30 |
| P01833 | PIGR | Polymeric immunoglobulin receptor (PIgR) (Poly-Ig receptor) (Hepatocellular carcinoma-associated protein TB6) [Cleaved into: Secretory component] | 0.25 | 0.30 |
| P02745 | C1QA | Complement C1q subcomponent subunit A | 0.11 | 0.31 |
| P19320_3 | Vascular cell adhesion protein 1 (V-CAM 1) (VCAM-1) (INCAM-100) (CD antigen CD106) | 0.13 | 0.31 |
| P22352 | GPX3 | Glutathione peroxidase 3 (GPx-3) (GSHPx-3) (EC 1.11.1.9) (Extracellular glutathione peroxidase) (Plasma glutathione peroxidase) (GPx-P) (GSHPx-P) | -0.07 | 0.32 |
| Q08380 | LGALS3B | Galectin-3-binding protein (Basement membrane autoantigen p105) (Lectin galactoside-binding soluble 3-binding protein) (Mac-2-binding protein) (MAC2BP) (Mac-2 BP) (Tumor-associated antigen 90K) | 0.14 | 0.32 |
| P12259 | F5 | Coagulation factor V (Activated protein C cofactor) (Proaccelerin, labile factor) [Cleaved into: Coagulation factor V heavy chain; Coagulation factor V light chain] | 0.03 | 0.32 |
| P07900 | HSP90AA1 | Heat shock protein HSP 90-alpha (Heat shock 86 kDa) (HSP 86) (HSP86) (Lipopolysaccharide-associated protein 2) (LAP-2) (LPS-associated protein 2) (Renal carcinoma antigen NY-REN-38) | 0.14 | 0.34 |
| P05452 | CLEC3B | Tetratcin (TN) (C-type lectin domain family 3 member B) (Plasminogen kringle 4-
| Accession | Gene Symbol | Description                                                                                           | Log2 Fold Change | p-value |
|-----------|-------------|-------------------------------------------------------------------------------------------------------|-----------------|---------|
| P19827    | ITIH1       | Inter-alpha-trypsin inhibitor heavy chain H1 (ITI heavy chain H1) (ITI-HC1) (Inter-alpha-inhibitor heavy chain 1) (Inter-alpha-trypsin inhibitor complex component III) (Serum-derived hyaluronan-associated protein) (SHAP) | -0.13           | 0.36    |
| O43866    | CD5L        | CD5 antigen-like (Apoptosis inhibitor expressed by macrophages) (hAIM) (CT-2) (IgM-associated peptide) (SP-alpha) | 0.24            | 0.37    |
| P00918    | CA2         | Carbonic anhydrase 2 (EC 4.2.1.1) (Carbonate dehydratase II) (Carbonic anhydrase C) (CAC) (Carbonic anhydrase II) (CA-II) | -0.03           | 0.37    |
| P04075    | ALDOA       | Fructose-bisphosphate aldolase A (EC 4.1.2.13) (Lung cancer antigen NY-LU-1) (Muscle-type aldolase)       | 0.10            | 0.39    |
| P07357    | C8A         | Complement component C8 alpha chain (Complement component 8 subunit alpha)                           | -0.01           | 0.39    |
| H0YD13    | CD44        | CD44 antigen                                                                                            | -0.01           | 0.41    |
| P02751_10 |             | Fibronectin (FN) (Cold-insoluble globulin) (CIG) [Cleaved into: Anastellin; Ugl-Y1; Ugl-Y2; Ugl-Y3]     |                 |         |
| P01877    | IGHA2       | Immunoglobulin heavy constant alpha 2 (Ig alpha-2 chain C region) (Ig alpha-2 chain C region BUT) (Ig alpha-2 chain C region LAN) | 0.32            | 0.41    |
| P01042    | KNG1        | Kininogen-1 (Alpha-2-thiol proteinase inhibitor) (Fitzgerald factor) (High molecular weight kininogen) (HMWK) (Williams-Fitzgerald-Flaujeac factor) [Cleaved into: Kininogen-1 heavy chain; T-kinin (Ile-Ser-Bradykinin); Bradykinin (Kallidin I); Lysyl-bradykinin (Kallidin II); Kininogen-1 light chain; Low molecular weight growth-promoting factor] | -0.15           | 0.42    |
| B0YI2W2   | APOC3       | Apolipoprotein C-III (Apolipoprotein C-III variant 1)                                                  | 0.17            | 0.43    |
| E7EPZ9    |             | Deleted.                                                                                               | 0.05            | 0.43    |
| P04114    | APOB        | Apolipoprotein B-100 (Apo B-100) [Cleaved into: Apolipoprotein B-48 (Apo B-48)]                         | 0.06            | 0.43    |
| P51884    | LUM         | Lumican (Keratan sulfate proteoglycan lumican) (KSPG lumican)                                          | -0.14           | 0.43    |
| Q5VY30    | RBP4        | Retinol-binding protein                                                                                 | -0.11           | 0.43    |
| Q9Y5Y7    | LYVE1       | Lymphatic vessel endothelial hyaluronic acid receptor 1 (LYVE-1) (Cell surface retention sequence-binding protein 1) (CRSBP-1) (Extracellular link domain-containing protein 1) (Hyaluronic acid receptor) | 0.14            | 0.44    |
| P04264    | KRT1        | Keratin, type II cytoskeletal 1 (67 kDa cytokeratin) (Cytokeratin-1) (CK-1) (Hair)                     | 0.18            | 0.44    |
| Accession | Gene Symbol | Description | Score 1 | Score 2 |
|-----------|-------------|-------------|---------|---------|
| P04004    | VTN         | Vitronectin (VN) (S-protein) (Serum-spreading factor) (V75) [Cleaved into: Vitronectin V65 subunit; Vitronectin V10 subunit; Somatomedin-B] | 0.01    | 0.46    |
| C9JF17    | APOD        | Apolipoprotein D (Fragment) | -0.12   | 0.46    |
| Q16610    | ECM1        | Extracellular matrix protein 1 (Secretory component p85) | 0.03    | 0.46    |
| B1AKG0    | CFHR1       | Complement factor H-related protein 1 | 0.08    | 0.47    |
| P02679_2  | Fibrinogen gamma chain | 0.16 | 0.48    |
| P01591    | JCHAIN      | Immunoglobulin J chain (Joining chain of multimeric IgA and IgM) | 0.21    | 0.49    |
| P35542    | SAA4        | Serum amyloid A-4 protein (Constitutively expressed serum amyloid A protein) (C-SAA) | -0.07   | 0.50    |
| P02675    | FGB         | Fibrinogen beta chain [Cleaved into: Fibrinopeptide B; Fibrinogen beta chain] | 0.15    | 0.51    |
| G3V2W1    | SERP10      | Protein Z-dependent protease inhibitor (Serpin peptidase inhibitor, clade A (Alpha-1 antiproteinase, antitrypsin), member 10, isoform CRA_a) | 0.03    | 0.51    |
| P01008    | SERPINC1    | Antithrombin-III (ATIII) (Serpin C1) | -0.11   | 0.51    |
| P60709    | ACTB        | Actin, cytoplasmic 1 (Beta-actin) [Cleaved into: Actin, cytoplasmic 1, N-terminally processed] | 0.07    | 0.53    |
| P15151_3  | Poliovirus receptor (Nectin-like protein 5) (NECL-5) (CD antigen CD155) | 0.02    | 0.54    |
| P61626    | LYZ         | Lysozyme C (EC 3.2.1.17) (1,4-beta-N-acetylmuramidase C) | 0.11    | 0.54    |
| V9GYE7    | CFHR2       | Complement factor H-related protein 2 | 0.02    | 0.54    |
| G3XAM2    | CFI         | Complement factor I (Complement factor I, isoform CRA_b) | -0.03   | 0.55    |
| P69905    | HBA1        | Hemoglobin subunit alpha (Alpha-globin) (Hemoglobin alpha chain) | 0.00    | 0.55    |
| P69905    | HBA2        | Hemoglobin subunit alpha (Alpha-globin) (Hemoglobin alpha chain) | 0.00    | 0.55    |
| O95445    | APOM        | Apolipoprotein M (Apo-M) (ApoM) (Protein G3a) | -0.08   | 0.55    |
| P02790    | HPX         | Hemopexin (Beta-1B-glycoprotein) | -0.10   | 0.56    |
| P68871    | HBB         | Hemoglobin subunit beta (Beta-globin) (Hemoglobin beta chain) [Cleaved into: LVV-hemorphin-7; Spinorphin] | 0.01    | 0.60    |
| P00747    | PLG         | Plasminogen (EC 3.4.21.7) [Cleaved into: Plasmin heavy chain A; Activation peptide; Angiotatin; Plasmin heavy chain A, short form; Plasmin light chain B] | -0.09   | 0.61    |
| P00748    | F12         | Coagulation factor XII (EC 3.4.21.38) (Hageman factor) (HAF) [Cleaved into: | 0.02    | 0.61    |
| PDB ID | Description |
|--------|-------------|
| Q92954_4 | Coagulation factor XIIa heavy chain; Beta-factor XIIa part 1; Coagulation factor XIIa light chain (Beta-factor XIIa part 2) |
| P01023 | Alpha-2-macroglobulin (Alpha-2-M) (C3 and PZP-like alpha-2-macroglobulin domain-containing protein 5) |
| P06396 | Gelsolin (AGEL) (Actin-depolymerizing factor) (ADF) (Brevin) |
| P0DJ18 | Serum amyloid A-1 protein (SAA) [Cleaved into: Amyloid protein A (Amyloid fibril protein AA); Serum amyloid protein A(2-104); Serum amyloid protein A(3-104); Serum amyloid protein A(2-103); Serum amyloid protein A(2-102); Serum amyloid protein A(4-101)] |
| P22891 | Vitamin K-dependent protein Z |
| P27169 | Serum paraoxonase/arylesterase 1 (PON 1) (EC 3.1.1.2) (EC 3.1.1.81) (EC 3.1.8.1) (Aromatic esterase 1) (A-esterase 1) (K-45) (Serum aryldialkylphosphatase 1) |
| D6RF35 | Vitamin D-binding protein |
| K7ERG9 | Complement factor D |
| P00742 | Coagulation factor X (EC 3.4.21.6) (Stuart factor) (Stuart-Prower factor) [Cleaved into: Factor X light chain; Factor X heavy chain; Activated factor Xa heavy chain] |
| P01024 | Complement C3 (C3 and PZP-like alpha-2-macroglobulin domain-containing protein 1) [Cleaved into: Complement C3 beta chain; C3-beta-c (C3bc); Complement C3 alpha chain; C3a anaphylatoxin; Acylation stimulating protein (ASP) (C3adesArg); Complement C3b alpha' chain; Complement C3c alpha' chain fragment 1; Complement C3d fragment; Complement C3g fragment; Complement C3d fragment; Complement C3f fragment; Complement C3c alpha' chain fragment 2] |
| P01861 | Immunoglobulin heavy constant gamma 4 (Ig gamma-4 chain C region) |
| P02042 | Hemoglobin subunit delta (Delta-globin) (Hemoglobin delta chain) |
| P02760 | Protein AMBP [Cleaved into: Alpha-1-microglobulin (Protein HC) (Alpha-1 microglycoprotein) (Complex-forming glycoprotein heterogeneous in charge); Inter- |
| P02765 | AHSG | Alpha-2-HS-glycoprotein (Alpha-2-Z-globulin) (Ba-alpha-2-glycoprotein) (Fetuin-A) [Cleaved into: Alpha-2-HS-glycoprotein chain A; Alpha-2-HS-glycoprotein chain B] | -0.07 | 0.72 |
|--------|------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------|--------|
| P08185 | SERPI6 | Corticosteroid-binding globulin (CBG) (Serpin A6) (Transcortin) | 0.00 | 0.72 |
| P0C0L5 | C4B | Complement C4-B (Basic complement C4) (C3 and PZP-like alpha-2-macroglobulin domain-containing protein 3) [Cleaved into: Complement C4 beta chain; Complement C4-B alpha chain; C4a anaphylatoxin; C4b-B; C4d-B; Complement C4 gamma chain] | -0.01 | 0.72 |
| P0C0L5 | C4B_2 | Complement C4-B (Basic complement C4) (C3 and PZP-like alpha-2-macroglobulin domain-containing protein 3) [Cleaved into: Complement C4 beta chain; Complement C4-B alpha chain; C4a anaphylatoxin; C4b-B; C4d-B; Complement C4 gamma chain] | -0.01 | 0.72 |
| Q9UGM5 | FETUB | Fetuin-B (16G2) (Fetu-in-like protein IRL685) (Gugu) | -0.12 | 0.72 |
| B4E2S7 | cDNA FLJ58780, highly similar to Homo sapiens lysosomal-associated membrane protein 2 (LAMP2), transcript variant LAMP2B, mRNA | 0.00 | 0.74 |
| E9PFZ2 | CP | Ceruloplasmin | 0.11 | 0.74 |
| F5GXS5 | Deleted. | 0.04 | 0.74 |
| P14151 | SELL | L-selectin (CD62 antigen-like family member L) (Leukocyte adhesion molecule 1) (LAM-1) (Leukocyte surface antigen Leu-8) (Leukocyte-endothelial cell adhesion molecule 1) (LECAM1) (Lymph node homing receptor) (TQ1) (gp90-MEL) (CD antigen CD62L) | 0.01 | 0.74 |
| P36955 | SERPINF1 | Pigment epithelium-derived factor (PEDF) (Cell proliferation-inducing gene 35 protein) (EPC-1) (Serpin F1) | 0.00 | 0.74 |
| Q04756 | HGFAC | Hepatocyte growth factor activator (HGF activator) (HGFA) (EC 3.4.21.-) [Cleaved into: Hepatocyte growth factor activator short chain; Hepatocyte growth factor activator long chain] | -0.02 | 0.74 |
| P05543 | SERPI7 | Thyroxine-binding globulin (Serpin A7) (T4-binding globulin) | -0.08 | 0.77 |
| F8W1Q3 | Biotinidase (Biotinase) (EC 3.5.1.12) | -0.01 | 0.78 |
| O75882_3 | Attractin (DPPT-L) (Mahogany homolog) | -0.02 | 0.78 |
| P05155 | SERPING | Plasma protease C1 inhibitor (C1 Inh) | 0.02 | 0.78 |
| Name      | Accession | Description                                                                 | EA  | PA  |
|-----------|-----------|-----------------------------------------------------------------------------|-----|-----|
| Q961Y4    | CPB2      | Carboxypeptidase B2 (EC 3.4.17.20) (Carboxypeptidase U) (CPU) (Plasma carboxypeptidase B) (pCPB) (Thrombin-activable fibrinolysis inhibitor) (TAFl) | -0.04 | 0.78 |
| P07996    | THBS1     | Thrombospondin-1 (Glycoprotein G)                                           | -0.08 | 0.79 |
| P08603    | CFH       | Complement factor H (H factor 1)                                            | 0.01  | 0.79 |
| P10909_4  |           | Clusterin (Aging-associated gene 4 protein) (Apolipoprotein J) (Apo-J) (Complement cytolysis inhibitor) (CLI) (Complement-associated protein SP-40,40) (Ku70-binding protein 1) (NA1/NA2) (Sulfated glycoprotein 2) (SGP-2) (Testosterone-repressed prostate message 2) (TRPM-2) [Cleaved into: Clusterin beta chain (ApoJalpha) (Complement cytolysis inhibitor a chain); Clusterin alpha chain (ApoJbeta) (Complement cytolysis inhibitor b chain)] | -0.04 | 0.82 |
| I3L145    | SHBG      | Sex hormone-binding globulin (Sex hormone-binding globulin, isoform CRA_a)   | -0.08 | 0.83 |
| K7ERI9    | APOC1     | Apolipoprotein C-I (Fragment)                                               | 0.09  | 0.83 |
| P00736    | C1R       | Complement C1r subcomponent (EC 3.4.21.41) (Complement component 1 subcomponent r) [Cleaved into: Complement C1r subcomponent heavy chain; Complement C1r subcomponent light chain] | -0.05 | 0.83 |
| P02775    | PPBP      | Platelet basic protein (PBP) (C-X-C motif chemokine 7) (Leukocyte-derived growth factor) (LDGF) (Macrophage-derived growth factor) (MDGF) (Small-inducible cytokine B7) [Cleaved into: Connective tissue-activating peptide III (CTAP-III) (LA-PF4) (Low-affinity platelet factor IV); TC-2; Connective tissue-activating peptide III(1-81) (CTAP-III(1-81)); Beta-thromboglobulin (Beta-TG); Neutrophil-activating peptide 2(74) (NAP-2(74)); Neutrophil-activating peptide 2(73) (NAP-2(73)); Neutrophil-activating peptide 2 (NAP-2); TC-1; Neutrophil-activating peptide 2(1-66) (NAP-2(1-66)); Neutrophil-activating peptide 2(1-63) (NAP-2(1-63))] | -0.08 | 0.83 |
| Q13103    | SPP2      | Secreted phosphoprotein 24 (Spp-24) (Secreted phosphoprotein 2)             | -0.02 | 0.83 |
| P0C0L4    | C4A       | Complement C4-A (Acidic complement C4) (C3 and PZP-like alpha-2-macroglobulin domain-containing protein 2) [Cleaved into: Complement C4 beta chain; Complement C4-A alpha chain; C4a anaphylatoxin; C4b-A; | 0.02  | 0.84 |
| ID    | Gene   | Description                                                                 |
|-------|--------|-----------------------------------------------------------------------------|
| E9PGP2|        | C4d-A; Complement C4 gamma chain                                            |
| P15169| CPN1   | Carboxypeptidase N catalytic chain (CPN) (EC 3.4.17.3) (Anaphylatoxin inactivator) (Arginine carboxypeptidase) (Carboxypeptidase N polypeptide 1) (Carboxypeptidase N small subunit) (Kininase-1) (Lysine carboxypeptidase) (Plasma carboxypeptidase B) (Serum carboxypeptidase N) (SCPN) |
| P08697| SERPINF2| Alpha-2-antiplasmin (Alpha-2-AP) (Alpha-2-plasmin inhibitor) (Alpha-2-Pi) (Serpin F2) |
| H0YAC1| KLKB1  | Plasma kallikrein (Fragment)                                                |
| P00740_2|        | Coagulation factor IX (EC 3.4.21.22) (Christmas factor) (Plasma thromboplastin component) (PTC) [Cleaved into: Coagulation factor IXa light chain; Coagulation factor IXa heavy chain] |
| P04180| LCAT   | Phosphatidylcholine-sterol acyltransferase (EC 2.3.1.43) (Lecithin-cholesterol acyltransferase) (Phospholipid-cholesterol acyltransferase) |
| P00739| HPR    | Haptoglobin-related protein                                                 |
| P02743| APCS   | Serum amyloid P-component (SAP) (9.5S alpha-1-glycoprotein) [Cleaved into: Serum amyloid P-component(1-203)] |
| P09871| C1S    | Complement C1s subcomponent (EC 3.4.21.42) (C1 esterase) (Complement component 1 subcomponent s) [Cleaved into: Complement C1s subcomponent heavy chain; Complement C1s subcomponent light chain] |
| P00450| CP     | Ceruloplasmin (EC 1.16.3.1) (Ferrooxidase)                                  |
| P00734| F2     | Prothrombin (EC 3.4.21.5) (Coagulation factor II) [Cleaved into: Activation peptide fragment 1; Activation peptide fragment 2; Thrombin light chain; Thrombin heavy chain] |
| P02774| GC     | Vitamin D-binding protein (DBP) (VDB) (Gc protein-derived macrophage activating factor) (Gc-MAF) (GcMAF) (Gc-globulin) (Group-specific component) (Gc) (Vitamin D-binding protein-macrophage activating factor) (DBP-maf) |
| P04217| A1BG   | Alpha-1B-glycoprotein (Alpha-1-B glycoprotein)                              |
| P06727| APOA4  | Apolipoprotein A-IV (Apo-AIV) (ApoA-IV) (Apolipoprotein A4)                |
| P13671| C6     | Complement component C6                                                     |
| P43652| AFM    | Afamin (Alpha-albumin) (Alpha-Alb)                                          |
| P69891| HBG1   | Hemoglobin subunit gamma-1 (Gamma-1)                                        |
globin) (Hb F Agamma) (Hemoglobin gamma-1 chain) (Hemoglobin gamma-A chain)

*log fold change - positive value for log fold change indicate higher enrichment in HIV(+);
false-discovery rate adjusted after logistic regression using inverse probability weighting accounting for recruitment site, sex, age, HIV status, mid-upper arm circumference, and presence of oedema