

Caractérisation de la maladie de Parkinson à l'aide de biomarqueurs sanguins: une analyse protéomique multicohort

Marijan Posavi, Maria Diaz-Ortiz, Benjamine Liu, Christine R. Swanson, R. Tyler Skrinak, Pilar Hernandez-Con, Defne A. Amado, Michelle Fullard, Jacqueline Rick, Andrew Siderowf, Daniel Weintraub, Leo McCluskey, John Q. Trojanowski, Richard B. Dewey Jr, Xuemei Huang, Alice S. Chen-Plotkin

Publié: 11 octobre 2019 • <https://doi.org/10.1371/journal.pmed.1002931>

Abstrait

Contexte

La maladie de Parkinson (MP) est une maladie neurodégénérative évolutive qui touche environ 5 millions de personnes dans le monde sans aucun traitement de fond. Nous avons recherché des biomarqueurs sanguins afin de fournir une caractérisation moléculaire des personnes atteintes de MP pour la confirmation diagnostique et la prédiction de la progression.

Méthodes et résultats

Dans 141 échantillons de plasma (96 PD, 45 individus témoins neurologiquement normaux [NC]; 45,4% de femmes, âge moyen 70,0 ans) d'une cohorte Discovery suivie longitudinalement basée à l'Université de Pennsylvanie (UPenn), nous avons mesuré les niveaux de 1129 protéines à l'aide d'une plateforme basée sur aptamère. Nous avons modélisé la concentration plasmatique des protéines (\log_{10} des unités de fluorescence relatives [RFU]) comme l'effet du groupe de traitement (PD versus NC), l'âge au moment de la collecte de plasma, le sexe et la dose quotidienne équivalente de lévodopa (LEDD), dérivant le candidat de premier passage biomarqueurs protéiques basés sur p -valeur pour PD par rapport à NC. Ces protéines candidates ont ensuite été classées par sélection de stabilité. Nous avons confirmé les résultats de notre cohorte de découverte dans une cohorte de réplication de 317 individus (215 PD, 102 NC; 47,9% de femmes, âge moyen 66,7 ans) du programme multisite suivi longitudinalement du National Institute of Neurological Disorders and Stroke Parkinson's Disease Biomarker Program (PDBP) Cohorte. L'approche analytique de la cohorte de réplication a reflété l'approche de la cohorte de découverte: chaque concentration plasmatique de protéine (\log_{10} de RFU) a été modélisée comme l'effet du groupe (PD par rapport à NC), l'âge à la collecte de plasma, le sexe, le site clinique et le lot. Parmi les 10 principales protéines de la cohorte Discovery classées par sélection de stabilité, quatre associations ont été répliquées dans la cohorte de réplication. Ces biomarqueurs sanguins étaient la sialoprotéine osseuse (BSP, Discovery false discovery rate [FDR]-corrected $p = 2,82 \times 10^{-2}$, Replication FDR-corrected $p = 1,03 \times 10^{-4}$), l'ostéomoduline (OMD, Discovery FDR-corrected $p = 2,14 \times 10^{-2}$, réplication corrigée FDR $p = 9,14 \times 10^{-5}$), aminoacylase-1 (ACY1, Discovery FDR corrigée $p = 1,86 \times 10^{-3}$, Réplication FDR corrigé $p = 2,18 \times 10^{-2}$) et récepteur de l'hormone de croissance (GHR, Discovery FDR corrigé $p = 3,49 \times 10^{-4}$, réplication FDR corrigé $p = 2,97 \times 10^{-3}$). Les mesures de ces protéines n'étaient pas significativement affectées par les différences dans la manipulation des échantillons, et elles n'ont pas changé en comparant les échantillons de plasma de 10 participants PD échantillonnés à la fois sur et hors médicaments dopaminergiques. Les mesures plasmatiques de l'OMD, ACY1 et GHR différaient en PD par rapport à NC, mais ne différaient pas entre les individus atteints de sclérose latérale amyotrophique (SLA, $n = 59$) par rapport à NC. Dans la cohorte Discovery, les personnes dont les niveaux de base de GHR et d'ACY1 se situaient dans le tertile le plus bas étaient plus susceptibles de progresser vers une déficience cognitive légère (ICM) ou une démence dans les analyses des risques proportionnels de Cox ajustés en fonction de l'âge, du sexe et de la durée de la maladie (rapport de risque [HR] 2,27 [IC à 95% 1,04–5,0, $p = 0,04$] pour GHR, et HR 3,0 [IC à 95% 1,24–7,0, $p = 0,014$] pour ACY1). L'association de la GHR avec le déclin cognitif a été confirmée dans la cohorte de réplication (HR 3,6 [IC à 95% 1,20–11,1, $p = 0,02$]). Les principales limites de cette étude étaient sa dépendance sur la plate-forme basée sur les aptamères pour la mesure des protéines et le temps de suivi limité disponible pour certaines cohortes.

Conclusions

Dans cette étude, nous avons constaté que les biomarqueurs sanguins BSP, OMD, ACY1 et GHR sont solidement associés à la MP sur plusieurs sites cliniques. Nos résultats suggèrent que des biomarqueurs basés sur un échantillon de sang périphérique peuvent être développés à la fois pour la caractérisation de la maladie et la prédiction de la progression future de la maladie dans la MP.

Résumé de l'auteur

Pourquoi cette étude a-t-elle été réalisée?

- Il n'existe actuellement aucun test sanguin qui distingue les personnes atteintes de la maladie de Parkinson (MP) des individus neurologiquement normaux ou qui prédisent le taux de progression de la maladie chez les personnes ayant déjà reçu un diagnostic de MP.
- Les tests sanguins qui distinguent les personnes atteintes de MP seraient utiles pour la confirmation du diagnostic (biomarqueurs diagnostiques), tandis que les tests sanguins qui prédisent le taux de progression de la maladie (biomarqueurs pronostiques) seraient utiles pour les essais cliniques et les soins cliniques.

Qu'ont fait et trouvé les chercheurs?

- Nous avons criblé plus de 1000 protéines sanguines de 527 personnes atteintes de MP, de sclérose latérale amyotrophique (SLA) ou d'aucune maladie neurologique afin de découvrir de nouveaux biomarqueurs diagnostiques et pronostiques. Nous avons utilisé un groupe de participants pour identifier les biomarqueurs potentiels, puis utilisé un groupe distinct de participants pour confirmer ces biomarqueurs.
- Nous avons constaté que les taux sanguins de quatre protéines - la sialoprotéine osseuse (BSP), l'ostéomoduline (OMD), l'acétylcholinestérase-1 (ACY1) et le récepteur de l'hormone de croissance (GHR) - différaient constamment chez les personnes atteintes de MP par rapport aux personnes sans MP.

- Nous avons constaté que des niveaux inférieurs de GHR au départ prédisaient un taux plus rapide de déclin cognitif chez les personnes atteintes de MP.

Que signifient ces résultats?

- Les niveaux de certaines protéines sanguines diffèrent constamment entre les personnes atteintes de la maladie de Parkinson et celles qui n'en ont pas, et certaines de ces protéines prédisent également les personnes atteintes de la maladie de Parkinson qui pourraient avoir une progression plus rapide de la maladie.
- Il peut être possible de développer des tests sanguins pour aider à confirmer le diagnostic de la MP et à prédire la progression de la maladie.

Référence: Posavi M, Diaz-Ortiz M, Liu B, Swanson CR, Skrinak RT, Hernandez-Con P, et al. (2019) Caractérisation de la maladie de Parkinson à l'aide de biomarqueurs sanguins: une analyse protéomique multicohort. PLoS Med 16 (10): e1002931. <https://doi.org/10.1371/journal.pmed.1002931>

Rédacteur académique: Rodolfo Savica, Université de l'Utah, ÉTATS-UNIS

Reçu: 4 avril 2019; **Accepté:** 5 septembre 2019; **Publié:** 11 octobre 2019

Il s'agit d'un article en libre accès, libre de tout droit d'auteur, et peut être librement reproduit, distribué, transmis, modifié, développé ou utilisé par quiconque à des fins légales. Le travail est rendu disponible dans le cadre de la dédicace du domaine public [Creative Commons CC0](https://creativecommons.org/licenses/by/4.0/).

Disponibilité des données: Toutes les données sont contenues dans le manuscrit et ses fichiers d'informations complémentaires.

Financement: Cette étude a été financée par les NIH / NINDS UO1 NS082134 et UO1 NS097056 (ASC-P), UO1 NS082148 (RBD) et P50 NS053488 (ASC-P, JQT, DW). ASC-P détient le poste de professeur doté de la famille Parker et a également été soutenu par le Fonds Benaroya. Les bailleurs de fonds n'ont joué aucun rôle dans la conception de l'étude, la collecte et l'analyse des données, la décision de publier ou la préparation du manuscrit.

Intérêts concurrents: J'ai lu la politique de la revue et les auteurs de ce manuscrit ont les intérêts concurrents suivants: L'Université de Pennsylvanie a une demande de brevet sur les biomarqueurs de la maladie de Parkinson basée sur les résultats de ce manuscrit. De plus, RBD déclare participer au programme NINDS de biomarqueurs de la maladie de Parkinson à partir duquel ce travail a été développé (NIH-NINDS-UO1 NS082148).

Abréviations: ACY1, aminoacylase-1; AD, maladie d'Alzheimer; SLA, sclérose latérale amyotrophique; BSP, sialoprotéine osseuse; LCR, liquide céphalo-rachidien; CTCC, Centre de coordination des essais cliniques; CV, coefficient de variation; DRS, Mattis Dementia Rating Scale-2; FDA, US Food and Drug Administration; FDR, taux de fausses découvertes; GH, hormone de croissance; GHR, récepteur de l'hormone de croissance; HR, rapport de risque; HSD, différence significative honnête; IGF-1, facteur de croissance 1 analogue à l'insuline; IGFBP-1, protéine 1 de liaison au facteur de croissance analogue à l'insuline; IRB, comité d'examen institutionnel; LASSO, opérateur de retrait et de sélection le moins absolu; LEDD, dose quotidienne équivalente de lévodopa; LLOD, limite inférieure de détection; MCI, trouble cognitif léger; MoCA, Évaluation cognitive de Montréal; NC, contrôle neurologiquement normal; OMD, ostéomoduline; PCSK9, proprotéine convertase subtilisine / kexine de type 9; PD, maladie de Parkinson; PDBP, programme de biomarqueurs de la maladie de Parkinson; QC, contrôle de qualité; RFU, unité de fluorescence relative; SD, écart type; SOMAmer, aptamère modifié à faible débit; STROBE, Renforcement de la notification des études d'observations en épidémiologie; ULOD, limite supérieure de détection; UPenn, Université de Pennsylvanie; UTSW, University of Texas Southwestern Medical Center déficience cognitive légère; MoCA, Évaluation cognitive de Montréal; NC, contrôle neurologiquement normal; OMD, ostéomoduline; PCSK9, proprotéine convertase subtilisine / kexine de type 9; PD, maladie de Parkinson; PDBP, programme de biomarqueurs de la maladie de Parkinson; QC, contrôle de qualité; RFU, unité de fluorescence relative; SD, écart type; SOMAmer, aptamère modifié à faible débit; STROBE, Renforcement de la notification des études d'observations en épidémiologie; ULOD, limite supérieure de détection; UPenn, Université de Pennsylvanie; UTSW, University of Texas Southwestern Medical Center déficience cognitive légère; MoCA, Évaluation cognitive de Montréal; NC, contrôle neurologiquement normal; OMD, ostéomoduline; PCSK9, proprotéine convertase subtilisine / kexine de type 9; PD, maladie de Parkinson; PDBP, programme de biomarqueurs de la maladie de Parkinson; QC, contrôle de qualité; RFU, unité de fluorescence relative; SD, écart type; SOMAmer, aptamère modifié à faible débit; STROBE, Renforcement de la notification des études d'observations en épidémiologie; ULOD, limite supérieure de détection; UPenn, Université de Pennsylvanie; UTSW, University of Texas Southwestern Medical Center l'ostéomoduline; PCSK9, proprotéine convertase subtilisine / kexine de type 9; PD, maladie de Parkinson; PDBP, programme de biomarqueurs de la maladie de Parkinson; QC, contrôle de qualité; RFU, unité de fluorescence relative; SD, écart type; SOMAmer, aptamère modifié à faible débit; STROBE, Renforcement de la notification des études d'observations en épidémiologie; ULOD, limite supérieure de détection; UPenn, Université de Pennsylvanie; UTSW, University of Texas Southwestern Medical Center l'ostéomoduline; PCSK9, proprotéine convertase subtilisine / kexine de type 9; PD, maladie de Parkinson; PDBP, programme de biomarqueurs de la maladie de Parkinson; QC, contrôle de qualité; RFU, unité de fluorescence relative; SD, écart type; SOMAmer, aptamère modifié à faible débit; STROBE, Renforcement de la notification des études d'observations en épidémiologie; ULOD, limite supérieure de détection; UPenn, Université de Pennsylvanie; UTSW, University of Texas Southwestern Medical Center SD, écart type; SOMAmer, aptamère modifié à faible débit; STROBE, Renforcement de la notification des études d'observations en épidémiologie; ULOD, limite supérieure de détection; UPenn, Université de Pennsylvanie; UTSW, University of Texas Southwestern Medical Center

introduction

La maladie de Parkinson (MP) est caractérisée par une perte progressive des neurones dopaminergiques dans la substance noire, entraînant un syndrome clinique défini par la bradykinésie, la rigidité, les tremblements et l'instabilité posturale [1]. Au moment du diagnostic clinique, 50% des neurones dopaminergiques nigraux peuvent déjà être perdus [2], suggérant une longue phase prodromique au cours de laquelle une intervention peut être possible. La pratique médicale actuelle pour le diagnostic de la MP repose presque entièrement sur un examen clinique, sans tests en laboratoire disponibles. Bien qu'un test d'imagerie du transporteur de dopamine (DaTSCAN) basé sur des radioligands, approuvé par la Food and Drug Administration (FDA) des États-Unis, puisse confirmer la dégénérescence des neurones dopaminergiques [3], le temps et les frais ont empêché son adoption

généralisée en milieu clinique, et un résultat positif n'est pas diagnostique pour la MP, car d'autres maladies neurologiques dégénératives telles que l'atrophie de systèmes multiples présentent des résultats similaires. De plus, même au sein de la DP, il existe une hétérogénéité considérable dans la présentation clinique, avec des taux très variables de progression cognitive et motrice dans le temps [4]. À l'heure actuelle, il n'existe aucun test clinique ou basé sur la recherche pour prédire la progression de la maladie de Parkinson, malgré la reconnaissance largement répandue que de tels outils prédictifs sont vitaux dans le domaine [5]. Ainsi, l'avènement de marqueurs sanguins pour définir moléculairement les individus de la MP et pour prédire la progression longitudinale de la MP pourrait transformer la pratique clinique et le développement de thérapies modificatrices de la maladie.

À ce jour, les études sur les biomarqueurs dans la MP se sont largement concentrées sur des approches candidates, en mettant l'accent sur les mesures protéiques obtenues dans le liquide céphalo-rachidien (LCR) [6 , 7], qui est considérablement plus difficile à obtenir que le sang dans un contexte clinique chargé. Bien qu'une poignée de biomarqueurs nommés à l'aide de ces approches candidates diffèrent constamment en comparant la PD et les individus témoins (par exemple, les mesures du LCR de l'alpha-synucléine totale [8]), les tailles d'effet des marqueurs individuels sont petites et la rareté des biomarqueurs robustes limite la capacité de développement panneaux multimarqueurs pour un meilleur pouvoir discriminant. De plus, le domaine manque largement de biomarqueurs protéiques qui prédisent la progression cognitive ou motrice dans plusieurs cohortes [7]. Ainsi, nous avons cherché à découvrir de nouveaux biomarqueurs sanguins pour la différenciation des individus atteints de MP des individus témoins, ainsi que la prédiction du taux de progression de la MP. Nous avons abordé ce problème en criblant plus de 1 000 protéines plasmatiques à l'aide d'une plate-forme à base d'aptamères [9] dans un plan de découverte-réplication.

Méthodes

Pour identifier les biomarqueurs susceptibles de caractériser les personnes atteintes de MP, 1 129 (cohorte de découverte) et 1 305 (cohorte de réplication du programme de biomarqueurs de la maladie de Parkinson [PDBP]) ont été criblés chez 527 individus et leurs données cliniques ont été analysées ([figure 1](#)). Les données cliniques et les échantillons de plasma ont été acquis auprès de la cohorte Udall de l'Université de Pennsylvanie (UPenn) (cohorte découverte) et d'une cohorte de réplication PDBP multisite ([tableau 1](#)). L'étude comprenait trois étapes principales: (1) différenciation de la MP des participants au contrôle neurologique normal (NC) (96 PD, 45 NC) dans une cohorte de découverte à un seul site, en utilisant des niveaux de 968 protéines qui ont passé les mesures de contrôle de qualité (QC), obtenu par un test de plate-forme à base d'aptamères; (2) réplication indépendante des meilleurs biomarqueurs candidats de l'étape 1 dans une cohorte de réplication multisite (215 PD, 102 participants NC); et (3) prédiction de la progression de la MP à l'aide des biomarqueurs candidats qui se sont répliqués dans la cohorte de découverte à un seul site et la cohorte de réplication multisite. Ces étapes sont résumées dans la [figure 1](#), dont certaines parties ont été créées avec Biorender.com. Les données ont été normalisées, traitées et analysées à l'aide du logiciel statistique R [10]. Cette étude est rapportée conformément aux directives de renforcement de la déclaration des observations des études en épidémiologie (STROBE) ([liste de contrôle S1 STROBE](#)). Au début de l'étude, le plan d'analyse ([S1 Analysis Plan](#)) consistait à enquêter de manière flexible sur la cohorte de découverte, puis à effectuer une analyse reflétant l'analyse de la cohorte de découverte dans la cohorte de réplication, avec les deux covariables supplémentaires du site clinique et du lot, si des effets de site ou de lot ont été observés lors de la transition d'un site unique / phase de découverte mono-lot à une phase de réplication multisite / multi-lots. Le plan d'analyse ultime ne différait du plan prédéterminé qu'en ce que la dose quotidienne équivalente de lévodopa (LEDD) n'était pas incluse comme covariable dans la cohorte de réplication lorsque nous avons constaté que tous les participants PDBP PD n'avaient pas de données LEDD disponibles. Les méthodes détaillées sont décrites ci-dessous.

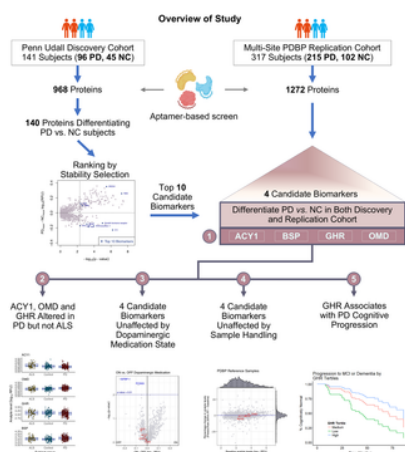


Fig 1. Aperçu de l'étude.

Nous avons utilisé les dosages SOMAScan 1.1k (panneau de gauche) et 1.3k (panneau de droite) pour quantifier les niveaux de protéines plasmatiques dans les cohortes de découverte et de réplication, respectivement. Des modèles de régression linéaire multiples ont révélé 140 protéines qui différaient entre les individus PD et NC dans la cohorte Discovery ($p < 0,005$). La sélection de stabilité avec LASSO a ensuite été utilisée pour classer les 140 protéines candidates de la cohorte Discovery, en obtenant le top 10 des protéines les plus robustes et stables. Quatre (ACY1, BSP, GHR, OMD) de ces 10 protéines différaient dans les échantillons PD par rapport aux échantillons NC dans la cohorte de découverte et ont également différencié PD par rapport à NC dans la cohorte de réplication (panneau du milieu, 1). Ceux-ci ont été testés pour la spécificité de la MP (2), l'influence des médicaments dopaminergiques ou la manipulation des échantillons (3, 4) et la capacité à prédire les taux ultérieurs de déclin cognitif (5). ACY1, aminoacylase-1; SLA, sclérose latérale amyotrophique; BSP, sialoprotéine osseuse; GHR, récepteur de l'hormone de croissance; LASSO, opérateur de retrait et de sélection le moins absolu; MCI, trouble cognitif léger; NC, contrôle neurologiquement normal; OMD, ostéomoduline; PD, maladie de Parkinson; PDBP, programme de biomarqueurs de la maladie de Parkinson.

<https://doi.org/10.1371/journal.pmed.1002931.g001>

Disease group	Penn Udall Discovery Cohort		PDBP Replication Cohort (Penn State and UTSW)		Penn MSK		BioFIND	
	PD	NC	PD	NC	ASP	ON	OFF	OFF
Participants total	NC	47	157	162	78	78	78	78
Female/Male	62/14	22/23	101/114	70/71	20/73	20/73	3/75	3/75
<i>p</i> -value	0.007		0.76		0.007		0.007	
Age at discovery (mean ± SD) (years)	48.9 ± 7.52	70.2 ± 10.06	68.4 ± 8.4	68.1 ± 10.3	63.9 ± 10.3	63.9 ± 10.3	73.9 ± 6.13	73.9 ± 6.13
<i>p</i> -value	0.007		0.007 (AS vs NC)		0.007 (AS vs NC)		0.007 (AS vs NC)	
Collection site number	N/A	N/A	101	162	N/A	N/A	N/A	N/A
From Penn UTSW	N/A	N/A	100/101	78/74	N/A	N/A	N/A	N/A
<i>p</i> -value	N/A		0.007		N/A		N/A	
Blood sample storage at -80 °C (mean ± SD) (months)	1.84 ± 0.88	1.43 ± 0.10	3.42 ± 0.34	3.68 ± 0.72	1.13 ± 2.01	1.28 ± 0.78	1.24 ± 0.78	1.24 ± 0.78
<i>p</i> -value	0.007		0.007 (AS vs NC)		0.007		0.007	
Sample plate number	N/A	N/A	101	162	N/A	N/A	N/A	N/A
Batch 1-2 (3x3)	N/A	N/A	89/88/88/82	21/20/24/24/13	N/A	N/A	N/A	N/A
<i>p</i> -value	N/A		0.007		N/A		N/A	
Baseline (MSK) (Penn) or (MSK, A, UTSW) (mean ± SD)	136.93 ± 9.96	N/A	26.47 ± 2.50*	N/A	N/A	N/A	N/A	N/A

Tableau 1. Caractéristiques démographiques des participants à l'étude.
<https://doi.org/10.1371/journal.pmed.1002931.t001>

Cohortes et prélèvement d'échantillons

Cohorte de découverte UPenn Udall.

Au cours de la période entre 2013 et 2015, des échantillons de plasma sanguin et des données cliniques ont été collectés auprès de 97 personnes PD et 45 CN inscrits pour participer à des recherches approuvées par l'UPenn Institutional Review Board (IRB). Toutes les personnes atteintes de la maladie de Parkinson remplissaient les critères diagnostiques de la banque britannique de cerveaux de la maladie de Parkinson [9] et faisaient partie d'une cohorte longitudinale largement caractérisée à l'UPenn [11]. Afin de contrôler les préjugés environnementaux, nous nous sommes assurés que les groupes PD et témoins ne différaient pas selon l'âge ou le sexe, et les NC ont été recrutés principalement parmi les conjoints non affectés de personnes PD de la même clinique. Les échantillons ont été acquis selon les protocoles approuvés par l'IRB comme décrit précédemment par Chen-Plotkin et ses collègues [12]. Le consentement éclairé écrit a été obtenu lors de l'inscription à l'étude. Un échantillon de PD avait des valeurs aberrantes dans les étapes de prétraitement et de normalisation sur le test à base d'aptamères et a été exclu des analyses ultérieures. Un total de 96/97 PD et tous les échantillons NC ont passé nos critères de CQ (voir Prétraitement et CQ des données sur les protéines SOMAScan pour une description du prétraitement et de la normalisation) et ont été inclus dans les analyses ultérieures.

Cohorte de réplication PDBP multisite.

Les résultats ont été testés dans une cohorte de réplication afin de tenir compte des biais environnementaux et techniques possibles dans notre analyse. Les échantillons sanguins de la cohorte de réplication (collectés entre 2013 et 2015) et les données cliniques ont été obtenus à partir de la cohorte PDBP [13], provenant de participants à la recherche vus sur deux sites du PDBP: Penn State University (Penn State, 100 PD et 78 NC) et University of Texas Southwestern Medical Center (UTSW, 115 PD et 24 NC). Un participant PD de l'UTSW et deux participants NC de Penn State ont été exclus des analyses en raison de mesures aberrantes pour une forte proportion de protéines SOMAScan. Un suivi longitudinal avec des tests cognitifs par le Montreal Cognitive Assessment (MoCA) a également été obtenu, et des associations avec des niveaux de biomarqueurs analysés, pour les participants à l'UTSW PD. L'IRB local de chaque centre PDBP a approuvé les protocoles d'étude, et tous les participants ont été consentants pour l'étude.

Déclaration d'éthique pour la recherche sur les participants humains.

Pour la cohorte Discovery, l'IRB de l'UPenn a approuvé la recherche sur les participants humains dans cette étude. Le consentement éclairé écrit a été obtenu des participants de la cohorte Discovery. Pour la cohorte de réplication, l'IRB local de chaque centre PDBP a approuvé les protocoles d'étude, et tous les participants ont fourni un consentement éclairé écrit pour participer au PDBP. Étant donné que l'un des objectifs du PDBP est de fournir un biorepository d'échantillons provenant d'un ensemble d'individus bien caractérisés, les participants ont consenti au partage d'échantillons et de données désidentifiées avec des enquêteurs approuvés par le Comité d'accès à l'examen des biospécimens au moment de l'inscription.

Échantillons de référence groupés.

Pour étudier l'effet de la manipulation du plasma sur le niveau de biomarqueur détecté, plusieurs aliquotes identiques d'échantillons de plasma regroupés de la découverte (pool de référence UPenn) et des cohortes de réplication (pool de référence PDBP) ont été utilisées. Le pool de référence UPenn a été préparé en mélangeant des échantillons de 450 individus PD et NC. Le pool de référence PDBP comprenait des échantillons de plasma provenant de 13 individus PD et NC. Pour ces études, une aliquote a été dosée directement par SOMAScan, tandis qu'une aliquote identique a été soumise à une incubation à température ambiante pendant 30 minutes, suivie d'un cycle supplémentaire de congélation-décongélation, avant d'être dosée par SOMAScan.

Échantillons de la cohorte BioFIND.

L'effet du traitement par la lévodopa sur la concentration de protéines plasmatiques a été testé sur 10 sujets PD sélectionnés au hasard (5 femmes et 5 hommes de plus de 50 ans) de la cohorte BioFIND ([tableau 1](#)) [14]. Des échantillons de plasma de la cohorte BioFIND ont été collectés lors de deux visites différentes: au départ (lorsque les personnes atteintes de DP avaient des échantillons prélevés tout en prenant leur médicament dopaminergique habituel - c'est-à-dire, un médicament ON) et 2 semaines après la visite de référence (lorsque les éliminations nocturnes des médicaments dopaminergiques, c'est-à-dire des médicaments OFF), comme décrit par Kang et ses collègues [14]. Tous les protocoles d'étude et stratégies de recrutement pour BioFIND ont été approuvés par les IRB du Centre de coordination des essais cliniques (CTCC) de l'Université de Rochester et des sites cliniques individuels.

Quantification des protéines

Les échantillons des cohortes de découverte et de réplication ont été analysés en utilisant les versions 1.1k et 1.3k Assay de la plate-forme SOMAScan (Somalogic, Boulder, CO, USA [9]) en deux séries distinctes, les opérateurs étant aveugles au statut de la maladie. Cette plate-forme est basée sur des aptamères modifiés à faible vitesse de capture de protéines (SOMAmers), qui sont

des oligonucléotides modifiés chimiquement avec une affinité spécifique pour les protéines cibles recombinantes, développés par sélection *in vitro* (SELEX) comme décrit précédemment [15 , 16].

Les étapes spécifiques du test SOMAScan ont été décrites en détail dans des publications antérieures [9 , 17 , 18], ainsi que dans des livres blancs techniques sur www.somalogic.com . En bref, des échantillons de plasma ont été incubés avec des mélanges de réactifs contenant des SOMAmères pour permettre la liaison à l'équilibre des aptamères marqués au fluorophore à leurs protéines cibles. Ensuite, une série d'étapes de partitionnement et de lavage ont été utilisées pour capturer uniquement les SOMAmers qui étaient liés à leurs protéines apparentées. Enfin, les oligonucléotides liés aux protéines ont été libérés du complexe protéique, capturés par complémentarité et quantifiés à l'aide de matrices d'hybridation d'ADN.

Pour ajuster les biais techniques, les matrices d'hybridation ont été normalisées et étalonnées en utilisant les données d'un ensemble de référence d'échantillons de plasma regroupés qui a été exécutée avec chaque lot. Les données somalogiques brutes en unités de fluorescence relative (RFU) ont été transformées en \log_{10} avant l'analyse. Un total de 142 échantillons de plasma (97 PD et 45 NC) de la cohorte Discovery ont été testés pour 1129 protéines (test 1,1k), et 320 échantillons de plasma (216 PD et 104 NC) de la cohorte de réplication ont été testés pour 1305 protéines (1,3 k Test). Les échantillons de la cohorte de réplication ont été analysés par lots de 85, répartis dans cinq plaques différentes, avec des échantillons de pool de référence de plasma, 59 échantillons de sclérose latérale amyotrophique (SLA) du biorepository UPenn et 20 échantillons du biorepository BioFIND.

Prétraitement et contrôle qualité des données sur les protéines SOMAScan

Les échantillons de plasma des cohortes de découverte et de réplication ont été dosés dans des séries Somalogic séparées. Des échantillons de plasma de la cohorte de découverte ont été analysés, ainsi que 13 échantillons de calibrateur de plasma, des contrôles d'hybridation et deux échantillons de contrôle de tampon. Les échantillons du pool de référence ($n = 8$), les échantillons Penn Udall ALS ($n = 59$) et les échantillons BioFIND ($n = 20$) provenant de 10 personnes atteintes de MP) ont été analysés avec des échantillons de plasma de la cohorte de réplication.

Les normes de contrôle de la qualité des essais ont été dérivées de mesures obtenues pendant le développement du test, et les méthodes de prétraitement et de normalisation sont décrites en détail dans un livre blanc technique [19]. En bref, les données d'échantillon ont d'abord été normalisées pour éliminer les artefacts d'hybridation, en utilisant des contrôles d'hybridation «enrichis». Une normalisation médiane a ensuite été appliquée pour chaque échantillon afin d'éliminer les autres biais intra-plaque. Pour le test SOMAScan, les facteurs de contrôle d'hybridation et d'échelle médiane devraient être compris entre 0,4 et 2,5 ($\pm 1,32$ sur une échelle \log_2). Tous les échantillons avaient des facteurs d'échelle d'hybridation dans la plage acceptable, à l'exception d'un échantillon PD de la cohorte Discovery, qui a été exclu des analyses en aval.

À l'étape suivante, deux critères de contrôle de la qualité ont été mis en œuvre pour filtrer les données sur les protéines SOMAScan. La variabilité technique globale intraplaque du test SOMAScan a été évaluée en utilisant trois échantillons QC (aliquotes identiques provenant de trois pools d'échantillons de référence différents) analysés en triple (pour un total de neuf échantillons). Ces trois ensembles de triplicats ont été placés au hasard dans les lots d'échantillons biologiques afin de capturer la variabilité intra-plaque. Les coefficients de variation (CV) de ces trois ensembles d'échantillons de CQ en triple ont été calculés (c.-à-d., Pour chaque protéine, trois CV ont été calculés) en utilisant les données brutes Somalogic (en RFU). Les protéines présentant des CV supérieurs à 0,2 de l'un quelconque des triplicats ont été exclues des analyses en aval. Il y avait respectivement 36 et 33 protéines dans les cohortes de découverte et de réplication, avec $CV > 0,51$ Fig). Le deuxième filtre, qui impliquait la suppression des protéines avec des mesures $> 25\%$ en dehors de la limite inférieure de détection (LLOD) ou de la limite supérieure de détection (ULOD), a été appliqué uniquement à la cohorte Discovery, car les limites de détection n'étaient pas fournies pour les versions ultérieures du SOMAScan. Cela a abouti à l'élimination de 125 protéines supplémentaires (Fig_S1A), laissant un total de 968 protéines de la cohorte de découverte pour les analyses en aval.

Traitement des données et analyses statistiques transversales

Nomination des protéines qui différaient dans le groupe PD par rapport au groupe NC.

Pour détecter les protéines dont la concentration plasmatique était associée de manière significative à la catégorie de la maladie (PD versus NC), de multiples modèles de régression linéaire ont été utilisés. Dans la cohorte Discovery, chaque concentration plasmatique de protéine (\log_{10} de RFU) a été modélisée comme l'effet du groupe de traitement (PD versus NC), l'âge à la collecte de plasma, le sexe et la LEDD. Un total de 140 biomarqueurs candidats avec une valeur p de l'effet de groupe $< 0,005$ (PD par rapport à NC) ont été nommés ([tableau S1](#)) pour les analyses en aval (regroupement hiérarchique et sélection de la stabilité) de la cohorte Discovery. Les valeurs p corrigées du taux de fausses découvertes (FDR) ont également été dérivées pour ces biomarqueurs en utilisant la méthode de Benjamini-Hochberg [20].

Clustering hiérarchique et génération de cartes thermiques.

Une carte thermique a été générée en utilisant la fonction `heatmap.2` du package R `gplots` [21]. Les données brutes Somalogic (RFU) ont été transformées en \log puis centrées et mises à l'échelle. Les participants et les protéines ont été classés hiérarchiquement par distance euclidienne et liaison moyenne utilisant la fonction `hclust` [10].

Classement de sélection de stabilité.

Nous avons effectué la sélection de stabilité [22] (sélection de variables basée sur le sous-échantillonnage en combinaison avec le retrait le moins absolu et l'opérateur de sélection [LASSO] [23]) sur les données de la cohorte de découverte (96 PD, 45 NC). Pour classer les biomarqueurs candidats, le package R `BioMark` sur 100 000 itérations jackknifed [24] a été utilisé. À chaque itération, 30% des protéines et 10% des échantillons ont été laissés hors du sac, et LASSO a été utilisé pour sélectionner des caractéristiques pour les variables sur les données restantes. La proportion d'itérations dans lesquelles LASSO a signalé un coefficient différent de zéro a été utilisée pour classer les protéines, générant une liste des 10 principales protéines à évaluer dans la cohorte de réplication.

Analyses de cohorte de réplication.

Dans la cohorte de réplication, chaque concentration plasmatique de protéine (\log_{10} de RFU) a été modélisée comme l'effet du groupe (PD contre NC), de l'âge à la collecte de plasma, du sexe, du site clinique (UTSW contre Penn State) et du lot (cinq plaques) . LEDD n'a pas été inclus comme covariable dans la cohorte de réplication lorsque nous avons constaté que tous les participants PDBP PD n'avaient pas de données LEDD disponibles dans la ressource de gestion de données PDBP. Les analyses

se sont concentrées sur les 10 principales protéines de la cohorte Discovery, classées par sélection de stabilité; Les valeurs p ont été corrigées pour les tests d'hypothèses multiples en utilisant la méthode de Benjamini-Hochberg [20]. Pour les quatre protéines validées, une analyse similaire a été répétée incluant 59 participants ALS. Concentration plasmatique de protéines (\log_{10} de RFU) a été modélisée comme l'effet du groupe de la maladie (NC, PD ou SLA), l'âge au moment de la collecte de plasma, le sexe et le site clinique. Les coefficients des groupes de maladies ont été extraits et les valeurs de p ont été ajustées en utilisant la méthode de Benjamini-Hochberg.

Test des effets du traitement par la lévodopa.

Pour tester l'effet du traitement par la lévodopa sur les taux de protéines plasmatiques (\log_{10} de RFU), un test t apparié a été appliqué à chacune des protéines testées. Les valeurs p nominales (non ajustées) du test t appariés sont présentées dans le [tableau S5](#) . De plus, les analyses ont été répétées et les valeurs p ont été obtenues par des tests de permutation par paires, ce qui évite les hypothèses de normalité. Les résultats sont inchangés ([Fig.S2](#)).

Associations avec la cognition.

La corrélation des rangs de Spearman a été calculée pour les niveaux de base de la Mattis Dementia Rating Scale-2 (DRS) et des biomarqueurs (\log_{10} de RFU) en utilisant la fonction R cor.test [10]. Les valeurs de p ont été ajustées pour FDR par la méthode de Benjamini-Hochberg en utilisant la fonction R p.adjust [10].

Analyse de la progression de la DP

Analyse de modèle linéaire à effets mixtes.

Des modèles linéaires à effets mixtes ont été ajustés pour déterminer l'effet du niveau de biomarqueur sur le taux de déclin cognitif en utilisant le package R nlme [25]. Dans la cohorte Discovery, seuls les participants avec des scores DRS mesurés dans les 6 mois suivant le prélèvement sanguin ainsi qu'au moins un score DRS ultérieur ont été inclus ($n = 91$), pour une période de suivi moyenne de 3,5 ans. Notre modèle a incorporé le DRS comme variable de réponse, avec l'âge, le sexe, la durée de la maladie, le DRS de base et l'interaction temps par protéine comme effets fixes et le participant comme effet aléatoire. La même analyse a été répétée en ajustant les années d'études comme un effet fixe supplémentaire. Les coefficients d'interaction temps par protéine ont été extraits, et le p -les valeurs du terme d'interaction ont été ajustées pour FDR par la méthode de Benjamini-Hochberg en utilisant la fonction R p.adjust [10].

Analyse de survie.

Dans les deux cohortes de découverte et de réplication, les individus ont été divisés en groupes de biomarqueurs faibles, moyens ou élevés en fonction des tertiles de biomarqueurs (\log_{10} de RFU). Pour la cohorte Discovery, nous avons extrait les diagnostics cognitifs basés sur le diagnostic de consensus clinique comme décrit précédemment [11], et les personnes atteintes de DP avec un diagnostic cognitif de démence au départ ou un diagnostic codé comme normal après un diagnostic de base de déficience cognitive légère (MCI) ont été exclues de l'analyse, laissant 86 personnes pour l'analyse de survie. Les événements ont été définis comme la conversion de la normale à la MCI, la normale à la démence ou la MCI à la démence, pour un total de 38 événements. Pour la cohorte de réplication, la catégorisation cognitive était basée sur les normes MoCA publiées (MoCA 26–30 = normal, MoCA 21–25 = MCI et MoCA 20 ou moins = démence) [26], et les événements ont été définis et les participants atteints de MP filtrés dans de la même manière que pour la Cohorte Découverte, pour un total de 26 événements observés chez 74 individus (suivis pendant une moyenne de 2,8 ans). L'analyse de survie a été réalisée dans R en utilisant la survie [27] paquet. Des analyses de risques proportionnels de Cox ont été réalisées à l'aide de la fonction coxph [27] pour tester si les groupes de biomarqueurs tertiles ont un effet sur la probabilité d'un événement, en ajustant l'âge, le sexe et la durée de la maladie. Les analyses ont été répétées, y compris les années de scolarité comme covariable supplémentaire. Les résultats ont été visualisés sous forme de courbes ajustées par régression de Cox ou de tracés forestiers utilisant les courbes ggadjustedcurves et les fonctions ggforest du package R survminer [28]. Les modèles avec et sans éducation ont été comparés à l'aide de la fonction anova.coxph du package de survie R [27].

Résultats

Écran de découverte pour les protéines plasmatiques différenciant PD de NC

À partir de l'ensemble original de 1 129 protéines analysées dans la cohorte de découverte à site unique, 968 (85,7%) répondaient aux normes de CQ ([Fig.S1](#) et [tableau S1](#)). Ces 968 protéines ont été retenues pour les analyses en aval.

Les protéines différenciant la PD des échantillons NC dans la cohorte Discovery ont été nommées à l'aide d'un modèle linéaire associant la concentration de chacune de ces 968 protéines à l'état pathologique, en ajustant la LEDD [29], l'âge au prélèvement de plasma et le sexe, générant une liste initiale de 140 candidats. biomarqueurs associés à une valeur p nominale $<0,005$ avec PD; la correction pour les tests d'hypothèses multiples utilisant la méthode de Benjamini-Hochberg [20] a démontré que les 140 protéines remplissaient le critère supplémentaire d'association à l'état pathologique avec une valeur p corrigée par FDR $<0,05$ ([tableau S2](#)).

Nous avons ensuite effectué un regroupement hiérarchique sur ces marqueurs candidats pour évaluer la structure de corrélation entre les groupes de protéines et l'état de la maladie. Le regroupement non supervisé a révélé une colinéarité entre les sous-ensembles de ces protéines, suggérant des redondances et des relations partagées possibles entre de nombreux biomarqueurs candidats ([figure 2A](#)). Nous avons donc utilisé la sélection de stabilité [22], un outil méta-statistique qui identifie les caractéristiques systématiquement importantes par sous-échantillonnage répété des données, afin d'identifier l'ensemble le plus robuste, stable et épars de protéines discriminatoires; nous avons classé les biomarqueurs candidats en utilisant la méthode LASSO sur 100 000 itérations jackknifed [23 , 24]. Les 10 principales protéines de la cohorte Discovery classées par sélection de stabilité, illustrées sur les [figures 2B et 2C](#) , ont été avancées pour la réplication.

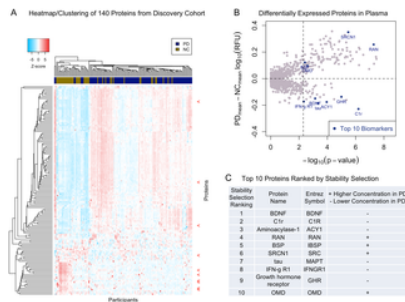


Fig 2. Identification des protéines différenciant les échantillons PD et NC dans la cohorte Discovery.

(A) Heatmap montrant les niveaux plasmatiques de 140 biomarqueurs candidats différemment exprimés ($p < 0,005$) entre PD ($n = 96$) et NC ($n = 45$) dans la cohorte Discovery (tableau S2). La couleur bleue indique une plus faible et la couleur rouge indique une expression protéique plus élevée, avec des intensités signifiant l'ampleur du changement. Le regroupement du dendrogramme sur l'axe des x montre la similitude des participants (PD = marine, NC = or), tandis que le regroupement sur l'axe des y regroupe les protéines en fonction de la similitude dans leurs profils d'expression. Les flèches rouges indiquent les 10 principaux biomarqueurs protéiques de la cohorte Discovery. (B) Graphique des protéines différemment exprimées entre le groupe PD et NC dans la cohorte Discovery. L'axe des abscisses correspond à la signification ($-\log_{10} p$ -valeur) de la différence des niveaux de protéines entre le groupe PD et NC, et l'axe des y affiche la différence moyenne de groupe entre PD et NC ($PD_{moyenne} \log_{10} RFU - NC_{moyenne} \log_{10} RFU$). Les 10 premières protéines classées par sélection de stabilité sont marquées en bleu. La ligne pointillée verticale représente la valeur p de 0,005. (C) Top 10 des protéines classées par sélection de stabilité dans la cohorte de découverte. BSP, sialoprotéine osseuse; PD, maladie de Parkinson; NC, contrôle neurologiquement normal; OMD, ostéomoduline; RFU, unité de fluorescence relative. <https://doi.org/10.1371/journal.pmed.1002931.g002>

Réplication des associations de biomarqueurs avec la MP dans la cohorte PDBP

Nous avons ensuite testé nos 10 premiers marqueurs de stabilité classés pour la robustesse dans une cohorte de réplication distincte de 215 PD et 102 participants NC tirés de la cohorte multicentrique PDBP [13] (tableau 1 et tableau S3). Les méthodes analytiques étaient identiques à celles utilisées dans la cohorte de découverte, sauf que l'analyse de la cohorte de réplication comprenait en outre le site clinique et le lot comme covariables et n'incluait pas la LEDD comme covariable, car les LEDD n'étaient pas universellement disponibles pour les participants au PDBP.

Malgré l'introduction inévitable de la variabilité d'une cohorte de réplication multisite et multibatch, avec de légères différences dans la disponibilité des données cliniques, quatre des 10 principales protéines qui différaient dans les échantillons PD par rapport aux échantillons NC dans la cohorte Discovery différaient également entre PD et NC, avec la même direction d'effet, dans la cohorte de réplication PDBP (Fig 3). Ces biomarqueurs protéiques étaient la sialoprotéine osseuse (BSP, Discovery FDR-corrected $p = 2,82 \times 10^{-2}$, Replication FDR-corrected $p = 1,03 \times 10^{-4}$), l'ostéomoduline (OMD, Discovery FDR-corrected $p = 2,14 \times 10^{-2}$, Réplication corrigée FDR $p = 9,14 \times 10^{-5}$), aminoacylase-1 (ACY1, Discovery FDR corrigé $p = 1,86 \times 10^{-3}$, réplication FDR corrigé $p = 2,18 \times 10^{-2}$) et récepteur de l'hormone de croissance (GHR, Discovery FDR corrigé $p = 3,49 \times 10^{-4}$, réplication corrigée par FDR $p = 2,97 \times 10^{-3}$, tableau S4).

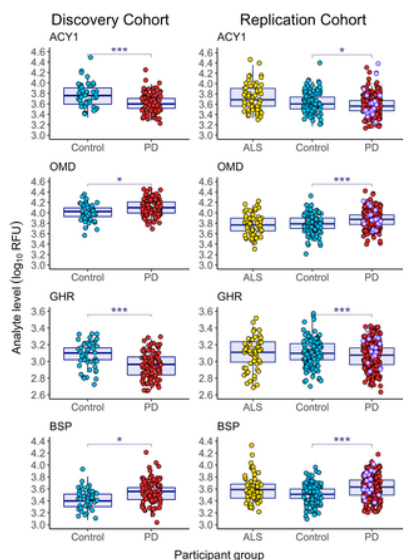


Fig 3. Biomarqueurs sanguins trouvés dans les cohortes de découverte et de réplication.

Boxplots (médiane et IQR) montrant la distribution des quatre principaux niveaux de protéines biomarqueurs (\log_{10} de RFU) trouvés dans les cohortes de découverte et de réplication en fonction du diagnostic. L'axe des y représente les niveaux de protéines plasmatiques (\log_{10} de RFU) et l'axe des x affiche le groupe de diagnostic, chaque point représentant un participant à la recherche. Dans la cohorte de réplication, 18 participants à la DP n'avaient jamais été traités avec des médicaments dopaminergiques (points roses). Les mesures des biomarqueurs pour ces participants PD jamais traités ne différaient pas des mesures des biomarqueurs pour les participants PD traités avec des médicaments dopaminergiques (test de Wilcoxon $p > 0,05$ nominal). FDR ajusté (méthode de Benjamini-Hochberg) * $p < 0,05$, *** $p < 0,005$. ACY1, aminoacylase-

1; SLA, sclérose latérale amyotrophique; BSP, sialoprotéine osseuse; FDR, taux de fausses découvertes; GHR, récepteur de l'hormone de croissance; OMD, ostéomoduline; PD, maladie de Parkinson; NC, contrôle neurologiquement normal; RFU, unité de fluorescence relative.
<https://doi.org/10.1371/journal.pmed.1002931.g003>

Mesures de biomarqueurs dans la SLA, une maladie neurodégénérative avec des caractéristiques motrices et cognitives

Pour déterminer si chacune de ces protéines plasmatiques caractérise spécifiquement la MP ou si elles sont observées dans de nombreux états de maladies neurodégénératives, nous avons également mesuré ces protéines chez 59 personnes atteintes de SLA ([tableau 1](#)), une maladie neurodégénérative qui, comme la MP, a à la fois des maladies motrices et cognitives. Caractéristiques. Comme le montre la [figure 3](#) et corroboré par une régression linéaire multiple ajustant pour l'âge, le sexe et le site clinique, à l'exception de la BSP, des changements protéiques ont été observés dans la MP mais pas dans la SLA.

Variabilité préanalytique et mesures des biomarqueurs

La plupart des personnes atteintes de MP sont traitées avec des médicaments dopaminergiques, ce qui soulève la crainte que les effets des médicaments sur le protéome plasmatique puissent être à l'origine de nos signaux de biomarqueurs. Nous avons abordé cette préoccupation de deux manières. Premièrement, dans notre cohorte de réplication PDBP, un sous-ensemble d'individus atteints de MP ($n = 18$) n'avait jamais été traité avec des médicaments dopaminergiques. Nous avons comparé les valeurs d'ACY1, OMD, GHR et BSP chez les personnes atteintes de MP traitées avec des médicaments dopaminergiques par rapport à celles jamais traitées avec des médicaments dopaminergiques, et nous n'avons trouvé aucune différence significative (test de Wilcoxon valeur nominale $p > 0,05$ pour les quatre protéines comparant la PD traitée versus non traitée avec des médicaments dopaminergiques, [Fig 3](#)). Deuxièmement, nous avons étudié des échantillons d'une cohorte multisite supplémentaire - l'étude BioFIND [[14](#)] - dans laquelle les participants à la DP ont été prélevés du sang dans deux contextes: (1) pendant qu'ils prenaient leurs médicaments dopaminergiques habituels et (2) après une nuit sans médicament ([tableau S5](#)). Bien que certaines protéines plasmatiques puissent être affectées par l'état du médicament, aucune de nos quatre protéines biomarqueurs n'a changé de manière substantielle lors de la comparaison des états médicamenteux ON et OFF chez le même individu ([Fig 4A](#) , [Fig S2](#)).

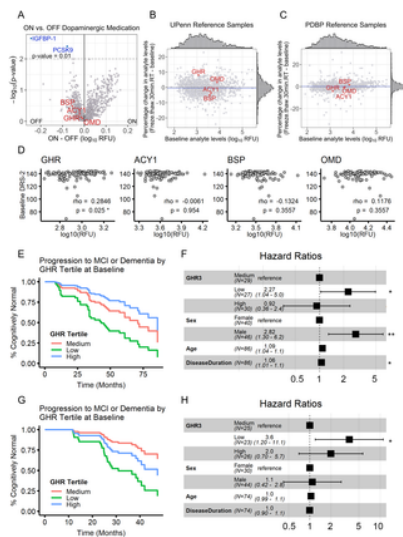


Fig 4. Les principaux biomarqueurs sont robustes et prédisent la trajectoire cognitive.

(A) Graphique du volcan montrant l'effet de la thérapie dopaminergique sur les niveaux de protéines plasmatiques, testé chez 10 personnes PD, comparant l'état de médicament ON à OFF chez les mêmes individus. Des différences nominale-ment significatives dans l'état ON par rapport à l'état OFF ont été trouvées pour seulement deux protéines (apparié test t nominal $p < 0,01$), et aucune de nos quatre principales protéines biomarqueurs n'a été sensiblement affectée. (B,C) Effet de la manipulation du plasma (échantillons comparés à une incubation RT de 30 minutes suivie d'un cycle supplémentaire de congélation-décongélation) dans deux pools de référence. La manipulation du plasma a provoqué des changements substantiels dans les niveaux de certaines protéines, mais aucun de nos quatre principaux biomarqueurs n'a été affecté. (D) Relation transversale entre les mesures des biomarqueurs PD (\log_{10} de RFU) et la fonction cognitive de base telle que mesurée par le score DRS (corrélation de Spearman, * FDR-ajusté $p < 0,05$). (E,H) Analyses des risques proportionnels de Cox examinant les différences dans les taux ultérieurs de conversion clinique en MCI ou en démence, stratifiées par les mesures de GHR au départ dans la cohorte Discovery (tertiles illustrés, panneaux E, F) et la cohorte de réplication (tertiles illustrés, panneaux G, H). Pour la cohorte de réplication, les affectations cognitives étaient basées sur les normes publiées pour le MoCA, évaluées longitudinalement sur le site UTSW. (E, G) Courbes de régression de Cox montrant les trajectoires ajustées pour chaque tertile des mesures de base de la GHR (dans chaque cohorte). (F, H) Parcelles forestières illustrant les rapports de risque pour les groupes tels que définis par le tertile des mesures de biomarqueurs au départ, le sexe, l'âge et la durée de la maladie. ** $p < 0,01$. ACY1, aminoacylase-1; BSP, sialoprotéine osseuse; DRS, Mattis Dementia Rating Scale-2; FDR, taux de fausses découvertes; GHR, récepteur de l'hormone de croissance; IGFBP-1, protéine 1 de liaison au facteur de croissance analogue à l'insuline; MCI, trouble cognitif léger; MoCA, Évaluation cognitive de Montréal; OMD, ostéomoduline; PCSK9, proprotéine convertase subtilisine / kexine de type 9; PD, maladie de Parkinson; PDBP, programme de biomarqueurs de la maladie de Parkinson; RFU, unité de fluorescence relative; RT, température ambiante; UPenn, Université de Pennsylvanie; UTSW, University of Texas Southwestern Medical Center.
<https://doi.org/10.1371/journal.pmed.1002931.g004>

Pour comprendre si les biomarqueurs de protéines candidates sont robustes aux sources communes de variabilité préanalytique, nous avons étudié des aliquotes identiques d'échantillons de plasma regroupés de la cohorte Discovery (pool d'échantillons de référence UPenn) et de la cohorte PDBP (pool d'échantillons de référence PDBP) qui ont été soumis à des différences dans l'échantillon. manipulation (avec ou sans 30 minutes à température ambiante suivie d'un gel-dégel supplémentaire de l'échantillon). Alors que certaines protéines ont changé leurs niveaux de > 30% en fonction des différences dans la manipulation des échantillons, aucune de nos principales protéines n'a changé substantiellement dans l'un ou l'autre pool ([figures 4B et 4C](#) , [tableau S5](#)).

GHR, ACY1 et OMD comme prédicteurs du déclin cognitif

Nous avons ensuite demandé si les niveaux de base de nos biomarqueurs candidats prédisaient la progression de la maladie. Parce que les symptômes cognitifs sont moins affectés par les médicaments dopaminergiques que les symptômes moteurs, et parce que le déclin de la cognition est variable mais cliniquement important dans la MP [[30](#)], nous avons examiné si les niveaux de base de BSP, OMD, ACY1 ou GHR prédisaient les taux ultérieurs de déclin cognitif.

Dans notre cohorte Discovery largement caractérisée, les participants qui ont été suivis en moyenne 3,5 ans après l'échantillonnage du plasma, des analyses transversales ont révélé une association minimale entre les niveaux de biomarqueurs plasmatiques et les scores cognitifs de base sur le DRS multidomaine, qui a été largement utilisé pour l'évaluation cognitive dans PD [[31](#)] ([Fig 4D](#)). Cependant, les taux plasmatiques de GHR, ACY1 et OMD prédisaient les taux ultérieurs de changement cognitif sur le DRS dans des modèles linéaires à effets mixtes ajustés pour l'âge, le sexe, la durée de la maladie et le score DRS de base, avec des coefficients d'interaction temps-protéine de 0,0905 (GHR, p corrigé FDR = $8,72 \times 10^{-6}$), 0,0478 (ACY1, p corrigé FDR = $2,574 \times 10^{-2}$) et -0,0457 (OMD, p corrigé FDR = $2,574 \times 10^{-2}$), respectivement. De plus, les personnes ayant des niveaux de base de GHR et d'ACY1 dans le tertile le plus bas étaient significativement plus susceptibles de progresser cliniquement vers l'ICM ou la démence dans les analyses des risques proportionnels de Cox ajustés en fonction de l'âge, du sexe et de la durée de la maladie (rapport de risque [HR] 2,27 [IC à 95% 1,04–5,0, $p = 0,04$] pour GHR [[Fig 4E et 4F](#)], et HR 3,0 [IC 95% 1,24–7,0, $p = 0,014$] pour ACY1 [[Fig S3](#)]). Enfin, la correction de l'éducation dans nos modèles n'a pas affecté nos résultats ([Tableau S6](#) , [Fig S3](#)).

La cohorte de réplication PDBP est moins mature dans le suivi que notre cohorte de découverte. En outre, les données des tests cognitifs sont plus limitées, avec une variabilité entre les sites PDBP en ce qui concerne leur collecte de données cognitives et le stade de la MP. Nonobstant ces limites, les scores au MoCA ont été obtenus à intervalles de 6 mois pour une période de suivi moyenne de 2,8 ans pour 74 personnes PD de la cohorte de réplication PDBP, suivies à l'UTSW. Chez ces participants, nous avons classé chaque individu comme ayant une cognition normale, un MCI ou une démence pour chaque point dans le temps selon les normes publiées pour le MoCA [[26](#)]. En utilisant les mêmes modèles de risques proportionnels de Cox (c.-à-d. Ajustés en fonction de l'âge, du sexe et de la durée de la maladie) que dans la cohorte Discovery, nous avons constaté que les personnes dont les niveaux de référence de GHR dans le tertile le plus bas étaient plus susceptibles de progresser vers les scores MoCA dans le MCI. ou de la démence (HR 3,6 [IC à 95% 1,20–11,1, $p = 0,02$]) dans la cohorte de réplication également ([Fig 4G et 4H](#)). De plus, tout comme dans la cohorte Discovery, une correction supplémentaire pour l'éducation dans notre modèle n'a pas affecté les résultats ([Fig. S4](#)).

Discussion

Dans cette étude, nous avons étudié plusieurs cohortes dans un plan de découverte-réplication pour développer de nouveaux biomarqueurs PD, à partir d'un écran non biaisé d'environ 1 000 protéines plasmatiques. Nous avons trouvé quatre meilleurs biomarqueurs candidats - ACY1, BSP, GHR et OMD - qui se répliquaient dans une cohorte de découverte à un seul site et dans une cohorte de réplication multisite, étaient robustes aux sources communes de variabilité préanalytique et ne diffèrent pas dans les échantillons appariés de participants PD. médicament dopaminergique activé ou désactivé. Dans des analyses de données longitudinales, nous avons montré que les niveaux de base d'ACY1 et de GHR - et, dans une moindre mesure, d'OMD - associés aux taux ultérieurs de déclin cognitif dans notre cohorte de découverte, le GHR de base prédisant également l'évolution cognitive ultérieure dans la cohorte de réplication .

Les biomarqueurs PD trouvés ici n'ont pas, à notre connaissance, été précédemment rapportés dans la littérature sur les maladies neurodégénératives. Cependant, les criblages non biaisés - le plus souvent illustrés par l'étude d'association à l'échelle du génome en génétique humaine - donnent souvent de nouvelles directions inattendues pour l'investigation [[6](#)]. Nous notons, cependant, que le GHR et le facteur de croissance analogue à l'insuline (IGF-1, un effecteur bien connu produit en réponse à la signalisation de l'hormone de croissance [GH]-GHR), sont exprimés dans le cerveau [[32](#) , [33](#)] et ont été impliqués dans des événements physiologiques et pathologiques dans le cerveau. La signalisation GH-GHR-IGF-1 a été impliquée dans la différenciation et la prolifération des cellules souches neurales au cours du développement embryonnaire [[33](#) - [35](#)], neurogenèse adulte chez les rongeurs [[36](#) , [37](#)], déclin cognitif lié à l'âge [[38](#) , [39](#)] et neuroprotection contre les atteintes neurologiques telles que les lésions hypoxiques-ischémiques [[40](#) , [41](#)], indiquant des liens potentiels entre cette voie et la protection contre neurodégénérescence. Des études futures utilisant des techniques de randomisation mendélienne [[42](#)] ou la manipulation des niveaux de biomarqueurs dans des systèmes modèles sont cependant nécessaires pour vraiment élucider les mécanismes potentiels menant aux signatures de biomarqueurs décrites ici.

Les points forts de cette étude comprennent l'attention portée à la reproductibilité, ainsi que la prise en compte des facteurs du monde réel qui influencent le potentiel de traduction en aval. En ce qui concerne la reproductibilité, nous soulignons quatre aspects. Premièrement, le classement des meilleures protéines candidates de la cohorte Discovery par sélection de stabilité, plutôt qu'un ordre strict par valeur p , évite les inquiétudes concernant le surajustement. Deuxièmement, les biomarqueurs décrits ici avaient corrigé FDR p -valeurs <0,05 dans les cohortes de réplication PDBP à site unique et multisite, attestant de la robustesse de nos résultats. Troisièmement, la stratégie d'analyse dans la cohorte de réplication a été prédéfinie pour refléter celle de la cohorte de découverte, avec seulement deux différences: (1) l'inclusion du site et du lot comme covariables supplémentaires, justifiée par le passage du site unique / lot unique à les phases d'analyse multisite / multibatch et (2) la suppression de la LEDD en tant que covariable, rendue nécessaire par le manque de ces données de manière uniforme chez tous les participants de la cohorte de réplication. Quatrièmement, nous notons que les niveaux de base de GHR, ACY1 et, dans une moindre mesure, OMD prédisaient le futur déclin cognitif chez les personnes atteintes de MP de la cohorte Discovery. De plus, malgré les différences d'échelle cognitive et de site clinique utilisé, ainsi que le stade de DP évalué (tous les facteurs connus pour affecter les mesures de la cognition au fil du temps), des niveaux inférieurs de GHR ont également prédit un déclin cognitif plus rapide dans la cohorte de réplication. En plus de répondre à un besoin clair de biomarqueurs prédisant la progression de la MP [[7](#)], l'association des mêmes protéines avec la classe de la maladie ainsi que la progression de la maladie augmente la confiance dans ces biomarqueurs candidats, car la gradation des niveaux au sein de la MP selon une mesure de la gravité physiopathologique (taux de déclin cognitif) suggère que les différences entre la MP et NC ne sont pas dus à une variable confondante cachée différenciant ces deux groupes. En ce qui concerne les considérations de translation en aval, nous avons étudié les aspects de la variabilité du monde

réel, démontrant que les quatre principaux biomarqueurs rapportés ici ne sont pas substantiellement affectés par l'état des médicaments dopaminergiques ou les sources communes de bruit liées à la manipulation des échantillons. Nous soulignons également le fait que nos candidats biomarqueurs sont mesurés à partir du plasma sanguin, ce qui permet une collecte dans n'importe quel contexte de phlébotomie de routine.

Notre étude a également des limites. Tout d'abord, notre étude s'appuie sur une plate-forme basée sur les aptamères [9] pour les mesures des protéines plasmatiques. Bien qu'il s'agisse d'une approche puissante pour le criblage à grande échelle, la traduction en aval nécessitera probablement le développement de tests de protéines alternatifs qui (1) donnent des quantités de protéines absolues plutôt que les RFU analysées ici et (2) confirment la spécificité du test. Deuxièmement, bien que nous ayons ajusté les effets des médicaments dopaminergiques lorsque cela était possible et analysé directement l'effet de l'état des médicaments dopaminergiques sur les mesures protéiques, notre étude ne peut pas exclure les petits effets des médicaments dopaminergiques sur les mesures des protéines candidates, car le lavage nocturne des médicaments dopaminergiques n'atténue pas complètement les effets des médicaments. Ainsi, l'évaluation des biomarqueurs de protéines candidates dans des cohortes à haut risque symptomatiques précoces ou même présymptomatiques non médicamentées est une voie d'avenir fructueuse. Troisième, bien que nous ayons démontré que les taux plasmatiques d'ACY1, de GHR et d'OMD sont dans une certaine mesure spécifiques de la MP, en ce qu'ils ne sont pas modifiés de la même manière dans la SLA, il est toujours possible que certains de ces biomarqueurs protéiques présentent des changements similaires dans d'autres maladies neurodégénératives qui n'ont pas été testés. Nous notons, cependant, que l'appréciation croissante du chevauchement de la pathologie à travers diverses maladies neurodégénératives - les personnes atteintes de MP, par exemple, sont très susceptibles d'avoir une neuropathologie concomitante de la maladie d'Alzheimer (MA) à l'autopsie [43] - les suggestions qui se chevauchent dans les biomarqueurs dans les catégories cliniques actuelles peuvent refléter un chevauchement des mécanismes physiopathologiques, plutôt qu'un mauvais biomarqueur. Quatrièmement, pour nos analyses longitudinales, nous avons évalué les changements cognitifs afin de comprendre si les biomarqueurs candidats prédisaient la progression de la maladie. Nous avons choisi d'étudier le déclin cognitif à la fois en raison de la morbidité majeure associée à cet aspect de la progression de la maladie et parce que la cognition n'est pas aussi affectée par les médicaments dopaminergiques que les performances motrices. Étant donné que la majorité des participants à la DP étudiés ici ont été évalués tout en prenant des médicaments dopaminergiques, les performances motrices devraient refléter non seulement l'état de la maladie sous-jacente au fil du temps (ce que nous visons à mesurer), mais également la réponse médicamenteuse et le moment de la dose la plus récente de médicament, ajoutant bruit considérable. Donc,

En résumé, nous présentons nos résultats du dépistage non biaisé de plus de 1000 protéines plasmatiques dans plusieurs cohortes de DP (une cohorte de découverte à un seul site, la cohorte de réplication PDBP multicentrique et la cohorte multicentrique BioFIND), ainsi que la maladie et les témoins normaux. En particulier, nous avons identifié quatre protéines plasmatiques - BSP, OMD, ACY1 et GHR - avec des altérations cohérentes de la MP, dont l'une (GHR) a également prédit un déclin cognitif ultérieur dans plusieurs cohortes, à travers plusieurs instruments de tests cognitifs. Nos résultats ouvrent de nouvelles voies pour l'investigation mécaniste, suggérant que le profilage «quasi-protéomique» du sang d'individus atteints de MP peut être une approche puissante à la fois pour le développement d'outils cliniques et pour un aperçu de la physiopathologie de cette maladie actuellement incurable.

Renseignements à l'appui

S1 Fig. QC mesures.

(A) Le diagramme de Venn de la cohorte de découverte montre le nombre de protéines qui ont passé les filtres QC (968 protéines) et le nombre de protéines qui ont échoué à l'un ou aux deux critères QC. Sur un total de 1 129 protéines, 161 ont été exclues en raison de CV élevés (> 20%) ou de proportions élevées (> 25%) de mesures en dehors des limites de détection du test. (B) Le diagramme de Venn de la cohorte de réplication montre que 33 protéines ont été éliminées en raison d'un CV supérieur à 20%, laissant un total de 1 272 protéines pour les analyses en aval. CV, coefficient de variation; QC, contrôle de qualité.

<https://doi.org/10.1371/journal.pmed.1002931.s001>

(TIF)

S2 Fig. Effet du traitement par la lévodopa sur les taux de protéines plasmatiques testés par un test de permutation appariée.

Graphique du volcan montrant l'effet de la thérapie dopaminergique sur les niveaux de protéines plasmatiques, testé chez 10 personnes PD, comparant l'état des médicaments ON à OFF chez les mêmes individus. Des différences nominalement significatives dans l'état ON par rapport à l'état OFF ont été trouvées pour une seule protéine (test de permutation appariée nominal $p < 0,01$). PD, maladie de Parkinson.

<https://doi.org/10.1371/journal.pmed.1002931.s002>

(TIF)

S3 Fig. Les niveaux plasmatiques de GHR et ACY1 prédisent le déclin cognitif chez les personnes atteintes de MP de la cohorte Discovery.

Les différences dans les taux ultérieurs de conversion clinique en MCI ou en démence dans la cohorte Discovery stratifiée par les niveaux de GHR ou d'ACY1 au départ (indiqués en tertiles) ne sont pas affectées par l'éducation. (A, C, E) Courbes de régression de Cox montrant les trajectoires ajustées pour chaque tertile des mesures de biomarqueurs de référence et les parcelles forestières (B, D, F) illustrant les rapports de risque pour les groupes tels que définis par le tertile des mesures de biomarqueurs au départ et les covariables. (A, B) Résultats pour ACY1 sans ajustement pour l'éducation. (C, F) Résultats des analyses de risques proportionnels de Cox ajustés pour l'éducation pour GHR (C, D) et ACY1 (E, F), respectivement. (G) Résultats de l'ANOVA (statistique χ^2 , p -value) comparant le modèle de risques proportionnels de Cox avec l'éducation (modèle 2) et sans éducation (modèle 1). ACY1, aminoacylase-1; GHR, récepteur de l'hormone de croissance; MCI, trouble cognitif léger; PD, maladie de Parkinson.

<https://doi.org/10.1371/journal.pmed.1002931.s003>

(TIF)

S4 Fig. Les niveaux plasmatiques de GHR prédisent le déclin cognitif chez les personnes atteintes de MP de la cohorte de réplication.

Les différences dans les taux ultérieurs de conversion clinique en MCI ou en démence dans la cohorte de réplication (UTSW) stratifiée par les niveaux de GHR au départ (indiqués en tertiles) ne sont pas affectées par l'éducation. (A) Courbe de régression de Cox montrant les trajectoires ajustées pour chaque tertile des mesures de référence de la GHR et (B) des parcelles forestières illustrant les rapports de risque pour les groupes tels que définis par le tertile de GHR, le sexe, l'âge, la durée de la maladie et la covariable supplémentaire de l'éducation. (C) Résultats de l'ANOVA (statistique χ^2 et valeur p) comparant le modèle de risques proportionnels de Cox avec l'éducation (modèle 2) et sans (modèle 1). GHR, récepteur de l'hormone de croissance; MCI, trouble cognitif léger; PD, maladie de Parkinson; UTSW, University of Texas Southwestern Medical Center.

<https://doi.org/10.1371/journal.pmed.1002931.s004>

(TIF)

Tableau S1. Niveaux SOMAScan (log₁₀ de RFU) de 1129 protéines et données démographiques pour 141 participants (96 PD et 45 NC) de la cohorte Discovery (Udall).

Les protéines présentant un coefficient de variation supérieur à 20% et plus de 25% des mesures en dehors du LLOD ou de l'ULOD ont été exclues des analyses en aval. Sur 1 129 protéines, 161 ont été exclues. LEDD, dose quotidienne équivalente de lévodopa; LLOD, limite inférieure de détection; NC, contrôle neurologiquement normal; PD, maladie de Parkinson; RFU, unité de fluorescence relative; ULOD, limite supérieure de détection.

<https://doi.org/10.1371/journal.pmed.1002931.s005>
(XLSX)

Tableau S2. Biomarqueurs PD candidats issus de l'analyse de la cohorte Discovery (140 protéines).

Les protéines qui différencient les individus PD et NC ($p < 0,005$) ont été identifiées en appliquant un modèle de régression linéaire multivariée aux données de la cohorte Discovery (96 PD et 45 NC). Les noms de protéines, les symboles ENTREZ, la valeur p du modèle de régression (différenciant le statut PD par rapport au statut NC, ajusté pour l'âge au moment de l'échantillonnage du plasma, le sexe et la LEDD) et la direction du changement de PD par rapport à NC sont indiqués. p -Valeur du modèle de régression différenciant le groupe PD et NC ajusté pour l'âge à la collecte de plasma, le sexe et la LEDD. FDR = Valeur p ajustée de Benjamini-Hochberg, 968 tests. FDR, taux de fausses découvertes; LEDD, dose quotidienne équivalente de lévodopa; NC, contrôle neurologiquement normal; PD, maladie de Parkinson.

<https://doi.org/10.1371/journal.pmed.1002931.s006>
(XLSX)

Tableau S3. Niveaux SOMAScan (log₁₀ de RFU) de 1305 protéines et données démographiques pour 376 participants à la recherche (59 ALS, 215 PD et 102 NC) de la cohorte de réplication.

Trois individus aberrants et 33 protéines présentant un coefficient de variation supérieur à 20% ont été exclus des analyses. SLA, sclérose latérale amyotrophique; NC, contrôle neurologiquement normal; PD, maladie de Parkinson; RFU, unité de fluorescence relative.

<https://doi.org/10.1371/journal.pmed.1002931.s007>
(XLSX)

Tableau S4. Régression multiple valeurs p ajustées par FDR pour les 10 principales protéines de la cohorte de découverte et de réplication.

Les 10 meilleurs candidats biomarqueurs plasmatiques (classés par sélection de stabilité) qui ont différencié les participants PD par rapport aux participants NC dans la cohorte Discovery; quatre associations de répliques dans la cohorte de réplication. Ici, nous rapportons les valeurs p du modèle de régression linéaire multiple différenciant le statut PD par rapport au statut NC, ajustées pour l'âge au moment de l'échantillonnage du plasma, le sexe et la LEDD dans la cohorte Discovery. Dans la cohorte de réplication, l'effet du groupe de traitement (PD versus NC) a été ajusté en fonction de l'âge au moment du prélèvement de plasma, du sexe, du site clinique et de l'effet du lot. L'ajustement pour le FDR a été effectué en utilisant la méthode Benjamini-Hochberg. FDR, taux de fausses découvertes; LEDD, dose quotidienne équivalente de lévodopa; NC, contrôle neurologiquement normal; PD, maladie de Parkinson.

<https://doi.org/10.1371/journal.pmed.1002931.s008>
(DOCX)

Tableau S5. Changement des niveaux de protéines à l'état dopaminergique ON par rapport à OFF, et après perturbation systématique d'échantillons de gel-dégel supplémentaire et d'exposition prolongée (30 minutes) à la température ambiante.

Pour l'analyse ON versus OFF, des échantillons de plasma ont été prélevés pendant l'état ON – OFF pour chacun des 10 participants à la DP (5 femmes et 5 hommes de plus de 50 ans) sélectionnés au hasard dans la base de données BioFIND. Le pourcentage de changement des RFU brutes et les différences de log₁₀ (RFU) entre l'état ON et OFF sont indiqués pour 1 272 protéines. Seules deux protéines (IGFBP-1 et PCSK9) ont différencié significativement l'état dopaminergique ON versus OFF. ** Test t apparié valeur p nominale $< 0,01$, bilatéral. IGFBP-1, protéine 1 de liaison au facteur de croissance analogue à l'insuline; PCSK9, proprotéine convertase subtilisine / kexine de type 9; PD, maladie de Parkinson; RFU, unité de fluorescence relative.

<https://doi.org/10.1371/journal.pmed.1002931.s009>
(XLSX)

Tableau S6. Résultats des modèles linéaires à effets mixtes de la cohorte Discovery.

Pour notre premier modèle (modèle 1), la variable de réponse DRS a été modélisée comme un effet fixe de l'interaction temps-protéine, de l'âge, du sexe, de la durée de la maladie et du DRS initial et comme un effet aléatoire du participant. Le modèle 2 s'adapte à l'éducation en l'incluant comme effet fixe supplémentaire. Les coefficients d'interaction temps-protéine ($\beta_{\text{temps} \cdot \text{prot}}$) et leurs valeurs p corrigées FDR correspondantes pour le modèle 1 et le modèle 2. DRS, Mattis Dementia Rating Scale-2; FDR, taux de fausses découvertes.

<https://doi.org/10.1371/journal.pmed.1002931.s010>
(DOCX)

Liste de contrôle S1 STROBE. STROBE, Renforcement de la notification des études d'observations en épidémiologie.

<https://doi.org/10.1371/journal.pmed.1002931.s011>
(DOCX)

Plan d'analyse S1.

<https://doi.org/10.1371/journal.pmed.1002931.s012>
(DOCX)

Remerciements

Nous remercions Mike Mehan et David Sterling pour leurs conversations utiles concernant les analyses statistiques. Nous remercions Tatiana Foroud pour son assistance dans la manipulation des biosamples. Les données brutes sur les protéines, ainsi que les informations cliniques associées, peuvent être téléchargées à partir de la ressource de gestion des données PDBP.

Les données et les échantillons biologiques utilisés dans la préparation de ce manuscrit ont été obtenus auprès du consortium PDBP, qui fait partie du National Institute of Neurological Disorders and Stroke des National Institutes of Health. Les enquêteurs comprennent Roger Albin, Roy Alcalay, Alberto Ascherio, Brad Boeve, DuBois Bowman, Alice Chen-Plotkin, Ted Dawson, Richard Dewey, Kirk Frey, Dwight German, Lawrence Honig, Xuemei Huang, Kejal Kantarci, Jim Leverenz, Karen Marder, Rachel Saunders-Pullman, Liana Rosenthal, Clemens Scherzer, Michael Schwarzschild, David Vaillancourt, David Walt, Andrew West et