

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30

Plasma MIA, CRP, and albumin predict cognitive decline in Parkinson’s Disease

by

Junchao Shen BA¹, Noor Amari BA¹, Rebecca Zack BA¹, R. Tyler Skrinak BA¹,
Travis L. Unger BA¹, Marijan Posavi PhD¹, Thomas F. Tropea DO¹, Sharon X. Xie² PhD,
Vivianna M. Van Deerlin MD PhD³, Richard B. Dewey, Jr. MD⁴, Daniel Weintraub MD^{5,6},
John Q. Trojanowski MD PhD¹, and Alice S. Chen-Plotkin MD¹

Departments of ¹Neurology, ²Biostatistics, Epidemiology and Informatics, ³Pathology and
Laboratory Medicine, ⁵Psychiatry, Perelman School of Medicine, University of Pennsylvania,
Philadelphia, PA. ⁴Department of Neurology, University of Texas Southwestern Medical Center,
Dallas, TX. ⁶Parkinson’s Disease Research, Education and Clinical Center (PADRECC),
Philadelphia Veterans Affairs Medical Center, Philadelphia, PA.

Corresponding Author:

Alice S. Chen-Plotkin MD
150 Johnson Pavilion
Philadelphia, PA 19104
215-573-7193
chenplot@pennmedicine.upenn.edu

Running head: Plasma biomarkers of PD cognitive decline

Character counts: Title 119 characters, Running head 36 characters

Word counts: Abstract 250 words, Introduction 443 words, Discussion 1331 words, Manuscript
4456 words

Figures and tables: 5 color figures, 2 tables

Keywords: Parkinson’s disease, cognitive decline, biomarkers, MIA

31 **ABSTRACT**

32

33 **Objective:** Using a multi-cohort, Discovery-Replication-Validation design, we sought new
34 plasma biomarkers that predict which PD individuals will experience cognitive decline.

35 **Methods:** In 108 Discovery Cohort PD individuals and 83 Replication Cohort PD individuals,
36 we measured 940 plasma proteins on an aptamer-based platform. Using proteins associating with
37 subsequent cognitive decline in both cohorts, we trained a logistic regression model to predict
38 which PD patients showed fast (≥ 1 point drop/year on Montreal Cognitive Assessment
39 (MoCA)) vs. slow (< 1 point drop/year on MoCA) cognitive decline in the Discovery Cohort,
40 testing it in the Replication Cohort. We developed alternate assays for the top three proteins and
41 confirmed their ability to predict cognitive decline – defined by change in MoCA or
42 development of incident Mild Cognitive Impairment (MCI) or dementia – in a Validation Cohort
43 of 118 PD individuals. We investigated the top plasma biomarker for causal influence by
44 Mendelian randomization.

45 **Results:** A model with only three proteins (Melanoma Inhibitory Activity Protein (MIA), C-
46 Reactive Protein (CRP), albumin) separated Fast vs. Slow cognitive decline subgroups with an
47 AUC of 0.80 in the Validation Cohort. Validation Cohort PD individuals in the top quartile of
48 risk for cognitive decline based on this model were 4.4 times more likely to develop incident
49 MCI or dementia than those in the lowest quartile. Genotypes at *MIA* SNP rs2233154 associated
50 with MIA levels and cognitive decline, providing evidence for MIA's causal influence.

51 **Conclusions:** An easily-obtained plasma-based predictor identifies PD individuals at risk for
52 cognitive decline. MIA may participate causally in development of cognitive decline.

53

54 INTRODUCTION

55 Parkinson's disease (PD) is the second most common neurodegenerative disease affecting
56 more than 5 million people worldwide. While the progressive loss of dopaminergic neurons in
57 the substantia nigra pars compacta (SNc) results in PD motor symptoms of bradykinesia, tremor,
58 and rigidity,¹ cognitive impairment and dementia also develop in a large proportion of
59 individuals with PD,² exacting a high financial and emotional cost for patients, their families,
60 and the healthcare system.^{3,4,5} Among PD patients, there is marked heterogeneity in cognitive
61 trajectory, making prognostication difficult and creating barriers for clinical trials aimed at
62 modifying this important aspect of disease.⁶ As a consequence, biomarkers predictive of
63 cognitive decline in PD are urgently needed, particularly if they might also shed light on the
64 mechanisms underlying development of cognitive decline in some PD patients but not others.

65 To date, such biomarkers of PD cognitive decline are sparse.⁷ In the cerebrospinal fluid
66 (CSF), higher phosphorylated tau and lower amyloid β_{42} are associated with a higher risk of
67 dementia in PD,^{8,9} and higher neurofilament light chain (NfL) also predicts cognitive decline.¹⁰
68 In the plasma, both higher NfL and lower epidermal growth factor levels predict cognitive
69 decline in PD.^{11,12,13,14,15} Extracellular-vesicle-associated tau and amyloid β_{42} have also been
70 reported to correlate with cognition in PD.¹⁶ Additionally, genetic variants have been linked to
71 cognitive trajectory in PD,^{17,18,19} with the most well-replicated effects on cognition seen for the
72 *APOE E4* allele^{20,21,22,23} and PD-associated *GBA* variants.^{17,24,25,26}

73 Studies to date, however, have limitations that hamper translation of biomarkers of PD
74 cognitive decline into clinical contexts. First, while some biomarkers have been widely
75 replicated, emerging biomarkers are often described in relatively small cohorts without
76 replication and validation.²⁷ Second, statistical associations may need to be converted into risk

77 scores or other tools that can be meaningfully used for risk stratification at the individual level.
78 Third, and perhaps most difficult to address, biomarkers emerging from large-scale screens (vs.
79 biomarkers tested for reasons related to, for example, their known role in Alzheimer’s disease
80 pathogenesis or *GBA*-related pathways) often lack biological context, leading to difficulty
81 discerning whether a given biomarker candidate is simply correlated with a phenotype or
82 causally involved in its development.²⁸

83 Here, we aimed to address these gaps by utilizing a multi-cohort, multi-stage design
84 starting with a screen of 940 plasma proteins in 191 longitudinally-followed PD patients. From
85 these initial data, we validated a three-protein blood-based biomarker panel in an additional 118
86 longitudinally-followed PD patients, demonstrating that this protein panel enriches for PD
87 individuals who will experience rapid cognitive decline regardless of cohort studied, cognitive
88 measure used, or method of biomarker measurement. Finally, we perform Mendelian-
89 randomization-based analyses to probe one of these newly-discovered biomarkers for evidence
90 of causality.

91
92

93 **METHODS**

94

95 ***Overview of study design***

96 108 PD patients were enrolled at the University of Texas Southwestern (UTSW Discovery
97 Cohort) as part of the NIH-NINDS Parkinson’s Disease Biomarker Program (PDBP). 83 PD
98 patients were enrolled at the University of Pennsylvania (UPenn Replication Cohort). The
99 Discovery and Replication Cohorts were screened for 1,129 and 1,305 proteins, respectively,

100 using an aptamer-based platform assay called SOMAScan.²⁹ A total of 940 proteins that passed
101 quality control (QC) metrics, as previously described,³⁰ in both cohorts, were retained for further
102 analysis. We then (1) assigned individuals to fast vs. slow cognitive decline subgroups based on
103 the rate of decline in MoCA score, and (2) identified which proteins differentiated fast vs. slow
104 cognitive decline subgroups in the Discovery and Replication Cohorts, to (3) develop a multi-
105 protein model for predicting cognitive course in PD, before (4) validating top biomarker
106 candidates with alternative assays (ELISA and BCP), and (5) testing the final predictive model in
107 a Validation Cohort (118 patients from University of Pennsylvania, **Fig. 1**).

108

109 *Cohorts and sample collection*

110 A total of 309 PD participants who were non-demented at the baseline visit, with blood samples
111 collected between 2013 and 2019 through the University of Pennsylvania (UPenn) and NIH-
112 NINDS Parkinson's Disease Biomarker Program (PDBP, with UTSW as the collection site) were
113 included in the analysis.^{31,32} All individuals met diagnostic criteria of the United Kingdom
114 Parkinson's Disease Brain Bank for PD. Demographics are summarized in **Table 1**, and details
115 are in the Supplementary Methods.

116

117 *Protein quantification*

118 Plasma proteins were quantified by SOMAScan in the Discovery and Replication Cohorts, then
119 validated by ELISA or BCP assay in the Validation Cohort.

120

121 **SOMAScan.** Plasma samples from both the Discovery and Replication Cohorts were assayed
122 using the 1.1k and 1.3k Assay versions of the SOMAScan platform (Somalogic, Boulder, CO,

123 USA) – based on aptamer capture of protein targets – as previously described.³⁰ SOMAScan
124 proteins that did not meet previously-described quality control metrics³⁰ were eliminated, leaving
125 940 plasma protein candidates. Protein measures are in relative fluorescence units (RFUs), log₁₀
126 transformed for downstream analyses.

127

128 **Enzyme-linked immunosorbent assay (ELISA).** For C-reactive protein, or CRP, and
129 melanoma inhibitory activity protein, or MIA, ELISAs were used in the Validation Cohort. For
130 CRP, we used the Human CRP Quantikine ELISA Kit (R&D CDRP00), diluting plasma samples
131 1:200. For MIA, we used the MIA ELISA Kit (Roche 11976826001), diluting plasma samples
132 1:2.

133

134 **Bromocresol Purple Assay (BCP).** For albumin, we employed the widely-used bromocresol
135 purple assay, using the BCP Albumin Assay Kit (Sigma MAK125), diluting plasma samples 1:5.

136

137 ***Comparison of protein measures across assay platforms***

138 For 15 individuals, we obtained two different aliquots of plasma sampled and banked on the
139 same day. For one aliquot, proteins were assayed by SOMAScan. For the other aliquot, proteins
140 were assayed by MIA ELISA, CRP ELISA, and BCP albumin assay. We compared measures for
141 these 15 duplicate samples across platforms, obtaining Pearson’s r and p-values for correlation.

142

143 ***Categorization of PD participants into fast and slow cognitive decline groups***

144 Participants in the Discovery, Replication, and Validation Cohorts were categorized into two
145 subgroups (Fast vs. Slow Cognitive Decline) based on rates of change in MoCA score. The rate

146 of MoCA decline was calculated using the MoCA score difference between the patient's last
147 follow-up visit and baseline visit divided by duration of time (years) between the two visits [i.e.,
148 (MoCA at last visit - MoCA at baseline visit)/years]. Patients with a MoCA score decline rate of
149 1 or more points/year were categorized as a fast cognitive decline subgroup, while the remaining
150 patients were assigned to a slow cognitive decline subgroup.

151

152 ***Statistical Analyses***

153 All analyses were performed in R. R-scripts are available in the Supplement, along with details
154 of statistical analyses.

155

156 ***Mendelian randomization-based analyses of MIA***

157 We used Mendelian randomization (MR)³³ to test the hypothesis that the top biomarker MIA is
158 causally related to development of cognitive decline in PD. Querying the Genotype-Tissue
159 Expression (GTEx) database,³⁴ we identified 4 linked single nucleotide polymorphisms (SNPs)
160 at the *MIA* locus demonstrating significant expression quantitative locus (eQTL) effects. We
161 tested these for association with plasma MIA protein levels, finding that all were protein
162 expression quantitative trait loci (pQTLs), with rs2233154 showing the strongest correlation.
163 Next, we tested for association between rs2233154 genotypes and cognitive decline in 180 PD
164 patients (all individuals from Replication and Validation Cohorts combined for which
165 genotyping data was available). We assessed the effect of rs2233154 genotype on cognitive
166 change over five years of follow-up using linear mixed-effects models as well as Cox
167 proportional hazards models adjusted for age, sex, and disease duration.

168

169 RESULTS

170

171 *Subgroups of Parkinson's Disease differing by rate of cognitive decline*

172 In 108 longitudinally-followed PD patients from the UTSW-based Discovery Cohort
173 (**Table 1**), 30 individuals (28%) declined by 1 or more points per year on the MoCA and were
174 assigned to the Fast Cognitive Decline subgroup. In 83 longitudinally-followed PD patients from
175 the UPenn-based Replication Cohort (**Table 1**), 22 individuals (27%) were assigned to the Fast
176 Cognitive Decline subgroup. Remaining PD patients were assigned to the Slow Cognitive
177 Decline subgroup.

178 To ensure that age, sex, or disease duration was not driving the rate of cognitive decline
179 in the Fast vs. Slow Cognitive Decline subgroups, we compared these subgroups in both the
180 Discovery and Replication Cohorts using linear mixed-effects models adjusted for these
181 variables. As shown in **Figures 2A** and **2B**, Fast and Slow Cognitive Decline subgroups differed
182 significantly in rate of change in MoCA even after adjustment for age, sex, and disease duration
183 ($p < 0.001$ for both cohorts).

184 Rates of motor change (captured in the UPDRS-III scores over time) did not differ for
185 Fast vs. Slow Cognitive Decline subgroups in the Discovery Cohort (**Figure 2C**). In the
186 Replication Cohort, which had a longer disease duration at the time of plasma sampling, the Fast
187 Cognitive Decline subgroup had a faster rate of motor decline ($p < 0.001$, **Figure 2D**).

188 The UPenn-based Replication Cohort has a battery of neuropsychological assessments
189 performed longitudinally, as well as longitudinal clinical cognitive consensus determination, as
190 previously described.³⁵ Since our method of assigning PD individuals to Fast vs. Slow Cognitive
191 Decline subgroups was chosen for ease and applicability across all cohorts, it is possible that

192 these groupings do not reflect meaningfully different trajectories. Thus, to verify our assignments
193 of Fast vs. Slow Cognitive Decline subgroups, we compared rates of change in the Mattis
194 Dementia Rating Scale-2 (DRS, a comprehensive measure of global cognition) for Fast vs. Slow
195 Cognitive Decline subgroups in the UPenn cohort, using linear mixed effects models adjusted for
196 age, sex, and disease duration. We found that the two subgroups differed significantly, with only
197 the Fast Cognitive Decline subgroup showing decline in DRS scores over time ($p < 0.001$, **Figure**
198 **2E**). We also compared rates of incident mild cognitive impairment (MCI) or dementia, as
199 clinically determined, in the Fast vs. Slow Cognitive Decline subgroups in the UPenn PD cohort.
200 In survival analyses, the Fast Cognitive Decline subgroup was more than twice as likely to
201 develop incident MCI or dementia than the Slow Cognitive Decline subgroup (67.4% developed
202 incident MCI or dementia in the Fast subgroup vs. 31.8% in the Slow subgroup, $p = 0.006$,
203 **Figure 2F**).

204 Taken together, Fast Cognitive Decline subgroups assigned based on longitudinal MoCA
205 scores comprised similar proportions of PD patients in both the Discovery and Replication
206 cohorts, despite differences in clinical site and disease duration for these two cohorts. Moreover,
207 subgroups assigned based on change in MoCA score had clinical significance; the Fast subgroup
208 had greater rates of incident MCI or dementia. Thus, we used these Fast vs. Slow Cognitive
209 Decline subgroup designations to develop and test biomarkers for cognitive decline in PD.

210

211 ***Plasma proteins associating with rates of cognitive change in PD***

212 In both the Discovery and Replication Cohorts, we measured levels of 940 plasma
213 proteins using an aptamer-based platform.³⁰ We tested each protein for association with Fast vs.
214 Slow Cognitive Decline using a linear model adjusted for age, sex, and disease duration,

215 generating an initial candidate list of nine proteins that associated with rates of cognitive decline
216 at a p-value of <0.10 , with the same directionality, in both the Discovery and Replication
217 Cohorts (**Table 2 and Supplementary Table 2**).

218 These nine proteins, along with clinical variables of age, sex, and disease duration, were
219 used to develop a logistic regression-based classifier to assign individuals to either the Fast or
220 Slow Cognitive Decline subgroups in the Discovery Cohort. In five-fold cross-validation within
221 the Discovery Cohort, the area under the receiver operating curve (AUC) for this model was
222 0.81. Applying the same logistic regression model to the Replication Cohort, which was never
223 used to train the model, we obtained an AUC of 0.82, demonstrating that the model was not
224 overfitted to the Discovery Cohort. In contrast, a model that used only the clinical variables of
225 age, sex, and disease duration, and did not use any plasma biomarkers, only reached an AUC of
226 0.65 (**Figure 3A**). Finally, we considered including baseline MoCA score as an additional
227 variable in the nine-protein model. However, this led to an increase in AUC for the Discovery
228 Cohort (AUC = 0.86), with a decrease in AUC for the Replication Cohort (AUC = 0.77),
229 suggesting overfitting with inclusion of this additional variable (**Supplementary Figure 1**).

230 Among the nine proteins in the model were three proteins with high potential for
231 downstream clinical biomarker translation based on existing translational uses and assay
232 availability: C-reactive protein (CRP), albumin, and melanoma inhibitory activity protein (MIA).
233 Specifically, CRP/albumin ratios are commonly used in the clinical setting to assess
234 inflammatory status,³⁶⁻³⁷ and reagents for measuring MIA are readily available since MIA has
235 been proposed as a biomarker for the common skin cancer melanoma.³⁸⁻³⁹ We thus evaluated the
236 performance of a Fast vs. Slow Cognitive Decline classifier that required only age, sex, disease
237 duration, and plasma levels of MIA, CRP, and albumin as input variables. As shown in **Figure**

238 **3B**, this three-protein model performed less well than the nine-protein model, but considerably
239 better than the clinical variable-only model, with an AUC of 0.73 in the Discovery Cohort and
240 0.75 in the Replication Cohort. Moreover, in both the Discovery and Replication Cohorts, MIA
241 plasma levels were higher in the Fast Cognitive Decline group ($p = 0.026$ for Discovery, $p =$
242 0.003 for Replication Cohort); the CRP/albumin ratio was lower in the Fast Cognitive Decline
243 group in the Replication cohort ($p = 0.004$, with a similar trend in the Discovery Cohort ($p =$
244 0.065 , **Figure 3C**).

245

246 *Validation of the top plasma biomarkers for cognitive decline*

247 For downstream clinical translation, biomarkers need to be robust to changes in cohort
248 and measurement platform. Thus, we sought to confirm our top biomarkers for cognitive decline
249 in PD in an additional validation cohort, using a different method of measurement.

250 We first evaluated enzyme-linked immunosorbent assays (ELISAs) for quantitation of
251 MIA and CRP in the plasma, using duplicate samples to compare measures obtained by ELISA
252 vs. the original aptamer-based platform. MIA measures were moderately correlated (Pearson $r =$
253 0.79 , **Figure 4A**) and CRP measures highly correlated (Pearson $r = 0.98$, **Figure 4B**) across
254 platforms. In contrast, plasma albumin measures obtained on the aptamer-based platform did not
255 correlate well with measures obtained with the widely-used bromocresol purple (BCP) assay
256 (Pearson $r = 0.29$, **Figure 4C**).

257 In a Validation Cohort of 118 longitudinally-followed PD patients from Penn (non-
258 overlapping with the 83 Penn PD patients in the Replication Cohort), we measured plasma MIA
259 and CRP by ELISA. Despite discrepancies between the aptamer-based platform and BCP
260 measures for albumin, we also measured plasma albumin in the Validation Cohort with the BCP

261 assay, as it is an assay commonly used in clinical settings, and our sample size of 15 for our
262 cross-platform comparison might be underpowered.

263 Using the same criteria as the Discovery and Replication Cohorts, we first characterized
264 PD individuals from the Validation Cohort as Fast vs. Slow Cognitive Decline, assigning 24/118
265 (20%) individuals to the Fast Cognitive Decline subgroup. In this group, plasma levels for MIA
266 were significantly higher ($p=0.022$, **Figure 4D**), but the CRP/albumin ratio did not differ
267 between Fast and Slow Cognitive Decline subgroups (**Figure 4E**).

268 Finally, we sought to validate a classifier for Fast vs. Slow Cognitive Decline in PD using
269 these three plasma proteins – MIA, CRP, and albumin – as well as age, sex and disease duration
270 in the 118-individual Validation Cohort. Despite differences in patient cohort and biomarker
271 assays used, performance for this three-protein classifier was just as strong in the Validation
272 Cohort, with an AUC of 0.81 (**Figure 4F**). Furthermore, when comparing a risk score based on
273 this three-protein model to clinical outcomes, we found significantly greater rates of incident
274 MCI or dementia among individuals with higher risk scores (**Figure 4G**), with a HR of 4.3
275 ($p=0.03$, 95% CI 1.15-15.9) for those in the third quartile of risk and a HR of 4.4 ($p=0.03$, 95%
276 CI 1.19-16.5) for those in the highest quartile of risk score, compared to individuals in the lowest
277 quartile of risk score.

278 Taken together, a three-protein predictor including MIA, CRP, and albumin, emerging
279 from our unbiased screen of 940 plasma proteins, significantly enriched for PD individuals most
280 likely to experience clinical cognitive decline in the near term. Moreover, results were not
281 affected by differences in patient cohort or biomarker testing platform.

282

283 ***MIA and cognitive decline in PD***

284 Among the three proteins in our newly-validated risk predictor for cognitive decline in
285 PD, CRP/albumin may indicate inflammatory status, but plasma MIA is harder to interpret
286 biologically. We thus investigated this signal further in two ways. First, we sought to understand
287 whether plasma MIA values, used alone, might risk-stratify PD patients in terms of future
288 clinical cognitive decline. Second, we used Mendelian-randomization (MR)-based techniques to
289 investigate MIA's causal influence on the development of cognitive decline in PD.

290 Since both the UPenn-based Replication Cohort, and the UPenn-based Validation Cohort
291 have clinical consensus diagnoses of normal cognition, MCI, or dementia, we asked whether
292 individuals with higher plasma MIA were more likely to develop incident MCI or dementia over
293 five years of follow-up. In both the Replication Cohort (**Figure 5A**), and the Validation Cohort
294 (**Figure 5B**), we observed non-significant trends towards higher rates of incident MCI or
295 dementia among PD patients with higher plasma MIA.

296 For our MR-based analyses of causal inference for MIA's role in PD cognitive decline,
297 we first identified expression quantitative trait locus (eQTL) single nucleotide polymorphisms
298 (SNP) for *MIA* mRNA expression from the Genotype-Tissue Expression database (GTEx).³⁴
299 Carriers of one or more minor (T) alleles at rs2233154 demonstrated higher *MIA* expression in
300 multiple tissues, including several brain regions (cortex, amygdala, hypothalamus, pituitary), the
301 tibial nerve, several vascular tissues (aorta, left ventricle of the heart), and several gastro-
302 intestinal tissues (sigmoid colon, esophagus) in GTEx data. Thus, we compared levels of plasma
303 MIA among carriers of different rs2233154 genotypes. As shown in **Figures 5C and 5D**,
304 rs2233154 genotypes associated significantly with MIA plasma protein levels in both the
305 Replication Cohort (rs2233154^{CC vs. CT} $p = < 0.0001$, rs2233154^{CC vs. TT} $p = 0.0076$) and the
306 Validation Cohort (rs2233154^{CC vs. CT} $p = < 0.0001$).

307 We next asked whether carriers of rs2233154 genotypes differed in rates of cognitive
308 decline. Since the Replication and Validation Cohort are both UPenn-based, and they were both
309 assessed for cognition and assigned cognitive diagnoses in the same way, we combined them for
310 a total set of n=180 PD individuals for these analyses.

311 In a linear mixed-effects model adjusted for age, sex, disease duration, and baseline
312 MoCA score, rs2233154 genotypes associated significantly with rates of decline in the MoCA (p
313 = 0.0004, **Figure 5E**), with carriers of one or more T alleles declining more rapidly. Specifically,
314 for each additional T allele carried, MoCA scores declined by an additional 0.445 points/year,
315 compared to trajectories for individuals without T alleles. Moreover, in a Cox proportional
316 hazards model adjusting for age, sex, and disease duration, rs2233154 genotypes trended towards
317 association with incident clinical MCI or dementia, with the same direction of effect (HR 1.8 for
318 carriers of one or more T alleles [95% CI 0.94-3.4, $p = 0.077$], **Figure 5F**).

319 Taken together, these analyses suggest that MIA may not only serve as a biomarker for,
320 but also play a causal role in, the development of cognitive decline in PD.

321

322 **DISCUSSION**

323 In this study, we investigated three different PD cohorts from two different clinical sites,
324 comprising a total of 309 longitudinally-followed individuals with PD, in order to discover,
325 replicate, and validate biomarkers predictive of cognitive decline. In 191 individuals with PD, we
326 first screened 940 plasma proteins for association with the rate of change in MoCA scores,
327 nominating top proteins for downstream investigation. We then developed two models – each
328 featuring multiple proteins as well as clinical variables – and tested them for ability to predict the
329 rate of cognitive decline. For the simpler model, requiring measures of just three proteins, we

330 developed alternate assays for these proteins and tested their performance in predicting both
331 decline in cognitive test scores and incident MCI or dementia in an additional 118 PD
332 individuals, demonstrating that our model can substantially enrich for individuals at high risk for
333 rapid cognitive decline. Finally, for one of the novel proteins discovered in this process, MIA,
334 we present evidence for a causal role from Mendelian randomization analyses.

335 Our findings have relevance for clinical prognostication as well as planning of clinical
336 trials aimed at modifying the rate of cognitive decline in PD. Specifically, while prior studies
337 have addressed the question of cognitive decline in PD, most studies that incorporate
338 biochemical biomarkers focus on demonstrating significant associations between a given
339 biomarker and cognitive trajectory,^{2,4,22} rather than developing tools that might be applied at an
340 individual level for risk stratification. A few studies have built predictors applicable to
341 individuals. For example, Liu *et al.* reported a clinicogenetic predictor with an AUC of 0.85 for
342 prediction of incident cognitive impairment within 10 years.⁴⁰ More recently, Tang *et al.* used
343 both clinical and radiographic data from the Parkinson's Progression Markers Initiative (PPMI)
344 cohort to predict time to cognitive progression, employing a training-test design similar to the
345 one used in our study and finding that clinical variables, in particular, were highly predictive of
346 incident MCI.⁴¹ Neither the Tang or Liu reports, however, focus on biochemical biomarkers. As
347 such, our findings contribute to knowledge in the field with respect to what plasma proteins may
348 add to ability to predict individual cognitive decline. Given the routine use of biochemical
349 biomarkers for risk stratification in other areas of medicine (*e.g.* lipid levels in cardiology, tumor
350 markers in oncology), development and validation of easily-accessed protein biomarkers that can
351 be used at an individual level in PD is a priority.²⁷

352 Our study also yields insight into the mechanisms that may lead to cognitive decline in
353 PD. First, our finding that plasma CRP levels may be predictive of rate of cognitive change in
354 PD should be considered in light of prior reports demonstrating that CRP levels are higher in PD
355 compared to neurologically normal controls.⁴² In both our Discovery and Replication cohorts,
356 however, PD individuals with higher CRP levels experience *less* subsequent cognitive decline.
357 Thus, our CRP findings contextualize the prior literature in several ways. They show that CRP
358 emerges from a screen of nearly 1000 proteins as an informative biomarker of disease trajectory,
359 adding confidence to a signal that has largely been evaluated in candidate-protein studies.
360 Additionally, they suggest that a more nuanced view is needed with respect to whether
361 therapeutic approaches aimed at reducing inflammation – and consequently CRP levels – would
362 be beneficial in PD, since higher-CRP individuals tended to maintain their cognitive status in our
363 study. Second, our study highlights plasma MIA as a biomarker predicting cognitive decline in
364 PD, with evidence to support a causal role for MIA-related pathways in the development of these
365 cognitive features. MIA is a secreted protein most well-characterized for its role as a biomarker
366 in the skin cancer melanoma. Known to be highly expressed in malignant melanocytes, MIA has
367 been used as a blood-based biomarker in cancer, with higher levels indicative of risk for
368 metastatic melanoma.^{43,44} To our knowledge, MIA has not been previously linked to PD.
369 Intriguingly, however, melanoma and PD have long been known to co-occur at rates that are
370 significantly higher than expected,⁴⁵ and melanoma patients without PD have been reported to
371 have a 10.5-fold relative risk of death from metastatic melanoma compared to melanoma patients
372 with PD.⁴⁶ Taken together, both these epidemiological studies and our current causal inference
373 analysis of MIA support mechanistic investigation of MIA-related pathways in PD.

374 Strengths of this study include (1) the large-scale screening of 940 plasma proteins as
375 potential biomarkers predicting cognitive decline in PD, through which novel leads might be
376 discovered, (2) the discovery-replication-validation design, ensuring that no single cohort,
377 method of ascertaining cognitive impairment, or protein measurement platform may be
378 responsible for our findings, and (3) the incorporation of analyses aimed at understanding
379 whether MIA as a top biomarker might be *causal* for the development of cognitive decline in
380 PD. Few reported biomarker studies in PD have incorporated protein screening at the scale used
381 here, with even fewer incorporating detailed longitudinal follow-up of patients for clinically-
382 relevant outcomes. Moreover, while some discovery screening studies offer promising leads,⁴⁷ to
383 date, these biomarker candidates largely lack replication.

384 Limitations of our study should also be considered. First, while our study incorporated
385 three different cohorts of PD patients, with each phase designed to replicate (or fail to replicate)
386 the findings of the prior phase, two-thirds of the PD individuals studied here were from one site
387 (UPenn). Future studies evaluating our cognitive decline predictor in larger, multi-site cohorts at
388 various stages of disease are needed to translate our findings into the most useful real-world
389 applications. Second, some findings were only incompletely replicated across phases of our
390 study. In particular, the CRP/albumin ratio, lower in the Slow Cognitive Decline subgroup in the
391 UPenn-based Replication Cohort, did not differ comparing Slow vs. Fast Cognitive Decline
392 subgroups in the UPenn-based Validation Cohort. This inconsistency might be due to biological
393 noise. Alternately, the lack of replication might stem from the measurement methods used in the
394 Validation Cohort, especially as plasma albumin measures did not correlate well for the
395 SOMAScan vs. BCP assay. Nonetheless, it is reassuring that performance for the three-protein
396 predictor remained as high in the Validation Cohort (AUC 0.80) as in the other phases of our

397 study. Third, while both our nine-protein model and our three-protein model show moderate
398 performance (AUC 0.73-0.82), and certainly improve upon the ability to make predictions based
399 on clinical variables alone (AUC 0.65), ability to separate Fast vs. Slow Cognitive Decline
400 groups is not perfect. It is likely that the arbitrary division of PD individuals into two cognitive
401 decline groups based on change in MoCA scores precludes our ability to see the fuller separation
402 one might expect for groups that are truly biologically distinct. That said, our goal in this study
403 was to create a robust, easy-to-use tool that might risk-stratify PD patients at the individual level.
404 We point to the fact that PD individuals in the highest 25% of risk score were 4.4 times more
405 likely to develop incident MCI or dementia in four years than those in the lowest 25% of risk
406 score.

407 In summary, we present our findings from a study of 309 longitudinally-followed PD
408 individuals, where, starting from 940 plasma protein candidates, we develop a risk predictor for
409 cognitive decline in a four-year window. We find that a risk score based only on age, sex, and
410 plasma values of MIA, CRP, and albumin identifies a subgroup of PD individuals 4.4 times more
411 likely to develop cognitive decline in the near term, regardless of cohort studied, cognitive
412 measure used, or method for biomarker measurement. We furthermore link MIA causally to
413 development of cognitive impairment through Mendelian randomization. Taken together, our
414 study offers an easy-to-use tool for risk stratification for future cognitive impairment in PD
415 individuals, as well as a new lead for mechanistic investigation. The development of molecular
416 tools – such as the risk score calculator presented here – enables “precision medicine”
417 approaches to the care of PD patients. Moreover, the emergence of MIA from a 940-protein
418 screen illustrates an approach for deriving targets for downstream mechanistic experiments based
419 on biochemical profiling of patient-derived biofluids.

420

421 **ACKNOWLEDGMENTS**

422 We thank Maria Diaz-Ortiz for technical assistance. We additionally thank our patients
423 and their families for their generosity in contributing to this research. We thank the NINDS
424 PDBP and the BioSEND biorepository for assistance with samples.

425 This research was supported by the NIH (RO1 NS115139, U19 AG062418, P50
426 NS053488), and a Biomarkers Across Neurodegenerative Diseases (BAND) grant from the
427 Michael J. Fox Foundation/Alzheimer's Association/Weston Institute. Alice Chen-Plotkin is
428 additionally supported by the Parker Family Chair, the AHA/Allen Institute Brain Health
429 Initiative, and the Chan Zuckerberg Initiative Neurodegeneration Challenge.

430

431 **AUTHOR CONTRIBUTIONS**

432 Author contributions to this work are as follows: ACP designed the study, with substantial input
433 from JS, NA, RZ, RTS, TLU, MP, TFT, VVD, RBD, and DW acquired the data. JS, NA, SXX,
434 MP and ACP analyzed the data. JS, NA, and ACP drafted the majority of the manuscript, and all
435 authors edited and approved the final manuscript.

436

437

438 REFERENCES

439

440 [1] Hoehn MM, Yahr MD. Parkinsonism - Onset Progression and Mortality. *Neurology*.

441 1967;17(5):427-&.

442 [2] Hely MA, Reid WGJ, Adena MA, et al. The Sydney multicenter study of Parkinson's disease:

443 The inevitability of dementia at 20 years. *Movement Disord*. 2008 Apr 30;23(6):837-44.

444 [3] He L, Lee EY, Sterling NW, et al. The Key Determinants to Quality of Life in Parkinson's

445 Disease Patients: Results from the Parkinson's Disease Biomarker Program (PDBP). *J Parkinson*

446 *Dis*. 2016;6(3):523-32.

447 [4] Aarsland D, Larsen JP, Karlsen K, et al. Mental symptoms in Parkinson's disease are

448 important contributors to caregiver distress. *Int J Geriatr Psych*. 1999 Oct;14(10):866-74.

449 [5] Pressley JC, Louis ED, Tang MX, et al. The impact of comorbid disease and injuries on

450 resource use and expenditures in parkinsonism. *Neurology*. 2003 Jan 14;60(1):87-93.

451 [6] Greenland JC, Williams-Gray CH, Barker RA. The clinical heterogeneity of Parkinson's

452 disease and its therapeutic implications. *Eur J Neurosci*. 2019 Feb;49(3):328-38.

453 [7] Guo Y, Liu FT, Hou XH, et al. Predictors of cognitive impairment in Parkinson's disease: a

454 systematic review and meta-analysis of prospective cohort studies. *J Neurol*. 2021

455 Aug;268(8):2713-22.

- 456 [8] Siderowf A, Xie SX, Hurtig H, et al. CSF amyloid beta 1-42 predicts cognitive decline in
457 Parkinson disease. *Neurology*. 2010 Sep 21;75(12):1055-61.
- 458 [9] Mollenhauer B, Bibl M, Wiltfang J, et al. Total tau protein, phosphorylated tau (181p)
459 protein, beta-amyloid(1-42), and beta-amyloid(1-40) in cerebrospinal fluid of patients with
460 dementia with Lewy bodies. *Clin Chem Lab Med*. 2006 Feb;44(2):192-5.
- 461 [10] Lerche S, Wurster I, Roben B, et al. CSF NFL in a Longitudinally Assessed PD Cohort:
462 Age Effects and Cognitive Trajectories. *Movement Disord*. 2020 Jul;35(7):1138-44.
- 463 [11] Aamodt WW, Waligorska T, Shen JC, et al. Neurofilament Light Chain as a Biomarker for
464 Cognitive Decline in Parkinson Disease. *Movement Disord*. 2021 Sep 4.
- 465 [12] Ma LZ, Zhang C, Wang H, et al. Serum Neurofilament Dynamics Predicts Cognitive
466 Progression in de novo Parkinson's Disease. *J Parkinson Dis*. 2021;11(3):1117-27.
- 467 [13] Chen-Plotkin AS, Hu WT, Siderowf A, et al. Plasma Epidermal Growth Factor Levels
468 Predict Cognitive Decline in Parkinson Disease. *Ann Neurol*. 2011 Apr;69(4):655-63.
- 469 [14] Pellecchia MT, Santangelo G, Picillo M, et al. Serum epidermal growth factor predicts
470 cognitive functions in early, drug-naïve Parkinson's disease patients. *J Neurol*. 2013
471 Feb;260(2):438-44.
- 472 [15] Lim NS, Swanson CR, Cherng HR, et al. Plasma EGF and cognitive decline in Parkinson's
473 disease and Alzheimer's disease. *Ann Clin Transl Neur*. 2016 May;3(5):346-55.

- 474 [16] Chung CC, Chan L, Chen JH, et al. Plasma extracellular vesicles tau and beta-amyloid as
475 biomarkers of cognitive dysfunction of Parkinson's disease. *Faseb J*. 2021 Oct;35(10).
- 476 [17] Liu GQ, Peng JJ, Liao ZX, et al. Genome-wide survival study identifies a novel synaptic
477 locus and polygenic score for cognitive progression in Parkinson's disease. *Nat Genet*. 2021
478 Jun;53(6):787-+.
- 479 [18] Paul KC, Schulz J, Bronstein JM, et al. Association of Polygenic Risk Score With Cognitive
480 Decline and Motor Progression in Parkinson Disease. *Jama Neurol*. 2018 Mar;75(3):360-6.
- 481 [19] Tropea TF, Mak J, Guo MH, et al. TMEM106B Effect on cognition in Parkinson disease
482 and frontotemporal dementia. *Ann Neurol*. 2019 Jun;85(6):801-11.
- 483 [20] Chung J, Ushakova A, Doitsidou M, et al. The impact of common genetic variants in
484 cognitive decline in the first seven years of Parkinson's disease: A longitudinal observational
485 study. *Neurosci Lett*. 2021 Nov 1;764.
- 486 [21] Kim R, Park S, Yoo D, et al. Association of Physical Activity and APOE Genotype With
487 Longitudinal Cognitive Change in Early Parkinson Disease. *Neurology*. 2021 May
488 11;96(19):E2429-E37.
- 489 [22] Tan MMX, Lawton MA, Jabbari E, et al. Genome-Wide Association Studies of Cognitive
490 and Motor Progression in Parkinson's Disease. *Movement Disord*. 2021 Feb;36(2):424-33.
- 491 [23] Tropea TF, Xie SX, Rick J, et al. APOE, Thought Disorder, and SPARE-AD Predict
492 Cognitive Decline in Established Parkinson's Disease. *Movement Disord*. 2018 Feb;33(2):289-
493 97.

- 494 [24] Greuel A, Trezzi JP, Glaab E, et al. GBA Variants in Parkinson's Disease: Clinical,
495 Metabolomic, and Multimodal Neuroimaging Phenotypes. *Movement Disord.* 2020
496 Dec;35(12):2201-10.
- 497 [25] Petrucci S, Ginevrino M, Trezzi I, et al. GBA-Related Parkinson's Disease: Dissection of
498 Genotype-Phenotype Correlates in a Large Italian Cohort. *Movement Disord.* 2020
499 Nov;35(11):2106-11.
- 500 [26] Davis MY, Johnson CO, Leverenz JB, et al. Association of GBA Mutations and the E326K
501 Polymorphism With Motor and Cognitive Progression in Parkinson Disease. *Jama Neurol.* 2016
502 Oct 1;73(10):1217-24.
- 503 [27] Chen-Plotkin AS, Albin R, Alcalay R, et al. Finding useful biomarkers for Parkinson's
504 disease. *Sci Transl Med.* 2018 Aug 15;10(454).
- 505 [28] Diaz-Ortiz ME, Chen-Plotkin AS. Omics in Neurodegenerative Disease: Hope or Hype?
506 *Trends Genet.* 2020 Mar;36(3):152-9.
- 507 [29] Candia J, Cheung F, Kotliarov Y, et al. Assessment of Variability in the SOMAscan Assay.
508 *Sci Rep-Uk.* 2017 Oct 27;7.
- 509 [30] Posavi M, Diaz-Ortiz M, Liu B, et al. Characterization of Parkinson's disease using blood-
510 based biomarkers: A multicohort proteomic analysis. *Plos Med.* 2019 Oct;16(10).
- 511 [31] Rosenthal LS, Drake D, Alcalay RN, et al. The NINDS Parkinson's Disease Biomarkers
512 Program. *Movement Disord.* 2016 Jun;31(6):915-23.

- 513 [32] Gwinn K, David KK, Swanson-Fischer C, et al. Parkinson's disease biomarkers: perspective
514 from the NINDS Parkinson's Disease Biomarkers Program. *Biomark Med.* 2017 May;11(6):451-
515 73.
- 516 [33] Lawlor DA, Harbord RM, Sterne JAC, Timpson N, Smith GD. Mendelian randomization:
517 Using genes as instruments for making causal inferences in epidemiology. *Stat Med.* 2008 Apr
518 15;27(8):1133-63.
- 519 [34] Lonsdale J, Thomas J, Salvatore M, et al. The Genotype-Tissue Expression (GTEx) project.
520 *Nat Genet.* 2013 Jun;45(6):580-5.
- 521 [35] Pigott K, Rick J, Xie SRX, et al. Longitudinal study of normal cognition in Parkinson
522 disease. *Neurology.* 2015 Oct 13;85(15):1276-82.
- 523 [36] Kaplan M, Ates I, Akpınar MY, et al. Predictive value of C-reactive protein/albumin ratio in
524 acute pancreatitis. *Hepatob Pancreat Dis.* 2017 Aug;16(4):424-30.
- 525 [37] Basile A, Lago AF, Menegheti MG, et al. The use of APACHE II, SOFA, SAPS 3, C-
526 reactive protein/albumin ratio, and lactate to predict mortality of surgical critically ill patients A
527 retrospective cohort study. *Medicine.* 2019 Jun;98(26).
- 528 [38] Feuerer L, Lamm S, Henz I, et al. Role of melanoma inhibitory activity in melanocyte
529 senescence. *Pigm Cell Melanoma R.* 2019 Nov;32(6):777-91.
- 530 [39] Bosserhoff AK. Melanoma inhibitory activity (MIA): an important molecule in melanoma
531 development and progression. *Pigm Cell Res.* 2005 Dec;18(6):411-6.

532 [40] Liu G, Locascio JJ, Corvol JC. Prediction of cognition in Parkinson's disease with a clinical-
533 genetic score: a longitudinal analysis of nine cohorts (vol 16, pg 620, 2017). *Lancet Neurol*. 2017
534 Sep;16(9):683-.

535 [41] Tang CY, Zhao XY, Wu W, et al. An individualized prediction of time to cognitive
536 impairment in Parkinson's disease: A combined multi-predictor study. *Neurosci Lett*. 2021 Sep
537 25;762.

538 [42] Qiu X, Xiao Y, Wu J, et al. C-Reactive Protein and Risk of Parkinson's Disease: A
539 Systematic Review and Meta-Analysis. *Front Neurol*. 2019;10:384.

540 [43] Kluger HM, Hoyt K, Bacchiocchi A, et al. Plasma markers for identifying patients with
541 metastatic melanoma. *Clin Cancer Res*. 2011 Apr 15;17(8):2417-25.

542 [44] Bosserhoff AK, Lederer M, Kaufmann M, et al. MIA, a novel serum marker for progression
543 of malignant melanoma. *Anticancer Res*. 1999 Jul-Aug;19(4A):2691-3.

544 [45] Bose A, Petsko GA, Eliezer D. Parkinson's Disease and Melanoma: Co-Occurrence and
545 Mechanisms. *J Parkinsons Dis*. 2018;8(3):385-98.

546 [46] Dalvin LA, Damento GM, Yawn BP, et al. Parkinson Disease and Melanoma: Confirming
547 and Reexamining an Association. *Mayo Clin Proc*. 2017 Jul;92(7):1070-9.

548 [47] Pan C, Zhou Y, Dator R, et al. Targeted discovery and validation of plasma biomarkers of
549 Parkinson's disease. *J Proteome Res*. 2014 Nov 7;13(11):4535-45.

550

551

552

553

554

555

556

557

558

TABLES AND FIGURE LEGENDS

559

Variable	Discovery Cohort			Replication Cohort			Validation Cohort		
	Fast Progressor (n = 30) (27.8%)	Slow Progressor (n = 78) (72.2%)	P Val	Fast Progressor (n = 22) (26.5%)	Slow Progressor (n = 61) (73.5%)	P Val	Fast Progressor (n = 24) (20.3%)	Slow Progressor (n = 94) (79.7%)	P Val
Age At Sample (yrs) Mean (SD)	68.9 (7.1)	64.5 (8.4)	0.013 [‡]	70.2 (6.4)	69.0 (7.6)	0.511 [‡]	72.6 (5.7)	64.6 (7.3)	<0.001 [‡]
Disease Duration (yrs) Mean (SD)	5.73 (4.6)	5.03 (7.5)	0.630 [‡]	10.3 (4.6)	8.2 (5.4)	0.101 [‡]	8.5 (6.2)	5.6 (3.9)	0.005 [‡]
Sex n (%)									
Female	13.0 (44.4%)	35.0 (44.9%)	1.000 ^{‡‡}	8 (36.4%)	30 (49.2%)	0.330 ^{‡‡}	7 (29.2%)	32 (34.0%)	0.809 ^{‡‡}
Male	17.0 (55.6%)	43.0 (55.1%)		15 (63.6%)	31 (50.8%)		17 (70.8%)	62 (66.0%)	
Baseline Montreal Cognitive Assessment Mean (SD)	27.3 (2.0)	26.5 (2.2)	0.086 [‡]	25.7 (3.1)	26.2 (2.4)	0.442 [‡]	26.0 (2.3)	26.0 (2.5)	1.000 [‡]
UPDRS-III	17.9 (8.7)	16.0 (8.2)	0.291 [‡]	28.2 (12.2)	22.6 (10.7)	0.046 [‡]	29.1 (9.8)	21.9 (11.9)	0.007 [‡]
Education				15.9 (2.3)	16.7 (2.1)	0.139 [‡]	16.8 (2.5)	16.3 (2.5)	0.384 [‡]

560

561 ‡ Unpaired t-test

562 ‡‡ Fisher's Exact Test

563 N/A = not available

564

565 **Table 1. Demographic characteristics of the Discovery (UTSW), Replication (UPenn), and**
 566 **Validation (UPenn) cohorts.**

567 Within each cohort, individuals defined as having a faster rate of cognitive decline vs. a slower
 568 rate of cognitive decline based on MoCA score change over time are compared as indicated with
 569 the symbol, based on distribution of the data. Years of education were not available for the
 570 Discovery Cohort.

571

572

573

574

		Discovery Cohort		Replication Cohort	
Protein Symbol	Protein Name	+ Higher in Rapid Group	P Value	+ Higher in Rapid Group	P Value
		- Higher in Slow Group		- Higher in Slow Group	
MIA	Melanoma Inhibitory Activity	+	0.027	+	0.013
CRP	C-Reactive Protein	-	0.061	-	0.002
Albumin	Albumin	+	0.058	+	0.052
MIP.5	Chemokine (C-C Motif) Ligand 15	+	0.032	+	0.037
FGF.5	Fibroblast Growth Factor 5	+	0.048	+	0.051
Aggrecan	Aggrecan	+	0.049	+	0.037
ROBO3	Roundabout Guidance Receptor 3	+	0.053	+	0.050
IL-4.sR	Soluble IL-4 receptor	+	0.015	+	0.022
RUXF	Small Nuclear Ribonucleoprotein Polypeptide F	+	0.083	+	0.036

575

576 **Table 2. Top 9 SOMAScan Proteins**

577 Nine proteins were associated with the rate of cognitive decline, in the same direction, at a p-
578 value cut-off of 0.1 in both the Discovery and Replication Cohorts.

579

580

581

582 **Fig 1. Study overview.** An aptamer-based platform was used to quantify the plasma levels of
583 940 proteins in the Discovery Cohort (left panel) and Replication Cohort (right panel). In each
584 cohort, PD patients were assigned into a fast or a slow cognitive decline group based on change
585 in MoCA score over time. A linear regression model was used to identify proteins differentiating
586 fast versus slow progressors in both cohorts, generating the top 9 proteins. Next, these identified
587 proteins were used to train two logistic regression-based models that predict whether individual
588 PD patients subsequently have fast vs. slow cognitive decline. Finally, in an additional validation
589 cohort of 118 PD patients, we measured top biomarker proteins using alternative assays, testing
590 for their performance in separating fast vs. slow cognitive decline subgroups.

591

592 **Fig 2. Characterization of cognitive decline subgroups.** Longitudinal cognitive and motor
593 performance in the fast versus slow cognitive decline subgroups was assessed using linear
594 mixed-effect models adjusting for age, sex and disease duration. Subgroups are indicated by
595 color; the band represents the 95% confidence interval. (A-B) In both the Discovery and
596 Replication Cohorts, MoCA scores decrease over time in the fast cognitive decline subgroup,
597 while remaining stable in the slow cognitive decline subgroup. (C-D) In the Discovery Cohort,
598 fast and slow cognitive decline subgroups do not differ in rate of motor change (UPDRS-III
599 score) over time. However, in the Replication Cohort, the fast cognitive decline subgroup also
600 experiences more rapid change in motor symptoms. (E) In the Replication Cohort, the fast
601 cognitive decline subgroup has a faster rate of decline in the DRS score as well. (F) In the
602 Replication Cohort, the fast cognitive decline subgroup has higher rates of incident MCI or
603 dementia over 4 years of follow-up.

604

605 **Fig 3. Identification of top biomarkers differentiating fast vs. slow cognitive decline**
606 **subgroups in both Discovery and Replication Cohorts.** (A-B) Performance characteristics of
607 the logistic regression model for predicting whether an individual PD patient falls in the fast vs
608 slow cognitive decline PD subgroup, trained using the measurements of all 9 proteins (panel A)
609 or only 3 proteins (MIA, CRP, albumin, panel B), together with age, sex and disease duration.
610 The model was trained using Discovery Cohort data and tested on the Replication Cohort. In the
611 Discovery cohort, area under the receiver operating curve (AUC) was derived by five-fold cross-
612 validation over 50 iterations. (C) Boxplots (median) showing the distribution of top biomarkers –
613 MIA, CRP/Albumin Ratio levels – in \log_{10} of RFU by PD cognitive decline subgroups. Mann-
614 Whitney test was used to compare biomarker differences between fast vs. slow cognitive decline
615 subgroups.

616
617 **Fig 4. Validation of top biomarkers using alternative assays in the Validation Cohort.** (A)
618 Comparison of the values for 3 top biomarkers (MIA, CRP, Albumin) obtained on SOMAScan
619 vs. Enzyme-linked immunosorbent (ELISA) or Bromocresol Purple (BCP) assay in 15 duplicate
620 plasma samples. Pearson's r is shown. (D-E) Boxplots (median) showing the distribution of MIA
621 and CRP/Albumin ratio in the Validation Cohort by PD cognitive decline subgroups. Mann-
622 Whitney test was used to compare the difference between subgroups. (F) Performance
623 characteristics of the logistic-regression model (incorporating MIA, CRP, albumin, age, sex, and
624 disease duration) for predicting fast vs slow cognitive decline subgroup in 118 PD patients from
625 the Validation Cohort. (G) Time to incident MCI or dementia for PD patients in each quartile of
626 risk score generated by the 6-parameter (3 protein, age, sex, disease duration) logistic regression
627 model.

628

629 **Fig 5. MIA as a novel blood biomarker for cognitive function decline in PD** (A-B) Cox

630 proportional hazards model adjusted for age, sex and disease duration showing incident MCI or

631 dementia rate for each quartile of baseline MIA measures in the (A) Discovery Cohort and (B)

632 Validation Cohort over 5 years of follow-up. (C-D) Boxplot showing the association between

633 genotypes at the *MIA* locus SNP rs2233154 and MIA expression in the plasma in the (C)

634 Discovery Cohort and (D) Validation Cohort. There were no individuals with the TT genotype in

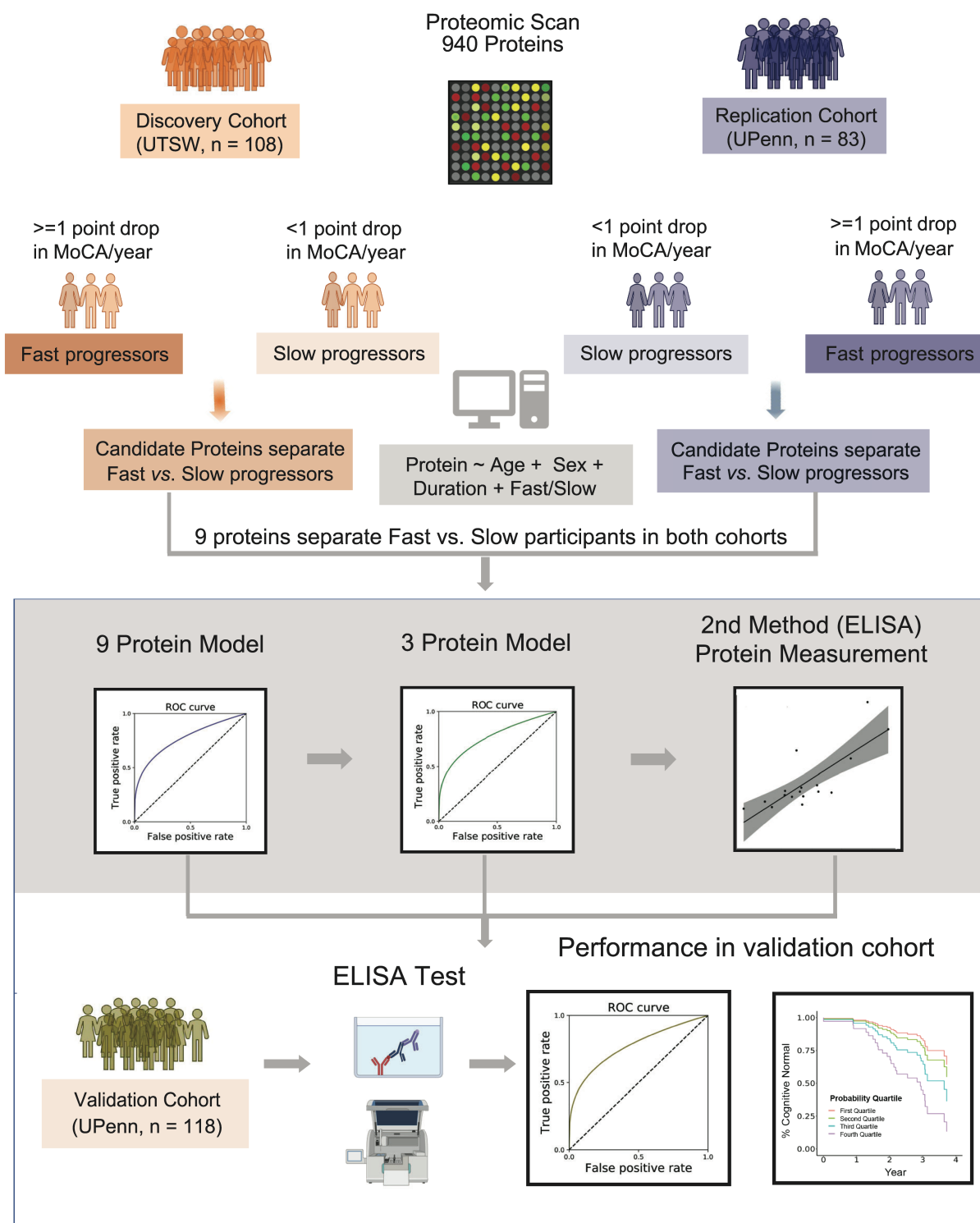
635 the Validation Cohort. (E) Effect of rs2233154 genotype on longitudinal MoCA performance

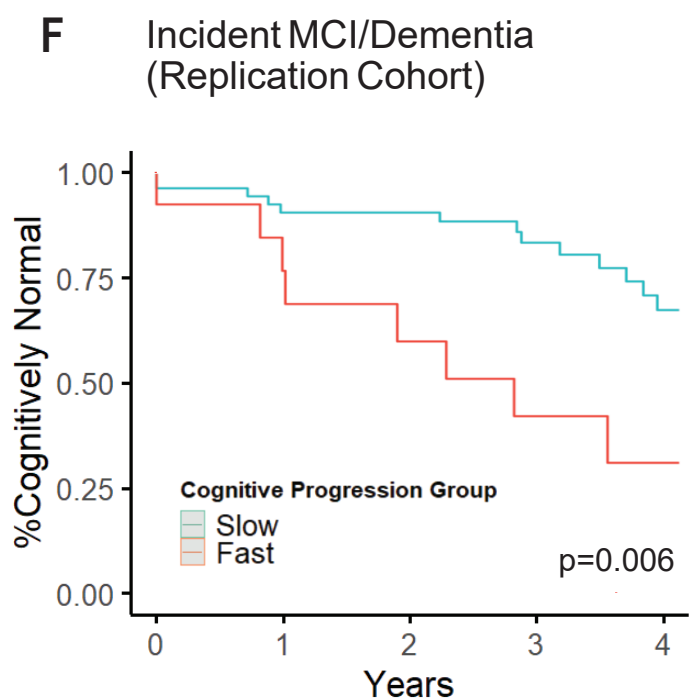
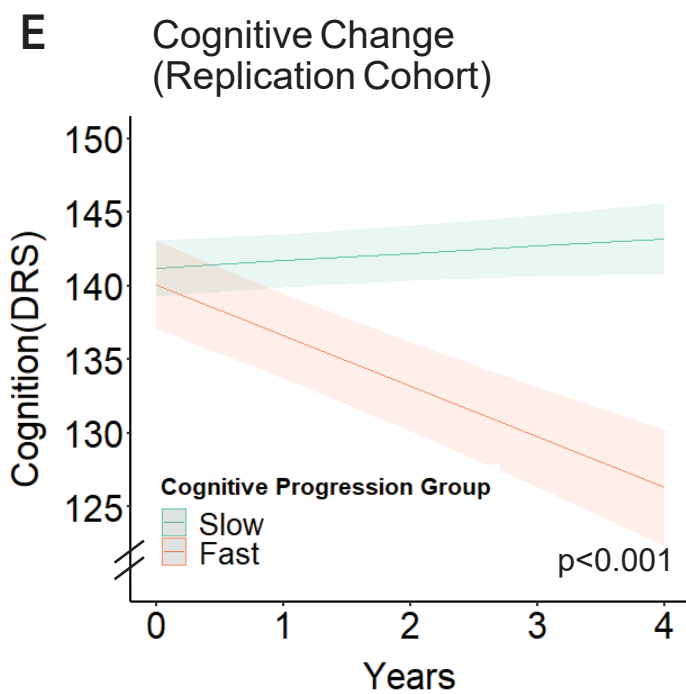
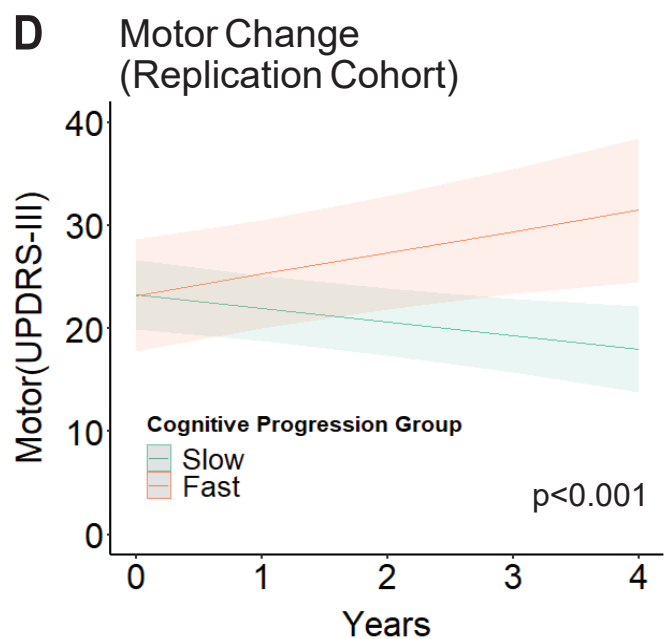
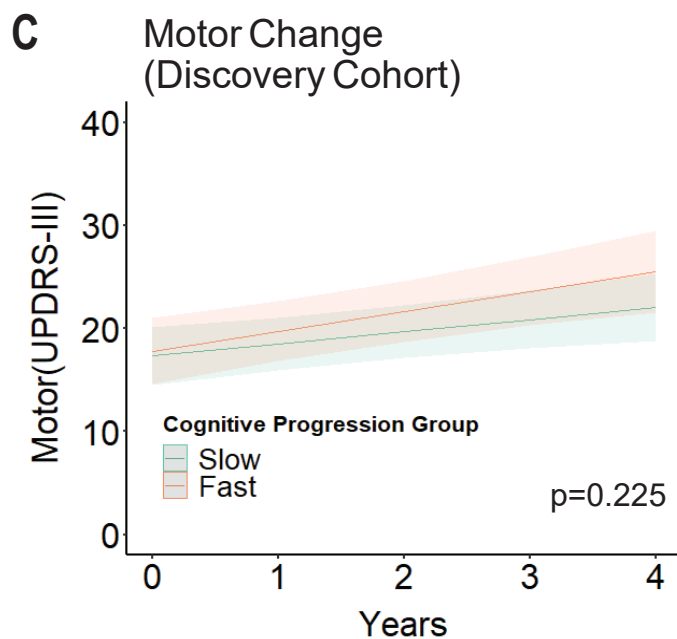
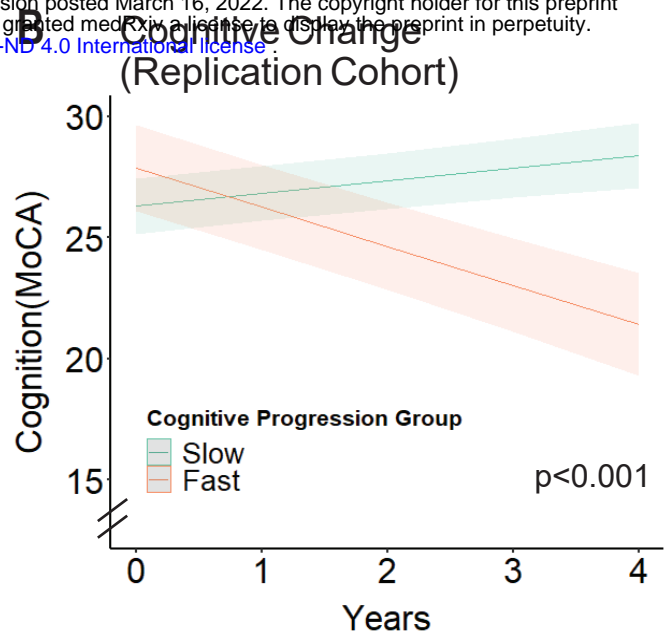
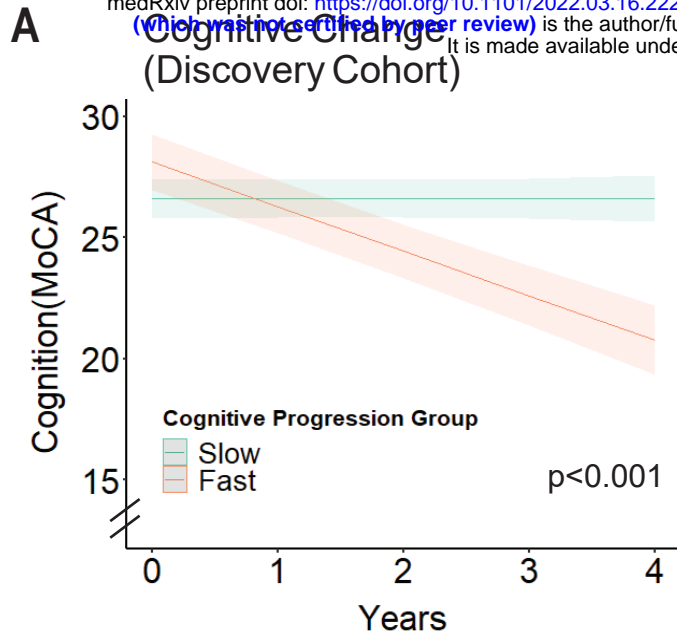
636 assessed using linear-mixed effects models adjusting for age, sex and disease duration. (F) Cox

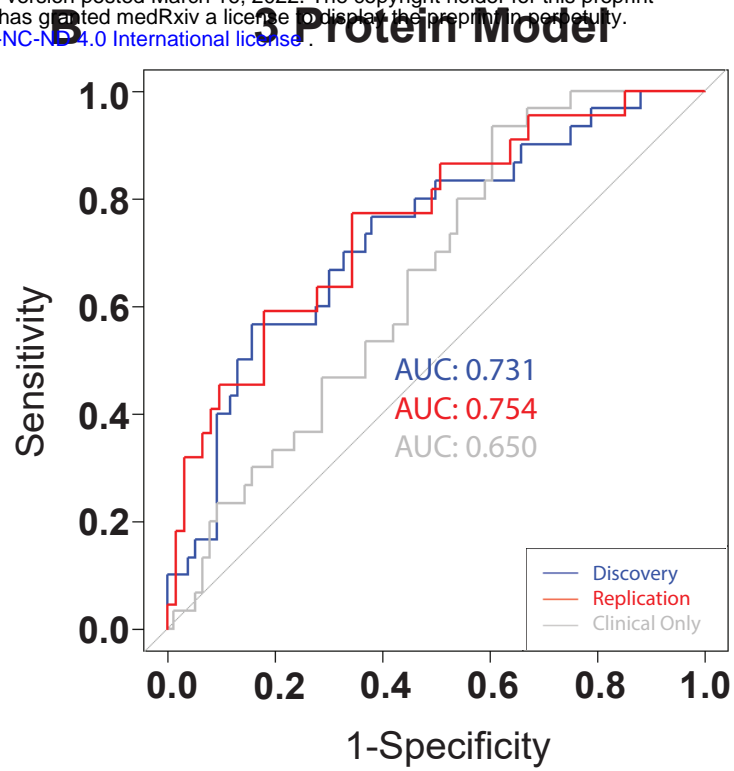
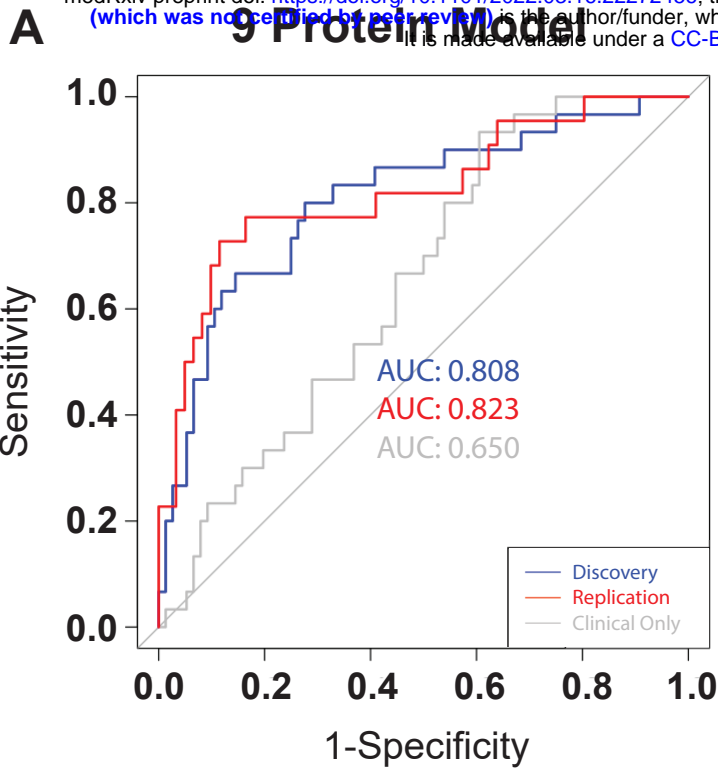
637 proportional hazards model adjusting for age, sex and disease duration shows a differential rate

638 of incident MCI or dementia comparing carriers of different rs2233154 genotypes over 5 years of

639 follow-up.

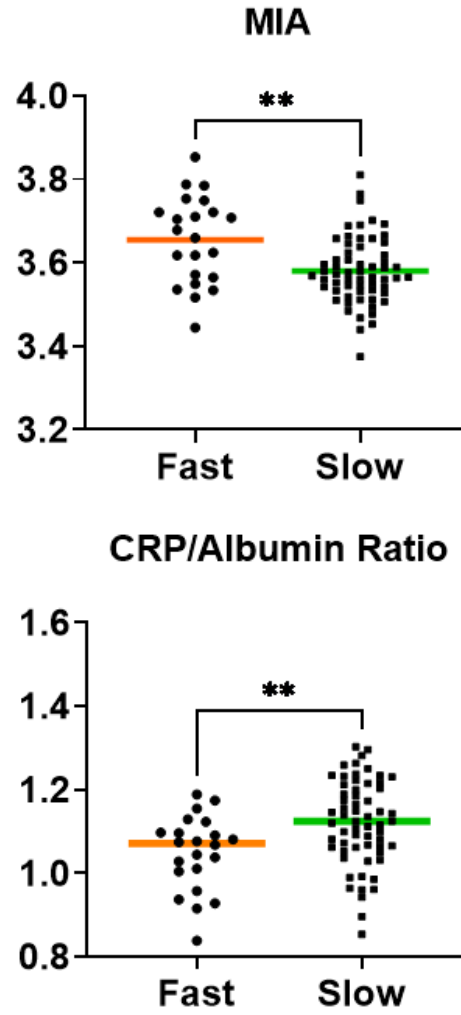
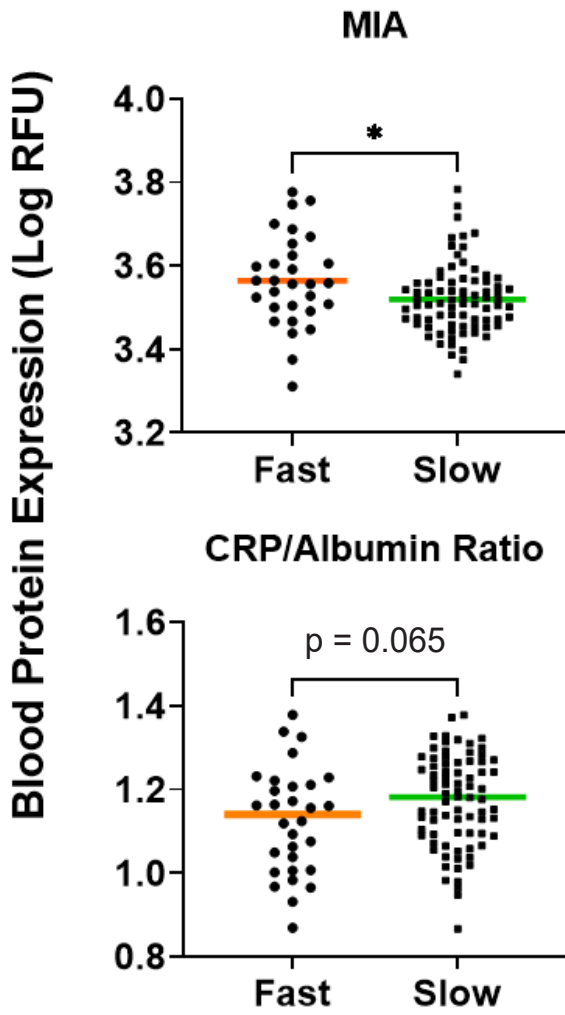






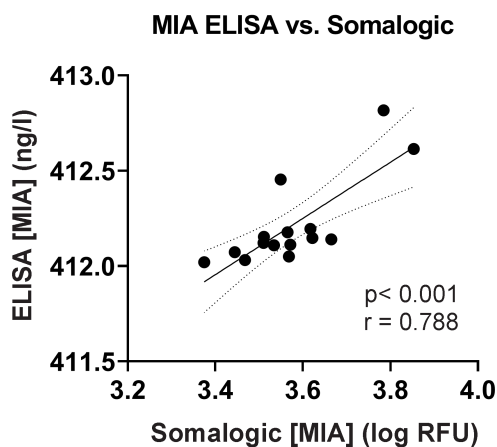
C Discovery Cohort

Replication Cohort

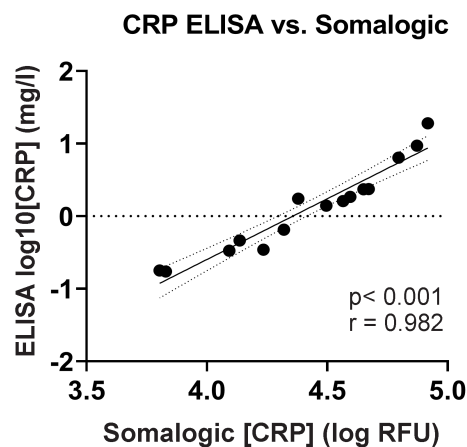


Progression Group

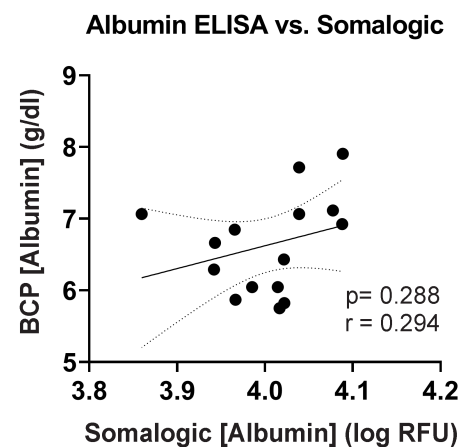
A



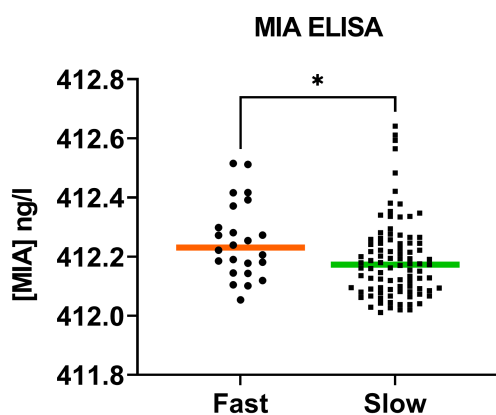
B



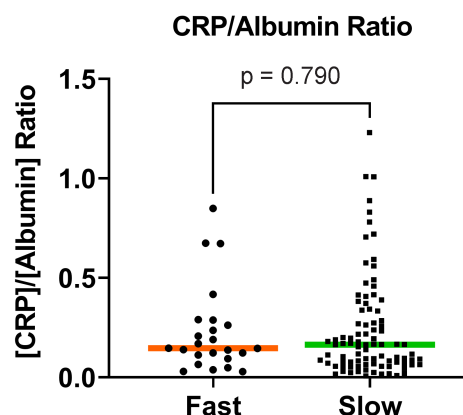
C



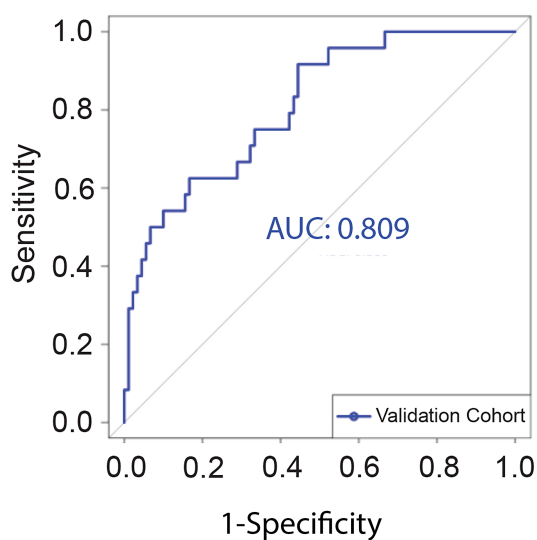
D



E

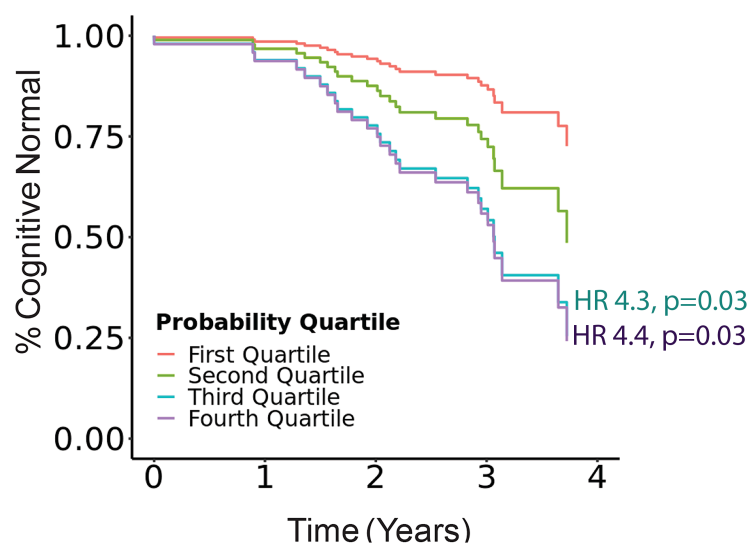


F 3 Protein Model (Validation Cohort)

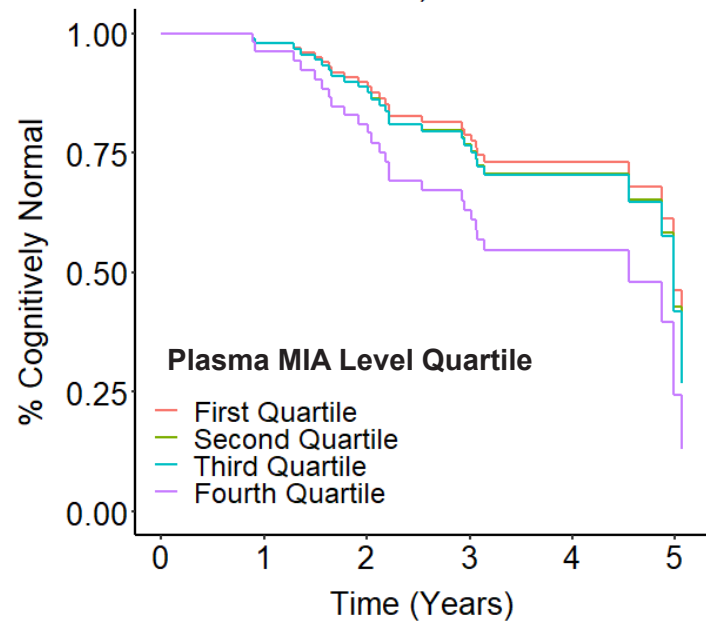
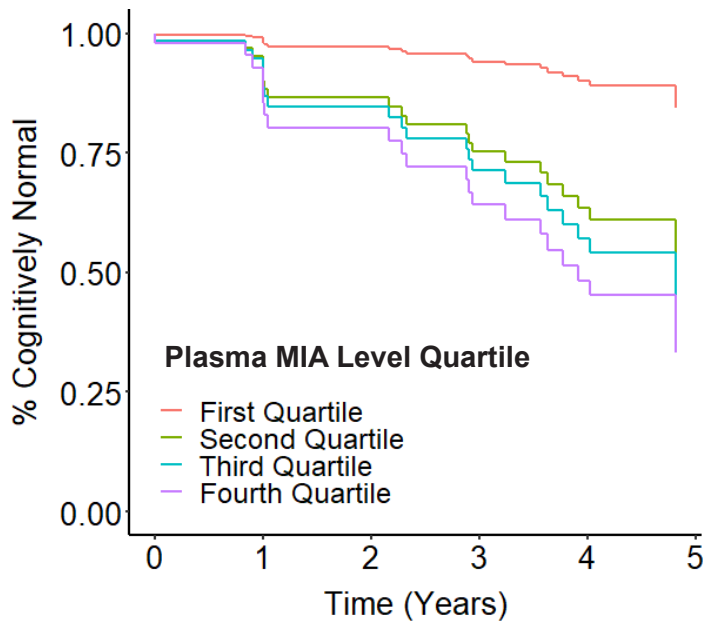


G

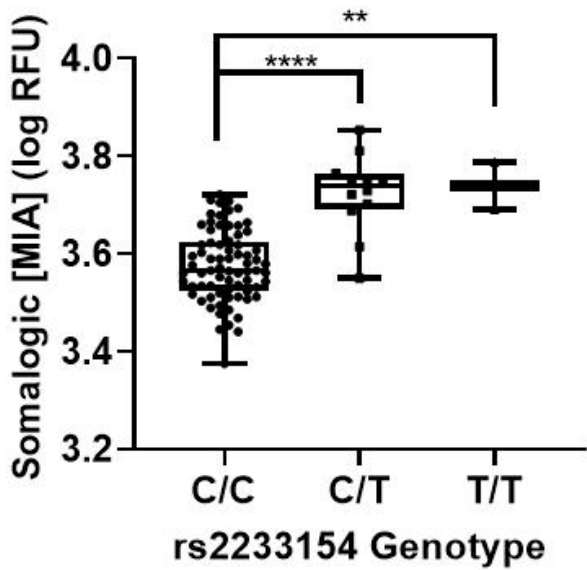
Incident MCI/Dementia by Risk Score Quartile



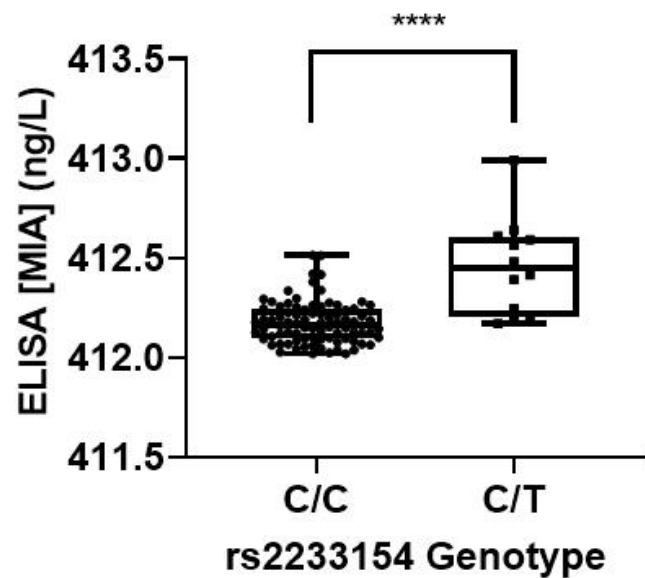
A Incident MCI/Dementia by MIA subgroup (Replication Cohort) **B** Incident MCI/Dementia by MIA subgroup (Validation Cohort)



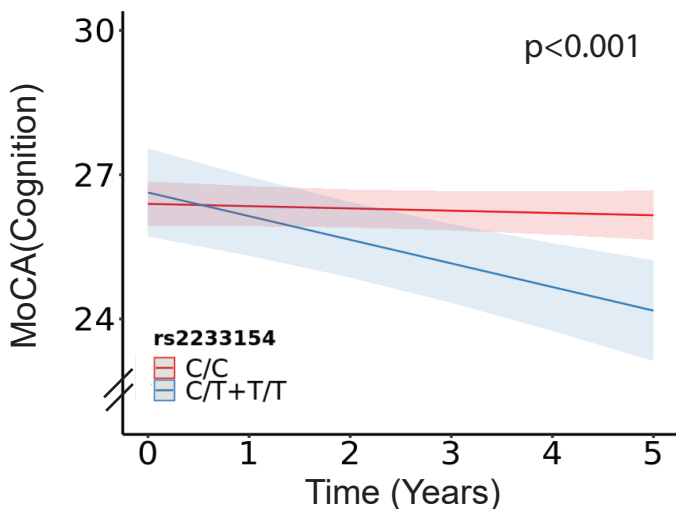
C MIA pQTL (Replication Cohort)



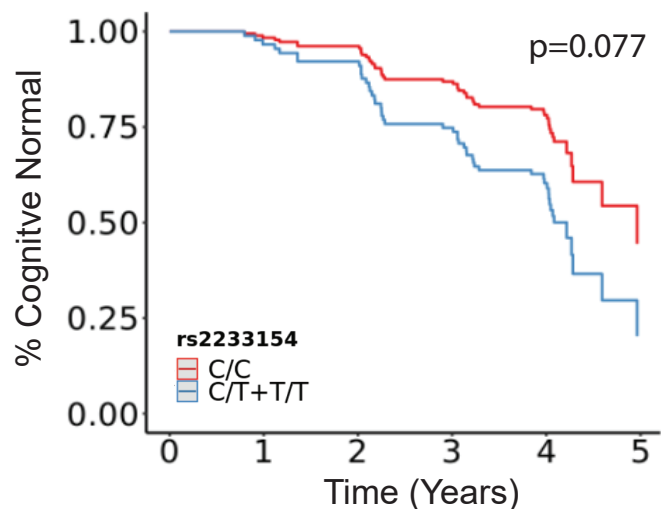
D MIA pQTL (Validation Cohort)



E Cognitive Change by MIA Genotype (Replication + Validation Cohort)



F Incident MCI/Dementia by MIA Genotype (Replication + Validation Cohort)



SUPPLEMENTARY MATERIALS

Shen et al, “Unbiased screen of 940 proteins identifies MIA, CRP, and albumin as plasma biomarker predicting cognitive decline in Parkinson’s Disease.”

INCLUDES:

(1) Supplementary Methods

(2) Supplementary Figure 1

Supplementary Methods

Cohorts

University of Texas Southwestern (UTSW) Discovery Cohort. The NIH-NINDS PDBP is the parent study into which the UTSW Discovery Cohort was enrolled. All individuals were followed on an annual basis; cognitive testing by the Montreal Cognitive Assessment (MoCA) was obtained longitudinally, and blood was collected under standard operating procedures.¹ Out of 115 PD individuals for which SOMAScan data was previously obtained, 108 individuals had a MoCA score of 19 or greater at the time of plasma sampling, as well as longitudinal cognitive testing data, and these individuals were included in the UTSW Discovery Cohort.^{1,2} The maximum follow-up time for the UTSW Discovery Cohort is four years (interquartile range [IQR] is 1.0-3.0 years with a median follow-up of 3.0 years). The UTSW institutional review board (IRB) approved study protocols, and all participants were consented for the study.

UPenn Replication Cohort and UPenn Validation Cohort. The parent study of both the UPenn Replication Cohort and the UPenn Validation Cohort is the Clinical Core of the UPenn NIA U19 (Center on alpha-synuclein strains in Alzheimer's disease and related dementias at the Perelman School of Medicine at the University of Pennsylvania, formerly the Morris K. Udall Center). Over the past 13 years, the Clinical Core of the UPenn NIA U19 has recruited PD subjects to participate in a longitudinal study that includes (1) serial assessment of cognition, and (2) biofluid sampling for biomarker and DNA analysis, with approximately 180 active participants at any given time. 201 PD participants from the UPenn NIA U19 cohort had plasma

samples available for biomarker testing, were non-demented at the time of plasma sampling, and had at least 4 years of cognitive testing and determination of cognitive diagnosis by clinical consensus after the time of sampling. Of these 201, SOMAscan measures were previously obtained on 83, and these individuals formed the UPenn Replication Cohort. The remaining 118 PD individuals comprised the UPenn Validation Cohort. Written informed consent was obtained at study enrollment, and the UPenn IRB approved study protocols. Although the UPenn NIA U19 PD participants have up to 13 years of follow-up, for our protein biomarker analyses, we used cognitive data from only four or five years of follow-up after plasma sampling (time period is indicated in the text) in this study for two reasons. First, we sought to match the Replication Cohort duration of follow-up to that of the Discovery Cohort. Second, as the vast majority of PD individuals will develop dementia over their disease course, we reasoned that the development of incident MCI or dementia in a fixed, relatively short period of time is more clinically meaningful for the purposes of our survival analyses.

Statistical Analyses

Nomination of proteins that differentiated PD individuals with Fast vs. Slow Cognitive

Decline. A linear regression model was used to identify proteins whose plasma concentration

associated significantly with cognitive subgroup (fast vs. slow cognitive decline) in both the

Discovery and Replication Cohorts. Age and disease duration at the time of plasma sampling,

sex, as well as cognitive subgroup (fast vs. slow) were included as independent variables, while

the individual protein concentrations (n=940) were used as outcome variables. Proteins that

associated with cognitive decline subgroup in both the Discovery and Replication Cohorts (1) at

a nominal p-value of 0.1 or less and (2) with the same direction of association in both cohorts were selected for downstream investigation.

Development of models to predict fast vs. slow cognitive decline. The nine proteins selected for their nominal association with fast vs. slow cognitive decline subgroups were incorporated in a logistic regression model for binary classification. Specifically, our model predicted whether an individual would belong to the fast or slow cognitive decline subgroup based on measures for these nine proteins, age, sex, and disease duration, using the Discovery Cohort to train the model (obtaining weights for all variables). We then tested the exact model developed in the Discovery Cohort on the Replication Cohort, in order to evaluate its performance in a cohort whose data were never used to train the model. Performance was assessed with Receiver Operating Curve (ROC) analyses; we obtained area under the curve (AUC) measures by fivefold cross-validation over 50 iterations. In addition, we also developed a simpler model incorporating only three top proteins, age, sex, and disease duration, using the same methods as for the nine-protein model.

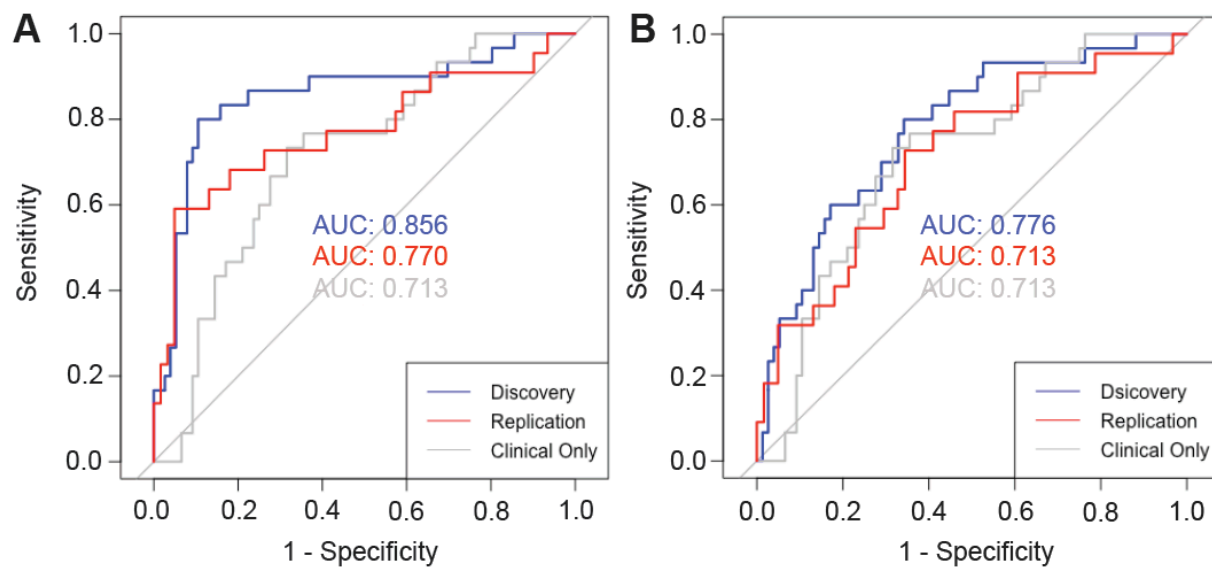
Linear mixed-effect model analyses. Linear mixed-effects models were used to evaluate the effect of biomarker levels on change in cognitive scores over time. A random intercept was introduced to the mixed effects model to account for correlations among multiple repeated measurements. Fixed effects included the interaction between plasma protein measure stratified by quartile and time, baseline MoCA score, age at sample collection, sex, and disease duration.

Cox proportional hazard models and survival analyses. To understand whether biomarkers associated with clinical outcomes, we performed survival analyses based on whether individuals

developed incident MCI or dementia (converting from normal cognition to MCI, normal to dementia, or MCI to dementia, as determined by expert clinical consensus). In some analyses, individuals were binned into quartiles based on (1) measures for candidate protein biomarkers or (2) calculated risk score for fast cognitive decline. Hazard ratios for development of incident MCI or dementia were then calculated for subgroups of patients as indicated in the text. For comparison of subgroups based on quartiles of protein measures, we performed Cox proportional hazards analyses, adjusting for age at sample, sex and disease duration. For comparison of subgroups based on calculated risk score for fast cognitive decline, we did not adjust for covariates, as clinical variables were already incorporated in the model for calculating cognitive decline risk scores.

REFERENCES FOR SUPPLEMENTARY METHODS

1. Rosenthal LS, Drake D, Alcalay RN, et al. The NINDS Parkinson's Disease Biomarkers Program. *Movement Disord.* 2016 Jun;31(6):915-23.
2. Gwinn K, David KK, Swanson-Fischer C, et al. Parkinson's disease biomarkers: perspective from the NINDS Parkinson's Disease Biomarkers Program. *Biomark Med.* 2017 May;11(6):451-73.



Supplementary Figure 1. (A-B) Performance characteristics of the logistic regression model for predicting whether PD patient falls in the fast vs. slow cognitive decline subgroup, trained using the measurements of 9 proteins (panel A) or only 3 proteins (MIA, CRP, Albumin, panel B) together with age, sex, disease duration, and baseline MoCA.