

1 Plasma proteomics reveals markers of metabolic stress in HIV infected children
2 with severe acute malnutrition

3
4 Gerard Bryan Gonzales^{1,2*}, James M Njunge^{3,4}, Bonface M Gichuki^{3,4}, Bijun Wen^{5,6}, Isabel
5 Potani³, Wieger Voskuil^{3,7,8}, Robert H J Bandsma^{3,5,6}, James A Berkley^{3,4,9}

6 1. Department of Gastroenterology, Faculty of Medicine and Health Sciences, Ghent
7 University, Ghent, Belgium

8 2. VIB Inflammation Research Centre, Ghent, Belgium

9 3. The Childhood Acute Illness & Nutrition (CHAIN) Network, Nairobi, Kenya

10 4. KEMRI/Wellcome Trust Research Programme, Kilifi, Kenya

11 5. Centre for Global Child Health, The Hospital for Sick Children, Toronto, Ontario,
12 Canada

13 6. Department of Nutritional Sciences, Faculty of Medicine, University of Toronto,
14 Toronto, Canada

15 7. Global Child Health Group, Emma Children's Hospital, Amsterdam University
16 Medical Centres, Amsterdam, the Netherlands

17 8. Amsterdam Institute for Global Health and Development, Department of Global
18 Health, Amsterdam University Medical Centres, Amsterdam, The Netherlands

19 9. Centre for Tropical Medicine & Global Health, Nuffield Department of Medicine,
20 University of Oxford, Oxford, UK

21

22 *Correspondence to:

23 Gerard Bryan Gonzales, PhD

24 Email: Gerard.gonzales@ugent.be

25 **Abstract (word count = 195)**

26

27 HIV infection affects up to 30% of children presenting with severe acute malnutrition (SAM)
28 in Africa and is associated with increased mortality. Children with SAM are treated similarly
29 regardless of HIV status, although mechanisms of nutritional recovery in HIV and/or SAM
30 are not well understood. We performed a secondary analysis of a clinical trial and plasma
31 proteomics data among children with complicated SAM in Kenya and Malawi. Compared to
32 children with SAM without HIV (n = 113), HIV-infected children (n = 54) had evidence
33 (false discovery rate (FDR) corrected $p < 0.05$) of metabolic stress, including enriched
34 pathways related to inflammation and lipid metabolism. Moreover, we observed reduced
35 plasma levels of zinc- α -2-glycoprotein, butyrylcholinesterase, and increased levels of
36 complement C2 resembling findings in metabolic syndrome, diabetes and other non-
37 communicable diseases. HIV was also associated (FDR corrected $p < 0.05$) with higher plasma
38 levels of inflammatory chemokines. Considering evidence of biomarkers of metabolic stress,
39 it is of potential concern that our current treatment strategy for SAM regardless of HIV status
40 involves a high-fat therapeutic diet. The results of this study suggest a need for clinical trials
41 of therapeutic foods that meet the specific metabolic needs of children with HIV and SAM.

42

43 **Keywords:** HIV, SAM, lipid metabolism, molecular nutrition

44 **Introduction**

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

54

55 Infection with the human immunodeficiency virus (HIV) is a common co-morbidity of SAM
56 in sub-Saharan Africa affecting up to 30% of admissions among SAM cases ⁹. HIV-infected
57 or exposed children are significantly more likely to be stunted, wasted, and underweight ¹⁰.
58 They also more often present with other clinical complications and greater susceptibility to
59 infections, thus further complicating their clinical management, which may include providing
60 more aggressive antimicrobial therapy and higher caloric nutritional intervention ¹¹.
61 Moreover, response to clinical management is also less predictable and less well-understood
62 in HIV-infected children compared to their uninfected counterparts ¹². Although acute
63 opportunistic infections play a key role in the outcome of these children, intestinal pathology
64 including inflammation and malabsorption, and metabolic perturbations may also be present.
65 However, mechanisms driving poor nutritional recovery of children with HIV even when
66 detected co-morbidities are treated remain poorly understood ¹².

67

68 We hypothesised that inflammatory, metabolic and other pathways which are likely to be
69 involved in the response to infection, survival and nutritional recovery differ between children
70 with SAM with and without HIV. We conducted a secondary analysis of clinical data and
71 biological samples from a randomised clinical trial in Kenya and Malawi ¹³.

72

73 **Results**

74 **Patient characteristics**

75 Table 1 presents the baseline characteristics of the children in the randomised trial. A total of
76 843 complicated SAM children were recruited for the randomised trial, of which 179 (22%)
77 patients were HIV(+). Age was higher and MUAC was lower in HIV(+) children than HIV(-)
78 counterparts. Most HIV cases were found in Malawi. Sex and the presence of oedema were
79 not associated with HIV status. Mortality was more than 2 times higher among in HIV(+)
80 compared to HIV(-) ($p < 0.001$). Children whose HIV status were unknown had the highest
81 mortality of 34%, which indicates bias due to frequent death before testing could be
82 undertaken or refusal of testing when a child was more severely ill.

83

84 Among HIV(+), 33% were already receiving an anti-retroviral treatment (ART) regime:
85 53/179 (30%) on highly active antiretroviral therapy (HAART), and 7/179 (4%) on
86 Nevirapine only. About half of the children (90/179) were naïve for ART whereas HIV
87 treatment status was unknown for 16% (29/179). Mortality was not significantly different
88 among children on HAART, ART naïve and children with unknown HIV treatment status
89 (Supplementary Table 1).

90

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

93 Among the children included in the proteomics study, 54 were HIV (+) and 113 were HIV(-)
94 (Table 2). In this sub-population, age, sex and the presence of oedema were not significantly
95 associated with HIV. HIV(+) children also had significantly lower MUAC and higher
96 mortality than HIV(-) children.

97

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

117

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

126

127 The weighted EN model extracted 73 circulating proteins (Figure 2A) that are associated with
128 HIV status with AUROC = 0.97 [95% CI: 0.95 – 0.99] (Figure 2B) and misclassification error
129 rate of 2.4%. Optimism-adjusted validated AUROC after bootstrapping was 0.90 [95% CI:
130 0.90 – 0.902], indicating a robust model. Pathway enrichment analysis highlighted that apart
131 from increased immune response, HIV(+) children with SAM had increased proteolysis and
132 lipid mobilisation, specifically increased very low-density lipoprotein assembly, indicating
133 metabolic dysregulation related to cholesterol and triglyceride metabolism among HIV(+)
134 patients (Figure 2D).

135

136 After 2000 bootstrap iterations during bootstrap validation, 3 proteins were consistently
137 extracted by the EN model >80% of the time (Figure 2D), namely: butyrylcholinesterase
138 (BChE), complement C2 and zinc- α -2-glycoprotein (ZAG), indicating that these 3 proteins
139 are likely to be the most important features associated with HIV in children with complicated
140 SAM. Weighted logistic regression model of these 3 proteins showed good discrimination of
141 HIV status (AUROC = 0.80 [95% CI: 0.74 – 0.87]) (Figure 2E).

142

143 **Discussion**

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

151

152 An earlier metabolomics study in Uganda reported reduced serum levels of adiponectin and
153 leptin, whereas serum triglycerides, ketones and even-chain acylcarnitines were higher in
154 HIV(+) children with SAM indicating perturbed lipid metabolism¹⁴. Our current study
155 therefore concurs with this finding, as we also found reduced plasma levels of adiponectin in
156 HIV(+) SAM children compared to HIV(-) SAM children, along with upregulation of
157 pathways involved in lipid transport and metabolism, specifically very low-density lipoprotein
158 assembly.

159

160 Using optimism-adjusted bootstrap validation of the EN model, we found 3 proteins:
161 complement c2, BChE and ZAG robustly distinguished HIV(+) from HIV(-) in children with
162 SAM, demonstrating the ability of multivariate analysis techniques, such as EN, to uncover
163 underlying relationships between protein markers which would be difficult to identify when
164 analysed individually. The activation of the complement system during HIV infection has
165 been previously discussed at length, which is associated with the increased cellular invasion
166 of HIV in cells¹⁵⁻¹⁷.

167

168 On the other hand, BChE is a protein synthesized in the liver and abundant in plasma, which
169 hydrolyses acetylcholine. Although very similar to its sister protein, acetylcholinesterase,
170 biological functions of BChE appear to be more varied but less understood ¹⁸. In a recent
171 study in China, low circulating BChE was found to be highly associated with HIV severity,
172 was predictive of mortality in adults, and was proposed as a plausible strategy for severity
173 classification among adults with HIV ¹⁹. BChE is also reported to be reduced in SAM, stress
174 and inflammation ²⁰. In animal studies, BChE deficiency was found to strongly affect fat
175 metabolism and promotes hepatic lipid accumulation ²¹. Serum BChE levels have been found
176 to have a significant negative correlation with serum total cholesterol and serum low-density-
177 lipoprotein cholesterol among people with diabetes mellitus ²².

178

179 ZAG is a newly characterized adipokine that is involved in lipolysis, body weight regulation
180 and may also be involved in the development of insulin resistance ²³. Reduction in plasma
181 levels of ZAG was previously reported to be implicated in dyslipidaemia in HIV(+) adults
182 under ART treatment ²³. Reduced circulating levels of ZAG has also been found among adults
183 with clinically diagnosed metabolic syndrome, based on guidelines of the United States
184 National Cholesterol Education Program (NCEP) Expert Panel Adult Treatment Panel (ATP)
185 III criteria ²⁴. Serum ZAG levels have been reported lower among adults with impaired
186 glucose tolerance and type 2 diabetes mellitus ²⁵. Taken together, our results therefore suggest
187 that children with both HIV and SAM manifest hallmarks of metabolic stress similar to those
188 occurring in metabolic syndrome and other non-communicable diseases (NCD).

189

190 This study is the first proteomics investigation on the interaction between HIV and SAM. In
191 summary, our results, which together with the previously published metabolomics study ¹⁴,
192 strengthens evidence on the increased metabolic stress and altered metabolic response among

193 children living with both HIV and SAM. Our results also concur with previous studies that
194 reported elevated metabolic stress among non-malnourished adults living with HIV leading to
195 increased prevalence or risk for metabolic syndrome, cardiovascular diseases, diabetes and
196 other non-communicable diseases ²⁶⁻³².

197

198 Metabolic abnormalities have previously been reported to be attributed HAART use among
199 HIV(+) patients ³³. In a recent systematic review, use of 2 classes of HAART, protease
200 inhibitors and nonnucleoside reverse transcriptase inhibitors, has been found to be associated
201 with abnormalities in plasma lipid profiles ³⁴. However, dysregulation in lipid metabolism has
202 also been reported in HAART-naïve patients, which indicates that HIV infection alone cause
203 lipid metabolism perturbations. An earlier longitudinal study of 50 men in the USA reported
204 notable declines in serum total cholesterol after HIV infection compared to results of blood
205 analysis from last seronegative visit. Large increases in total cholesterol and low-density
206 lipoproteins (LDL) were detected after HAART initiation ³⁵. However, many other studies
207 reported increases in total cholesterol among HIV-infected patients naïve to HAART. For
208 instance, in a study of ART-naïve HIV-infected adults in Ethiopia, malnutrition and lipid
209 abnormalities (specifically total cholesterol) were associated with CD4+ T cell counts ³⁶. In in
210 vitro studies, transfection of a T-cell (RH9) with HIV led to the enhanced production of free
211 fatty acids and LDL ³⁷. Furthermore, monocytes isolated from HIV-infected patients both
212 taking HAART and HAART-naïve, were found to have altered expression patterns of receptors
213 linked with lipid metabolism (i.e. FXR, PXR, PPAR α , GR, RAR α and RXR) compared to
214 monocytes of HIV-uninfected controls ³⁸. For our study however, we are unable to ascertain
215 whether the lipid metabolism dysregulation we observed is due primarily on the viral load
216 itself or the use of HAART due to lack of power for this sub-analysis. Majority of the
217 participants subjected to proteomics analysis were HAART-naïve (50%), where 26% were on

218 HAART, 6% were on Nevirapine alone and we had no data on treatment of 18% of the
219 patients (Table 2). In all these studies cited, authors argue to need for monitoring of lipid
220 profiles in HIV-infected populations. Hence, lipid monitoring may also inform nutritional and
221 clinical recovery of children with SAM and HIV and could be implemented to improve
222 clinical care for these children.

223

224 However, despite our knowledge that HIV-infected populations have altered metabolic
225 requirements compared to HIV-uninfected counterparts, WHO guidelines for the nutritional
226 management for SAM are globally the same regardless of HIV status, which is summarized in
227 Table 3³⁹. Nutritional management for in-patient children with SAM involves provision of a
228 low-protein, low-fat milk-based food, F75, every three hours. F75 is used during clinical
229 stabilization occurring during the first few days after admission and is not intended for weight
230 gain. Once the children are clinically stabilized and are able to tolerate the milk/solute load,
231 children are transitioned to F100, a higher-calorie, high-fat milk intended to boost weight gain
232 or to Ready-to-Use Therapeutic Food (RUTF), a peanut-based calorie-dense diet. Upon
233 discharge from in-patient care, children are referred to community based nutritional
234 therapeutic centres where they are provided with RUTF on a 2 weekly basis.

235

236 Considering evidence of biomarkers of metabolic syndrome and NCD in HIV(+) children
237 with SAM, it is of potential concern that our current treatment strategy involves a high-fat
238 therapeutic diet. About 50% of much needed calories during the growth catch-up phase are
239 supplied as lipids, which HIV(+) children may not be able to efficiently assimilate.
240 Alterations in lipid metabolism in HIV(+) children with SAM may also mean that the high
241 amounts of dietary lipids could be deposited as ectopic fat in the liver and muscle,
242 predisposing to insulin resistance, diabetes, cardiovascular problems and other NCDs later in

243 life. Although long-term metabolic follow-up studies could be done for HIV(+) children
244 previously treated for either complicated and uncomplicated SAM, significant barriers are the
245 high mortality rate in earlier studies of HIV(+) children with SAM, cost and difficulty tracing
246 them years later. The results of this study indicate a need for clinical trials of F100 or RUTF
247 modified to meet the expected metabolic needs of HIV(+) children with SAM. This could
248 initially be done in relatively small groups with outcomes that include measuring metabolic
249 stress.

250

251 Several studies on nutritional intervention strategies among HIV-infected adults have been
252 reported. For instance, a study in the USA showed that dietary fat intake, specifically
253 saturated fats, was significantly associated with hypertriglyceridemia among HIV-infected
254 adults (18 – 60 years)⁴⁰. Moreover, in a preclinical model, high saturated fat consumption
255 was found to accelerate immunodeficiency virus disease progression in macaques, specifically
256 increased mortality hazard and circulating levels of pro-inflammatory cytokines, especially
257 IL8⁴¹, which has been previously reported to be associated with lipodystrophy among HIV
258 patients⁴². In our study, we also found a significant association between high plasma IL8
259 concentration and HIV in SAM children. Hence, modifying the saturated fat composition of
260 the milk-based F75 and F100 could potentially lower metabolic stress.

261

262 The European Society for Parenteral and Enteral Nutrition (ESPEN) have given a grade A
263 recommendation for the use of medium-chain triglyceride (MCT)-based diet on HIV(+)
264 patients with diarrhoea and severe undernutrition in its 2006 ESPEN Guidelines on Enteral
265 Nutrition⁴³. Grade A recommendations are given to strategies based on meta-analysis or at
266 least one randomised control trial. In this case, the recommendation was based on a
267 prospective, randomized double-blind comparative trial on 24 adult patients with HIV and

268 diarrhoea of more than 4-wk duration, fat malabsorption, and loss of 10–20% of ideal body
269 weight⁴⁴. In this study, the authors found improved outcomes from diarrhoea and fat
270 malabsorption from MCT than long-chain triglyceride-based diet among HIV(+) adults.

271

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

277

278 Limitations of this study include absence of data on viral load and CD4+ counts of the
279 patients, which could provide a deeper understanding of the results. Furthermore, in this
280 study, we did not find association between oedematous malnutrition and HIV status, although
281 several studies have found higher HIV prevalence among non-oedematous children with
282 SAM⁴⁵⁻⁴⁷. In our study however, we found high in-patient mortality rate (16/46, 34%) among
283 children with unknown HIV status, where 39/46 (85%) had non-oedematous SAM.
284 Considering the high rate of mortality, these children may have been HIV(+). This highlights
285 the need for earlier HIV screening among children with SAM.

286

287 **Conclusion**

288 Plasma proteomics reveals that HIV(+) children with SAM manifest hallmarks of metabolic
289 stress similar to those observed in non-communicable diseases. This could be related to the
290 poor nutritional recovery and high mortality of HIV(+) children with SAM despite clinical
291 and nutritional intervention. The results of this study indicate a need for clinical trials
292 modifying the composition of F100 or RUTF to meet the specific metabolic needs of HIV(+)

293 children with SAM during rehabilitation phase. This could initially be done in relatively small
294 groups with outcomes that include measuring metabolic stress.

295

296 **Methods**

297 **Patient recruitment and study design**

298 This is a secondary analysis of a nested case control study from a randomised controlled trial
299 (NCT02246296), which tested the effect of a lactose-free, low-carbohydrate F75 milk to limit
300 carbohydrate malabsorption, diarrhoea and refeeding syndrome among children hospitalized
301 for complicated SAM at Queen Elizabeth Central Hospital in Blantyre, Malawi, Kilifi County
302 Hospital and Coast General Hospital, Mombasa, Kenya ¹³. Children aged 6 months to 13
303 years were eligible for enrolment into the trial at admission to hospital if they had SAM,
304 defined as: mid-upper arm circumference (MUAC) <11.5cm or weight-for-height Z score <-3
305 if younger than 5 years of age, BMI Z score <-3 if older than 5 years, or oedematous
306 malnutrition at any age and had medical complications or failing an appetite test, as defined
307 by WHO guidelines ⁴⁸. Children were excluded if they had a known allergy to milk products
308 and did not provide consent. Biological samples were obtained before the children received
309 the randomised treatment irrespective of HIV status. Unless a child's HIV positive status was
310 documented, HIV status was assessed by offering an antibody test at admission plus
311 appropriate counseling. For this analysis, patients that tested positive on an HIV antibody test
312 were considered HIV(+) and children with missing or declined HIV test were excluded.

313

314 The nested case-control study was designed to investigate inpatient mortality for which
315 proteomic, cytokine, and chemokine data was generated using plasma samples collected at
316 admission during enrolment to the trial. To compare the proteomic profiles between HIV
317 affected and non-affected children with SAM, we used data from a nested case-control study

318 to investigate inpatient mortality during the trial including 79 deaths and 88 children that were
319 clinically stabilized and discharged from the hospital matched by site of recruitment. For this
320 analysis, all observations from the proteomics dataset were included and the analysis was
321 designed to help overcome selection bias, as given in detail below.

322

323 **Proteomics, cytokine and chemokine analysis**

324 Untargeted proteomics and targeted cytokines and chemokines analysis of plasma samples
325 were performed following methods described previously⁴⁹. The targeted protein panel
326 included: epidermal growth factor (EGF); eotaxin; granulocyte-colony stimulating factor
327 (GCSF); granulocyte-macrophage colony-stimulating factor (GMCSF); interferon alpha-2
328 (IFNa2); interferon gamma (IFNg); interleukins 10, 12p40, 12p70, 13, 15, 17A, 1a, 1b, 1RA,
329 2 to 8; interferon gamma-induced protein 10 (IP10); monocyte chemoattractant protein 1
330 (MCP1), macrophage inflammatory protein 1 alpha and beta (MIP1a & b); tumour necrosis
331 factor alpha (TNFa) and beta (TNFb); and vascular endothelial growth factor (VEGF).

332

333 **Data analysis**

334 Data analyses were performed using R v3.5⁵⁰. Analysis of the prevalence of HIV(+),
335 nutritional status and their associations with inpatient mortality utilised the entire trial dataset
336 (N=843). Analysis of categorical data was performed using Fisher's test and generalised
337 linear models for continuous outcomes. Logistic regression was used to analyse binary
338 outcomes adjusting for age, sex, presence of oedema, and site of recruitment. These
339 associations were also adjusted for MUAC. As a sensitivity analysis to address the possibility
340 of confounding due to HIV maternal antibodies in younger children, a test of interaction
341 between age above or below 18 months and individual proteins towards HIV status was
342 performed.

343

344 The proteomics, cytokines and chemokines analyses were secondary analyses of data
345 collected from a nested case-control study with inpatient mortality as its primary outcome,
346 hence with strong selection bias. The analysis for the association between HIV status and
347 individual proteins was therefore performed using logistic regression analysis with inverse
348 probability weighting (IPW) to correct for selection bias⁵¹⁻⁵⁴. Weights (w) were calculated as
349 suggested by Samuelsen⁵² wherein the weight for each observation selected into the nested
350 case-control study was computed as the inverse of the probability of being selected for the
351 nested study from the main clinical trial. The probability of inclusion was therefore calculated
352 as:

353

$$354 \quad p(i) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x_1 + \beta_2 x_2 \dots + \beta_n x_n)}};$$

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

359

$$360 \quad w(i) = \frac{1}{p(i)}$$

361

362 Differences in individual proteins abundances were considered statistically significant when
363 $p < 0.05$ after adjustment for multiple comparisons using Benjamini-Hochberg false discovery
364 rate (FDR)⁵⁵.

365

366 Multivariate analysis was undertaken in order to determine several proteins that are
367 collectively associated with HIV status, some of which may not be significantly associated to

368 HIV independently. This was performed using a weighted elastic net (EN) model
369 implemented using the “glmnet” package in R⁵⁶. EN is a penalized regression approach that
370 was developed to help overcome problems caused by high dimensional data. It is an
371 integration of two regularized approaches, ridge regression and least absolute shrinkage and
372 selection operator (LASSO), wherein the contribution of each of these models to the final EN
373 model is controlled by the α parameter^{56,57}. The strong penalization imposed by LASSO
374 draws coefficients to zero thereby eliminating non-predictive proteins features, whereas ridge
375 regression addresses potential multi-collinearity problems in high-dimensional data^{56,57}.

376

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

385

386 Proteins with significant association with HIV status after correction for false discovery and
387 those extracted by the EN model were then uploaded to The Database for Annotation,
388 Visualization and Integrated Discovery (DAVID) v6.8 Bioinformatics Resource⁵⁸ to assess
389 the Gene ontology (GO) enriched pathways of the differentially expressed proteins.

390

391 EN model validity was judged based on the AUROC and misclassification error rate. The
392 fitted EN model performance measured as optimism-corrected AUC was validated using

393 bootstrap, following the procedure of Smith et al.⁵⁹. Bootstrapping was performed on 2000
394 iterations using the “BootValidation” package in R. Protein features extracted at least 80% of
395 all iterations by the bootstrap EN model were then considered to be the most relevant protein
396 biomarkers. To test how well these proteins can discriminate HIV status, they were then fitted
397 on a weighted logistic regression with HIV as outcome.

398

399 **Visualisation of significantly enriched GO terms**

400 Bubble plots were used to visualise the significantly enriched pathways ($p < 0.05$ after
401 adjustment for FDR) obtained from DAVID. The p -values in DAVID were obtained using a
402 modified Fisher’s exact test⁶⁰. The y -axis represents the fold enrichment which indicates the
403 magnitude of the enrichment, as calculated in DAVID. Fold enrichment is defined as:

$$404 \quad \text{fold enrichment} = \frac{(m/n)}{M/N},$$

405 where m is the number of proteins significantly associated with HIV status or proteins
406 extracted by the EN model that belong to a particular pathway, while M is the total number of
407 proteins belonging to the same pathway. Variable n is the number of all proteins significantly
408 associated with HIV status or extracted by the EN model and N is the total number of all
409 proteins in the human background. Therefore, a fold enrichment of 10 indicates that 10% of
410 the proteins significantly associated with HIV status belong to a particular pathway, and 1%
411 of all annotated proteins in the human background belongs to the same pathway⁶⁰. However,
412 the proponents of this metric warn that big fold enrichments could be obtained from a small
413 number of proteins, which could be due to small n or pathways with fewer members.

414

415 The x -axis on the hand represents the enrichment z -score for a particular pathway⁶¹, which is
416 calculated as follows:

$$417 \quad z - \text{score} = \frac{(up-down)}{\sqrt{count}};$$

418 where *up* is the total number of proteins upregulated, *down* is the total number of proteins
419 downregulated, and *count* is the total number of proteins in the input which belongs to a
420 particular pathway. Variables *up* and *down* were based on the weighted logistic regression for
421 each individual protein. Hence, if 5 proteins belonging to pathway *x* were upregulated and 2
422 were downregulated, the z-score for pathway *x* would be: $(5-2)/\sqrt{7} = 1.13$. A positive z-score
423 indicates that the particular pathway is overall upregulated in HIV(+), whereas a negative z-
424 score indicates an overall downregulation⁶¹.

425

426 **Funding Sources**

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

435

436 **Declaration of Interests**

437 The authors declare that we do not have any conflicts of interest.

438

439 **Study approvals**

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

443

444 **Author contributions**

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

449

450 **References**

- 451 1 Black, R. E. *et al.* Maternal and child undernutrition and overweight in low-income
452 and middle-income countries. *The Lancet* **382**, 427-451, doi:10.1016/S0140-
453 6736(13)60937-X (2013).
- 454 2 Heikens, G. T. *et al.* Case management of HIV-infected severely malnourished
455 children: challenges in the area of highest prevalence. *Lancet* **371**, 1305-1307,
456 doi:10.1016/s0140-6736(08)60565-6 (2008).
- 457 3 Irena, A. H., Mwambazi, M. & Mulenga, V. Diarrhea is a Major killer of Children
458 with Severe Acute Malnutrition Admitted to Inpatient Set-up in Lusaka, Zambia. *Nutr.*
459 *J.* **10**, 110-110, doi:10.1186/1475-2891-10-110 (2011).
- 460 4 Maitland, K. *et al.* Children with Severe Malnutrition: Can Those at Highest Risk of
461 Death Be Identified with the WHO Protocol? *PLoS Med.* **3**, e500,
462 doi:10.1371/journal.pmed.0030500 (2006).
- 463 5 Hossain, M. *et al.* Efficacy of World Health Organization guideline in facility-based
464 reduction of mortality in severely malnourished children from low and middle income
465 countries: A systematic review and meta-analysis. *J. Paediatr. Child Health* **53**, 474-
466 479, doi:10.1111/jpc.13443 (2017).
- 467 6 Kerac, M. *et al.* Follow-Up of Post-Discharge Growth and Mortality after Treatment
468 for Severe Acute Malnutrition (FuSAM Study): A Prospective Cohort Study. *PLOS*
469 *ONE* **9**, e96030, doi:10.1371/journal.pone.0096030 (2014).
- 470 7 Berkley, J. A. *et al.* Daily co-trimoxazole prophylaxis to prevent mortality in children
471 with complicated severe acute malnutrition: a multicentre, double-blind, randomised
472 placebo-controlled trial. *The Lancet. Global health* **4**, e464-473, doi:10.1016/s2214-
473 109x(16)30096-1 (2016).
- 474 8 Nemetchek, B. *et al.* Paediatric postdischarge mortality in developing countries: a
475 systematic review. *BMJ Open* **8**, e023445, doi:10.1136/bmjopen-2018-023445 (2018).
- 476 9 Fergusson, P. & Tomkins, A. HIV prevalence and mortality among children
477 undergoing treatment for severe acute malnutrition in sub-Saharan Africa: a
478 systematic review and meta-analysis. *Trans. R. Soc. Trop. Med. Hyg.* **103**, 541-548,
479 doi:10.1016/j.trstmh.2008.10.029 (2009).
- 480 10 Ndirangu, M., Wariero, J. O., Sachs, S. E., Masibo, P. & Deckelbaum, R. J.
481 Nutritional status of under-five children in HIV-affected households in western Kenya.
482 *Food Nutr. Bull.* **32**, 159-167, doi:10.1177/156482651103200208 (2011).

- 483 11 Rose, A. M., Hall, C. S. & Martinez-Alier, N. Aetiology and management of
484 malnutrition in HIV-positive children. *Arch. Dis. Child.* **99**, 546-551,
485 doi:10.1136/archdischild-2012-303348 (2014).
- 486 12 Trehan, I., O'Hare, B. A., Phiri, A. & Heikens, G. T. Challenges in the Management of
487 HIV-Infected Malnourished Children in Sub-Saharan Africa. *AIDS Res. Treat.* **2012**,
488 8, doi:10.1155/2012/790786 (2012).
- 489 13 Bandsma, R. H. J. *et al.* A reduced-carbohydrate and lactose-free formulation for
490 stabilization among hospitalized children with severe acute malnutrition: A double-
491 blind, randomized controlled trial. *PLoS Med.* **16**, e1002747,
492 doi:10.1371/journal.pmed.1002747 (2019).
- 493 14 Mody, A. *et al.* Effects of HIV infection on the metabolic and hormonal status of
494 children with severe acute malnutrition. *PLoS One* **9**, e102233,
495 doi:10.1371/journal.pone.0102233 (2014).
- 496 15 Tacnet-Delorme, P. *et al.* In Vitro Analysis of Complement-Dependent HIV-1 Cell
497 Infection Using a Model System. *The Journal of Immunology* **162**, 4088 (1999).
- 498 16 Stoiber, H., Kacani, L., Speth, C., Würzner, R. & Dierich, M. P. The supportive role of
499 complement in HIV pathogenesis. *Immunol. Rev.* **180**, 168-176, doi:10.1034/j.1600-
500 065X.2001.1800115.x (2001).
- 501 17 Stoiber, H., Speth, C. & Dierich, M. P. Role of complement in the control of HIV
502 dynamics and pathogenesis. *Vaccine* **21**, S77-S82, doi:[https://doi.org/10.1016/S0264-
503 410X\(03\)00203-2](https://doi.org/10.1016/S0264-410X(03)00203-2) (2003).
- 504 18 Oliveira, J. d. *et al.* Effect of BCHE single nucleotide polymorphisms on lipid
505 metabolism markers in women. *Genet. Mol. Biol.* **40**, 408-414, doi:10.1590/1678-
506 4685-GMB-2016-0123 (2017).
- 507 19 Xu, L. *et al.* Butyrylcholinesterase Levels on Admission Predict Severity and 12-
508 Month Mortality in Hospitalized AIDS Patients. *Mediators Inflamm.* **2018**, 5201652,
509 doi:10.1155/2018/5201652 (2018).
- 510 20 Santarpia, L., Grandone, I., Contaldo, F. & Pasanisi, F. Butyrylcholinesterase as a
511 prognostic marker: a review of the literature. *Journal of cachexia, sarcopenia and
512 muscle* **4**, 31-39, doi:10.1007/s13539-012-0083-5 (2013).
- 513 21 Chen, V. P. *et al.* Butyrylcholinesterase Deficiency Promotes Adipose Tissue Growth
514 and Hepatic Lipid Accumulation in Male Mice on High-Fat Diet. *Endocrinology* **157**,
515 3086-3095, doi:10.1210/en.2016-1166 (2016).
- 516 22 Sridhar, G. *et al.* Serum butyrylcholinesterase in type 2 diabetes mellitus: a
517 biochemical and bioinformatics approach. *Lipids Health Dis.* **4**, 18, doi:10.1186/1476-
518 511x-4-18 (2005).
- 519 23 Ceperuelo-Mallafre, V. *et al.* Zinc alpha-2 glycoprotein is implicated in dyslipidaemia
520 in HIV-1-infected patients treated with antiretroviral drugs. *HIV Med.* **13**, 297-303,
521 doi:10.1111/j.1468-1293.2011.00976.x (2012).
- 522 24 Lei, L. *et al.* Circulating zinc- α 2-glycoprotein levels are low in newly diagnosed
523 patients with metabolic syndrome and correlate with adiponectin. *Nutr. Metab. (Lond.)*
524 **14**, 53, doi:10.1186/s12986-017-0210-6 (2017).
- 525 25 Yang, M. *et al.* Zinc- α 2-glycoprotein is associated with insulin resistance in humans
526 and is regulated by hyperglycemia, hyperinsulinemia, or liraglutide administration:
527 cross-sectional and interventional studies in normal subjects, insulin-resistant subjects,
528 and subjects with newly diagnosed diabetes. *Diabetes care* **36**, 1074-1082,
529 doi:10.2337/dc12-0940 (2013).
- 530 26 Bastard, J. P. *et al.* Diabetes and dyslipidaemia are associated with oxidative stress
531 independently of inflammation in long-term antiretroviral-treated HIV-infected
532 patients. *Diabetes Metab.*, doi:10.1016/j.diabet.2019.02.008 (2019).

- 533 27 Nix, L. M. & Tien, P. C. Metabolic syndrome, diabetes, and cardiovascular risk in
534 HIV. *Curr. HIV/AIDS Rep.* **11**, 271-278, doi:10.1007/s11904-014-0219-7 (2014).
- 535 28 Sobieszczyk, M. E. *et al.* Metabolic Syndrome After HIV Acquisition in South
536 African Women. *J. Acquir. Immune Defic. Syndr.* **73**, 438-445,
537 doi:10.1097/qai.0000000000001123 (2016).
- 538 29 Policarpo, S., Rodrigues, T., Moreira, A. C. & Valadas, E. Cardiovascular risk in HIV-
539 infected individuals: A comparison of three risk prediction algorithms. *Rev. Port.*
540 *Cardiol.* **38**, 463-470, doi:10.1016/j.repc.2019.08.002 (2019).
- 541 30 Lagathu, C. *et al.* Metabolic complications affecting adipose tissue, lipid and glucose
542 metabolism associated with HIV antiretroviral treatment. *Expert Opin. Drug Saf.* **18**,
543 829-840, doi:10.1080/14740338.2019.1644317 (2019).
- 544 31 Sears, S. *et al.* Metabolic Syndrome Among People Living with HIV Receiving
545 Medical Care in Southern United States: Prevalence and Risk Factors. *AIDS Behav.*,
546 doi:10.1007/s10461-019-02487-8 (2019).
- 547 32 Seth, A. & Sherman, K. E. Fatty liver disease in persons with HIV infection. *Top.*
548 *Antivir. Med.* **27**, 75-82 (2019).
- 549 33 Palios, J., Kadoglou, N. P. E. & Lampropoulos, S. The pathophysiology of HIV-
550 /HAART-related metabolic syndrome leading to cardiovascular disorders: the
551 emerging role of adipokines. *Exp. Diabetes Res.* **2012**, 103063-103063,
552 doi:10.1155/2012/103063 (2012).
- 553 34 Silva, B., Peixoto, G., da Luz, S., de Moraes, S. & Peres, S. Adverse effects of chronic
554 treatment with the Main subclasses of highly active antiretroviral therapy: a systematic
555 review. *HIV Med.* **20**, 429-438, doi:10.1111/hiv.12733 (2019).
- 556 35 Riddler, S. A. *et al.* Impact of HIV Infection and HAART on Serum Lipids in Men.
557 *JAMA* **289**, 2978-2982, doi:10.1001/jama.289.22.2978 (2003).
- 558 36 Adal, M., Howe, R., Kassa, D., Aseffa, A. & Petros, B. Malnutrition and lipid
559 abnormalities in antiretroviral naïve HIV-infected adults in Addis Ababa: A cross-
560 sectional study. *PLOS ONE* **13**, e0195942, doi:10.1371/journal.pone.0195942 (2018).
- 561 37 Rasheed, S., Yan, J. S., Lau, A. & Chan, A. S. HIV Replication Enhances Production
562 of Free Fatty Acids, Low Density Lipoproteins and Many Key Proteins Involved in
563 Lipid Metabolism: A Proteomics Study. *PLOS ONE* **3**, e3003,
564 doi:10.1371/journal.pone.0003003 (2008).
- 565 38 Renga, B. *et al.* HIV-1 infection is associated with changes in nuclear receptor
566 transcriptome, pro-inflammatory and lipid profile of monocytes. *BMC Infect. Dis.* **12**,
567 274, doi:10.1186/1471-2334-12-274 (2012).
- 568 39 Trehan, I. & Manary, M. J. Management of severe acute malnutrition in low-income
569 and middle-income countries. *Arch. Dis. Child.* **100**, 283-287,
570 doi:10.1136/archdischild-2014-306026 (2015).
- 571 40 Joy, T. *et al.* Dietary fat intake and relationship to serum lipid levels in HIV-infected
572 patients with metabolic abnormalities in the HAART era. *AIDS (London, England)* **21**,
573 1591-1600, doi:10.1097/QAD.0b013e32823644ff (2007).
- 574 41 Mansfield, K. G. *et al.* A diet high in saturated fat and cholesterol accelerates simian
575 immunodeficiency virus disease progression. *J. Infect. Dis.* **196**, 1202-1210,
576 doi:10.1086/521680 (2007).
- 577 42 Lindegaard, B. *et al.* Adipose tissue expression of IL-18 and HIV-associated
578 lipodystrophy. *AIDS* **18**, 1956-1958, doi:10.1097/00002030-200409240-00013 (2004).
- 579 43 Ockenga, J. *et al.* ESPEN Guidelines on Enteral Nutrition: Wasting in HIV and other
580 chronic infectious diseases. *Clin. Nutr.* **25**, 319-329, doi:10.1016/j.clnu.2006.01.016
581 (2006).

- 582 44 Wanke, C. A. *et al.* A medium chain triglyceride-based diet in patients with HIV and
583 chronic diarrhea reduces diarrhea and malabsorption: A prospective, controlled trial.
584 *Nutrition* **12**, 766-771, doi:[https://doi.org/10.1016/S0899-9007\(96\)00221-3](https://doi.org/10.1016/S0899-9007(96)00221-3) (1996).
- 585 45 Amadi, B. *et al.* Intestinal and systemic infection, HIV, and mortality in Zambian
586 children with persistent diarrhea and malnutrition. *J. Pediatr. Gastroenterol. Nutr.* **32**,
587 550-554, doi:10.1097/00005176-200105000-00011 (2001).
- 588 46 Kessler, L., Daley, H., Malenga, G. & Graham, S. The impact of HIV infection on the
589 clinical presentation of severe malnutrition in children at QECH. *Malawi medical*
590 *journal : the journal of Medical Association of Malawi* **13**, 30-33 (2001).
- 591 47 Musoke, P. M. & Fergusson, P. Severe malnutrition and metabolic complications of
592 HIV-infected children in the antiretroviral era: clinical care and management in
593 resource-limited settings. *The American journal of clinical nutrition* **94**, 1716S-1720S,
594 doi:10.3945/ajcn.111.018374 (2011).
- 595 48 WHO. *Management of severe malnutrition: a manual for physicians and other senior*
596 *health workers*, <<http://apps.who.int/iris/bitstream/10665/41999/1/a57361.pdf>>
597 (1999).
- 598 49 Njunge, J. M. *et al.* Biomarkers of post-discharge mortality among children with
599 complicated severe acute malnutrition. *Sci. Rep.* **9**, 5981, doi:10.1038/s41598-019-
600 42436-y (2019).
- 601 50 R Core Team. *R: A language and environment for statistical computing*. R Foundation
602 *for Statistical Computing*, <<https://www.R-project.org/>> (2017).
- 603 51 Kim, R. S. & Kaplan, R. C. Analysis of secondary outcomes in nested case-control
604 study designs. *Stat. Med.* **33**, 4215-4226, doi:10.1002/sim.6231 (2014).
- 605 52 Samuelsen, S. O. A Pseudolikelihood Approach to Analysis of Nested Case-Control
606 Studies. *Biometrika* **84**, 379-394 (1997).
- 607 53 Pan, Y., Cai, J., Longnecker, M. P. & Zhou, H. Secondary outcome analysis for data
608 from an outcome-dependent sampling design. *Stat. Med.* **37**, 2321-2337,
609 doi:10.1002/sim.7672 (2018).
- 610 54 Jiang, Y., Scott, A. J. & Wild, C. J. Secondary analysis of case-control data. *Stat. Med.*
611 **25**, 1323-1339, doi:10.1002/sim.2283 (2006).
- 612 55 Benjamini, Y. & Hochberg, Y. Controlling the False Discovery Rate: A Practical and
613 Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society.*
614 *Series B (Methodological)* **57**, 289-300 (1995).
- 615 56 Friedman, J. H., Hastie, T. & Tibshirani, R. Regularization Paths for Generalized
616 Linear Models via Coordinate Descent. *2010* **33**, 22, doi:10.18637/jss.v033.i01
617 (2010).
- 618 57 Zou, H. & Hastie, T. Regularization and variable selection via the elastic net. *J. Roy.*
619 *Stat. Soc. Ser. B. (Stat. Method.)* **67**, 301-320, doi:10.1111/j.1467-9868.2005.00503.x
620 (2005).
- 621 58 Huang, D. W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis
622 of large gene lists using DAVID bioinformatics resources. *Nat. Protocols* **4**, 44-57,
623 doi:http://www.nature.com/nprot/journal/v4/n1/supinfo/nprot.2008.211_S1.html
624 (2008).
- 625 59 Smith, G. C. S., Seaman, S. R., Wood, A. M., Royston, P. & White, I. R. Correcting
626 for Optimistic Prediction in Small Data Sets. *Am. J. Epidemiol.* **180**, 318-324,
627 doi:10.1093/aje/kwu140 (2014).
- 628 60 Huang da, W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis
629 of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4**, 44-57,
630 doi:10.1038/nprot.2008.211 (2009).

631 61 Walter, W., Sanchez-Cabo, F. & Ricote, M. GOpot: an R package for visually
632 combining expression data with functional analysis. *Bioinformatics* **31**, 2912-2914,
633 doi:10.1093/bioinformatics/btv300 (2015).
634

635 Figure Legend

636

637 **Figure 1. Univariate analysis of plasma proteome and individual plasma cytokines**
638 **associated with HIV.** (A) Volcano plot showing several significantly different (FDR adjusted
639 p value < 0.05) proteins and their log₂ HIV(+) versus HIV(-) fold change. Red points
640 represent those significantly higher in plasma of HIV(-), blue points significantly enriched in
641 plasma of HIV(+) and orange points significantly higher than 1.5 folds in HIV(+) compared
642 to HIV(-) SAM children. Vertical lines indicate significance level at $p = 0.05$ and 0.01 ;
643 horizontal lines indicate more than 1.5 folds enrichment. (B) Log odds plots showing
644 association of chemokine markers analysed using Luminex platform and HIV status. Points
645 indicate log odds ratio for every log increase in plasma protein concentration; bars indicate
646 95% confidence interval.
647

648 **Figure 2. Multivariate analysis of plasma proteome associated with HIV.** (A) Elastic net
649 (EN) regularized regression lambda parameter optimization curve, optimal lambda parameter
650 was chosen based on the highest area under the receiver operator curve (AUROC); (B)
651 AUROC (0.97 [95% CI: 0.95 – 0.99]) of the EN model generated using the lambda parameter,
652 alpha parameter was set to 0.75, final model extracted 34 protein features, optimism-adjusted
653 bootstrap validation of the generated EN model, validated AUROC = 0.90 [95% CI: 0.90 –
654 0.90] using 2000 iterations; (C) Gene ontology (GO-terms) enrichment analysis of proteins
655 extracted by the EN model. X-axis represents z-scores; y-axis, fold enrichment, and size of
656 the spheres represent the number of proteins involved in the particular pathway. Gold circles
657 represent pathways enriched in HIV(+) whereas blue circles are pathways more associated
658 with HIV(-). The grey circle indicate that there are as much proteins in this pathway that are
659 significantly upregulated and downregulated in HIV. Only significantly enriched pathways (p
660 < 0.05 after FDR adjustment) are plotted. See main text for explanation of the plots. Pathways
661 enriched are identified in the table. (D) Log odds ratio plot of the 3 proteins extracted after
662 bootstrap validation with log odds on the x-axis and bars indicating 95% confidence interval
663 obtained using weighted logistic regression with HIV as outcome variable and the 3 proteins
664 as covariates. Weights used were obtained by inverse probability of treatment weights; (E)
665 Predictive ability of the weighted logistic regression model using the 3 bootstrap validated
666 proteins with HIV as outcome variable, AUROC = 0.80 [95% CI: 0.73 – 0.87]
667

668

669 Tables

670 Table 1. Descriptive characteristics of the study participants

	All	HIV (+)	HIV (-)	Unknown HIV status	<i>p</i> *
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 ^o
% 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)	<i>Reference</i>
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

671 *comparison between HIV(+) and HIV(-); ^oadjusted for age, sex and site of recruitment

672

673

674 Table 2. Patient characteristics of those subjected to proteomics analysis

	All	HIV (+)	HIV (-)	<i>p</i> *
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
n oedematous (%)	49 (29%)	15 (28%)	34 (30%)	0.856
n mortality (%)	79 (47%)	36 (67%)	43 (38%)	<0.001 ^o
<i>Use of antiretroviral medication</i>				
Naïve		27 (50%)		
highly active antiretroviral therapy (HAART)		14 (26%)		
Nevirapine only		3 (6%)		
Unknown		10 (18%)		

675 *comparison between HIV(+) and HIV(-), ^oadjusted for age, sex, site of recruitment, oedema

676

677

678 Table 3. Nutritional management protocol for children with severe acute malnutrition ³⁹

	Stabilization phase	In-patient Rehabilitation phase	Out-patient Rehabilitation phase
	<i>Days 1 - 7</i>	<i>Weeks 2 - 6</i>	<i>Lengths vary depending on site</i>
Complicated SAM	F75	F100	RUTF
Uncomplicated SAM	-	-	RUTF
<i>Composition</i>			
Energy (kcal per 100 mL F75/F100 or 100 g RUTF)	75	100	5.2 – 5.5
Protein (% total energy)	5	12	10 – 12
Fat (% total energy)	32	53	45 - 60

679

680
681

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 (%) ^o	53 (30%)	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] p=0.63 [#]	-	Reference	1.07 [0.88 - 1.30] p=0.45 [#]

682 *HAART - highly active antiretroviral therapy; NVP – Nevirapine; [#] adjusted for age, sex, site and oedema
683

684 Supplementary table 2. Association between individual proteins to HIV status
685

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

		chain V-VI region EB4) (Ig lambda chain V-VI region NIG-48) (Ig lambda chain V-VI region SUT) (Ig lambda chain V-VI region WLT)		
P18428	LBP	Lipopolysaccharide-binding protein (LBP)	0.31	0.01
P01834	IGKC	Immunoglobulin kappa constant (Ig kappa chain C region) (Ig kappa chain C region AG) (Ig kappa chain C region CUM) (Ig kappa chain C region EU) (Ig kappa chain C region OU) (Ig kappa chain C region ROY) (Ig kappa chain C region TI)	0.69	0.02
P0CG05		Immunoglobulin lambda constant 3 (Ig lambda chain C region DOT) (Ig lambda chain C region NEWM) (Ig lambda-3 chain C regions)	0.67	0.02
P0CG05		Immunoglobulin lambda constant 2 (Ig lambda chain C region Kern) (Ig lambda chain C region NIG-64) (Ig lambda chain C region SH) (Ig lambda chain C region X) (Ig lambda-2 chain C region)	0.67	0.02
B9A064	IGLL5	Immunoglobulin lambda-like polypeptide 5 (G lambda-1) (Germline immunoglobulin 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) (Ig kappa chain V-I region Hau) (Ig kappa chain V-I region Mev) (Ig kappa chain V-I region OU) (Ig kappa chain V-I region Walker)	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) (Hucolin) (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

		lambda-7 chain C region)		
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	PI-16	Peptidase inhibitor 16 (PI-16) (Cysteine-rich secretory protein 9) (CRISP-9) (PSP94-binding protein) (CD antigen CD364)	-0.22	0.05
P01876	IGHA1	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	HRG	Histidine-rich glycoprotein (Histidine-proline-rich glycoprotein) (HPRG)	-0.26	0.05
P04433	IGKV3-11	Immunoglobulin kappa variable 3-11 (Ig kappa chain V-III region VG)	0.57	0.05
P06310	IGKV2-30	Immunoglobulin kappa variable 2-30 (Ig kappa chain V-II region RPMI 6410)	0.52	0.05
G8JLA8		Deleted.	0.11	0.07
P01034	CST3	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	IGHV1-46	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/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)	0.34	0.08
P07225	PROS1	Vitamin K-dependent protein S	0.04	0.08
P29622	SERPI4	Kallistatin (Kallikrein inhibitor) (Peptidase inhibitor 4) (PI-4) (Serpine A4)	-0.16	0.08
H0Y755	FCGR3A	Low affinity immunoglobulin gamma Fc	0.18	0.08

		region receptor III-A (Fragment)		
P10643	C7	Complement component C7	0.06	0.08
P01859	IGHG2	Immunoglobulin heavy constant gamma 2 (Ig gamma-2 chain C region) (Ig gamma-2 chain C region DOT) (Ig gamma-2 chain C region TIL) (Ig gamma-2 chain C region ZIE)	0.36	0.09
P01625		Immunoglobulin kappa variable 4-1 (Ig kappa chain V-IV region B17) (Ig kappa chain V-IV region JI) (Ig kappa chain V-IV region Len) (Ig kappa chain V-IV region STH)	0.38	0.09
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 (Serp1n 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	0.34	0.13

		chain V-III region POM)		
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	SERPI3	Alpha-1-antichymotrypsin (ACT) (Cell growth-inhibiting gene 24/25 protein) (Serpins 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) (Carbonate dehydratase I) (Carbonic anhydrase B) (CAB) (Carbonic anhydrase 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) (GPIb-alpha) (GPIbA) (Glycoprotein Ibalph) (Antigen CD42b-alpha) (CD antigen CD42b) [Cleaved into: Glycocalicin]	0.06	0.17
P02749	APOH	Beta-2-glycoprotein 1 (APC inhibitor) (Activated protein C-binding protein) (Anticardiolipin cofactor) (Apolipoprotein H) (Apo-H) (Beta-2-glycoprotein I) (B2GPI) (Beta(2)GPI)	-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) (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	SERPIND1	Heparin cofactor 2 (Heparin cofactor II) (HC-II) (Protease inhibitor leuserpin-2) (HLS2) (Serpins D1)	-0.22	0.19
P06276	BCHE	Cholinesterase (EC 3.1.1.8) (Acylcholine	-0.19	0.21

		acylhydrolase) (Butyrylcholine esterase) (Choline esterase II) (Pseudocholinesterase)		
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	LGALS3BP	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	Tetranectin (TN) (C-type lectin domain family 3 member B) (Plasminogen kringle 4-	-0.14	0.35

		binding protein)		
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]		0.41
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
B0YIW2	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

		alpha protein) (Keratin-1) (K1) (Type-II keratin Kb1)		
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	SERPI10	Protein Z-dependent protease inhibitor (Serpine peptidase inhibitor, clade A (Alpha-1 antitrypsin, antitrypsin), member 10, isoform CRA_a)	0.03	0.51
P01008	SERPINC1	Antithrombin-III (ATIII) (Serpine 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; Angiostatin; 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

		Coagulation factor XIIa heavy chain; Beta-factor XIIa part 1; Coagulation factor XIIa light chain (Beta-factor XIIa part 2)]		
Q92954_4		Proteoglycan 4 (Lubricin) (Megakaryocyte-stimulating factor) (Superficial zone proteoglycan) [Cleaved into: Proteoglycan 4 C-terminal part]	-0.10	0.61
P01023	A2M	Alpha-2-macroglobulin (Alpha-2-M) (C3 and PZP-like alpha-2-macroglobulin domain-containing protein 5)	-0.06	0.63
P06396	GSN	Gelsolin (AGEL) (Actin-depolymerizing factor) (ADF) (Brevin)	-0.08	0.67
P0DJ18	SAA1	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)]	0.11	0.68
P22891	PROZ	Vitamin K-dependent protein Z	-0.05	0.68
P27169	PON1	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)	-0.17	0.69
D6RF35	GC	Vitamin D-binding protein	-0.09	0.69
K7ERG9	CFD	Complement factor D	0.02	0.72
P00742	F10	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]	0.03	0.72
P01024	C3	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 C3dg fragment; Complement C3g fragment; Complement C3d fragment; Complement C3f fragment; Complement C3c alpha' chain fragment 2]	-0.08	0.72
P01861	IGHG4	Immunoglobulin heavy constant gamma 4 (Ig gamma-4 chain C region)	0.06	0.72
P02042	HBD	Hemoglobin subunit delta (Delta-globin) (Hemoglobin delta chain)	-0.03	0.72
P02760	AMBP	Protein AMBP [Cleaved into: Alpha-1-microglobulin (Protein HC) (Alpha-1 microglycoprotein) (Complex-forming glycoprotein heterogeneous in charge); Inter-	-0.10	0.72

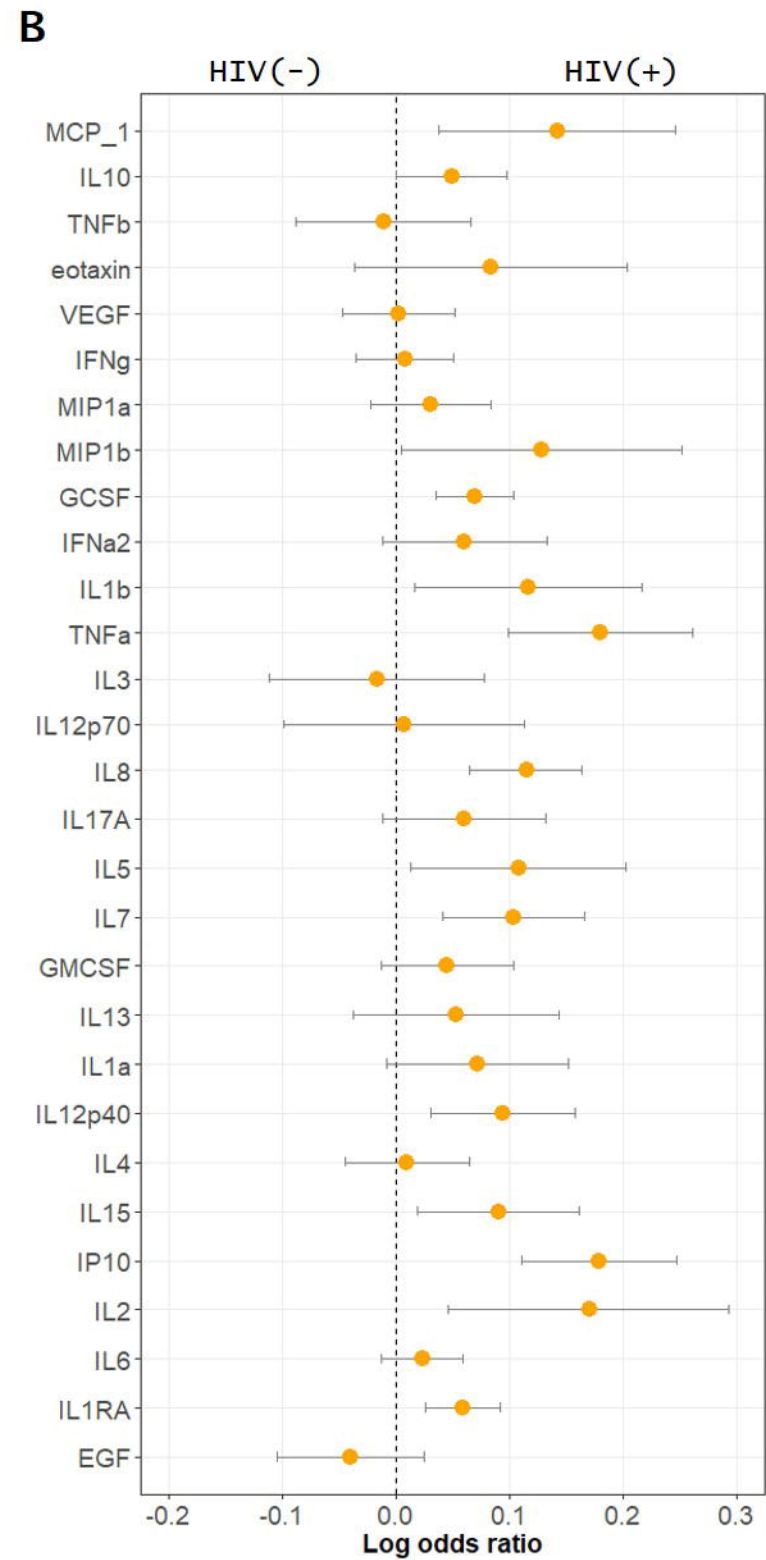
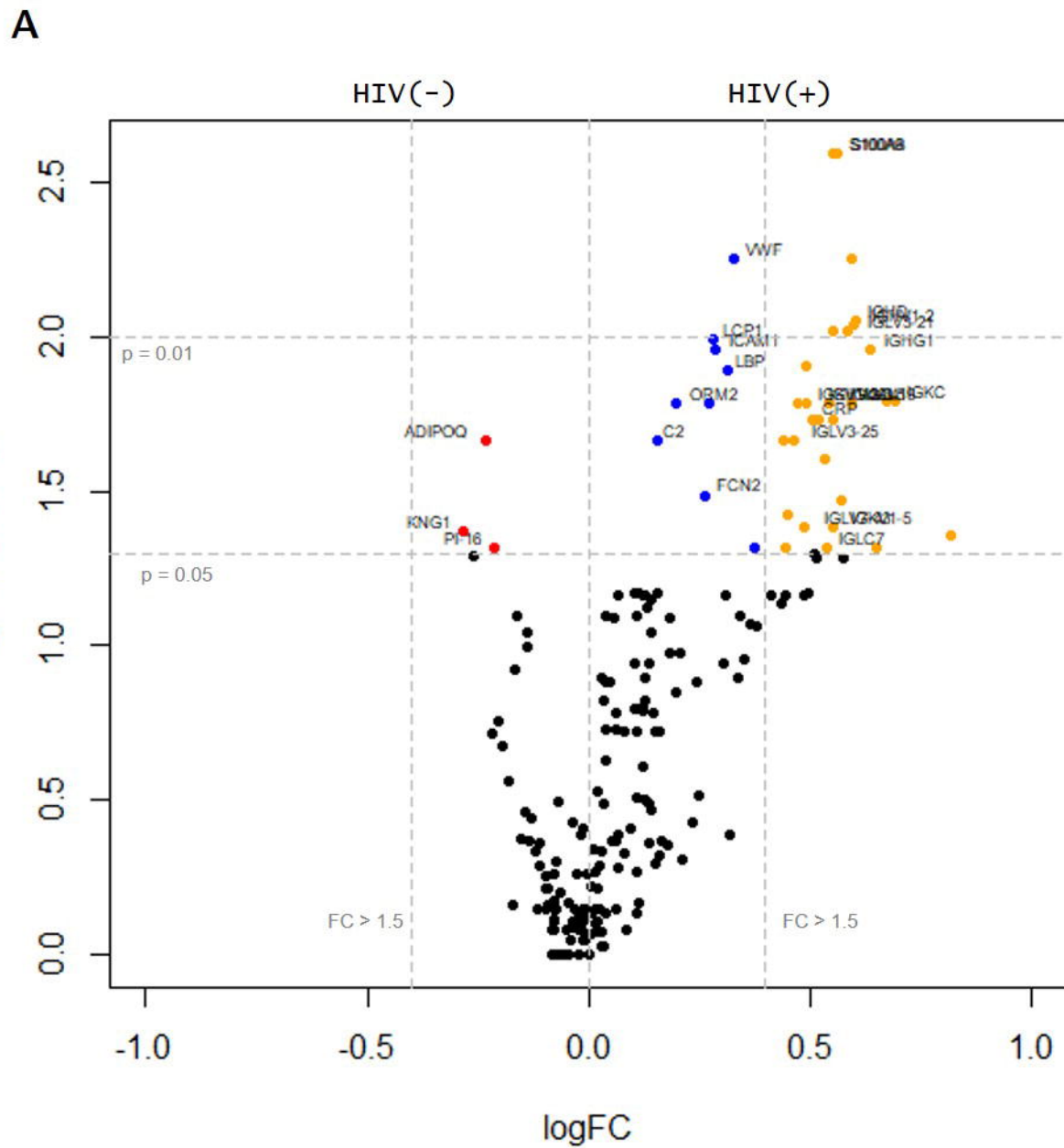
		alpha-trypsin inhibitor light chain (ITI-LC) (Bikunin) (EDC1) (HI-30) (Uronic-acid-rich protein); Trypstatin]		
P02765	AHSG	Alpha-2-HS-glycoprotein (Alpha-2-Z-globulin) (Ba-alpha-2-glycoprotein) (Fetuin-A) [Cleaved into: Alpha-2-HS-glycoprotein chain A; Alpha-2-HS-glycoprotein chain B]	-0.07	0.72
P08185	SERPI6	Corticosteroid-binding globulin (CBG) (Serpins 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) (Fetuin-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) (Serpins 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 (Serpins 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

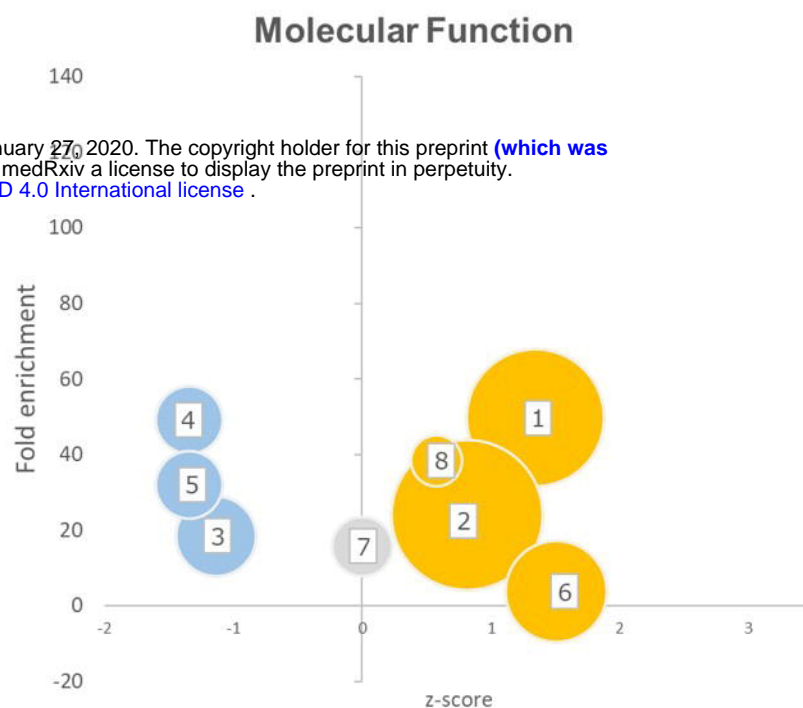
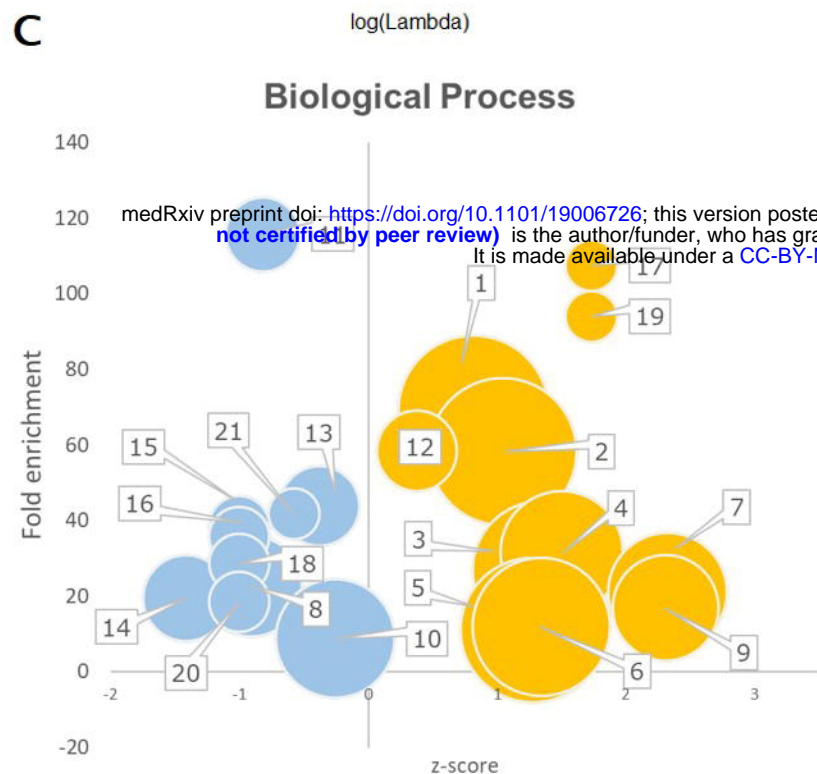
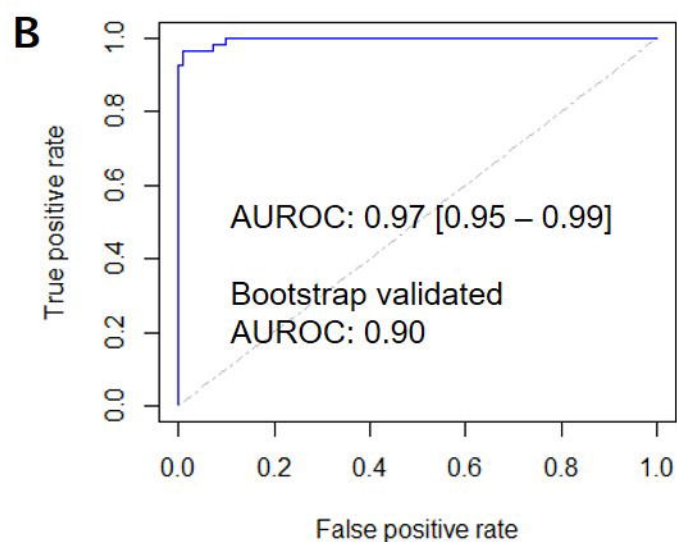
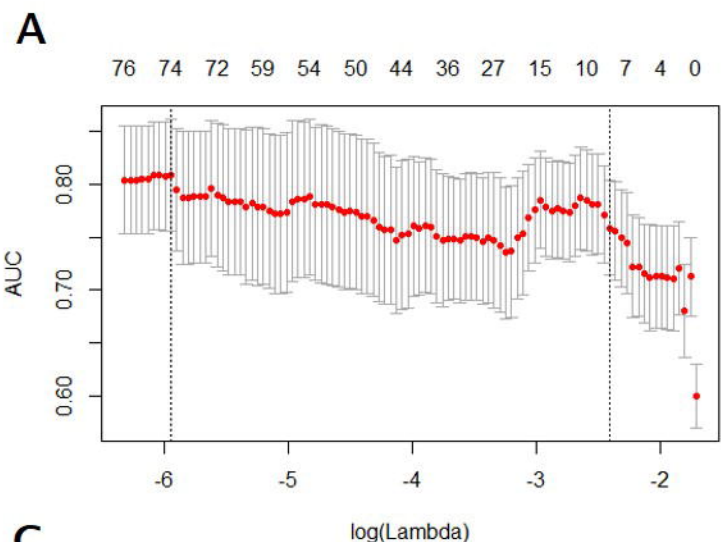
	1	(C1Inh) (C1 esterase inhibitor) (C1-inhibiting factor) (Serpin G1)		
Q96IY4	CPB2	Carboxypeptidase B2 (EC 3.4.17.20) (Carboxypeptidase U) (CPU) (Plasma carboxypeptidase B) (pCPB) (Thrombin-activable fibrinolysis inhibitor) (TAFI)	-0.04	0.78
P07996	THBS1	Thrombospondin-1 (Glycoprotein G)	-0.08	0.79
P08603	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
P0COL4	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

		C4d-A; Complement C4 gamma chain]		
E9PGP2		Deleted.	0.03	0.85
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)	-0.01	0.85
P08697	SERPINF2	Alpha-2-antiplasmin (Alpha-2-AP) (Alpha-2-plasmin inhibitor) (Alpha-2-PI) (Serpins F2)	0.01	0.85
H0YAC1	KLKB1	Plasma kallikrein (Fragment)	-0.04	0.90
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]	-0.01	0.90
P04180	LCAT	Phosphatidylcholine-sterol acyltransferase (EC 2.3.1.43) (Lecithin-cholesterol acyltransferase) (Phospholipid-cholesterol acyltransferase)	-0.01	0.90
P00739	HPR	Haptoglobin-related protein	0.03	0.93
P02743	APCS	Serum amyloid P-component (SAP) (9.5S alpha-1-glycoprotein) [Cleaved into: Serum amyloid P-component(1-203)]	0.03	0.94
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]	-0.08	1.00
P00450	CP	Ceruloplasmin (EC 1.16.3.1) (Ferroxidase)	-0.05	1.00
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]	-0.05	1.00
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)	-0.06	1.00
P04217	A1BG	Alpha-1B-glycoprotein (Alpha-1-B glycoprotein)	-0.02	1.00
P06727	APOA4	Apolipoprotein A-IV (Apo-AIV) (ApoA-IV) (Apolipoprotein A4)	0.00	1.00
P13671	C6	Complement component C6	-0.05	1.00
P43652	AFM	Afamin (Alpha-albumin) (Alpha-Alb)	-0.07	1.00
P69891	HBG1	Hemoglobin subunit gamma-1 (Gamma-1-	-0.07	1.00

		globin) (Hb F Agamma) (Hemoglobin gamma-1 chain) (Hemoglobin gamma-A chain)		
--	--	---	--	--

686 *log fold change - positive value for log fold change indicate higher enrichment in HIV(+);
687 °false-discovery rate adjusted after logistic regression using inverse probability weighting
688 accounting for recruitment site, sex , age, HIV status, mid-upper arm circumference, and
689 presence of oedema





Code	Pathway
1	complement activation
2	complement activation, classical pathway
3	receptor-mediated endocytosis
4	Fc-gamma receptor signalling pathway involved in phagocytosis
5	Proteolysis
6	immune response
7	regulation of immune response
8	negative regulation of endopeptidase activity
9	Fc-epsilon receptor signalling pathway
10	innate immune response
11	complement activation, alternative pathway
12	regulation of complement activation
13	retina homeostasis
14	platelet degranulation
15	positive regulation of B cell activation
16	phagocytosis, recognition
17	complement activation, lectin pathway
18	phagocytosis, engulfment
19	very-low-density lipoprotein particle assembly
20	B cell receptor signaling pathway
21	blood coagulation, intrinsic pathway

Code	Pathway
1	antigen binding
2	serine-type endopeptidase activity
3	serine-type endopeptidase inhibitor activity
4	immunoglobulin receptor binding
5	endopeptidase inhibitor activity
6	identical protein binding
7	glycoprotein binding
8	lipid transporter activity

