

27 **Abstract**

28 **Background**

29 The number of carotid plaques independently predicts incident atherosclerotic cardiovascular
30 disease (ACVD).

31 However, performing vascular imaging in apparently healthy subjects is challenging, owing
32 organizational/economical barriers. Plasma proteomics can offer an alternative approach to identify
33 individuals with carotid plaques, at high risk of eventually developing ACVD.

34
35 **Methods**

36 In this observational study, we studied by Normalized Protein eXpression (NPX; OlinkTM), the
37 plasma levels of 368 proteins in 664 subjects from the PLIC study, who were screened by
38 ultrasound for the presence of carotid plaques. We clustered, by artificial intelligence, the proteins
39 that more accurately identified subjects, stratifying them according to the number of plaques they
40 presented with. We also study prediction of occurring events over 22 years.

41 **Results**

42 299/664 subjects had at least 1 carotid plaque. Among those, 77 subjects presented with only one
43 plaque, 101 with 2 plaques and 121 with ≥ 3 plaques (3+). The remaining 365 subjects with no
44 plaques acted as controls. The proteins differently expressed versus controls increased as a function
45 of the number of plaques. 32 proteins were shared among the groups of subjects with plaques, but
46 87, significantly associated with the presence of 3+ plaques, improved the AUC of the ROC,
47 together with the ACVD risk factors, to discriminate subjects with 3+ plaques versus the AUC of
48 the ROC considering the ACVD risk factors only (AUC= 0.918 (0.887-0.943) vs AUC= 0.760
49 (0.716-0.801) respectively, $p < 0.001$). The ACVD risk factors barely predicted the 198 occurring
50 events (AUC= 0.559 (0.521-0.598)), but proteomics associated with plaques improved the
51 prediction (AUC= 0.739 (0.704-0.773), $p < 0.001$).

52 By analyzing the biological processes, we identified that chemotaxis/migration of leukocytes and
53 the signaling of interleukins/cytokines were the top pathways involved.

54
55 **Conclusions**

56 Plasma proteomics helps to identify apparently healthy subjects with higher number of carotid
57 plaques more accurately and to predict occurring ACVDs in those individuals.

58

59 Introduction

60 Atherosclerotic Cardiovascular Disease (ACVD) is the leading cause of death worldwide¹. In
61 “apparently healthy” population a subset of subjects with subclinical ACVD can be detected and
62 these subjects are identified as high, or very high, risk in contemporary cardiovascular guidelines,
63 and carotid plaques have been suggested as a hallmark of increased ACVD risk²⁻⁴.

64 However, applying noninvasive vascular imaging procedures to detect subclinical ACVD, such as
65 those identifying carotid plaques, is challenging in “apparently healthy” population, owing
66 organizational, economical barriers, and given the substantial shortcomings in sensitivity and inter-
67 operator variability^{5,6}.

68 Plasma biomarkers have been recently used to identify people at high risk of developing⁷ or
69 reoccurring cardiovascular events⁸ and accumulating evidence indicates that plasma proteins, or
70 specific subsets of them, outperform the currently available algorithms to predict increased ACVD
71 risk in “apparently healthy” individuals⁷. Yet evidence regarding their ability to inform the presence
72 of vascular plaques is still limited, as in few studies the number of plasma proteins that have been
73 investigated is limited to 86^{9,10}.

74 Inflammation is known to play an important role in the onset and progression of plaques¹¹. We
75 therefore sought to address the question whether we could more accurately identify “apparently
76 healthy individuals” carrying asymptomatic carotid plaques among 664 subjects from a free-living
77 population of the northern area of Milan. To this end we studied, by Normalized Protein eXpression
78 (NPX) (an index of relative abundance¹²), the plasma levels of 368 proteins from the targeted
79 Cardio-metabolic, Inflammation, Cardiovascular II, and III panels (OlinkTM). New omics techniques
80 can inform on vascular activation of pathway related to atherosclerosis¹³⁻¹⁵ and the power of
81 proteomics as an independent marker of carotid atherosclerosis has been still proposed, albeit in a
82 limited number of subjects and by addressing a maximum of 82 plasma proteins linked to
83 cardiovascular disease^{16,17}. Therefore, we searched for proteins whose expressions in plasma,
84 recapitulating biologically relevant pathways, can improve the accuracy to identify apparently
85 healthy subjects with asymptomatic atherosclerotic disease at the carotids and if they also could
86 predict the occurrence of clinically overt atherosclerotic stages.

87

88

89 Materials and methods

90 Study population

91 The population was from a subgroup of subjects from the northern area of Milan, randomly selected
92 during the basal clinical evaluation of the Progressione delle Lesioni Intimali Carotidee (“PLIC”)
93 Study, which initially included 2.606 participants from 2001 to 2003^{18,19}. Targeted plasma
94 proteomics was available for 702 participants, but only 664 subjects, who did not develop ACVD
95 before the collection of plasma, were considered for the purpose of the study (**Supplemental**
96 **Figure 1**). Out of them, 365 were free of carotid plaques detected by ultrasound and acted as
97 controls. The subjects’ characteristics agreed with the Sex and Gender Equity in Research (SAGER)
98 guidelines²⁰. The PLIC Study has been approved by the Ethics Committee of the University of
99 Milan (SEFAP/Pr.0003). Written informed consent was obtained from subjects (all over 18 years-
100 old), in accordance with the Declaration of Helsinki. This work is an observational study and it was
101 conducted following the standards of the STrengthening the Reporting of OBservational studies in
102 Epidemiology (STROBE) initiative²¹.

103 Clinical and biochemical data

104 Participants in this study were screened at the Center for the Study of Atherosclerosis at E. Bassini
105 Hospital (Cinisello Balsamo, Milan, Italy) for the personal, familial, pathological, and

106 pharmacological clinical history. Subjects with malignancies (any type), chronic liver disease (self-
107 reported or clinically documented in the records of the Hospital), chronic kidney disease
108 (Glomerular Filtration Rate, GFR < 60 ml/min or documented albuminuria > 30 mg/g) and pregnant
109 women were excluded from the study.

110 Information about lifestyle habits (physical activity, former or active smoking habit) was also
111 collected by validated questionnaire used in the PLIC Study as described elsewhere²². Systolic and
112 diastolic blood pressure, body mass index (BMI), waist and hip circumferences were measured. We
113 followed the recent EAS/ESC guidelines for the diagnosis of hypertension, for the diagnosis of
114 dyslipidemia and for the definition of the cut-off of waist circumference (>94 cm for men and >88
115 cm for women)²³. Blood samples were collected from antecubital vein after 12 hours fasting on
116 NaEDTA tubes (BD Vacuette), centrifuged at 3,000 rpm for 12 minutes (Eppendorf 580r,
117 Eppendorf, Hamburg, Germany) aliquoted in multiple 200 ul samples, stored at -80 degrees for the
118 subsequent profiling of biochemical parameters (which included: total cholesterol, High-Density
119 Lipoprotein cholesterol (HDL-C), triglycerides, ApoB, ApoA-I, glucose, Lp(a), and C-Reactive
120 Protein (CRP)). All the measurements were performed on the aliquots, thawed for the first time,
121 using immuno-turbidimetric and enzymatic methods thorough automatic analyzers (Randox,
122 Crumlin, UK). Low-Density Lipoprotein cholesterol (LDL-C) was derived from Friedewald
123 formula. Lipoprotein(a) (Lp(a)) was measured by immunoassay (Randox, Crumlin, UK).

124

125 **Carotid ultrasound**

126 Carotid ultrasound analysis was performed by a single expert sonographer, blinded on subject's
127 identity and clinical history, with a linear ultrasound probe (4.0e13.0 MHz frequency, 14X48 mm
128 footprint, 38 mm field of view, Vivid S5 GE Healthcare®, Wauwatosa, WI, USA). A detailed
129 description of the ultrasound methodology, based on validated protocol for the PLIC study, has
130 been published elsewhere¹⁹.

131 Carotid plaques were defined if (i) focal loci were identified and/or (ii) if the measured carotid
132 Intima Media Thickness (IMT) was equal or above than 1.3 mm in any vascular tract (common,
133 bulb, bifurcation, internal and external tracts) of one or both carotid arteries²⁴.

134

135 **Plasma proteomics**

136 368 proteins were measured in plasma by Proximity Extension Assay (PEA) technology. The
137 proteins are included in the Cardiovascular II, Cardiovascular III, Inflammation and Cardio-
138 metabolic panels of OlinkTM platform. The complete list of the proteins that are included in the
139 panels, their Uniprot ID and the Olink ID, are provided by the company (<http://www.olink.com>).
140 Proteins were measured from 200 uL of plasma samples by Proximity Extension Assay (PEA).
141 Plasma samples from each subject was separated from whole blood by centrifugation (3,000 rpm,
142 12 minutes at 4 degrees), aliquoted in 96-wells plates during the first clinical evaluation of each
143 subject (between 1998 and 2000) and immediately stored at -80 degrees up to 2018, when they were
144 sent to OlinkTM (Uppsala, Sweden) for proteomic characterization⁷. The final assay read-out of
145 each determination is given in Normalized Protein eXpression (NPX), which is an arbitrary unit on
146 log₂-scale where a high value corresponds to a higher protein plasma level. Each PEA measurement
147 has a lower detection limit (LOD) calculated versus negative controls that are included in each run,
148 and measurements of individual proteins below LOD for each panel were removed from further
149 analysis (2 samples from Cardiometabolic panel, 29 samples from Cardiovascular II; 31 samples
150 from the Inflammation panel; all the samples were with measurement > LOD in the Cardiovascular
151 III panel). All assay characteristics including detection limits, the measurements of assay

152 performance, and validations are available from the manufacturer webpage (<http://www.olink.com>).
153 All raw data can be provided upon reasonable written request to ALC.

154 **Statistical analysis**

155 Descriptive statistics

156 Clinical and biochemical data are presented as median with interquartile ranges (the lower 25th
157 percentile and the upper 75th percentile). After checking for the type of distribution of linear
158 variables (Shapiro-Wilk Test), we compared them between either two groups (Kolmogorov-
159 Smirnov test) or four groups (controls, subjects with 3+, with 2 and with 1 carotid plaque), either by
160 Analysis of Variances (ANOVA) and Bonferroni post-hoc analysis (if variables were normally
161 distributed variables), or by Kruskal-Wallis test (if non-normally distributed variables).
162 Dichotomous variables were compared by Fisher's chi-square test. Grubb's test for the detection of
163 the outliers was performed for all the linear variables before running tests. P values less than 0.05
164 were considered significant for all the statistical analyses. SPSSTM v.24 for Windows (IBM
165 corporation®) was used for all statistical analyses. Graphs were prepared using Graphpad® (v.9.0)
166 and Excel (Microsoft®).
167

168 Machine Learning model

169 Proteomics data were analyzed by Machine Learning (ML) model, previously published for
170 proteomics data in PLIC cohort²⁵.
171 In brief, the ML model consists of an eight binary outcome gradient boosting prediction model
172 (XGboost), including the 368 proteins measured in the cohort. The XGboost model is a statistical
173 learning technique, which produces a non-linear model in the form of an ensemble of weak
174 prediction tree-based models. The extreme gradient boosting classification algorithm optimizes a
175 cost function²⁶ by iteratively choosing a weak hypothesis that points in the negative gradient
176 direction. Using a fivefold cross-validation by random reshuffling of the training set, overfitting
177 was avoided. The XGBoost classifier model was trained in the derivation cohort with 1000 iteration
178 round and 0.300 learning rate. Hyperparameter optimization was performed by k-fold iteration. To
179 avoid overfitting of the models, the dataset was randomly split into a training set (70% of total
180 sample; "train set") and the algorithm was thereafter applied to the internal validation set (30% of
181 total sample; "test set"). We applied a permutation (randomization test) to evaluate statistical
182 validity of the results, since standard univariate significance tests cannot be applied to the used
183 models due to the large number of tests. By evaluating the distribution of all the results obtained in
184 these simulations and comparing it to the true outcomes, we computed statistical significance
185 associated with the joint panel of the selected markers. We also reported importance scores for each
186 of the proteins that demonstrate preferences of model when constructing non-linear prediction
187 function based on the selected biomarkers. (Random Forest classifiers). Sensitivity and specificity
188 to predict the outcome, were performed by running analysis of Receiver Operating Curves (ROC).
189 Comparison of AUC was performed according to the method of Hanley-McNeil. All the ML
190 models were constructed by using Python 6.4.5 with pandas, scikit-learn, NumPy, XGboost.
191

192 Gene Ontology (GO) and KEGG pathway enrichment analyses

193 We conducted enrichment analysis of biological processes with the proteins that emerged as
194 significantly associated with carotid atherosclerosis following in silico strategies, as previously
195 published^{27,28}. The DAVID (The Database for Annotation, Visualization, and Integrated
196 Discovery, NIAID, North Bethesda, MD, USA) platform was used for gene ontology (GO)
197 enrichment analyses. The significant GO biological processes (GO_bp) were selected for

198 Pvalue<0.05. Then, for each GO biological process (GO_bp) we annotated the fold of enrichment,
199 an index of the percentage of proteins belonging to a pathway, and the false discovery rate (FDR) to
200 indicate how likely the enrichment is by chance (FDR<0.05 indicates a statistically significant
201 enrichment of proteins in that pathway).

202

203 **Results**

204 **Increased number of plasma proteins associate with a higher number of carotid plaques.**

205

206 Out of 664 apparently healthy subjects, 299 presented with at least one plaque in one of the
207 carotids, while 365 did not present any plaque and acted as “controls” (the clinical and biochemical
208 parameters are reported in **Supplemental Table 1**). Out of the 299 subjects presenting at least one
209 plaque, 77 had one, 101 subjects had 2 and 121 subjects presented with 3 or more plaques (by 3 up
210 to 7 plaques as a maximum; herein defined as “3+”; the ultrasonographic results are reported in
211 **Supplemental Table 2**). The analysis of the prevalence of the main ACVD risk factors showed a
212 significant trend for active smoking, hypertension, dyslipidemia, increased waist circumference and
213 reduced physically activity associated with an increased number of plaques (**Supplemental Table**
214 **3**). Of note, the plasma levels of a widely used marker of chronic low-grade inflammation, CRP, did
215 not differ among all the groups ($P_{\text{trend}}=0.427$ **Supplemental Table 3**), while the number of plasma
216 proteins with significantly different NPX increased as a function of the number of plaques vs
217 controls. In fact, 17 proteins were significantly different in subjects with 1 carotid plaque compared
218 to controls (1 reduced and 16 increased; **Figure 1A, Supplemental Table 4A**), 47 proteins were
219 differently expressed in the plasma of the subjects with 2 plaques (4 proteins reduced and 43
220 increased; **Figure 1B, Supplemental Table 4B**) and 97 proteins were differently expressed in
221 subjects with 3+ plaques versus controls (9 reduced and 88 increased; **Figure 1C, Supplemental**
222 **Table 4C**).

223 Overall, 32 shared proteins were differently expressed among subjects with increasing number of
224 plaques versus controls. Indeed, 6 proteins, increased in subjects with 1 plaque, were also higher in
225 subjects with 2 and in those with 3+ plaques versus controls, the NPX of 2 additional proteins were
226 also higher in subjects with 3+ plaques only (**Table 1**) and, finally, 24 proteins, 2 reduced and 22
227 increased in subjects with 2 plaques, were present also in subjects with 3+ plaques (**Table 1**).

228

229

230 **The shared proteins associated with increasing number of plaques recapitulate biological** 231 **pathways related with atherosclerosis.**

232

233 We analyzed whether the plasmatic expressions of the 32 shared proteins recapitulate biological
234 pathways associated with the development of carotid atherosclerosis. To this end, we performed a
235 protein-protein network (STRING v.11²⁹) and gene ontology (GO) enrichment analyses (see
236 methods). By interrogating *in silico* dataset, we predicted that almost all the 32 proteins could be
237 linked into a network of possible biologically relevant pathways, with Interleukin-6 (IL6), Matrix-
238 Metalloproteinase-9, Interleukin-18 (IL18) and plasminogen activator inhibitor 1 (SERPINE1)
239 being central connectors (**Figure 2A**). We then deciphered this complex scenario by gene ontology
240 (GO) enrichment analysis and, although up to 51 representative biological processes were annotated
241 (“GO_bp” in **Supplemental Table 5**), only 9 of them were principally represented by most of the
242 32 proteins with a significant probability (FDR<0.05; the fold of enrichment for each GO_bp is
243 reported in **Supplemental Table 5**). These GO_bp included pathways biologically relevant for the
244 development of atherosclerosis, such as “proteolysis”, “immune response”, “cytokine-mediated
245 signaling pathway”, “signal transduction”, “cell-cell signaling”, “positive regulation of

246 inflammatory response”, “chemotaxis” and “defense response to Gram positive bacteria” (**Figure**
247 **2B**).

248

249 **A higher number of biological pathways associates with the presence of a higher number of**
250 **carotid plaques.**

251 We also sought to investigate whether the additional proteins that were not shared among the
252 groups of subjects with increasing number of plaques, inform on non-redundant biologically
253 relevant pathways.

254 Indeed, 9 proteins (out of the 17) were differently expressed only in subjects with 1 plaque;
255 **Supplemental Table 4A**), 17 (out of the 47) only in subjects with 2 plaques; **Supplemental Table**
256 **4B**) and 65 (out of the 97) only in subjects with 3+ plaques (**Supplemental Table 4C**).

257 We found 5 GO_bp that, out of 13 annotated, barely represented the presence of one carotid plaque
258 only compared to controls ($-\log_{10}p\text{-value} \geq 1.3$; **Supplemental Table 6A**) and included “viral entry
259 into host cells” and other T-cell mediated pathways (**Figure 3A**). By contrast, the number of
260 GO_bp significantly associated with the presence of 2 plaques only increased up to 22 (out of 35
261 annotated; **Figure 3B**), with only one of them being also found to be associated with the presence of
262 1 plaque (“positive regulation of NF-kappaB transcription factor activity”; in **Supplemental Table**
263 **6B**).

264 Finally, we found that the number of GO_bp significantly associated with the presence of 3+
265 plaques markedly increased up to 80 (out of 104 annotated; **Figure 3C**). It is of note that only 6 of
266 them were also found among the pathways associated with the presence of one plaque only
267 (“response to hypoxia”, “positive regulation of protein phosphorylation”, “signal transduction”, “T
268 cell co-stimulation”, “positive regulation of peptidyl-serine phosphorylation”, “viral entry into host
269 cell”) and 16 were also found among the pathways associated with the presence of 2 plaques
270 (including pathways involved in leukocytes chemotaxis, migration, cell adhesion, inflammatory
271 response, immune response, and other more specific, such as “positive regulation of MAP kinase
272 activity”, “positive regulation of protein kinase B signaling”, “response to lipopolysaccharide”,
273 “defense response to Gram- bacterium” and “antimicrobial humoral immune response mediated by
274 antimicrobial peptide”) (**Supplemental Table 6C**).

275 This observation indicates that, especially in the presence of 3+ plaques, an increased number of
276 proteins is linked to a larger number of pathways involved in the development of atherosclerosis.

277

278 **Plasma proteins improve the accuracy of predicting the number of carotid plaques.**

279

280 We looked for the performance of plasma proteomics to improve the capability of the classical
281 ACVD risk factors to distinguish the apparently healthy subjects who are still exposed to carotid
282 atherosclerosis.

283 Only the 97 proteins that were differently expressed in subjects with 3+ plaques versus controls
284 contributed to an improvement of the ROC (AUC= 0.683 (0.601-0.785), $p < 0.001$) (**Table 2**), while
285 neither the 17 proteins that were differently expressed in subjects with only one plaque, nor the 47
286 differently expressed in subjects with 2 plaques, exerted this performance (AUC=0.504 (0.461-
287 0.591), $p=0.948$ and AUC= 0.589 (0.503-0.671), $p=0.083$, respectively (**Table 1**)).

288 Next, to find the proteins, out of the 97, whose expression in plasma can significantly improve the
289 accuracy to identify subjects with higher number of carotid plaques on top of the factors commonly
290 considered for the assessment of ACVD risk, we harnessed machine learning model which listed 87
291 important contributors for a significantly increased performance of the ROC curve. Of note, these
292 87 proteins improved the accuracy of the classical ACVD risk factors (hypertension, active smoking
293 habit, increased waist circumference, dyslipidemia and reduced physical activity) to identify the
294 subjects with 3+ plaques from controls (**Supplemental Table 7** lists the proteins by descending
295 importance factor derived from random forest classifier and their mean NPX (\pm standard error),
296 being reduced for 5 and increased for 82 proteins in subjects with 3+ plaques vs controls). Indeed,

297 the AUC of the model including these 87 proteins together with the risk factors was markedly
298 superior to the AUC of the model including the same ACVD risk factors alone (AUC= 0.918
299 (0.887-0.943) vs AUC= 0.760 (0.716-0.801) respectively, $p<0.001$; **Figure 4A**).

300 These findings support the concept that plasma proteomics independently predict the presence of a
301 higher number of carotid plaques and improves our capability to identify apparently healthy
302 subjects with diffuse carotid atherosclerosis (three or more plaques).

303

304 **Plasma proteins associated with number of carotid plaques also predict occurring ACVDs.**

305

306 Finally, to study whether the proteins that identify subjects with carotid plaques, and that
307 recapitulate biologically relevant pathways in plasma, could also improve the performance of the
308 commonly used risk factors to predict occurring ACVDs, we followed the population over 22 (21-
309 23) years.

310 198 events occurred over time (29.81% in the population). Overall, ACVD classical risk factors
311 modestly predicted the occurrence of the events (AUC= 0.559 (0.521-0.598)) and were not
312 prevalent among subjects with an incurring event. By contrast, 87 proteins that were associated with
313 the presence of 3+ plaques (including also the 32 shared proteins) markedly improved the ROC to
314 predict the event vs the ROC of the model with the same classical risk factors only (AUC=0.739
315 (0.704-0.773), $p<0.001$; **Figure 4B**).

316 Of note, the AUC of the ROC considering the 32 shared proteins only was also significantly higher
317 vs the ROC considering the ACVD risk factors only ($p=0.001$), although its performance was
318 attenuated (AUC= 0.660 (0.623-0.696); **Figure 4C**).

319

320 **Discussion**

321 In this study we searched for proteins, out of a set of 368 measured in plasma by PEA technology,
322 whose expressions can recapitulate biologically relevant pathways and can also improve the
323 accuracy to identify apparently healthy subjects presenting with carotid plaques and to predict
324 occurring ACVDs.

325 We detected proteins whose plasmatic expression significantly discriminates subjects who, although
326 being apparently healthy according to the analysis of the classical ACVD risk factors, present with
327 carotid atherosclerosis versus subjects who do not display atherosclerosis and that, in our study,
328 were considered as “controls”. Moreover, we found that the significance of this association
329 increased as a function of the number of carotid plaques analyzed by ultrasound, and we also
330 provide a specific set of proteins that were associated with diffuse atherosclerosis (3+ plaques).
331 Indeed, we identified 87 proteins, which associated with the presence of three or more plaques,
332 presenting either a reduced (5) or increased (82) plasmatic expression compared to controls. By
333 contrast, plasma proteomics was not able to significantly discriminate neither the subjects
334 presenting with 2 or with just one carotid plaque. The finding that proteomics did not discriminate
335 the subjects with lower number of plaques could perhaps be due to the smaller number of subjects
336 in these groups. Anyhow, we cannot rule out the effect of the clinical conditions that mostly
337 contribute to hemorheological turbulence and endothelial damage, such as hypertension and
338 smoking habit, increased waist circumference, dyslipidemia and reduced physical activity, which
339 were prevalent in subjects with 3+ plaques but not in the subjects with 2 or with 1 plaque only
340 versus controls and which were also associated with higher NPX values for the majority of proteins
341 analyzed (**Supplemental Figure 2A-E**).

342 The AUC of the model combining the information of these 87 proteins with that of the classical
343 ACVD risk factors to discriminate subjects with 3+ plaques compared to controls was markedly
344 elevated versus the AUC of the model considering the risk factors alone. Also, this significant
345 improvement held true between a subset of the two groups, matched for prevalence of hypertension,
346 active smoking habit, dyslipidemia, frequency of physical activity and similar waist circumference

347 (AUC= 0.840 (0.771-0.895) of the model with proteins and ACVD risk factors vs AUC= 0.684
348 (0.602-0.757) with ACVD risk factors alone, $p < 0.001$; not shown). Therefore, this observation
349 further supports that plasma proteomics can improve our capability to classify, with higher
350 accuracy, individuals among the apparently healthy population towards higher degrees of ACVD
351 risk. Anyhow, we acknowledge that larger and independent studies are needed to validate this
352 hypothesis.

353 We also reported comparable levels of CRP, a commonly considered marker of low grade
354 inflammation in clinics, in subjects with 3+ plaques versus controls. We do not rule out the
355 relevance of CRP, which is generally elevated in subjects at higher ACVD risk and co-morbid
356 patients, but we are more likely to support the possibility that pre-clinical inflammation in
357 “apparently healthy” individuals could be recaptured more significantly by alternative sets of
358 proteins in plasma.

359 Furthermore, by combining a high-sensitive gradient boosting machine learning strategy together
360 with the analysis of pathways, we also provided a biological association for the large number of
361 proteins that were predicted to be linked to the number of carotid plaques. For instance, Growth
362 Differentiation Factor 15 (GDF-15), whose NPX was increased in subjects with 3+ plaques versus
363 controls, ranged first out of the 87 proteins in the random forest classifier (**Supplemental Table 7**),
364 confirming independent epidemiological studies that previously associated its increased plasma
365 levels to ACVD risk^{8,30}, to carotid atherosclerosis, increased carotid IMT^{7,9,10} and high-risk
366 coronary plaques defined by computed tomography angiography³¹. However, our additional GO
367 analysis found that alternative set of proteins, including the spectrum of β -chemokines and other
368 interleukins/cytokines, but not GDF-15, significantly clustered into biological pathways involved in
369 the development of atherosclerotic process (**Supplemental Table 5** and **Supplemental table 6C**).
370 From a translational point of view, these findings outline further mechanisms, involving immune
371 cells chemotaxis in atherosclerotic lesions³² and lymphocytes activation in the atheroma³³, taking
372 place yet during the subclinical stages of atherosclerosis.

373 Finally, the robustness of our predictions is also supported by our follow-up analysis. Indeed, the 87
374 proteins (including the 32 shared proteins), identifying subjects with higher number of plaques,
375 significantly predicted incident ACVDs over a large twenty-years follow-up. From a clinical
376 standpoint, the finding that proteomics also markedly outperformed the accuracy of ACVD risk
377 factors, further supports the possibility to overcome, by identifying individual proteomic
378 fingerprints, the clinical algorithms which often underestimate the residual risk in apparently
379 healthy subjects, owing a ten-year perspective for most of them.

380 Our study has some limitations, and some technical aspects of our work should be discussed.

381 First, our machine learning model, although previously demonstrated to provide superior accuracy
382 compared to common statistical tools²⁵, was trained and applied in internal sets of our population
383 and, consequently, our findings need to be validated in independent cohorts.

384 Second, the methodological assessment of our proteomics analysis, although providing an optimal
385 sensitivity to target low proteins abundance in plasma, otherwise not reachable by common
386 techniques of molecular biology³⁴, relies on PCR-based amplification reactions with antibodies-
387 linked oligonucleotides. A quantitative validation of our results will add further strength; However,
388 for some of the proteins we have identified, a validation by immunoassays has been provided³⁴. All
389 in all, compared existing literature using this strategy^{9,10,16}, we have provided information from one
390 of the largest panels so far reported.

391 In conclusion, we found plasma proteins that, informing on pathways that are biologically relevant
392 for the development of atherosclerosis, can identify apparently healthy subjects with higher number
393 of plaques and that outperform the classical ACVD risk factors to predict occurring atherosclerotic
394 events.

395
396
397

Acknowledgments

398 The work of A.B and A.L.C. is supported by a grant Ricerca Corrente from the Ministry of Health
399 to IRCCS Multimedica; PRIN 2017H5F943; ERANET ER-2017-2364981 and a grant A.L.C. RF-
400 2019-12370896
401

402 **Conflicts of Interest:** The authors declare no Conflicts of Interest relevant to the submitted work.

403

404

405 **Figures and Tables Legends**

406

407 **Figure 1. Variations of plasma proteomics between subjects with one, with 2 and with 3+** 408 **carotid plaques compared to controls.**

409 Volcano plots showing the variation of the 368 plasma proteins in (A) subjects with 1 carotid
410 plaque, in (B) subjects with 2 carotid plaques and (C) in subjects with 3+ plaques. All the variations
411 are referred to controls and are calculated as fold of changes (x axis). In the y axis, the p value is
412 reported indicating the significance of the variation.

413 **Figure 2. The shared proteins associated with increasing number of plaques recapitulate** 414 **biological pathways involved in atherosclerosis.**

415 (A) STRING protein-protein association network reporting the predicted connections between the
416 32 shared proteins associated with the number of carotid plaques. (B) Bubble plot showing the most
417 significant GO_bp processes represented by the 32 shared plasma proteins. Each pathway is plotted
418 for its fold of enrichment and the significance in the intensity (-log₁₀ FDR) of the signal from
419 proteome analysis (see methods). Significant signals were considered for FDR<0.05.

420 **Figure 3. A higher number of biological pathways associates with the presence of a higher** 421 **number of carotid plaques.**

422 Bubble plot showing the most significant GO_bp processes represented by (A) the unshared
423 proteins associated the presence of one plaque only, (B) the unshared proteins associated the
424 presence of two plaques and (C) the unshared proteins associated the presence of 3+ plaques.

425 **Figure 4. The plasma proteomics outperforms the ACVD risk factors in identifying subjects** 426 **with 3+ plaques and in predicting occurring atherosclerotic events.**

427 Comparison of the accuracy (Receiver Operating Curve “ROC”) of the model to discriminate: (A)
428 subjects with 3+ plaques from the controls, when including the ACVD risk factors together with the
429 87 proteins (including the 32 shared proteins) differently expressed in subjects with 3+ plaques (red
430 curve) versus the accuracy of the model that included the ACVD risk factors only (blue curve); (B)
431 subjects with incident ACVD, when including the ACVD risk factors and the 87 proteins associated
432 with carotid plaques (including the 32 shared proteins) (red curve), versus the accuracy of the model
433 that included the ACVD risk factors only (blue curve); (C) subjects with incident ACVD, when
434 including the ACVD risk factors and the 32 shared proteins only (red curve), versus the accuracy of
435 the model that included the ACVD risk factors only (blue curve); P values are derived from Hanley-
436 McNeil test.

437 **Supplemental Figure 1. Study design.**

438 **Supplemental Figure 2. The plasmatic expression of several proteins associated with the**
439 **ACVD risk factors.**

440 Volcano plots showing the variation of the 368 plasma proteins in **(A)** subjects with hypertension
441 compared to subjects without hypertension, in **(B)** active smokers versus non-active smokers, **(C)** in
442 subjects with higher waist circumference versus subjects with lower waist circumference, **(D)** in
443 subjects with dyslipidemia and **(E)** in subjects who declare to practice less physical activity versus
444 subjects who declared to practice more physical activity. All the variations are referred to controls
445 and are calculated as fold of changes (x axis). In the y axis, the p value is reported indicating the
446 significance of the variation.

447

448 **Table 1. Shared proteins found significantly associated among the subjects with one, two or**
449 **with 3+ carotid plaques compared to controls.**

450 List of proteins with their name, their Uniprot-ID and their fold change in **(A)** subjects with one
451 plaque only, in **(B)** subjects with two plaques and **(C)** in subjects with 3+ plaques. All the fold
452 changes are referred compared to controls. Proteins that are shared between subjects with 2 and
453 subjects with 3+ plaques only are underlined.

454

455 **Table 2. Targeted plasma proteomics outperforms the identification of carotids with higher**
456 **number of plaques.**

457 Table reporting the parameters of the Receiving Operating Curves (ROC) for the models including
458 the proteins differently expressed either in subjects with one plaque only, or in subjects with 2 or in
459 subjects with 3+ plaques compared to controls. The mean Areas Under the Curve (AUC), the 95%
460 confidence intervals (C.I.), the z-statistics and the p-value for each ROC are reported.

461

462 **References**

- 463 1. Vaduganathan M, Mensah GA, Turco JV, Fuster V, Roth GA. The Global Burden of Cardiovascular
464 Diseases and Risk: A Compass for Future Health. *J Am Coll Cardiol* [Internet]. 2022 [cited 2023 Feb
465 12];80:2361–2371. Available from: <https://www.jacc.org/doi/10.1016/j.jacc.2022.11.005>
- 466 2. Fernández-Friera L, Peñalvo JL, Fernández-Ortiz A, Ibañez B, López-Melgar B, Laclaustra M, Oliva B,
467 Moco-roa A, Mendiguren J, De Vega VM, et al. Prevalence, Vascular Distribution, and Multiterritorial
468 Extent of Subclinical Atherosclerosis in a Middle-Aged Cohort: The PESA (Progression of Early
469 Subclinical Atherosclerosis) Study. *Circulation* [Internet]. 2015 [cited 2022 Sep 14];131:2104–2113.
470 Available from: <https://pubmed.ncbi.nlm.nih.gov/25882487/>
- 471 3. Ibanez B, Fernández-Ortiz A, Fernández-Friera L, García-Lunar I, Andrés V, Fuster V. Progression of
472 Early Subclinical Atherosclerosis (PESA) Study: JACC Focus Seminar 7/8. *J Am Coll Cardiol* [Internet].
473 2021 [cited 2022 Sep 14];78:156–179. Available from: <https://pubmed.ncbi.nlm.nih.gov/34238438/>
- 474 4. Nicolaidis AN, Panayiotou AG, Griffin M, Tyllis T, Bond D, Georgiou N, Kyriacou E, Avraamides C,
475 Martin RM. Arterial Ultrasound Testing to Predict Atherosclerotic Cardiovascular Events. *J Am Coll*
476 *Cardiol* [Internet]. 2022;79:1969–1982. Available from: <https://pubmed.ncbi.nlm.nih.gov/35589158/>

- 477 5. McKenna MA, Bonfield MC, Robinson T. Levels of agreement in the measurements of carotid artery
478 ultrasound across a regional vascular network. *Ultrasound* [Internet]. 2018 [cited 2023 Apr
479 17];26:101–109. Available from: <https://pubmed.ncbi.nlm.nih.gov/30013610/>
- 480 6. Polak JF, Pencina MJ, Herrington D, O’Leary DH. Associations of edge-detected and manual-traced
481 common carotid intima-media thickness measurements with Framingham risk factors: the multi-
482 ethnic study of atherosclerosis. *Stroke* [Internet]. 2011 [cited 2023 Apr 17];42:1912–1916. Available
483 from: <https://pubmed.ncbi.nlm.nih.gov/21546477/>
- 484 7. Hoogeveen RM, Pereira JPB, Nurmohamed NS, Zampoleri V, Bom MJ, Baragetti A, Boekholdt SM,
485 Knaapen P, Khaw K-T, Wareham NJ, et al. Improved cardiovascular risk prediction using targeted
486 plasma proteomics in primary prevention. *Eur Heart J* [Internet]. 2020 [cited 2020 Sep 9];Available
487 from: <https://pubmed.ncbi.nlm.nih.gov/32808014/>
- 488 8. Nurmohamed NS, Belo Pereira JP, Hoogeveen RM, Kroon J, Kraaijenhof JM, Waissi F, Timmerman N,
489 Bom MJ, Hoefler IE, Knaapen P, et al. Targeted proteomics improves cardiovascular risk prediction in
490 secondary prevention. *Eur Heart J* [Internet]. 2022 [cited 2022 Feb 15];Available from:
491 <https://pubmed.ncbi.nlm.nih.gov/35139537/>
- 492 9. Lind L, Gigante B, Borne Y, Mälärstig A, Sundström J, Ärnlov J, Ingelsson E, Baldassarre D, Tremoli E,
493 Veglia F, et al. The plasma protein profile and cardiovascular risk differ between intima-media
494 thickness of the common carotid artery and the bulb: A meta-analysis and a longitudinal evaluation.
495 *Atherosclerosis* [Internet]. 2020 [cited 2023 Apr 17];295:25–30. Available from:
496 <https://pubmed.ncbi.nlm.nih.gov/31981948/>
- 497 10. Lind L, Gigante B, Borné Y, Feldreich T, Leppert J, Hedberg P, Ostgren CJ, Nystrom FH, Sundstrom J,
498 Arnlov J, et al. Plasma Protein Profile of Carotid Artery Atherosclerosis and Atherosclerotic
499 Outcomes: Meta-Analyses and Mendelian Randomization Analyses. *Arterioscler Thromb Vasc Biol*
500 [Internet]. 2021 [cited 2021 Dec 26];41:1777–1788. Available from:
501 <https://www.ahajournals.org/doi/abs/10.1161/ATVBAHA.120.315597>
- 502 11. Libby P. Inflammation in Atherosclerosis - No Longer a Theory. *Clin Chem*. 2021;67:131–142.
- 503 12. What is NPX? - Olink [Internet]. [cited 2022 Sep 2];Available from: [https://www.olink.com/faq/what-](https://www.olink.com/faq/what-is-npx/)
504 [is-npx/](https://www.olink.com/faq/what-is-npx/)
- 505 13. Joshi A, Rienks M, Theofilatos K, Mayr M. Systems biology in cardiovascular disease: a multiomics
506 approach. *Nat Rev Cardiol* [Internet]. 2021;18:313–330. Available from:
507 <https://pubmed.ncbi.nlm.nih.gov/33340009/>
- 508 14. Nurmohamed NS, Kraaijenhof JM, Mayr M, Nicholls SJ, Koenig W, Catapano AL, Stroes ESG.
509 Proteomics and lipidomics in atherosclerotic cardiovascular disease risk prediction. *Eur Heart J*
510 [Internet]. 2023;Available from: <https://pubmed.ncbi.nlm.nih.gov/36988179/>
- 511 15. Borrelli A, Marrano N, Biondi G, Rella M, Roberto L, Cignarelli A, Perrini S, Laviola L, Giorgino F,
512 Natalicchio A, et al. Acute ischemic stroke: how to investigate the association between disease
513 etiology and gene expression profiles: Selected Abstract – Spring Meeting 2023. *European*
514 *Atherosclerosis Journal* [Internet]. 2023 [cited 2023 Sep 8];2:20–20. Available from:
515 <https://www.eathj.org/index.php/eaj/article/view/32>
- 516 16. Sandstedt J, Vargmar K, Björkman K, Ruetschi U, Bergström G, Hultén LM, Skiöldebrand E. COMP
517 (Cartilage Oligomeric Matrix Protein) Neoepitope: A Novel Biomarker to Identify Symptomatic
518 Carotid Stenosis. *Arterioscler Thromb Vasc Biol* [Internet]. 2021;41:1218–1228. Available from:
519 <https://pubmed.ncbi.nlm.nih.gov/33472398/>

- 520 17. Lind L, Ärnlöv J, Lindahl B, Siegbahn A, Sundström J, Ingelsson E. Use of a proximity extension assay
521 proteomics chip to discover new biomarkers for human atherosclerosis. *Atherosclerosis* [Internet].
522 2015;242:205–210. Available from: <https://pubmed.ncbi.nlm.nih.gov/26204497/>
- 523 18. Olmastroni E, Baragetti A, Casula M, Grigore L, Pellegatta F, Pirillo A, Tragni E, Catapano AL.
524 Multilevel Models to Estimate Carotid Intima-Media Thickness Curves for Individual Cardiovascular
525 Risk Evaluation. *Stroke* [Internet]. 2019 [cited 2019 Jul 8];50:1758–1765. Available from:
526 <http://www.ncbi.nlm.nih.gov/pubmed/31164073>
- 527 19. Baragetti A, Mattavelli E, Grigore L, Pellegatta F, Magni P, Catapano AL. Targeted Plasma Proteomics
528 to Predict the Development of Carotid Plaques. *Stroke* [Internet]. 2022 [cited 2022 Sep 1];53.
529 Available from: <https://pubmed.ncbi.nlm.nih.gov/35775410/>
- 530 20. Heidari S, Babor TF, De Castro P, Tort S, Curno M. Sex and Gender Equity in Research: rationale for
531 the SAGER guidelines and recommended use. *Res Integr Peer Rev* [Internet]. 2016 [cited 2022 Dec
532 15];1:1–9. Available from:
533 <https://researchintegrityjournal.biomedcentral.com/articles/10.1186/s41073-016-0007-6>
- 534 21. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening the
535 Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting
536 observational studies. *Lancet* [Internet]. 2007;370:1453–1457. Available from:
537 <https://pubmed.ncbi.nlm.nih.gov/18064739/>
- 538 22. Baragetti A, Severgnini M, Olmastroni E, Dioguardi C, Mattavelli E, Angius A, Rotta L, Cibella J,
539 Caredda G, Consolandi C, et al. Gut Microbiota Functional Dysbiosis Relates to Individual Diet in
540 Subclinical Carotid Atherosclerosis. *Nutrients* [Internet]. 2021;13:1–19. Available from:
541 <https://pubmed.ncbi.nlm.nih.gov/33494335/>
- 542 23. Mach F, Baigent C, Catapano AL, Koskinas KC, Casula M, Badimon L, Chapman MJ, Backer GG De,
543 Delgado V, Ference BA, et al. 2019 ESC/EAS Guidelines for the management of dyslipidaemias: Lipid
544 modification to reduce cardiovascular risk. *Eur Heart J*. 2020;41:111–188.
- 545 24. Johri AM, Nambi V, Naqvi TZ, Feinstein SB, Kim ESH, Park MM, Becher H, Sillesen H.
546 Recommendations for the Assessment of Carotid Arterial Plaque by Ultrasound for the
547 Characterization of Atherosclerosis and Evaluation of Cardiovascular Risk: From the American
548 Society of Echocardiography. *Journal of the American Society of Echocardiography*. 2020;33:917–
549 933.
- 550 25. Unterhuber M, Kresoja KP, Rommel KP, Besler C, Baragetti A, Klötting N, Ceglarek U, Blüher M, Scholz
551 M, Catapano AL, et al. Proteomics-Enabled Deep Learning Machine Algorithms Can Enhance
552 Prediction of Mortality. *J Am Coll Cardiol* [Internet]. 2021 [cited 2021 Dec 26];78:1621–1631.
553 Available from: <https://pubmed.ncbi.nlm.nih.gov/34649700/>
- 554 26. <https://www.javatpoint.com/cost-function-in-machine-learning>.
- 555 27. Baragetti A, Da Dalt L, Moregola A, Svecla M, Terenghi O, Mattavelli E, De Gaetano LN, Uboldi P,
556 Catapano AL, Norata GD. Neutrophil aging exacerbates high fat diet induced metabolic alterations.
557 *Metabolism*. 2023;144.
- 558 28. Nour J, Moregola A, Svecla M, Da Dalt L, Bellini R, Neyrolles O, Fadini GP, Rombouts Y, Albiero M,
559 Bonacina F, et al. Mannose Receptor Deficiency Impacts Bone Marrow and Circulating Immune Cells
560 during High Fat Diet Induced Obesity. *Metabolites*. 2022;12.

- 561 29. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, Simonovic M, Doncheva NT,
562 Morris JH, Bork P, et al. STRING v11: Protein-protein association networks with increased coverage,
563 supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.*
564 2019;47:D607–D613.
- 565 30. Wallentin L, Eriksson N, Olszowka M, Grammer TB, Hagström E, Held C, Kleber ME, Koenig W, März
566 W, Stewart RAH, et al. Plasma proteins associated with cardiovascular death in patients with chronic
567 coronary heart disease: A retrospective study. *PLoS Med* [Internet]. 2021 [cited 2023 Apr 18];18.
568 Available from: <https://pubmed.ncbi.nlm.nih.gov/33439866/>
- 569 31. Bom MJ, Levin E, Driessen RS, Danad I, Van Kuijk CC, van Rossum AC, Narula J, Min JK, Leipsic JA,
570 Belo Pereira JP, et al. Predictive value of targeted proteomics for coronary plaque morphology in
571 patients with suspected coronary artery disease. *EBioMedicine* [Internet]. 2019 [cited 2022 Sep
572 1];39:109–117. Available from: <https://pubmed.ncbi.nlm.nih.gov/30587458/>
- 573 32. de Jager SCA; BI, KAOKSJABM van SPJ van BTJCKJBEAL. Leukocyte-specific CCL3 deficiency inhibits
574 atherosclerotic lesion development by affecting neutrophil accumulation.
- 575 33. Schaller TH, Batich KA, Suryadevara CM, Desai R, Sampson JH. Chemokines as adjuvants for
576 immunotherapy: implications for immune activation with CCL3. *Expert Rev Clin Immunol.*
577 2017;13:1049–1060.
- 578 34. Joshi A, Rienks M, Theofilatos K, Mayr M. Systems biology in cardiovascular disease: a multiomics
579 approach. *Nat Rev Cardiol* [Internet]. 2021 [cited 2021 Oct 23];18:313–330. Available from:
580 <https://pubmed.ncbi.nlm.nih.gov/33340009/>
- 581
- 582
- 583

584 **Table 1. Shared proteins found significantly associated among the subjects with one, two or**
 585 **with 3+ carotid plaques compared to controls.**

(A)

(B)

(C)

586

Protein	Uniprot ID	fold change	Protein	Uniprot ID	fold change	p value	Protein	Uniprot ID	fold change	p value
TFPI	P10646	0,01	TNFRSF10A	O00220	0,05	3,60	GDF15	Q99988	0,05	6,95
APN	P15144	0,02	HAVCR1	Q96D42	0,04	3,57	IL6	P05231	0,10	5,68
IL1RT2	P27930	0,02	CSTB	P04080	0,04	3,20	HAVCR1	Q96D42	0,05	5,62
TNFSF13B	Q9Y275	0,01	IL18	Q14116	0,04	3,14	PRSS8	Q16651	0,02	5,55
CPB1	P15086	0,03	IGFBP7	Q16270	0,02	2,75	CSTB	P04080	0,05	5,31
CPA1	P15085	0,04	GDF15	Q99988	0,03	2,62	TNFSF13B	Q9Y275	0,02	5,10
TNFRSF10A	O00220	0,03	PON3	Q15166	-0,06	2,53	HGF	P14210	0,03	4,85
IL16	Q14005	0,04	PAI1	P05121	0,05	2,46	CHI3L1	P36222	0,05	4,47
			CTSD	P07339	0,03	2,38	IL18	Q14116	0,04	4,26
			TRAP6	P13686	0,07	2,18	IGFBP7	Q16270	0,02	3,97
			PCOLCE	Q15113	0,04	2,04	TRAP6	P13686	0,09	3,69
			HGF	P14210	0,02	1,98	TNFRSF10A	O00220	0,04	3,25
			PRSS8	Q16651	0,01	1,87	PLAUR	Q03405	0,03	3,25
			IL6	P05231	0,06	1,86	CTSD	P07339	0,03	3,23
			PGF	P49763	0,01	1,83	PGF	P49763	0,02	3,11
			IL7R	P16871	-0,09	1,72	IL7R	P16871	-0,12	2,98
			CHI3L1	P36222	0,03	1,70	IL1RT2	P27930	0,02	2,85
			CPB1	P15086	0,03	1,65	TFPI	P10646	0,01	2,72
			APN	P15144	0,02	1,63	IL16	Q14005	0,04	2,45
			CCL20	P78556	0,04	1,61	TNFR1	P19438	0,02	2,30
			IL1RT2	P27930	0,01	1,61	MMP9	P14780	0,05	2,28

<u>PLAUR</u>	<u>Q0340</u>			<u>CEACA</u>	<u>P3199</u>		
	<u>5</u>	<u>0,02</u>	<u>1,53</u>	<u>M8</u>	<u>7</u>	<u>0,06</u>	<u>2,22</u>
<u>CEACA</u>	<u>P3199</u>			<u>APN</u>	<u>P1514</u>		
	<u>7</u>	<u>0,05</u>	<u>1,50</u>		<u>4</u>	<u>0,02</u>	<u>2,16</u>
<u>MMP9</u>	<u>P1478</u>			<u>CCL16</u>	<u>O1546</u>		
	<u>0</u>	<u>0,04</u>	<u>1,42</u>		<u>7</u>	<u>0,02</u>	<u>2,09</u>
<u>TNFR1</u>	<u>P1943</u>			<u>LOX1</u>	<u>P7838</u>		
	<u>8</u>	<u>0,01</u>	<u>1,41</u>		<u>0</u>	<u>0,03</u>	<u>2,02</u>
<u>IL16</u>	<u>Q1400</u>			<u>PCSK9</u>	<u>Q8NB</u>		
	<u>5</u>	<u>0,03</u>	<u>1,39</u>		<u>P7</u>	<u>0,03</u>	<u>2,01</u>
<u>LOX1</u>	<u>P7838</u>			<u>PAI</u>	<u>P0512</u>		
	<u>0</u>	<u>0,03</u>	<u>1,35</u>		<u>1</u>	<u>0,04</u>	<u>1,88</u>
<u>CCL16</u>	<u>O1546</u>			<u>CCL20</u>	<u>P7855</u>		
	<u>7</u>	<u>0,02</u>	<u>1,34</u>		<u>6</u>	<u>0,04</u>	<u>1,72</u>
<u>CPA1</u>	<u>P1508</u>			<u>CPA1</u>	<u>P1508</u>		
	<u>5</u>	<u>0,03</u>	<u>1,34</u>		<u>5</u>	<u>0,03</u>	<u>1,64</u>
<u>PCSK9</u>	<u>Q8NB</u>			<u>CPB1</u>	<u>P1508</u>		
	<u>P7</u>	<u>0,03</u>	<u>1,30</u>		<u>6</u>	<u>0,03</u>	<u>1,58</u>
				<u>PCOLC</u>	<u>Q1511</u>		
				<u>E</u>	<u>3</u>	<u>0,05</u>	<u>1,63</u>
				<u>PON3</u>	<u>Q1516</u>		
					<u>6</u>	<u>-0,04</u>	<u>1,36</u>

587

588

589 **Table 2. Targeted plasma proteomics outperforms the identification of carotids with higher**
 590 **number of plaques.**

Performance of the models to distinguish:	AUC	Standard error	95% C.I. (lower-upper)	z-statistic	p-value
Number of plaques					
1 plaque vs controls	0.504	0.0561	0.461-0.591	0.065	0.948
2 plaques vs controls	0.589	0.0513	0.503-671	1.735	0.083
3+ plaques vs controls	0.683	0.0471	0.601-0.785	3.892	<0.001

591

592

593

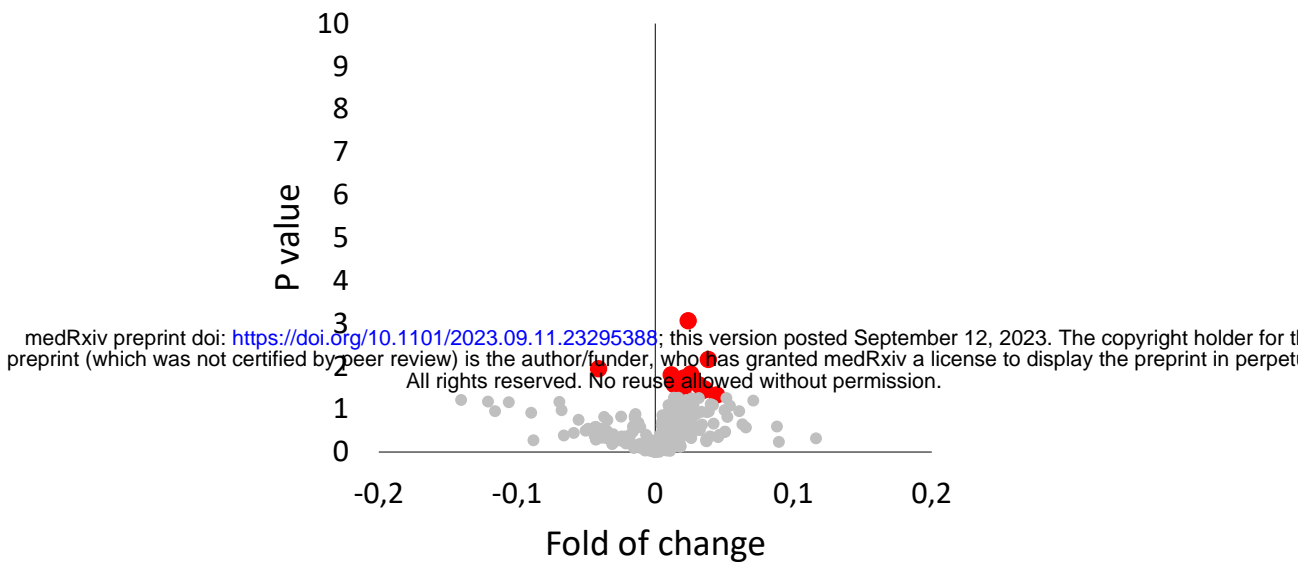
594

595

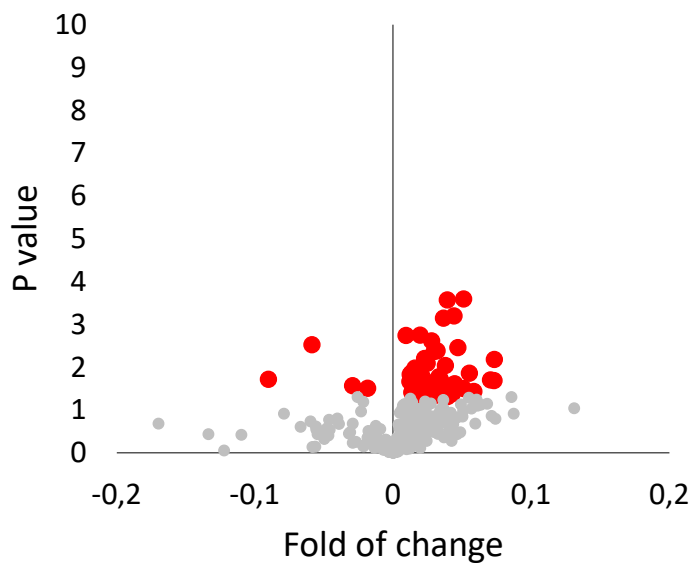
596

597

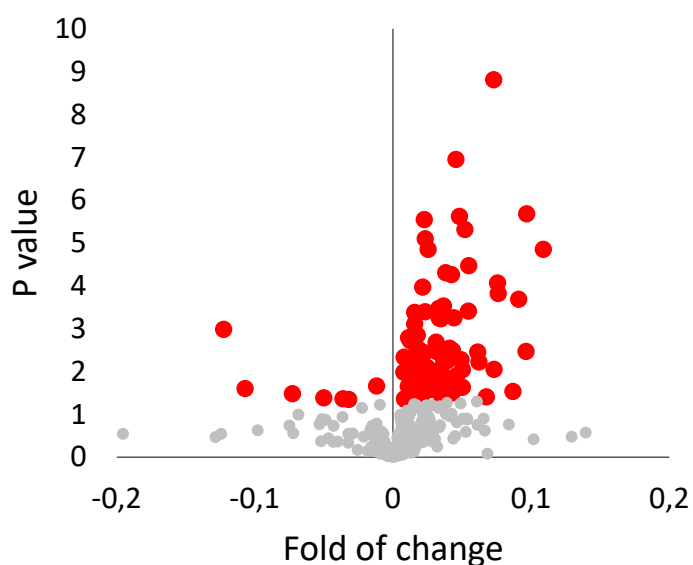
(A) 1 vs no carotid plaques



(B) 2 vs no carotid plaques



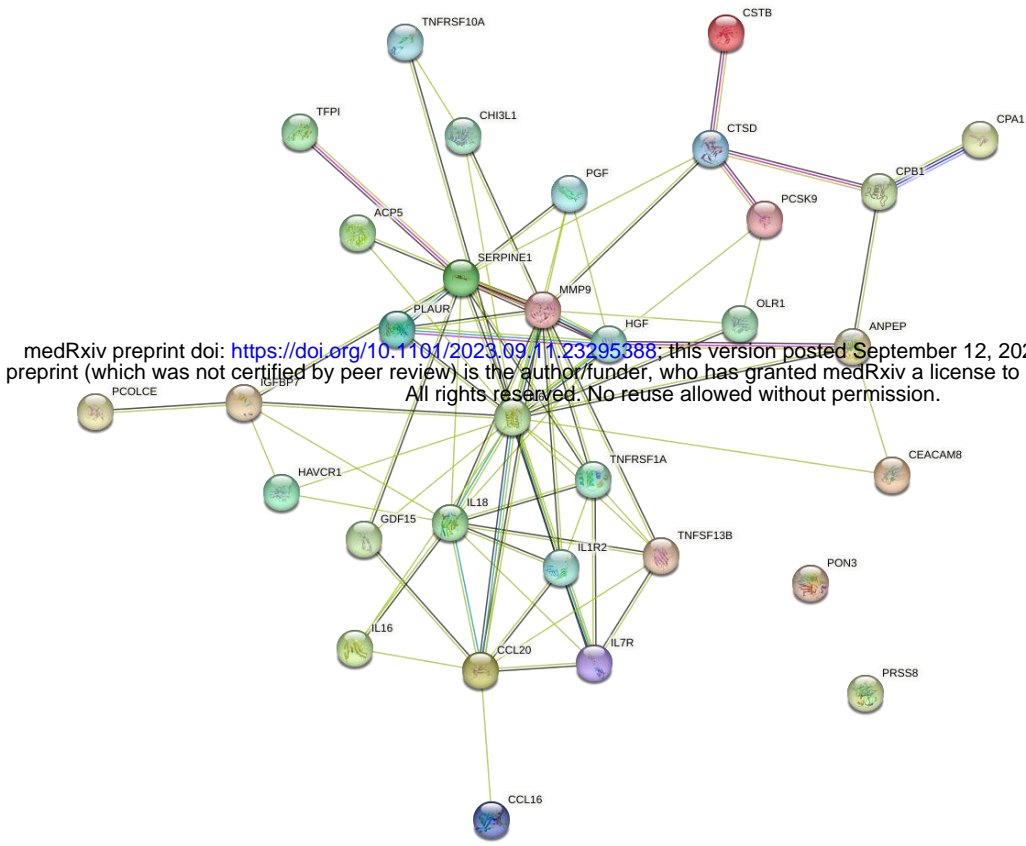
(C) 3+ vs no carotid plaques



- Proteins with significant variation**
- Proteins with no significant variation**

Figure 1.

(A)



(B)

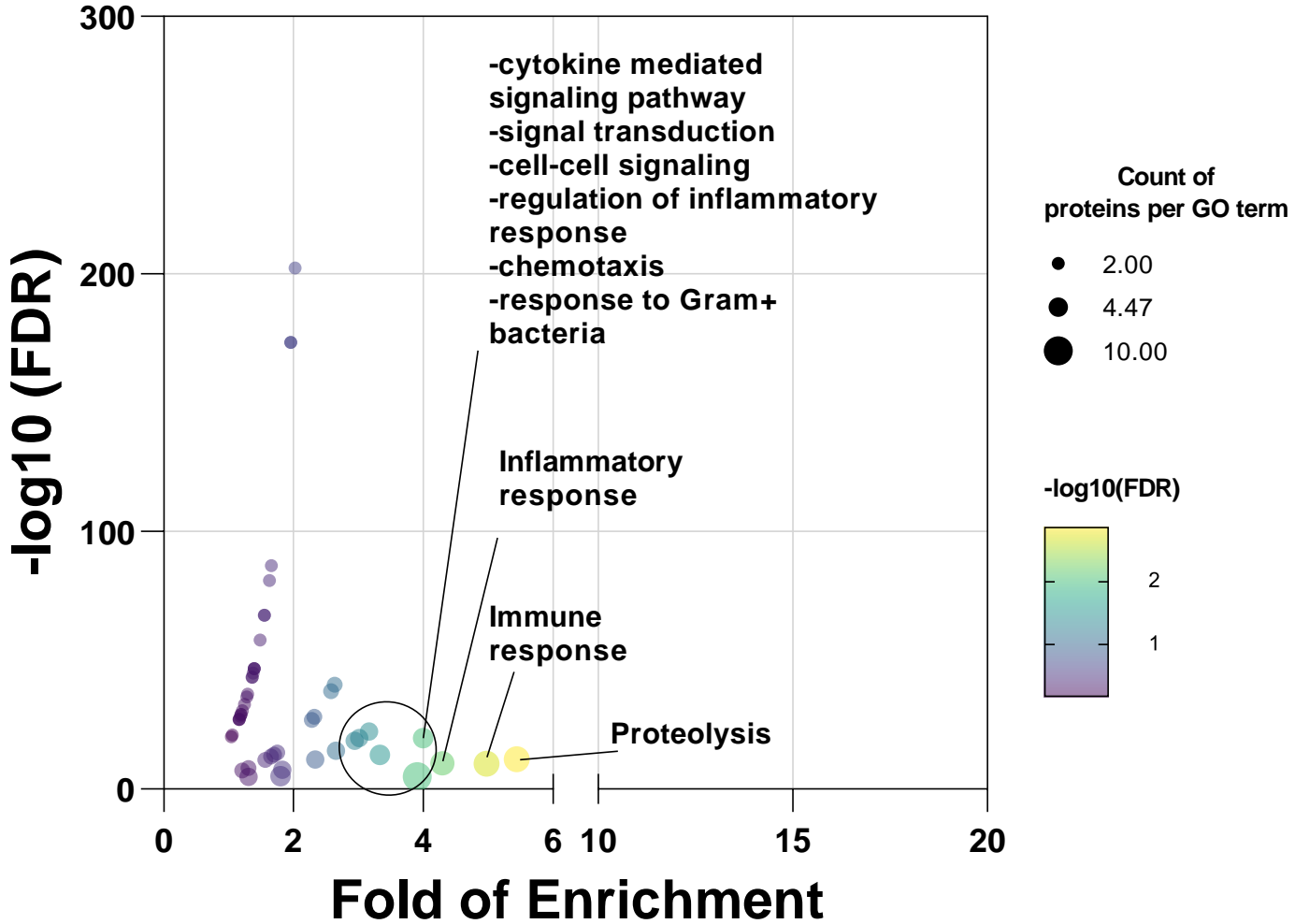
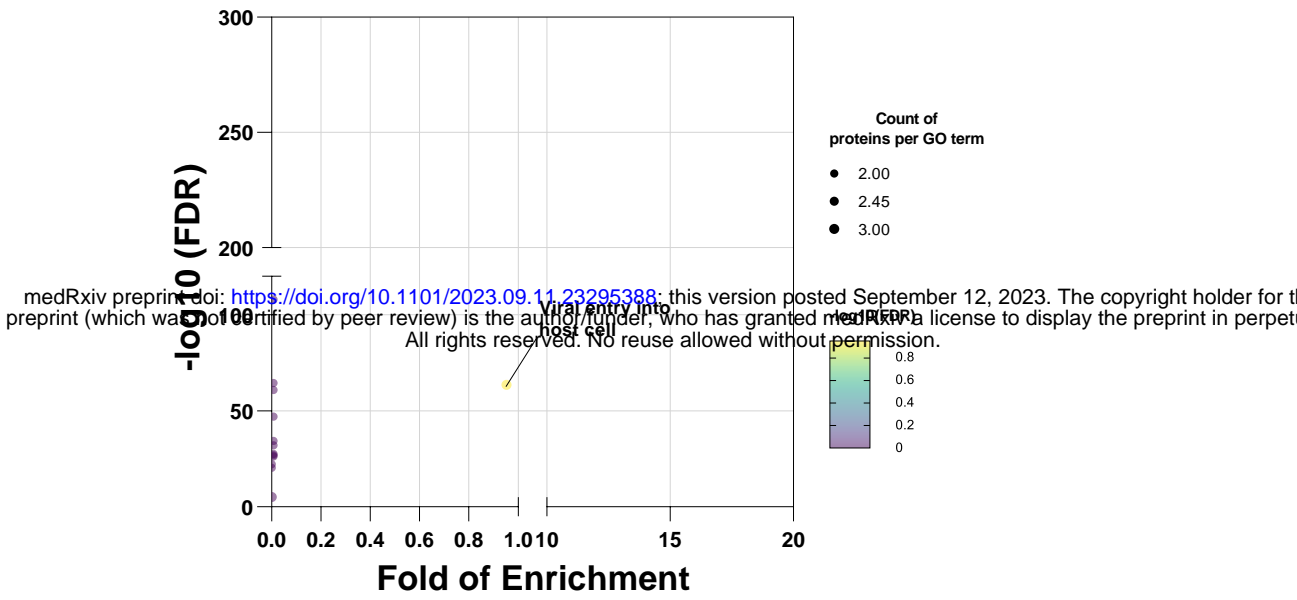
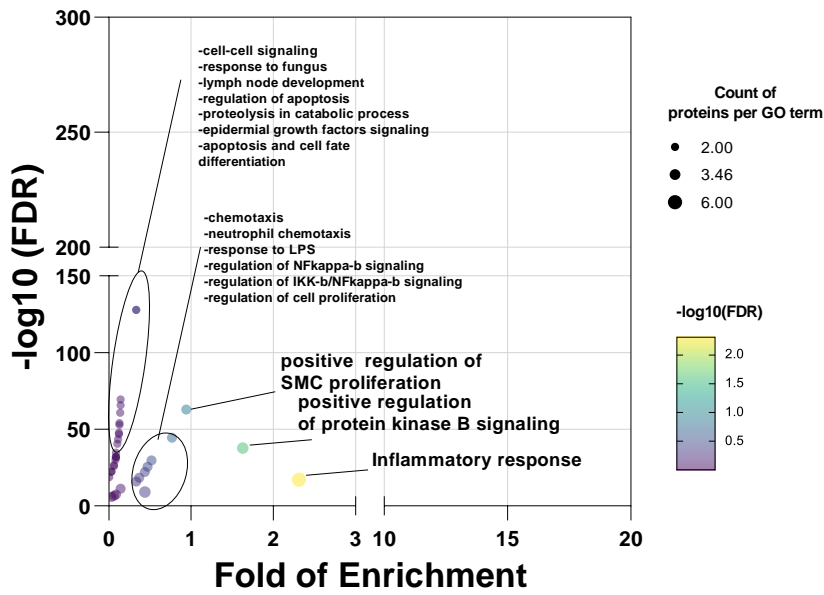


Figure 2.

(A) Biological pathways enriched with the proteins associated with the presence of 1 carotid plaque only



(B) Biological pathways enriched with the proteins associated with the presence of 2 carotid plaques only



(C) Biological pathways enriched with the proteins associated with the presence of 3+ carotid plaques only

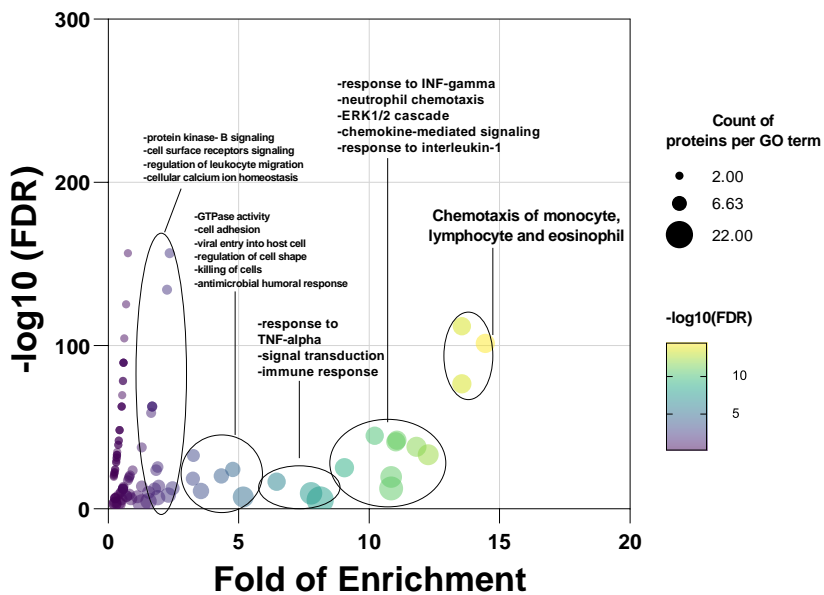
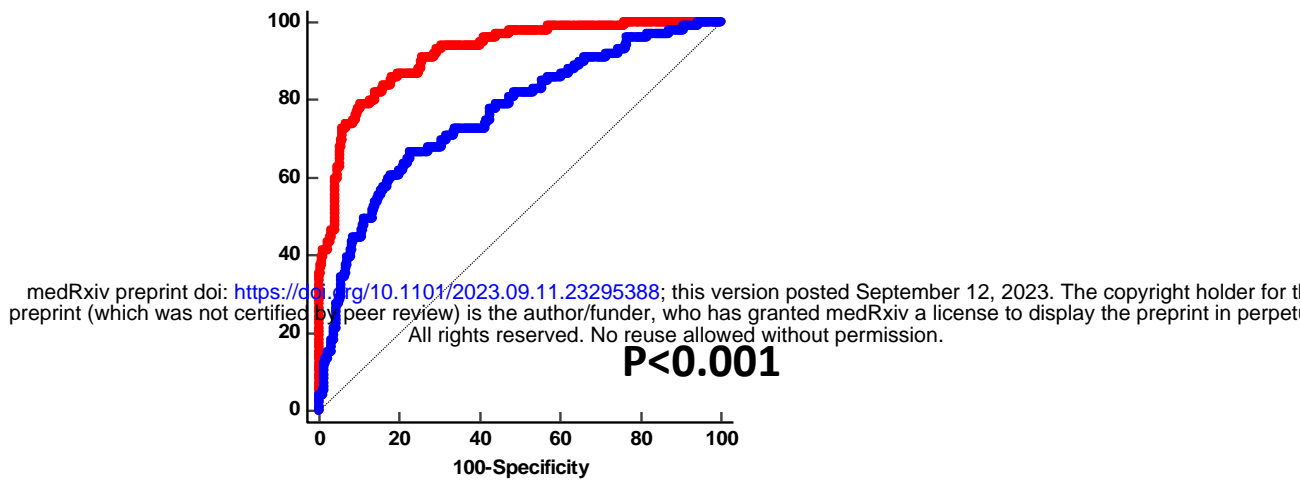
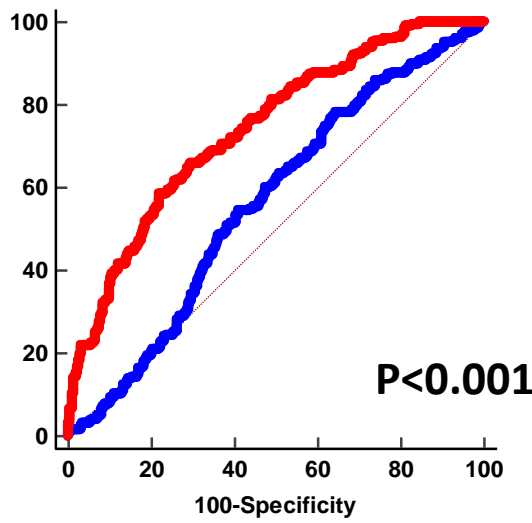


Figure 3.

(A) Performance of 87 proteins in identifying 3+ plaques



(B) Performance of all the proteins associated with plaques in predicting occurring ACVDs



(C) Performance of 32 shared proteins in predicting occurring ACVDs

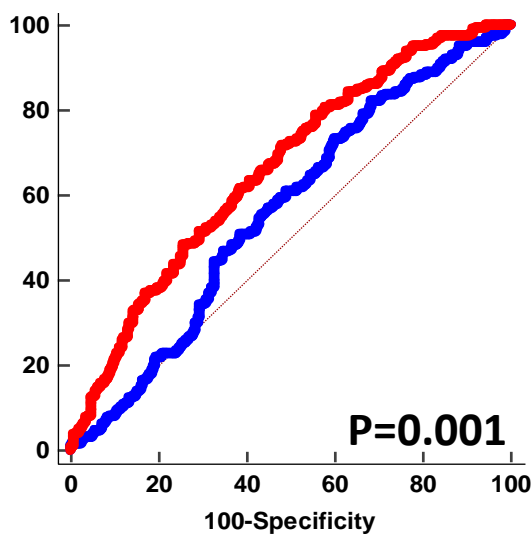
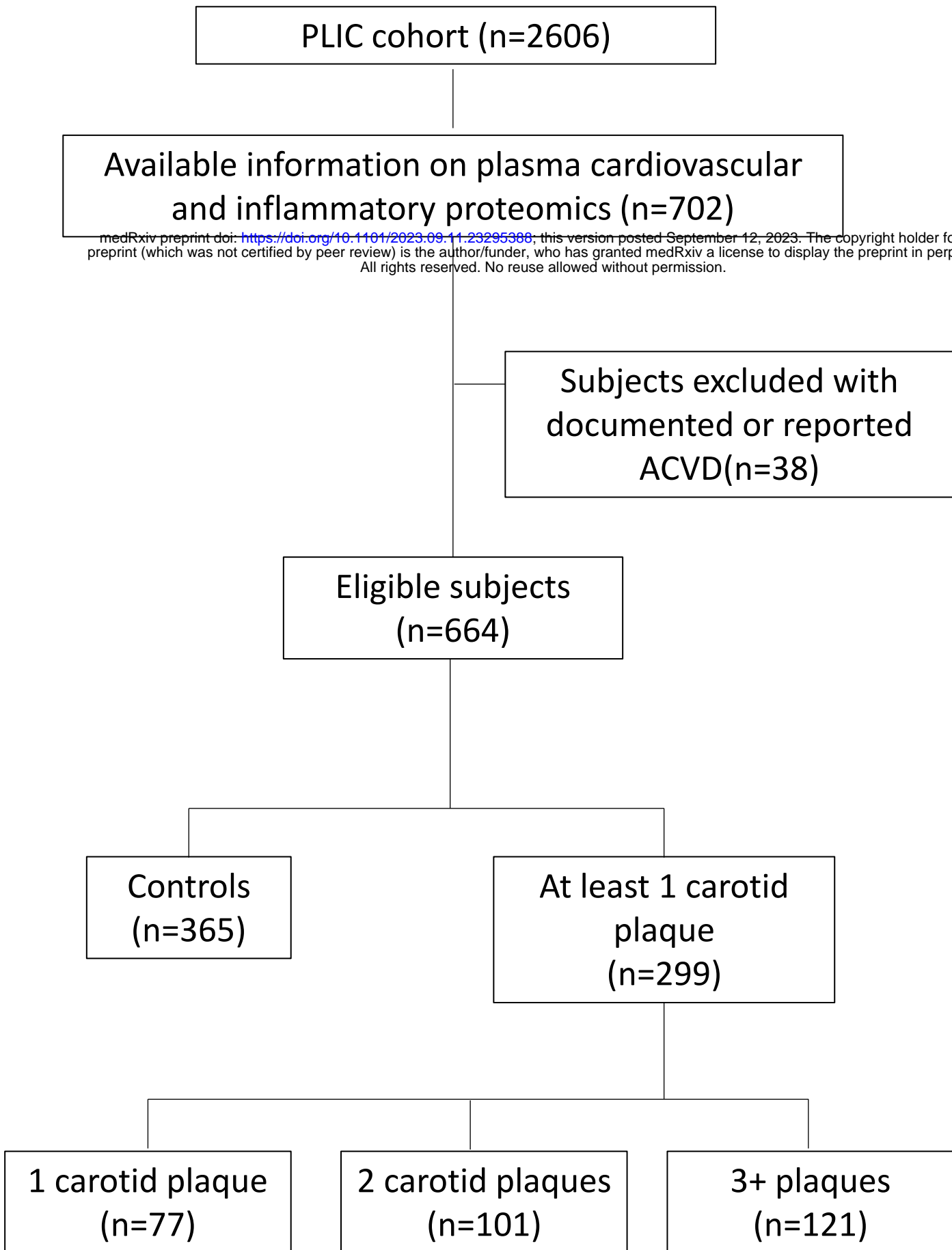
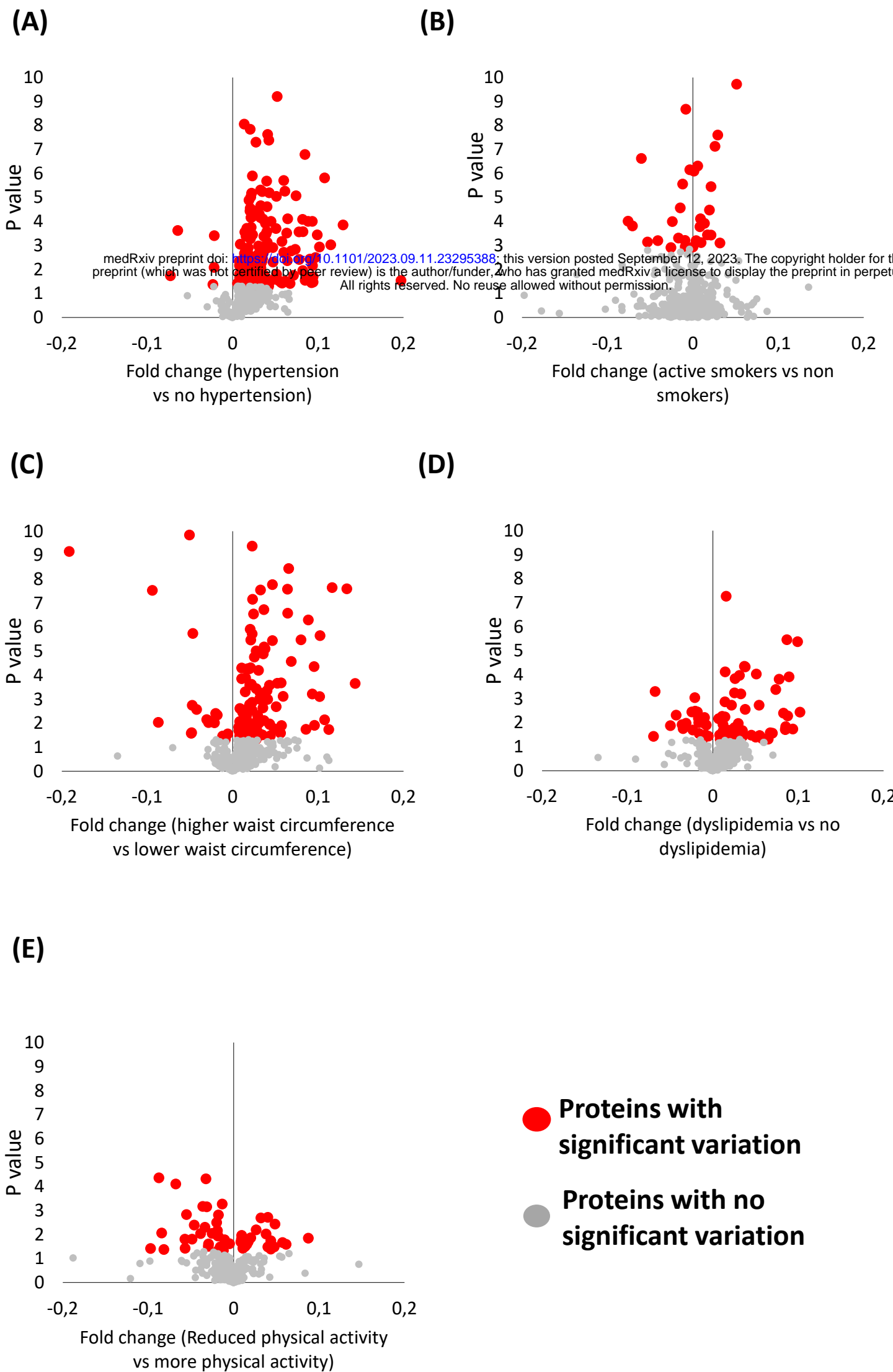


Figure 4.



Supplemental Figure 1.



Supplemental Figure 2.