Mitochondrial haplogroups and cognitive progression in Parkinson's disease

Ganqiang Liu, Chunming Ni, Jiamin Zhan, Weimin Li, Junfeng Luo, Zhixiang Liao, Joseph J. Locascio, Wenbiao Xian, Ling Chen, Zhong Pei, et al.

OXFORD UNIVERSITY PRESS | Brain

Mitochondrial haplogroups and cognitive progression in Parkinson's disease

Journal:	*Brain*
Manuscript ID	BRAIN-2021-02504.R2
Manuscript Type:	Report
Date Submitted by the Author:	n/a
Complete List of Authors:	Liu, Ganqiang; Sun Yat-Sen University, School of Medicine Ni, Chunming; Sun Yat-Sen University, School of medicine Zhan, Jiamin; Sun Yat-Sen University, School of medicine Li, Weimin; Sun Yat-Sen University, School of medicine Luo, Junfeng; Sun Yat-Sen University, School of medicine Liao, Zhixiang; Brigham and Women's Hospital Locascio, Joseph; Massachusetts General Hospital, Neurology; Harvard Medical School, Biostatistics Xian, Wenbiao; Sun Yat-sen University First Affiliated Hospital, Neurology Chen, Ling; Sun Yat-sen University First Affiliated Hospital, Department of Neurology Pei, Zhong; Sun Yat-sen University First Affiliated Hospital, Department of Neurology and Stroke Center Corvol, Jean-Christophe; Sorbonne Université, Centre d'Investigation Clinique Maple-Grødem , Jodi; Stavanger University College, Stavanger University Hospital Campbell, Meghan; Washington University School of Medicine in Saint Louis, Neurology Elbaz, Alexis; Paris-Saclay University, Neuroepidemiology; Institut national de la santé et de la recherche médicale Lesage, Suzanne; Institut national de la santé et de la recherche médicale; Sorbonne Université, Brice, Alexis; Sorbonne Université; Pitie-Salpetriere Hospital, Neurologie et de Génétique Hung, Albert; Massachusetts General Hospital, Department of Neurology Schwarzschild, Michael; Massachusetts General Hospital, Neurology Hayes, Michael; Brigham and Women's Hospital Wills, Anne-Marie; Massachusetts General Hospital, Neurology Clinical Trials Unit Ravina, Bernard; Praxis Institute, Praxis Precision Medicines Shoulson, Ira; University of Rochester Taba, Pille; University of Tartu, Department of Neurology and Neurosurgery Kõks, Sulev; Perron Institute for Neurological and Translational Science Beach, Thomas; Banner Sun Health Research Institute Cormier, Florence; Sorbonne Université, Institut du Cerveau – Paris Brain Institute - ICM Alves, Guido; Stavanger University College, Norwegian Centre for

	Movement Disorders Tysnes, Ole-Bjørn; Haukeland University Hospital, Department of Neurology Perlmutter, Joel; Washington University in Saint Louis School of Medicine, Neurology, Radiology, Neuroscience, Physical Therapy, and Occupational Therapy Heutink, Peter; German Centre for Neurodegenerative Diseases, Department of Clinical Genetics, Neuroscience Campus Amsterdam; van Hilten, Bob; Leiden University Medical Center, Neurology Barker, Roger; University of Cambridge, Cambridge Centre for Brain Repair Williams-Gray, Caroline; University of Cambridge, Brain Repair Centre, Department of Clinical Neurosciences Scherzer, Clemens; Brigham and Women's Hospital, Neurology; Harvard Medical School
Methodology:	GENETICS, GENOMICS & EPIGENETICS
Subject area:	MOVEMENT DISORDERS

Mitochondrial haplogroups and cognitive progression in Parkinson's disease

Ganqiang Liu, PhD[1*], Chunming Ni, BE[1], Jiamin Zhan, BS[1], Weimin Li, MS[1] , Junfeng Luo, MS[1], Zhixiang Liao, MS[2,3], Joseph J. Locascio, PhD[2,3,4], Wenbiao Xian, MD[5], Ling Chen, MD[5], Zhong Pei, MD[5], Jean-Christophe Corvol, MD[6], Jodi Maple-Grødem, PhD[7,8], Meghan C. Campbell, PhD[9], Alexis Elbaz, MD[10], Suzanne Lesage, PhD[6], Alexis Brice, MD[6], Albert Y. Hung, MD, PhD[4], Michael A. Schwarzschild, MD, PhD[4], Michael T. Hayes, MD[11], Anne-Marie Wills, MD[4], Bernard Ravina, MD[12], Ira Shoulson, MD[13], Pille Taba, MD, PhD[14,15], Sulev Kõks, MD, PhD[16,17], Thomas G. Beach, MD, PhD[18], Florence Cormier-Dequaire, MD[6], Guido Alves, PhD[7,8,19], Ole-Bjørn Tysnes, MD[20,21], Joel S. Perlmutter, MD[9,22,23], Peter Heutink, PhD[24], Jacobus J. van Hilten, MD[25], Roger A. Barker, PhD[26,27], Caroline H. Williams-Gray, PhD[26] and Clemens R. Scherzer, MD[2,3,4,11*] for the International Genetics of Parkinson Disease Progression (IGPP) Consortium[&]

Abstract

Mitochondria are a culprit in the onset of Parkinson's disease, but their role during disease progression is unclear. Here we used Cox proportional hazards models to exam the effect of variation in the mitochondrial genome on longitudinal cognitive and motor progression over time in 4,064 patients with Parkinson's disease. Mitochondrial macro-haplogroup was associated with reduced risk of cognitive disease progression in the discovery and replication population. In the combined analysis, patients with the super macro-haplogroup J, T, U[#] had a 41% lower risk of cognitive progression with $P = 2.42 \times 10^{-6}$ compared to those with macro-haplogroup H. Exploratory analysis indicated that the common mitochondrial DNA variant, m.2706A>G, was associated with slower cognitive decline with a hazard ratio of 0.68 (95% CI 0.56-0.81) and $P = 2.46 \times 10^{-5}$. Mitochondrial haplogroups were not appreciably linked to motor progression. This initial genetic survival study of the mitochondrial genome suggests

that mitochondrial haplogroups may be associated with the pace of cognitive progression in Parkinson's disease over time.

Author affiliations:

1 Neurobiology Research Center, School of Medicine, Shenzhen Campus of Sun Yat-sen University, Shenzhen, Guangdong, 518107, China

2 Center for Advanced Parkinson Research, Brigham and Women's Hospital and Harvard Medical School, Boston, MA 02115, USA

3 Neurogenomics Lab , Harvard Medical School, Brigham and Women's Hospital, Boston, MA 02115, USA

4 Department of Neurology, Massachusetts General Hospital and Harvard Medical School, Boston, MA 02114, USA

5 Department of Neurology, The First Affiliated Hospital, Sun Yat-Sen University, Guangzhou, China

6 Sorbonne Université, Institut du Cerveau – Paris Brain Institute - ICM, Institut National de la Santé et de la Recherche Médicale, Centre National de la Recherche Scientifique, Assistance Publique Hôpitaux de Paris, Département de Neurologie et de Génétique, Hôpital Pitié-Salpêtrière, F-75013, Paris, France

7 The Norwegian Centre for Movement Disorders, Stavanger University Hospital, 4068, Stavanger, Norway

8 Department of Chemistry, Bioscience and Environmental Engineering, University of Stavanger, 4021, Stavanger, Norway

9 Department of Neurology, Washington University School of Medicine, St. Louis, MO 63110, USA

10 Paris-Saclay University, UVSQ, Inserm, Gustave Roussy, "Exposome and heredity" team, CESP, F94805, Villejuif, France

11 Department of Neurology, Brigham and Women's Hospital, Boston, MA 02115, USA

12 Praxis Precision Medicines, Cambridge, MA 02142, USA

13 Department of Neurology, Center for Health + Technology, University of Rochester, Rochester, NY,14642, USA

14 Department of Neurology and Neurosurgery, Institute of Clinical Medicine, University of Tartu, Tartu, 50406, Estonia

15 Neurology Clinic, Tartu University Hospital, Tartu, 50406, Estonia

16 Centre for Molecular Medicine and Innovative Therapeutics, Murdoch University, Murdoch, Perth, 6150 WA, Australia.
17 Perron Institute for Neurological and Translational Science, Nedlands, WA 6009, Australia
18 Banner Sun Health Research Institute, Sun City, AZ 85351, USA
19 Department of Neurology, Stavanger University Hospital, 4068, Stavanger, Norway
20 Department of Neurology, Haukeland University Hospital, 5020, Bergen, Norway
21 Department of Clinical Medicine, University of Bergen, 5020, Norway
22 Departments of Radiology and Neuroscience, Washington University School of Medicine, St. Louis, MO 63110, USA
23 Program of Physical Therapy and Program of Occupational Therapy, Washington University School of Medicine, St. Louis, MO 63110, USA
24 German Center for Neurodegenerative diseases (DZNE), Tübingen, 72076, Germany
25 Department of Neurology, Leiden University Medical Center, Albinusdreef 2, 2333 ZA, Leiden, The Netherlands
26 John Van Geest Centre for Brain Repair, Department of Clinical Neurosciences, University of Cambridge, Cambridge, CB2 0PY, UK
27 Wellcome - MRC Cambridge Stem Cell Institute, University of Cambridge, Cambridge, CB2 0AW, UK

[&]International Genetics of Parkinson Disease Progression (IGPP) Consortium
Ganqiang Liu, PhD, Rebecca R. Valentino, PhD, Jiajie Peng, PhD, Zhixiang Liao, MS, Joseph J. Locascio, PhD, Jean-Christophe Corvol, MD, Xianjun Dong, PhD, Jodi Maple-Grødem, PhD, Meghan C. Campbell, PhD, Alexis Elbaz, MD, Suzanne Lesage, PhD, Alexis Brice, MD, Graziella Mangone, MD, John H. Growdon, MD, Albert Y. Hung, MD, PhD, Michael A. Schwarzchild, MD, PhD, Michael T. Hayes, MD, Anne-Marie Wills, MD, Todd M. Herrington, MD, Bernard Ravian, MD, Ira Shoulson, MD, Pille Taba, MD, PhD, Sulev Kõks, MD, PhD, Thomas G. Beach, MD, PhD, Florence Cormier-Dequaire, MD, Guido Alves, PhD, Ole-Bjørn Tysnes, MD, Joel S. Perlmutter, MD, Peter Heutink, PhD, Jacobus J. van Hilten, MD, Meike Kasten, MD, Brit Mollenhauer, MD, Claudia Trenkwalder, MD, Christine Klein, MD, Roger A. Barker, PhD, Caroline H. Williams-Gray, PhD, Johan Marinus, PhD and Clemens R. Scherzer, MD

*Correspondence to:

Ganqiang Liu, Ph.D.
School of Medicine, Shenzhen Campus of Sun Yat-sen University
88 Gongchang Road
Shenzhen, Guangdong, 518107, China
Phone: 86-020-83226059
Fax: 86-020-83271560
Email: liugq3@mail.sysu.edu.cn

Clemens R. Scherzer, M.D.
Center for Advanced Parkinson Research and
Precision Neurology Program
Harvard Medical School and Brigham & Women's Hospital
Hale Building for Transformative Medicine, 9002V
60 Fenwood Road
Boston, MA 02465
Phone: 857-307-5422
Fax: 857-307-5476
Email: cscherzer@rics.bwh.harvard.edu

Running title: mtDNA variants link to cognitive decline in PD

Keywords: Parkinson's disease, mitochondrial haplogroups, cognitive progression

Abbreviations: GCI = global cognitive impairment; HY = Hoehn and Yahr stage; MDS-UPDRS = Movement Disorder Society-Unified Parkinson's disease Rating Scale; MMSE = Serial Mini Mental State Exam; PD = Parkinson's disease

Introduction

Disability and quality of life of patients with Parkinson's disease (PD) is affected by progressive cognitive impairment[1]. Increasing numbers of cognitively impaired patients with PD pose a medical and socio-economic challenge in many countries[2]. The pace of cognitive changes during the disease course, however, varies substantially from patient to patient[3] and the genetic architecture accounting for this heterogeneity in disease progression has not been well established.

Genome-wide association studies (GWAS) during the past decade have delineated the genetic architecture of disease susceptibility with 90 association signals in 78 common autosomal loci in PD patients of European ancestry[4]. Our recent genome-wide survival study identified associations with longitudinal progression from PD to Lewy body dementia in five loci, *RIMS2, GBA,* and *APOE*, *WWOX,* and *TMEM108*[5]. This extends and confirms longitudinal studies implicating *GBA* variants[6,7] and *APOE* ε4[8] in cognitive decline in PD. These genome-wide and targeted sequencing efforts have paved the way for unravelling the genetic architecture of disease progression in PD, but have not yet investigated the second critical source of human DNA – the mitochondrial genome (mtDNA).

MtDNA mutations contribute to a spectrum of human diseases[9] and in PD there is accumulating genetic and environmental evidence that mitochondrial dysfunction may play a key role in the pathogenesis of the disease[10,11]. There are high level of somatic mtDNA mutations in substantia nigra neurons in early PD[12] and dysregulation of mtDNA homeostasis

in sporadic PD[13]. Mutations in the nuclear-encoded *PINK1* and *PRKN* cause autosomal recessive PD and disrupt mitophagy[14]. Moreover, there is a pervasive defect in *PGC-1alpha*-regulated mitochondrial bioenergetics gene expression in nigral dopamine neurons and substantia nigra even in prodromal, subclinical Lewy body neuropathology[15].

The diversity of modern human mtDNA haplogroups (variants) has provided valuable information to trace the history of human evolution, and many studies in recent years have reported links between specific mtDNA haplogroups and susceptibility for PD[16], however, the impact of mtDNA haplogroups or variants on progression in PD has not been defined. To characterize whether genetic variation in the mitochondrial genome influences the progression of PD, we performed a longitudinal, multi-cohort analysis, and identified specific mitochondrial haplogroups linked to cognitive decline in PD. Further exploratory analysis indicated two single nucleotide polymorphisms (SNPs) in mtDNA specifically associated with cognitive progression.

Materials and Methods

Study participants

The cohorts included in this study were described in previous work form the IGPP Consortium[5-7]. In brief, 4,491 patients with PD (with available genotyping data and quality control) were longitudinally assessed with 33,406 study visits in 15 cohorts from North America and Europe between 1986 and 2017 (**Supplementary Methods**). Written informed

consent for DNA collection and phenotypic data collection for secondary research use for each cohort was obtained from the participants with approval from the local ethics committees. The Institutional Review Board of Mass General Brigham and the Institutional Review Board of the School of Medicine, Sun Yat-sun University approved the current analyses. Patients whose longitudinal follow-up evaluations were not consistent with a diagnosis of PD were excluded. Fifteen cohorts were *a priori* assigned to discovery or replication cohorts as we previously described[5] (**Supplementary Fig. 2**). This achieves an approximately two-thirds to one-third split among the two stages and a balanced distribution of the distinct types of cohorts (for example, purpose-designed biomarkers studies, phase 3 clinical trials, population-based cohorts) across stages.

Polymorphism identification and haplogroup classification

We analyzed 763 mitochondrial SNPs in 4,491 patients with PD and predicted their mitochondrial haplogroup using Haplogrep2.0[17] with default parameters using mitochondrial rCRS reference (**Supplmentary Fig. 1**). We next simplified the sub-haplogroups (455 sub-haplogroups) to the 34 haplogroups (**Supplementary Table 1**). After quality control (**Supplementary Methods**), 4,064 subjects with 30,515 study visits were used for haplogroup analysis (including H, HV*(excluding H, V), I, J, K, T, U# (excluding K) haplogroups). Out of 763 mitochondrial SNPs, 102 SNPs with allele frequency > 1% were used for single SNP Cox regression analysis.

Statistical analysis

The Cox proportional hazards (Cox PH) analysis was used to estimate the influence of different mitochondrial haplogroups on time (years from onset of PD) to reaching the endpoint of global cognitive impairment (GCI) as indicated by a MMSE ≤ 25 according to the recommendation the International Parkinson and Movement Disorder Society (MDS) Task Force[18] and adjusting for the covariates of age at onset, gender, years of education, and polygenic hazard score (PHS) as fixed effect, and for a cohort term as random effect. A second endpoint was time to motor disability with postural instability as indicated by Hoehn and Yahr stage 3 (HY3) adjusting for age at onset, gender, *GBA* carrier status and the cohort term similar to Ref.[6] (see **Supplementary Methods** for detail). For the single nucleotide variants, a similar Cox PH analysis was used (using the same co-variants as mentioned above) to investigate the effect of each SNP on time to cognitive impairment.

Generalized longitudinal mixed fixed and random effects analysis (LMM) of cognitive decline was performed with MMSE scores longitudinally assessed at varying times (enrollment visit and multiple longitudinal follow-up visits) in the combined data set (**Supplementary Methods**). All analyses were conducted in the R statistical environment version 4.0.2.

Data availability

The genotype and clinical data for PPMI included in this study are publicly available upon request to ppmi@loni.usc.edu through a PPMI Whole Genome Sequencing Data Agreement. Clinical data for PDBP included in this study are publicly available through

https://pdbp.ninds.nih.gov. Clinical longitudinal data and genotyping data for the other cohorts included are accessible through appropriate data sharing agreements that protect patient privacy with the institutions that conducted or are conducting study consents and clinical assessments under local institutional review board approvals.

Results

Mitochondrial haplogroup is associated with cognitive decline in patients with PD.

The genotyped data of 4,491 patients with PD across 15 cohorts from North America and Europe were used to estimate their mitochondrial haplogroups. 4,447 patients with 33,068 longitudinal study visits passed quality control (**Supplementary Fig. 1A**) and were classified into eight groups: seven macro-haplogroups (H, HV*, I, J, T, K, U$^{\#}$) and a group comprising various other haplogroups (**Supplementary Fig. 1B, Supplementary Table 1**). 41.13%, (1,829) patients belonged to macro-haplogroup H, which is a common mtDNA clade in Europe and it is found in approximately 43.10% of UK Biobank individuals[19]. There were no significant differences in demographic and clinical characteristics of the patients in the various macro-haplogroups (**Supplementary Table 2**). The proportion of the seven macro-haplogroups was consistent with a previous survey in various European countries (**Supplementary Table 3**) and did not differ between the 15 cohorts ($P \approx 1$, Fisher exact test, **Supplementary Fig. 2**). For 4,064 patients within seven macro-haplogroups, we assigned

2,811 patients and 12,605 longitudinal visits to the discovery population. 1,253 patients and 17,910 visits comprised the replicate population.

We then investigated the effect of seven macro-haplogroups on the risk of cognitive and motor impairment during the progression of Parkinson's disease in discovery and replicate populations. "Haplogroup" was an unordered categorical variable in our Cox PH model. An omnibus test for haplogroup variation with 6 degrees of freedom showed that the 7 haplogroups in general were differed from each other in their association with cognitive progression (the null hypothesis is that the haplogroups have the same effect) with an "omnibus" test P value < 0.001 in the discovery stage. We followed up this omnibus test with pertinent post hoc likelihood ratio tests which are the pairwise comparisons of each of the haplogroups against the "reference" haplogroup H. J, T, U$^{\#}$ haplogroups were associated with a reduced risk for GCI (MMSE ≤ 25) compared to the common haplogroup H with a hazard ratios (HR) of 0.65 (95% CI 0.44-0.97) and P = 0.033, HR of 0.53 (95% CI 0.34-0.83) and P = 0.0052 and HR of 0.68 (95% CI 0.49-0.96) with P = 0.028 in discovery stage, respectively (**Fig. 1A**). We further confirmed these associations in a replicate population, where the HRs were 0.45 (95% CI 0.22-0.94), 0.54 (95% CI 0.29-0.99), 0.51 (95% CI 0.28-0.92) with P values of 0.033, 0.047 and 0.025 for J, T, U$^{\#}$ compared with Haplogroup H, respectively (**Fig. 1B**). Consistently, in the combined analysis, HR were 0.58 (95% CI 0.41-0.82) with P = 0.0023, 0.53 (95% CI 0.37-0.77) with P = 0.0007, and 0.63 (95% CI 0.47-0.85) with P = 0.0023, respectively (**Fig. 1C**). For each haplogroup compared to haplogroup H, the

Cochran's Q-test and the I^2 index showed HRs across studies were homogeneous (**Supplementary Table 4**).

There was no difference in HR for GCI among sub-haplogroups of H (**Supplementary Fig. 3**). There was no difference in HR for motor progression to Hoehn & Yahr stage 3 (motor disability with postural instability in PD) for each of the seven macro-haplotypes in discovery, replication or combined populations (**Supplementary Fig. 4**). A polygenic hazard score (PHS) based on five nuclear genetic loci exhibited a substantial aggregate association with progression to PD dementia in our recent study[5]. Here, we calculated the PHS for each patient and found no association between PHS and mtDNA haplogroups (Kruskal-Wallis rank sum test $P = 0.59$, **Supplementary Fig. 5**). This suggests that mitochondrial and nuclear genome variants may play independent roles in the cognitive progression of PD.

Since subjects with macro-haplogroups J, T, $U^{\#}$ showed a protective effect compared to haplogroup H, we combined these subjects into a super-group (n = 1,298) and showed reduced risk for GCI with a HR = 0.59 (95% CI 0.48-0.74) and $P = 2.42 \times 10^{-6}$ (**Fig. 1D**) (macro-haplogroup H as reference) after adjusting for covariates. A liner mixed model analysis confirmed serial MMSE scores in patients with macro-haplogroups J, T, $U^{\#}$ declined more slowly over time compared to patients in the common macro-haplogroup H ($P = 0.018$).

Exploratory analysis of single nucleotide polymorphisms in mtDNA and cognitive decline in PD.

We next carried out an exploratory analysis to investigate the effect of single nucleotide polymorphisms in mtDNA on cognitive impairment during the progression of PD in combined population (**Methods**). We observed that two variants, m.2706A>G and m.14766C>T, were associated with cognitive decline (**Fig. 2A**). The common m.2706A>G variant (G allele carriers, 58.3% in our cohorts) is located in the 16S rRNA locus. Patients with the m.2706G allele had a reduced risk of developing GCI with a HR = 0.68 (95% CI 0.56-0.81) and $P = 2.46\times10^{-5}$ compared to patients with the A allele (**Fig. 2B**). The common variant m.14766C>T (C allele carrier, 47.5% in our cohorts) codes for an amino acid substitution of an isoleucine for threonine at amino acid site 7 in *CYTB*. Patients with PD and m.14766T had a reduced risk of developing GCI with a HR = 0.70 (95% CI 0.58-0.84) and $P = 1.15\times10^{-4}$ compared to patients carrying the C allele. For m.2706A>G and m.14766C>T, proportional HRs across studies were homogeneous with $P = 0.46$ ($I^2 = 0\%$) and $P = 0.44$ ($I^2 = 0.96\%$) respectively, by Cochran's Q-test for heterogeneity. Associations of these two variants remained significant after considering multiple-testing with both *P* values lower than the Bonferroni-corrected significance threshold (0.05/102 variants tested = 4.9×10^{-4}). 12 additional variants were associated with cognitive decline during the course of PD with a $P < 0.05$ (**Fig. 2A, Table 1**).

Both m.2706A and m.14766C are largely specific to the H or HV* haplogroup. The alternative alleles m.2706G and m.14766T occur in other haplogroups (**Fig. 2C**). These results are consistent with our haplogroup analysis as patients within haplogroup J, T, $U^{\#}$ have a lower risk for cognitive progression compared to those with haplogroup H. We found

high correlation ($r^2 = 0.78$) of these two common variants in our cohorts and 94.1% of patients carried the same risk/protective alleles (m.2706A/m.14766C or m.2706G/m.14766T). After correcting for the effect of m.2706A>G, conditional Cox PH analysis no longer showed an association of m.14766C>T with cognitive decline (HR = 0.92 (95% CI 0.62-1.38); $P = 0.7$). Thus, m.14766C>T was dependent with m.2706A>G in our cohorts.

Age at disease onset, years of education, sex, MMSE at enrollment, MDS-UPDRS III score at enrollment, depression at enrollment are clinical variables associated with cognitive decline in PD[7]. 2,629 patients were included in both our previous[7] and current study, and we used these 2,376 patients (253 left censored patients were removed) for further analyses (**Supplementary Methods**). m.2706A and m.14766C carriers showed significant hazard ratios of 1.48 (95% CI 1.18-1.86, $P = 8.21\times10^{-4}$) and 1.38 (95% CI 1.09-1.74, $P = 7.23\times10^{-3}$) for risk of progression to GCI, respectively, adjusting for all six clinical predictors (**Supplementary Fig. 6**).

Consistent with our previous genome-wide survival analysis for progression from PD to PD dementia, *GBA* carriers had a HR of 1.91 (95% 1.39-2.64) with $P = 7.76\times10^{-5}$ and *APOE* ε4 carriers had a HR of 1.29 (95% 1.03-1.62) with $P = 0.028$ for cognitive decline (without accounting for mitochondrial variants, **Supplementary Fig. 6**). *GBA* carriers who carried the mitochondrial m.2706A allele (linked to relatively more *rapid* progression compared to the m.2706G allele) had a HR of 2.92 (95% CI, 1.87-4.55, $P = 2.23\times10^{-6}$). *GBA*-positive non-m.2706A carriers had the second highest HR of 1.84 (95% CI 1.15-2.93, *P*

= 0.011), and *GBA*-negative m.2706A carriers had a HR of 1.46 (95% CI 1.14-1.88, *P* = 0.0028) compared to patients carrying neither *GBA* variants nor the m.2706A variant (**Fig. 3**). Thus, m.2706A>G and *GBA* variants may have additive effects. Moreover, patients homozygous for the *APOE* ε4 allele and carrying m.2706A had a substantially elevated risk for longitudinal cognitive decline with HR = 5.09 (95% CI 2.04-12.56 *P* = 0.0005) compared to patients carrying neither *APOE* ε4 allele nor the m.2706A variant (**Supplementary Fig. 7)**.

Discussion

This genetic survival study overall indicates that mitochondrial macro-haplogroups are associated with reduced risk of cognitive disease progression. Post-hoc analyses identified the haplogroups J, T, U# as the haplogroups associated with reduced risk compared to the macro-haplogroup H in Parkinson's patients, but further research is required to definitively identify the contribution and statistical significance of each individual haplogroup. Previous meta-analyses found that the haplogroups J, K, T are associated with reduced susceptibility for PD and the haplogroup H is linked to elevated susceptibility for PD[16].

About 41% of patients with PD in this study belong to the macro-haplogroup H, the most common genotype in Europeans. The European mtDNA haplogroup H is associated with a higher survival ratio after sepsis[20], but is linked to higher risk of developing PD in late life[16]. On the flip side our findings are consistent with a relatively more deleterious effect of haplogroup H on the progression of PD compared to haplogrups J, T, U#. This may represent an evolutionary trade-offs[21] where genetic variants that increase the chance of surviving early-life illness such as sepsis might contribute to pathogenic events later in life[21].

Alzheimer's disease-associated plaques and tangles are found in a substantial proportion of brains with of patients with PD dementia, in addition to Lewy bodies[22]. Haplogroups H and HV are a risk haplogroup for AD[23], while the JT haplogroup was protective in a prior study[24]; evidence for the other haplogroups (K, J, T, U) is limited and controversial (e.g. J[23,25]; **Supplementary Table 5)**. This is also consistent with our study,

where H carriers had a relatively more rapid cognitive progression compared to the protective haplogroups J, T, U[#].

Two common significant mtDNA variants showed effects on the risk of global cognitive impairment and are related to the haplogroups (**Fig. 2C**). The common m.2706A>G variant, located at 16S rRNA gene, is close to the ribosomal peptidyl transferase center, and might be relevant to many diseases, such as MELAS, Alzheimer's and PD[26]. This variant can induce substantial alterations in the mitochondrial 16S rRNA secondary structure[27]. The m.14766C variant might increase the risk for late-onset Alzheimer's disease[24] consistent with our findings. Interestingly, contrary to our data, m.2706G was associated with faster cognitive aging in a large longitudinal cohort of African Americans, but not Caucasian Americans[28].

Our study is limited in sample size and statistical power. *P* values for individual haplogroups were not adjusted for multiple testing. Another limitation of this study is that we evaluated the effects of mitochondrial genetic variants during the progression of PD with European ancestry only. The mtDNA variants (m.2706A>G or m.14766C>T) are rare in populations from East Asia or Africa (**Table 1**). Further studies in other populations are urgently needed due to difference of mtDNA haplogroups, considering that more than 60% of PD patients are expected to live in the Western Pacific Region by 2030[29], most of them belonging to haplogroups A, B, C, D, F, G. Moreover, replication of our exploratory findings in additional longitudinal patient populations of European ancestry is needed.

This study suggests that mitochondrial genotypes may not be innocent bystanders in the progression of Parkinson's, but, might play an role in modulating disease progression. Our study provides evidence for the role of mitochondrial haplogroups in the progression of PD towards Lewy body dementia, and this association appears independent of *GBA* and *APOE*.

Mitochondrial dysfunction[15] and alpha-synuclein accumulation are two pathologically[30] and biologically[31] linked culprits of PD. alpha-synclein triplication causes mitochondrial bioenergetics dysfunction[32]. Conversely, the mitochondrial toxin rotenone leads to alpha-synclein accumulation[33]. Taken together with our new findings, this body of evidence suggests that mitochondria might play a role not only in the onset, but also in the progression of Parkinson's disease.

Acknowledgments

We thank Mr. Ofer Nemirovsky for his invaluable support, encouragement, insights, and dedication to accelerating PD research. We thank all study participants, their families, and friends for their support and participation, and our study coordinators for making this work possible. For each individual cohort, acknowledgements and fundings are listed in the **Supplementary material online.**

Funding

G.L.'s work is supported by the Fundamental Research Funds for the Central Universities, Sun Yat-sen University (22ykqb07), National Natural Science Foundation of China (31900475), Guangdong Basic and Applied Basic Research Foundation (2022A1515011440), Shenzhen Basic Research Project (JCYJ20190807161601692) and Young Talent Recruitment Project of Guangdong (2019QN01Y139). C.R.S.'s work is supported by NIH grants NINDS/NIA R01NS115144, U01NS095736, U01NS100603, and the American Parkinson Disease Association Center for Advanced Parkinson Research. The study was made possible in part by a philanthropic support for Illumina MEGA chip genotyping (to Brigham & Women's Hospital and C.R.S.).

Competing interests

G.L., C.N., J.Z., W.L., J.L., J.J.L., W.X., L. C., Z.P., A.E., J.H.G., A.Y.H., S.K., P.T., S.S.A., J.S.P., and M.C.C. report no relevant financial or other conflicts of interest in relation to this study.

Outside this work, C.R.S. has served as consultant, scientific collaborator or on scientific advisory boards for Sanofi, Berg Health, Pfizer, Biogen, and has received grants from NIH, U.S. Department of Defense, American Parkinson Disease Association, and the Michael J Fox Foundation (MJFF).

M.A.S. has no conflict of interest related to this work. Outside this work, M.A.S. has received grants from NINDS, DoD, MJFF, Parkinson's Foundation, Farmer Family Foundation, and has served as a consultant to commercial programs: Eli Lilly & Co (data monitoring committee), Prevail Therapeutics (scientific advisory board), and Denali Therapeutics (scientific advisory board); and via the Parkinson Study Group to nQ Medical (scientific advisory board), Chase Therapeutics (scientific advisory board) and Partner Therapeutics (scientific advisory board).

A.-M.W. has received research funding from the ALS Association, the Parkinson's Foundation, has participated in clinical trials funded by Acorda, Biogen, Bristol-Myers Squibb, Sanofi/Genzyme, Pfizer, Abbvie, and received consultant payments from Mitsubishi Tanabe and from Accordant.

J.-C.C. has no conflict of interest related to this work. Outside this work, J.C.C. received honoraria for consulting in advisory boards for Abbvie, Actelion, Air Liquide, Biogen, BMS, BrainEver, Clevexel, Denali, Pfizer, Theranexus, and Zambon.

B.R. is an employee of and holds equity in Praxis Precision Medicines and is advisor for Caraway Therapeutics and Brain Neurotherapy Bio.

I.S. is Principal investigator of a MJFF Computational Science Grant (2017-19).

S.K. is supported by Multiple Sclerosis of Western-Australia (MSWA) and the Perron Institute.

P.H. is a Scientific Advisor of Neuron23.

T.G.B has no conflict of interest related to this work. Outside this work, T.G.B. has received grants from NIA, NINDS, MJFF, and the State of Arizona, has served as a scientific advisory board member (with stock options) and consultant to Vivid Genomics, Inc. and has received honoraria from the World PD Coalition.

J.J.v.H. has no conflict of interest related to this work. Outside this work, J.J.v.H. has received grants from the Alkemade-Keuls Foundation, Stichting Parkinson Fonds, Parkinson Vereniging, The Netherlands Organisation for Health Research and Development, The Netherlands Organisation for Scientific Research, Hersenstichting, AbbVie, Michael J Fox Foundation, and research support from Hoffmann-La-Roche, Lundbeck and the Centre of Human Drug Research.

R.A.B. has no conflict of interest related to this work. Outside this work, R.A.B. received consultancy monies from LCT; FCDI; Novo Nordisk; Cellino; Sana; UCB; received royalties from Wiley and Springer-Nature; grant funding from CPT; NIHR Cambridge Biomedical Research Centre (146281); MRC; Wellcome (203151/Z/16/Z) and Rosetrees Trust (A1519 M654).

C.H.W.-G. has no conflict of interest related to this work. C.H.W.-G. is supported by a RCUK/UKRI Research Innovation Fellowship awarded by the Medical Research Council (MR/R007446/1) and the NIHR Cambridge Biomedical Research Centre, and received grant support from MJFF, the Evelyn Trust, the Cure Parkinson's Trust, Parkinson's UK, the Rosetrees Trust and the Cambridge Centre for Parkinson-Plus. C.H.W.-G. has received

honoraria from Lundbeck and Profile Pharma Ltd and consultancy payments from Modus Outcomes and Evidera.

Supplementary material

Supplementary material is available at *Brain* online.

References

1. Poewe W, Seppi K, Tanner CM, *et al*. Parkinson disease. *Nat Rev Dis Primers*. Mar 23 2017;3:17013.
2. Yang W, Hamilton JL, Kopil C, *et al*. Current and projected future economic burden of Parkinson's disease in the U.S. *NPJ Parkinsons Dis*. 2020;6:15.
3. Greenland JC, Williams-Gray CH, Barker RA. The clinical heterogeneity of Parkinson's disease and its therapeutic implications. *Eur J Neurosci*. Feb 2019;49(3):328-338.
4. Nalls MA, Blauwendraat C, Vallerga CL, *et al*. Identification of novel risk loci, causal insights, and heritable risk for Parkinson's disease: a meta-analysis of genome-wide association studies. *The Lancet Neurology*. Dec 2019;18(12):1091-1102.
5. Liu G, Peng J, Liao Z, *et al*. Genome-wide survival study identifies a novel synaptic locus and polygenic score for cognitive progression in Parkinson's disease. *Nature genetics*. May 6 2021;53(6):787-793.
6. Liu G, Boot B, Locascio JJ, *et al*. Specifically neuropathic Gaucher's mutations accelerate cognitive decline in Parkinson's. *Annals of neurology*. Nov 2016;80(5):674-685.
7. Liu G, Locascio JJ, Corvol JC, *et al*. Prediction of cognition in Parkinson's disease with a clinical-genetic score: a longitudinal analysis of nine cohorts. *The Lancet Neurology*. Aug 2017;16(8):620-629.
8. Iwaki H, Blauwendraat C, Leonard HL, *et al*. Genomewide association study of Parkinson's disease clinical biomarkers in 12 longitudinal patients' cohorts. *Mov Disord*. Dec 2019;34(12):1839-1850.
9. Tuppen HA, Blakely EL, Turnbull DM, Taylor RW. Mitochondrial DNA mutations and human disease. *Biochim Biophys Acta*. Feb 2010;1797(2):113-28.
10. Giannoccaro MP, La Morgia C, Rizzo G, Carelli V. Mitochondrial DNA and primary mitochondrial dysfunction in Parkinson's disease. *Mov Disord*. Mar 2017;32(3):346-363.
11. Gonzalez-Rodriguez P, Zampese E, Stout KA, *et al*. Disruption of mitochondrial complex I induces progressive parkinsonism. *Nature*. Nov 2021;599(7886):650-656.
12. Lin MT, Cantuti-Castelvetri I, Zheng K, *et al*. Somatic mitochondrial DNA mutations in early Parkinson and incidental Lewy body disease. *Annals of neurology*. Jun 2012;71(6):850-4.
13. Dolle C, Flones I, Nido GS, *et al*. Defective mitochondrial DNA homeostasis in the substantia nigra in Parkinson disease. *Nat Commun*. Nov 22 2016;7:13548.
14. Pickrell AM, Youle RJ. The roles of PINK1, parkin, and mitochondrial fidelity in Parkinson's disease. *Neuron*. Jan 21 2015;85(2):257-73.
15. Zheng B, Liao Z, Locascio JJ, *et al*. PGC-1alpha, a potential therapeutic target for early intervention in Parkinson's disease. *Sci Transl Med*. Oct 6 2010;2(52):52ra73.
16. Hudson G, Nalls M, Evans JR, *et al*. Two-stage association study and meta-

analysis of mitochondrial DNA variants in Parkinson disease. *Neurology*. May 28 2013;80(22):2042-8.

17. Weissensteiner H, Pacher D, Kloss-Brandstatter A, *et al*. HaploGrep 2: mitochondrial haplogroup classification in the era of high-throughput sequencing. *Nucleic Acids Res*. Jul 8 2016;44(W1):W58-63.

18. Dubois B, Burn D, Goetz C, *et al*. Diagnostic procedures for Parkinson's disease dementia: recommendations from the movement disorder society task force. *Mov Disord*. Dec 2007;22(16):2314-24.

19. Yonova-Doing E, Calabrese C, Gomez-Duran A, *et al*. An atlas of mitochondrial DNA genotype-phenotype associations in the UK Biobank. *Nature genetics*. May 17 2021;

20. Baudouin SV, Saunders D, Tiangyou W, *et al*. Mitochondrial DNA and survival after sepsis: a prospective study. *Lancet*. Dec 17 2005;366(9503):2118-21.

21. Benton ML, Abraham A, LaBella AL, Abbot P, Rokas A, Capra JA. The influence of evolutionary history on human health and disease. *Nat Rev Genet*. Jan 6 2021;

22. Aarsland D, Batzu L, Halliday GM, *et al*. Parkinson disease-associated cognitive impairment. *Nat Rev Dis Primers*. Jul 1 2021;7(1):47.

23. Maruszak A, Canter JA, Styczynska M, Zekanowski C, Barcikowska M. Mitochondrial haplogroup H and Alzheimer's disease--is there a connection? *Neurobiol Aging*. Nov 2009;30(11):1749-55.

24. Maruszak A, Safranow K, Branicki W, *et al*. The impact of mitochondrial and nuclear DNA variants on late-onset Alzheimer's disease risk. *J Alzheimers Dis*. 2011;27(1):197-210.

25. Swerdlow RH, Hui D, Chalise P, *et al*. Exploratory analysis of mtDNA haplogroups in two Alzheimer's longitudinal cohorts. *Alzheimer's & Dementia*. 2020;16(8):1164-1172.

26. Katzman SM, Strotmeyer ES, Nalls MA, *et al*. Mitochondrial DNA Sequence Variation Associated With Peripheral Nerve Function in the Elderly. *J Gerontol A Biol Sci Med Sci*. Nov 2015;70(11):1400-8.

27. Rovcanin B, Jancic J, Samardzic J, *et al*. In silico model of mtDNA mutations effect on secondary and 3D structure of mitochondrial rRNA and tRNA in Leber's hereditary optic neuropathy. *Exp Eye Res*. Dec 2020;201:108277.

28. Yen K, Wan J, Mehta HH, *et al*. Humanin Prevents Age-Related Cognitive Decline in Mice and is Associated with Improved Cognitive Age in Humans. *Sci Rep*. Sep 21 2018;8(1):14212.

29. Lim SY, Tan AH, Ahmad-Annuar A, *et al*. Parkinson's disease in the Western Pacific Region. *The Lancet Neurology*. Sep 2019;18(9):865-879.

30. Shahmoradian SH, Lewis AJ, Genoud C, *et al*. Lewy pathology in Parkinson's disease consists of crowded organelles and lipid membranes. *Nat Neurosci*. Jul 2019;22(7):1099-1109.

31. Scherzer CR, Grass JA, Liao Z, *et al*. GATA transcription factors directly

regulate the Parkinson's disease-linked gene alpha-synuclein. *Proc Natl Acad Sci U S A*. Aug 5 2008;105(31):10907-12.

32. Zambon F, Cherubini M, Fernandes HJR, *et al*. Cellular alpha-synuclein pathology is associated with bioenergetic dysfunction in Parkinson's iPSC-derived dopamine neurons. *Hum Mol Genet*. Jun 15 2019;28(12):2001-2013.

33. Betarbet R, Sherer TB, MacKenzie G, Garcia-Osuna M, Panov AV, Greenamyre JT. Chronic systemic pesticide exposure reproduces features of Parkinson's disease. *Nat Neurosci*. Dec 2000;3(12):1301-6.

Figure legends

Figure 1 Mitochondrial haplogroups and risk for global cognitive impairment over time in patients with PD. The forest plot shows hazard ratios for global cognitive impairment in specific types of macro-haplogroups compared to macro-haplogroup H in patients with PD from the discovery **(A)**, replication **(B)** and combined **(C)** populations. The squares represent point estimates, with the sides of the square inversely proportional to the standard error of the estimates. The horizontal lines indicate 95% confidence intervals of the estimates. **(D)** Covariate-adjusted survival curves for patients with PD in macro-haplogroups J T, U$^{\#}$ (cyan line) and those in macro-haplogroups H (magenta line).

Figure 2 mtDNA variants associated with cognitive progression in patients with PD. (A) Association plot of SNPs in mtDNA associated with risk of developing global cognitive impairment (dot) in the combined population. The outside labels indicate mitochondrial genes; circular axis from outside to inside represents the value of -log10(*P*) from 0 to 5; SNPs with $P < 0.05$ are shown in magenta, while SNPs with $P \geq 0.05$ are shown in grey. **(B)** Covariate-adjusted survival curves for patients with PD carrying mtDNA m.2706G (cyan line) and those with m.2706A (magenta line). m.2706A was used as a reference allele to calculate HR from the Cox PH analysis and *P* values from two-sided Wald tests. **(C)** Overlap between carriers of the m.2706A>G and the m.14766C>T variant. Out of 2,611 m.2706G allele carriers and 2,347 m.14766T allele carriers, 2,342 individuals carried both alleles. Out of 1,830 m.2706A allele carriers and 2,080 m.14766C allele carriers, 1,819 individuals carried both alleles.

Figure 3 Effects of *GBA* variants and mtDNA genotype in predicting global cognitive impairment for patients with PD. Covariate-adjusted survival curves for patients with PD

stratified into four subgroups: *GBA*-negative and non-m.2706A carriers (n = 1,257), *GBA*-negative and m.2706A carriers (n = 891), *GBA*-positive and non-m.2706G carriers (n = 132), *GBA*-positive and m.2706A carriers (n = 96). HR and P values were calculated adjusting for clinical covariates and study cohort as a random term. The group of *GBA*-negative and non-m.2706A carriers is denoted as reference group (REF) in Cox PH analysis.

Tables

Table 1 mtDNA SNPs were associated with developing global cognitive impairment in the progression of Parkinson's

rCRS	Effect allele	Alternative allele	*P*	*P*[a]	HR (95%CI)	EAF[b]	EAF in European	EAF in East Asian	EAF in African
m.2706A>G	G	A	2.46×10^{-5}	0.003	0.68 (0.56-0.81)	0.5826	0.5746	0.9960	0.9970
m.14766C>T	T	C	1.15×10^{-4}	0.012	0.70 (0.58-0.84)	0.5249	0.5169	0.9960	0.9985
m.11251A>G	G	A	0.002	0.204	0.67 (0.52-0.86)	0.1942	0.1610	0.0000	0.0015
m.15452C>A	A	C	0.002	0.204	0.67 (0.52-0.87)	0.1951	0.1610	0.0000	0.0000
m.15607A>G	G	A	0.017	1	0.65 (0.46-0.93)	0.0984	0.0875	0.0000	0.0015
m.16162A>G	G	A	0.019	1	1.95 (1.12-3.40)	0.0235	0.0199	0.0417	0.0015
m.15928G>A	A	G	0.021	1	0.66 (0.47-0.94)	0.0998	0.0875	0.0080	0.0000
m.11812A>G	G	A	0.029	1	0.65 (0.44-0.96)	0.0787	0.0696	0.0000	0.0045
m.4917A>G	G	A	0.030	1	0.66 (0.46-0.96)	0.0971	0.0875	0.0000	0.0000
m.9477G>A	A	G	0.031	1	0.66 (0.46-0.96)	0.0926	0.1392	0.0000	0.0061
m.10589G>A	A	G	0.041	1	1.89 (1.03-3.47)	0.0110	0.0060	0.0020	0.0530
m.16482A>G	G	A	0.043	1	1.62 (1.02-2.59)	0.0200	0.0139	0.0020	0.0000
m.15218A>G	G	A	0.045	1	0.55 (0.31-0.99)	0.0453	0.0437	0.0119	0.0000
m.10463T>C	C	T	0.048	1	0.71 (0.51-1.00)	0.1030	0.0875	0.0040	0.0000

rCRS = revised Cambridge Reference Sequence; HR = hazard ratio; CI = confidence interval; *P* from the Cox proportional hazards statistic to estimate the influence of SNP on time (years from onset of PD) to reaching the endpoint of global cognitive impairment (GCI) as indicated by a MMSE ≤ 25 in exploratory analysis using combined population; age at onset of PD, sex, years of education, and polygenic hazard score (PHS including *GBA* mutation status, *APOE* ε4 allele haplotype and novel 3 SNPs loci rs182987047, rs138073281 and rs8050111) were included as covariates in the Cox analyses. A "cohort" term was included as a random effect; *P*[a]: Bonferroni correction based on the result of 102 mtDNA SNPs from combined analysis was performed using p.adjust function with "bonferroni" method in R. Effect allele frequency (EAF)[b] was calculated based on 4,491 patients with PD across 15 cohorts. EAF in 503 European, 503 East Asian or 661 African was calculated based on dataset of Phase 1 and 3 of the 1000 Genome Project mitochondrial variants calling by the MToolBox pipeline.

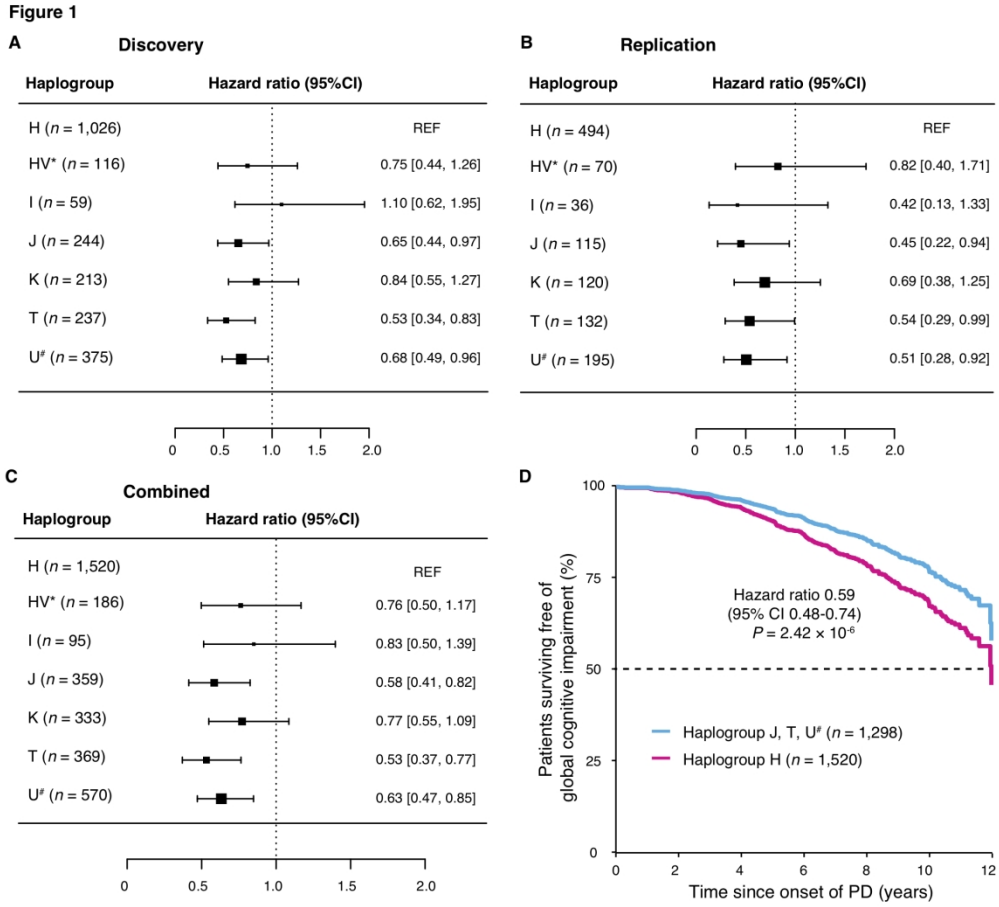


Figure 1 Mitochondrial haplogroups and risk for global cognitive impairment over time in patients with PD.

177x162mm (300 x 300 DPI)

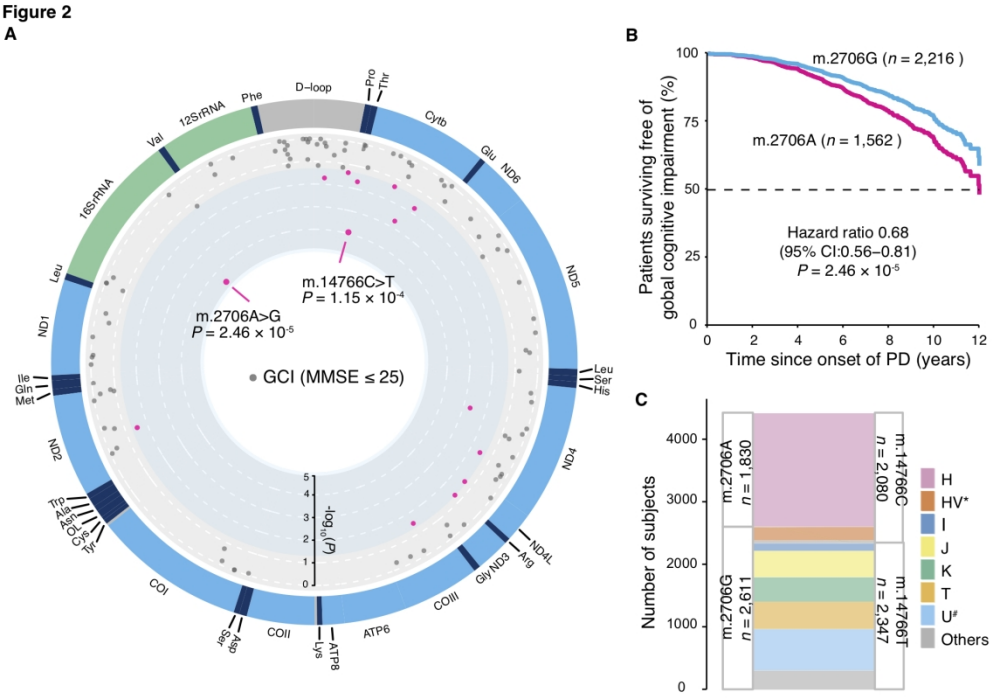


Figure 2 mtDNA variants associated with cognitive progression in patients with PD.

177x127mm (300 x 300 DPI)

Figure 3

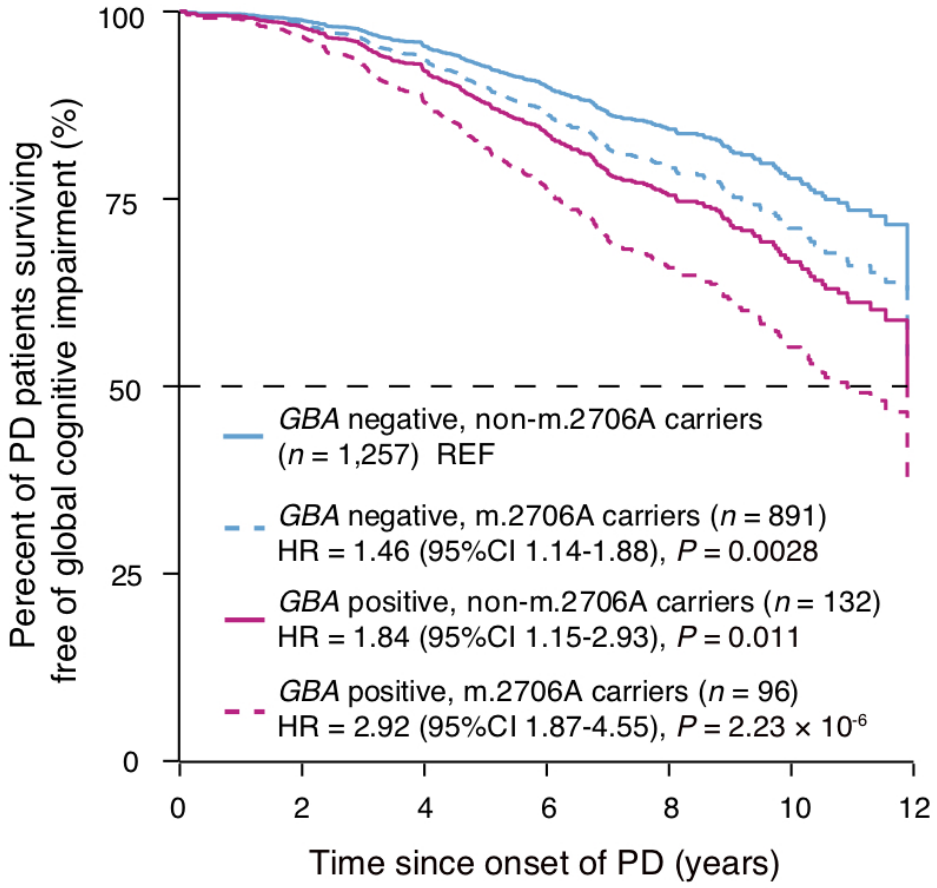


Figure 3 Effects of GBA variants and mtDNA genotype in predicting global cognitive impairment for patients with PD.

71x72mm (300 x 300 DPI)

Mitochondrial haplogroups and cognitive progression in Parkinson's disease

Supplementary material

Supplementary material 1

Study participants 7

Polymorphism identification and haplogroup classification 8

Statistical analysis 9

Comparison of models 10

Supplementary Table 1 Classification of mitochondrial haplogroups in patients with PD across the 15 cohorts 12

Supplementary Table 2 Clinical characteristics of patients with PD at enrollment with different macro-haplogroups across the 15 cohorts 13

Supplementary Table 3 The percentage (%) of different mitochondrial haplogroups in European population from literatures 14

Supplementary Table 4 Test for residual heterogeneity for each haplogroup compared to haplogroups of H in GCI combined analysis 15

Supplementary Table 5 The association of mitochondrial haplogroups in Alzheimer disease ... 16

Supplementary Figure 1 The classification of haplogroup in patients with PD across 15 cohorts. 17

Supplementary Figure 2 The stacking diagram for distribution of seven macro-haplogroups in patients with PD across 15 cohorts 18

Supplementary Figure 3 Patients with PD with major sub-haplogroups of H have similar risk of progression to global cognitive impairment 19

Supplementary Figure 4 Patients with PD in seven macro-haplogroups have similar risk of progression to HY3. 20

Supplementary Figure 5 Patients with PD in seven mitochondrial macro-haplogroups have similar polygenic hazard scores 21

Supplementary Figure 6 The exploratory analysis for global cognitive impairment models with different genetic factors. 22

Supplementary Figure 7 Exploratory analysis for global cognitive impairment models with APOE ε4 and m.2706A>G. 23

References 24

List of Acknowledgements

Harvard Biomarkers Study (including HBS and HBS2 parts). Co-Directors: Brigham and Women's Hospital: Clemens R. Scherzer, Massachusetts General Hospital: Bradley T. Hyman; Investigators and Study Coordinators: Brigham and Women's Hospital: Yuliya Kuras, Karbi Choudhury, Nada Laroussi, Daly Franco, Michael T. Hayes, Nutan Sharma, Vikram Khurana, Claudio Melo De Gusmao, Chizoba C. Umeh, Reisa Sperling; Massachusetts General Hospital: John H. Growdon, Michael A. Schwarzschild, Albert Y. Hung, Aleksandar Videnovic, Alice W. Flaherty, Deborah Blacker, Anne-Marie Wills, Steven E. Arnold, Ann L. Hunt, Nicte I. Mejia, Anand Viswanathan, Stephen N. Gomperts, Mark W. Albers, Maria Allora-Palli, David Hsu, Alexandra Kimball, Scott McGinnis, John Becker, Randy Buckner, Thomas Byrne, Maura Copeland, Bradford Dickerson, Matthew Frosch, Theresa Gomez-Isla, Steven Greenberg, Julius Hedden, Elizabeth Hedley-Whyte, Keith Johnson, Raymond Kelleher, Aaron Koenig, Maria Marquis-Sayagues, Gad Marshall, Sergi Martinez-Ramirez, Donald McLaren, Olivia Okereke, Elena Ratti, Christopher William, Koene Van Dij, Shuko Takeda, Anat Stemmer-Rachaminov, Jessica Kloppenburg, Catherine Munro, Rachel Schmid, Sarah Wigman, Sara Wlodarcsyk; Data Coordination: Brigham and Women's Hospital: Thomas Yi; Biobank Management Staff: Brigham and Women's Hospital: Idil Tuncali. We thank all study participants and their families for their invaluable contributions. HBS is made possible by generous support from the Harvard NeuroDiscovery Center, with additional contributions from the Michael J Fox Foundation, NINDS U01NS082157, U01NS100603, and the Massachusetts Alzheimer's Disease Research Center NIA P50AG005134.

DIGPD: Steering committee: Jean-Christophe Corvol (Pitié-Salpêtrière Hospital, Paris, principal investigator of DIGPD), Alexis Elbaz (CESP, Villejuif, member of the steering committee), Marie Vidailhet (Pitié-Salpêtrière Hospital, Paris, member of the steering committee), Alexis Brice (Pitié-Salpêtrière Hospital, Paris, member of the steering committee and PI for genetic analysis) ; Statistical analyses: Alexis Elbaz (CESP, Villejuif, PI for statistical analyses), Fanny Artaud (CESP, Villejuif, statistician); Principal investigators for sites (alphabetical order): Frédéric Bourdain (CH Foch, Suresnes, PI for site), Jean-Philippe Brandel (Fondation Rothschild, Paris, PI for site), Jean-Christophe Corvol (Pitié-Salpêtrière Hospital, Paris, PI for site), Pascal Derkinderen (CHU Nantes, PI for site), Franck Durif (CHU Clermont-Ferrand, PI for site), Richard Levy (CHU Saint-Antoine, Paris, PI for site), Fernando Pico (CH Versailles, PI for site), Olivier Rascol (CHU Toulouse, PI for site); Co-investigators (alphabtical order): Anne-Marie

Bonnet (Pitié-Salpêtrière Hospital, Paris, site investigator), Cecilia Bonnet (Pitié-Salpêtrière Hospital, Paris, site investigator), Christine Brefel-Courbon (CHU Toulouse, site investigator), Florence Cormier-Dequaire (Pitié-Salpêtrière Hospital, Paris, site investigator), Bertrand Degos (Pitié-Salpêtrière Hospital, site investigator), Bérangère Debilly (CHU Clermont-Ferrand, site investigator), Alexis Elbaz (Pitié-Salpêtrière Hospital, Paris, site investigator), Monique Galitsky (CHU de Toulouse, site investigator), David Grabli (Pitié-Salpêtrière Hospital, Paris, site investigator), Andreas Hartmann (Pitié-Salpêtrière Hospital, Paris, site investigator), Stephan Klebe (Pitié-Salpêtrière Hospital, Paris, site investigator), Julia Kraemmer (Pitié-Salpêtrière Hospital, site investigator), Lucette Lacomblez (Pitié-Salpêtrière Hospital, Paris, site investigator), Sara Leder (Pitié-Salpêtrière Hospital, Paris, site investigator), Graziella Mangone (Pitié-Salpêtrière Hospital, Paris, site investigator), Louise-Laure Mariani (Pitié-Salpêtrière Hospital, Paris, site investigator), Ana-Raquel Marques (CHU Clermont Ferrand, site investigator), Valérie Mesnage (CHU Saint Antoine, Paris, site investigator), Julia Muellner (Pitié-Salpêtrière Hospital, Paris, site investigator), Fabienne Ory-Magne (CHU Toulouse, site investigator), Violaine Planté-Bordeneuve (Henri Mondor Hospital, Créteil, site investigator), Emmanuel Roze (Pitié-Salpêtrière Hospital, Paris, site investigator), Melissa Tir (CH Versailles, site investigator), Marie Vidailhet (Pitié-Salpêtrière Hospital, Paris, site investigator), Hana You (Pitié-Salpêtrière Hospital, Paris, site investigator); Neuropsychologists: Eve Benchetrit (Pitié-Salpêtrière Hospital, Paris, neuropsychologist), Julie Socha (Pitié-Salpêtrière Hospital, Paris, neuropsychologist), Fanny Pineau (Pitié-Salpêtrière Hospital, Paris, neuropsychologist), Tiphaine Vidal (CHU Clermont-Ferrand, neuropsychologist), Elsa Pomies (CHU de Toulouse, neuropsychologist), Virginie Bayet (CHU de Toulouse, neuropsychologist); Genetic core: Alexis Brice (Pitié-Salpêtrière Hospital, Paris, PI for genetic studies), Suzanne Lesage (INSERM, ICM, Paris, genetic analyses), Khadija Tahiri (INSERM, ICM, Paris, lab technician) Hélène Bertrand (INSERM, ICM, Paris, lab technician), Graziella Mangone (Pitié-Salpêtrière Hospital, Paris, genetic analyses); Sponsor activities and clinical research assistants: Alain Mallet (Pitié-Salpêtrière Hospital, Paris, sponsor representative), Coralie Villeret (Hôpital Saint Louis, Paris, Project manager), Merry Mazmanian (Pitié-Salpêtrière Hospital, Paris, project manager), Hakima Manseur (Pitié-Salpêtrière Hospital, Paris, clinical research assistant), Mostafa Hajji (Pitié-Salpêtrière Hospital, Paris, data manager), Benjamin Le Toullec (Pitié-Salpêtrière Hospital, Paris, clinical research assistant), Vanessa Brochard (Pitié-Salpêtrière Hospital, Paris, project manager), Monica Roy (CHU de Nantes, clinical researh assistant), Isabelle Rieu (CHU Clermont-Ferrand, clinical research assistant), Stéphane Bernard (CHU Clermont-Ferrand, clinical research assistant), Antoine Faurie-Grepon (CHU de Toulouse, clnical research assistant). The study was sponsored

by the Assistance Publique Hôpitaux de Paris, and was funded by a grant from the Ministry of Health (PHRC AOR0810).

PreCEPT/PostCEPT Study: PreCEPT/PostCEPT Steering Committee: University of Rochester: David Oakes, Ira Shoulson; University of Toronto: Anthony E. Lang; Parlinson's Institute: Caroline Tanner; Institute for Neurodegenerative Disorders: Kenneth Marek; Voyager Therapeutics: Bernard Ravina; Brigham and Women's Hospital: Clemens Scherzer, University of Ottawa: Michael Schlossmacher, Avid Radiopharmaceuticals: Andrew Siderowf, We thank the Parkinson Study Group (PSG) PreCEPT/PostCEPT investigators for the acquisition of high-quality clinical data, careful follow up of study subjects and collection of blood samples.

DATATOP: We thank the investigators of the Parkinson Study Group (PSG) DATATOP for the acquisition of high-quality clinical data, careful follow-up of study subjects, and DNA collection in the DATATOP cohort.

PICNICS: Investigators:Roger Barker, Caroline Williams-Gray, David P Breen, Gemma Cummins, Jonathan Evans, Sophie Winder-Rhodes, Ruwani Wijeyekoon. The PICNICS study was sponsored by the University of Cambridge/Cambridge University Hospitals NHS Trust UK and received funding from the Cure Parkinson's Trust, the Van Geest Foundation, Parkinson's UK, and the NIHR Cambridge Biomedical Research Centre.

CamPaIGN: Investigators: Roger Barker, Tom Foltynie, Caroline Williams-Gray, Trevor Robbins, Carol Brayne, Sarah Mason, Sophie Winder-Rhodes, Ruwani Wijeyekoon. The CamPaIGN study was sponsored by the University of Cambridge/Cambridge University Hospitals NHS Trust UK and has received funding from the Wellcome Trust, the Medical Research Council, the Patrick Berthoud Trust, and the NIHR Cambridge Biomedical Research Centre.

PROPARK/PROPARK-C: The PROPARK study was headed by Jacobus J. van Hilten and Johan Marinus.

PDBP: Data and biospecimens used in preparation of this manuscript were obtained from the Parkinson's Disease Biomarkers Program (PDBP) Consortium, supported by the National Institute of Neurological Disorders and Stroke at the National Institutes of Health. Investigators include: Roger Albin, Roy Alcalay, Alberto Ascherio, Thomas Beach, Sarah Berman, Bradley

Boeve, F. DuBois Bowman, Shu Chen, Alice Chen-Plotkin, William Dauer, Ted Dawson, Paula Desplats, Richard Dewey, Ray Dorsey, Jori Fleisher, Kirk Frey, Douglas Galasko, James Galvin, Dwight German, Lawrence Honig, Xuemei Huang, David Irwin, Kejal Kantarci, Anumantha Kanthasamy, Daniel Kaufer, James Leverenz, Carol Lippa, Irene Litvan, Oscar Lopez, Jian Ma, Lara Mangravite, Karen Marder, Laurie Orzelius, Steven Gunzler, Vladislav Petyuk, Judith Potashkin, Liana Rosenthal, Rachel Saunders-Pullman, Clemens Scherzer, Michael Schwarzschild, Tanya Simuni, Andrew Singleton, David Standaert, Debby Tsuang, David Vaillancourt, David Walt, Andrew West, Cyrus Zabetian, Jing Zhang, and Wenquan Zou. The PDBP Investigators have not participated in reviewing the data analysis or content of the manuscript.

PPMI: Data used in the preparation of this article were obtained from the Parkinson's Progression Markers Initiative (PPMI) database (www.ppmi-info.org/data). For up-to-date information on the study, visit www.ppmi-info.org. PPMI, a public-private partnership, is funded by the Michael J. Fox Foundation for Parkinson's Research and funding partners, including Abbvie, AcureX, Allergan, Amathus, Asap, Avid, Bial Biotech, Biogen Idec, BioLegend, Bristol-Myers Squibb, Calico, Celgene, DACAPO Brainscience, DENALI, 4D pharma plc, EDMOND J. SAFRA, GE Healthcare, Genentech, GlaxoSmithKline, GOLUB CAPITAL, Handl Therapeutics, Insitro, Janssen Neuroscience, Lilly, Lundbeck, Merck, Meso Scale Discovery, Neurocrine, Pfizer, Piramal, Prevail Therapeutics, Roche, SANOFI GENZYME, SERVIER, Takeda, TEVA, UCB, Verily and Voyager Therapeutics.

Arizona Study of Aging/Brain and Body Donation Program: National Institute of Neurological Disorders and Stroke, U24 NS072026 National Brain and Tissue Resource for Parkinson's Disease and Related Disorders; National Institute on Aging, P30 AG19610 Arizona Alzheimer's Disease Core Center; Arizona Department of Health Services, Arizona Alzheimer's Consortium; Arizona Biomedical Research Commission, Arizona Parkinson's Disease Consortium; Michael J. Fox Foundation for Parkinson's Research.

NET-PD LS1: We would like to thank the patients and families who participate in the NET-PD LS1 study. The following additional NINDS grants supported the Net-PD LS-1 study: U01NS043127, U01NS043128, and U10NS44415-44555 from the National Institute of Neurologic Disorders and Stroke.

Tartu/Perron: We would like to thank all patients and families who participated in the study. Support by the following grants and funders: institutional research grants PRG957 and IUT20–46 of the Estonian Research Council, H2020 ERA-chair grant (agreement 668989, project Transgeno), MSWA, The Michael J. Fox Foundation, Shake It Up Australia and The Perron Institute.

ParkWest: Principal investigators: Guido Alves (Norwegian Centre for Movement Disorders, Stavanger University Hospital), Ole-Bjørn Tysnes (Haukeland University Hospital). Investigators and study coordinators: Karen Herlofson, Solgunn Ongre, Siri Bruun (Sørlandet Hospital Arendal); Ineke HogenEsch, Marianne Kjerandsen, Liv Kari Håland (Haugesund Hospital); Wenche Telstad, Aliaksei Labusau, Jane Kastet (Førde Hospital); Bernd Müller, Geir Olve Skeie, Charalampos Tzoulis (Haukeland University Hospital); Kenn Freddy Pedersen, Michaela Dreetz Gjerstad, Elin Bjelland Forsaa, Jodi Maple-Grødem, Johannes Lange, Veslemøy Hamre Frantzen, Anita Laugaland, Karen Simonsen, Ingvild Dalen (Stavanger University Hospital). The ParkWest study has received funding from the Western Norway Regional Health Authority (grant number 911218), and the Norwegian Parkinson's Disease Association, and the Research Council of Norway (grant number 287842).

PIB funding: NIH grants: NS075321 and NS097437, the American Parkinson Disease Association (APDA), the Greater St. Louis Chapter of the APDA, the Barnes Jewish Hospital Foundation (Elliot Stein Family Fund, Parkinson disease research fund).

Supplementary Methods

Study participants

4,491 patients with PD (with available genotyping data and quality control) were longitudinally assessed with 3,3406 study visits in 15 cohorts from North America and Europe between 1986 and 2017: Harvard Biomarkers Study (HBS)[1], Neuroprotection Exploratory Trials in PD- Long term Study-1 (NET-PD LS1)[2], Drug Interaction with Genes in PD (DIGPD)[3] , PROfiling PARKinson's disease (PROPARK) study[4], PROPARK-Cross sectional cohort; Cambridgeshire Parkinson's Incidence from GP to Neurologist (CamPaIGN)[5-7]; Parkinsonism: Incidence, Cognition and Non-motor heterogeneity in Cambridgeshire (PICNICS)[8]; Parkinson's Disease Biomarkers Program (PDBP)[9]; Banner Health study(Arizona Study of Aging/Brain and Body Donation Program)[10]; ParkWest[11] and PIB[12]; Deprenyl and Tocopherol Antioxidative Therapy of Parkinsonism (DATATOP)[13]; Parkinson Research Examination of CEP-1347 Trial/A Longitudinal Follow-up of the PRECEPT Study Cohort (PreCEPT/PostCEPT)[14] and Tartu[15], Parkinson's Progression Markers Initiative (PPMI)[16] . For PPMI, approval was obtained to download and analyze the publicly accessible WGS and clinical data. 13 cohorts enrolled patients with a diagnosis of PD established according to modified UK PD Society Brain Bank diagnostic criteria as previously reported[1-5,8,9,11,12,14-18]. In DATATOP, the eligibility criteria required a clinical diagnosis of early, idiopathic PD (HY stages 1 or 2) with patients not on anti-parkinsonian medications[17]. Banner Health study: all subjects have come to autopsy and have had full neuropathological examinations with diagnosis[10]. Diagnostic certainty was increased by confirming the clinical diagnosis of PD during longitudinal follow-up visits[19] in all cohorts.

Serial Mini Mental State Exam (MMSE) scores[20] were longitudinally collected in 10 cohorts. Montreal Cognitive Assessment (MoCA)[21] scores were collected in the PDBP, PPMI study and converted to MMSE scores according to a published formula[22]. SCOPA-COG were collected in PROPARK, PROPARK-C and NET-PD LS1 cohort and converted to MMSE scores.

Polymorphism identification and haplogroup classification

The genotyping data of the 4,491 subjects with Parkinson's disease were reported in Ref.[23]. Briefly, the samples (excluded PPMI with WGS) were genotyped with Illumina Multi-Ethnic Genotyping Array (MEGA, Illumina), which includes 810 SNP markers in mtDNA after quality control as described in Ref.[23] 810 mtDNA variants were converted from "plink" format to "vcf" format according to the rCRS reference alleles. We removed 25 mismatched SNPs and InDel SNPs, 11 duplicated SNP probes on the array, and 11 variants with highly discordant MAF (> 5%) compared to Phase 1 and 3 of the 1000 Genomes Project[24] mitochondrial variants (n = 503 European) as called by the MToolBox pipeline[25]. The remained 763 SNPs were used to predict mitochondrial haplogroups using Haplogrep2.0[26] with default parameter using rCRS reference. Haplotype quality-control was performed according to the Haplogrep2 instruction and 44 subjects with quality score < 0.8 were excluded (**Supplementary Fig. 1**). We next simplified the sub-haplogroups (455 sub-haplogroups) to the 34 haplogroups (**Supplementary Table 1**) according to the mtDNA tree http://www.phylotree.org/tree/index.htm. 4,447 subjects were successfully assigned haplogroup, and 24 of these patients had no clinical records of note so were removed from the analysis. We further removed the haplogroups with less than 100 subjects (**Supplementary Table 1**, total 359 subjects), and the remaining 4,064 subjects with 30,515 study visits were used for haplogroup analysis (including H, HV* (excluding H, V), I, J, K, T, U#

(excluding K) haplogroups). It should be note that in our work, U[#] denotes all U haplogroups (U1, U2, U3, U4, U5, U6, U7…) but not haplogroup K as in Ref[27]. Out of 763 SNPs, 102 SNPs with allele frequency > 1% were used for single SNP Cox regression analysis. Notably, common mtDNA haplogroup or mtDNA variants are often ancient and are usually homoplasmic[28]. We did not analyze heteroplasmic mtDNA mutations in this study.

Statistical analysis

The Cox proportional hazards statistic was used to estimate the influence of different mitochondrial haplogroups on time (years from onset of PD) to reaching the endpoint of motor disability with postural instability (Hoehn and Yahr stage HY 3) or global cognitive impairment (GCI) as indicated by a MMSE ≤ 25 according to the recommendation the International Parkinson and Movement Disorder Society (MDS) Task Force[29] as in Ref.[30]. For HY analysis, age at onset of PD, sex and *GBA* carrier status were included as covariates. In the GCI analysis, age at onset of PD, sex, years of education, and polygenic hazard score (PHS including *GBA* carrier status, *APOE* ε4 allele haplotype and three novel progression variants rs182987047, rs138073281 and rs8050111 from Ref.[23]) were included as covariates in the Cox analyses. A "cohort" term was included as a random effect (a random effects Cox model is often termed a "frailty" model). 29,115 (95.4 %) of the visits from 4,088 patients with PD occurred within 12 years of longitudinal follow-up from disease onset with a median follow-up time of 6.7 years (inter-quartile range, 4.2 years), thus we focused our survival analyses on the 12-year time frame from disease onset. Patients were left-censored and those with missing or non-quality clinical data were excluded. Cox proportional hazards analyses were performed using the coxph function in the Survival package (v2.38-1)[31] and the "breslow" method was used for handling observations that have tied

survival times in the analysis and *P* values less than or equal to 0.05 were considered as indicative of haplogroup significance.

For single polymorphism variants analysis, we used a similar Cox proportional hazards regression model (same co-variants as mentioned above) to investigate each SNP effect on motor and cognitive impairment. Bonferroni correction was performed using p.adjust function with "bonferroni" in R.

Generalized longitudinal mixed fixed and random effects analysis (LMM)[32] of cognitive decline was performed with serial Mini Mental State Exam (MMSE) scores longitudinally assessed at varying times (enrollment visit and multiple longitudinal follow-up visits) in the combined data set. Two cohorts (PROPARK-C and Tartu) were excluded from the LMM because no longitudinal MMSE scores were available. The MMSE score was the dependent variable and the primary predictors were mitochondrial haplogroup status, time in the study (years), and their interaction. An intercept term and linear rate of change across time per subject were the random terms (permitted to be correlated). Subject level fixed covariates were age at baseline, sex, years of education, duration of PD illness at baseline, as well as PHS score. A study term was included as a random effect. This analysis was performed using the lme4 package (v1.1-23). All analyses were conducted in the R statistical environment version 4.0.2.

Comparison of models

The original multivariable Cox model from a previous study[33] included age at Parkinson's disease onset, years of education, sex, MMSE at enrolment, MDS-UPDRS III score at enrolment, depression at enrolment and *GBA* carrier status, and a cohort term was included as a random effect (using a frailty Cox model). 2,629 patients in the original nine longitudinal cohorts with available

mitochondrial variants, and 2,376 patients (253 left censored patients were removed) with 22,617 visits within 12 years of longitudinal follow-up from disease onset were used for comparison of different Cox regression genetic models (*GBA* carrier, *APOE* ε4, m.2706A>G, m.14766C>T), adjusting by age at Parkinson's disease onset, years of education, sex, MMSE at enrolment, MDS-UPDRS III score at enrolment, depression at enrolment, and a cohort term was included as a random effect.

Combination analysis of two genetic risk (*GBA* carrier and m.2706A>G variant/*APOE* ε4 and m.2706A>G variant) was performed using 2,376 patients from the Cox regression model, adjusting by the same six clinical predictors as mentioned bove, and a cohort term was included as a random effect.

Supplementary Table 1 Classification of mitochondrial haplogroups in patients with PD across the 15 cohorts

Haplogroups	Number	Sub-haplogroups of H	Number
H	1,829		
U#	666	H1	599
T	440	H2	599
J	427	H3	155
K	393	H5	146
HV*	218	H6	94
I	115	H4	75
W	93	H13	47
X	75	H46	20
N1	65	H15	11
V	42	H26	11
R0[&]	17	H7	10
L2	11	H41	8
M1	9	H14	6
D	7	H79	6
A	5	H85	6
C	5	H28	5
L1	4	H44	5
R1	4	H24	4
L3e	3	H100	3
B	2	H22	3
L3b	2	H81	3
M9	2	H56	2
N2	2	H94	2
N3	2	H17	1
G	1	H30	1
L0	1	H33	1
M7	1	H34	1
M30	1	H42	1
M33	1	H49	1
M49	1	H50	1
N9	1	H73	1
Y	1	H77	1
Z	1		

Haplogroups according to the mtDNA tree http://www.phylotree.org/tree/index.htm. 4,447 subjects were successfully assigned haplogroup. HV*: not including H, HV; U[#]: including all U haplogroups (U1, U2, U3, U4, U5, U6, U7…) but not haplogroup K as in Ref[27]; R0[&]: not including HV, H, V.

Supplementary Table 2 Clinical characteristics of patients with PD at enrollment with different macro-haplogroups across the 15 cohorts

***n* = 4,423**	**Marco-haplogroup**								***P*[&]**
	H	**HV***	**I**	**J**	**K**	**T**	**U[#]**	**Others**	
Number of men (*n*, %)	1,819 (63.2)	217 (65.4)	115 (57.4)	427 (61.8)	391 (65.7)	435 (62.1)	660 (63.8)	359 (62.1)	0.76
Age at onset, mean (SD), years	60.6 (10.6)	61.4 (10.1)	61.0 (10.8)	61.6 (10.9)	61.1 (9.6)	60.9 (10.5)	60.7 (11.0)	61.3 (10.1)	0.45
Age at enrollment, mean (SD), years	64.2 (10.2)	64.8 (9.8)	64.5 (10.5)	65.1 (10.4)	64.2 (9.3)	64.4 (10.1)	64.3 (10.5)	64.8 (10.3)	0.62
Years of education, mean (SD), years	14.2 (3.8)	14.6 (3.7)	14.1 (4.1)	14.4 (3.9)	14.6 (3.8)	14.1 (3.9)	14.2 (3.6)	14.7 (3.6)	0.08
Study years, mean (range), years	3.8 (0-19.9)	3.7 (0-9.3)	3.9 (0-8.3)	3.6 (0-13.1)	3.8 (0-14.5)	3.6 (0-13.5)	3.6 (0-12.6)	3.5 (0-12.3)	0.26
Hoehn and Yahr, mean (SD)	1.9 (0.8)	1.9 (0.7)	1.8 (0.6)	1.9 (0.8)	1.9 (0.7)	2.0 (0.7)	2.0 (0.7)	1.9 (0.7)	0.28
MDS-UPDRS III, mean (SD)	28.4 (14.2)	27.2 (13.5)	27.4 (13.7)	27.9 (14.5)	26.9 (13.0)	28.4 (13.7)	28.4 (14.0)	27.9 (14.9)	0.72
MMSE, mean(SD)	28.2 (2.2)	28.2 (1.9)	28.5 (1.5)	28.1 (2.3)	28.3 (2.0)	28.2 (2.1)	28.2 (2.2)	28.3 (2.4)	0.47
LED, mean(SD)	436.6 (439.9)	402.3 (470.0)	428.6 (458.2)	415.2 (428.0)	373.6 (398.2)	399.9 (416.8)	433.2 (446.0)	418.7 (447.7)	0.34

24 subjects have no available clinic data, the table showed clinical characteristics of 4,423 patients with PD.

[&] Fisher exact test was used for the number of men in each group. Group comparisons were performed using Kruskal-Wallis test for age at onset, age at enrollment, years of education, study years, HY, MDS-UPDRS III, MMSE, LED.

HV*: The sub-haplogroups of haplogroup HV, not including haplogroup H, V.

U[#]: The sub-haplogroups of haplogroup U, not including haplogroup K.

Supplementary Table 3 The percentage (%) of different mitochondrial haplogroups in European population from literatures

Haplogroup	Latvia[34] n=299	Spain[35] n=312	Portugal[36] n=241	France[37] n=210	Norway[34] n=397	Czech[38] n=300	Germany[34] n=333	Iceland[27] n=467	Italy[39] n=124
H	44.5	42.3	40.7	41.9	45.1	40.7	47.7	47.6	41.1
HV*	2.3	NA	NA	NA	0.3	2.7	0.6	NA	1.6
I	4.3	1.6	0.8	2.9	2.3	2.0	1.8	4.7	NA
J	6.4	6.7	6.6	5.2	12.6	8.3	8.4	14.1	4.8
K	2.3	4.8	5.4	11.4	5.0	4.0	7.5	7.7	1.6
T	9.4	8.3	10.8	11.9	9.8	8.0	9.0	10.1	8.1
U[#]	23.1	16.0	17.4	17.6	16.9	21.3	13.5	11.8	30.6

HV*: The sub-haplogroups of haplogroup HV, not including haplogroup H, V;
H: sum of available sub-haplogroups of H;
J: sum of available sub-haplogroups of J;
K: sum of available sub-haplogroups of K;
T: sum of available sub-haplogroups of T;
U[#]: sum of available sub-haplogroups of haplogroup U, but not including haplogroup K.

Supplementary Table 4 Test for residual heterogeneity for each haplogroup compared to haplogroups of H in GCI combined analysis

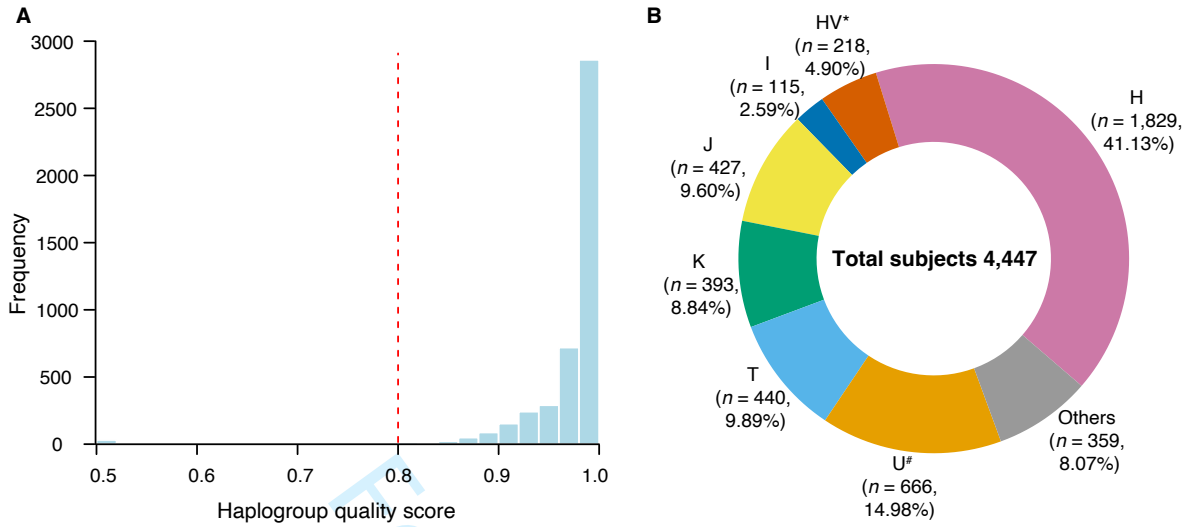
Haplogroups (H as reference)	Heterogeneity Q	*P* value[&]	I^2
HV*	5.32	0.87	0%
I	12.59	0.32	12.61%
J	4.88	0.96	0%
K	7.54	0.82	0%
T	5.05	0.96	0%
U[#]	15.90	0.20	24.54%

HV*: The sub-haplogroups of haplogroup HV, not including haplogroup H, V.
U[#]: The sub-haplogroups of haplogroup U, not including haplogroup K.
[&]The Cochran's Q-test was used to test for residual heterogeneity across studies via R metafor package (version 2.4-0). I^2 index (100%×(Q-df)/Q) was used to quantify the degree of heterogeneity.

Supplementary Table 5 The association of mitochondrial haplogroups in Alzheimer disease

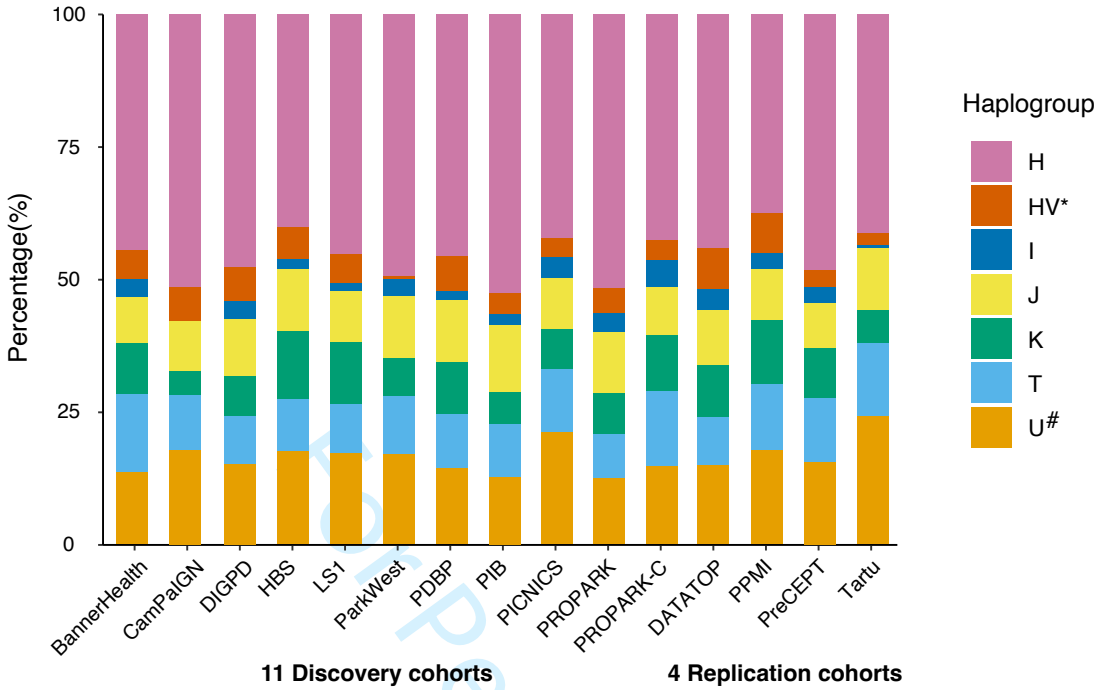
Haplogroup	Years	Effect	Ethnicity	Dataset size (case/control)	Dataset type
H	2009[40]	Risk	Poland	222/252	Whole mitochondrial genomics
	2011[41]	Risk	Caucasian	422/318	Control_region position (16624-576) + 9 coding SNPs
HV	2009[40]	Risk	Poland	222/252	Whole mitochondrial genomics
	2011[41]	Risk	Caucasian	422/318	Control_region position (16624-576) + 9 coding SNPs
K	2001[42]	Protective	Italian	213/389	10 restricted sites
	2011[41]	Protective	Caucasian	422/318	Control_region position (16624-576) + 9 coding SNPs
	2020[43]	Protective	American	309/507	Whole mitochondrial genomics
K1A1B	2013[44]	Risk	Caucasian	154/175	138SNPS
J	2009[40]	Protective in males	Poland	222/252	Whole mitochondrial genomics
	2020[43]	Risk	American	309/507	Whole mitochondrial genomics
T	2011[41]	Protective in females	Caucasian	422/318	Control_region position (16624-576) + 9 coding SNPs
JT	2011[41]	Protective in females	Caucasian	422/318	Control_region position (16624-576) + 9 coding SNPs
U	2001[42]	Protective	Italian	213/389	10 restricted sites
	2004[45]	Risk in males, protective in females	Caucasian	989/328	10 SNPs

Supplementary Figure 1 The classification of haplogroup in patients with PD across 15 cohorts.



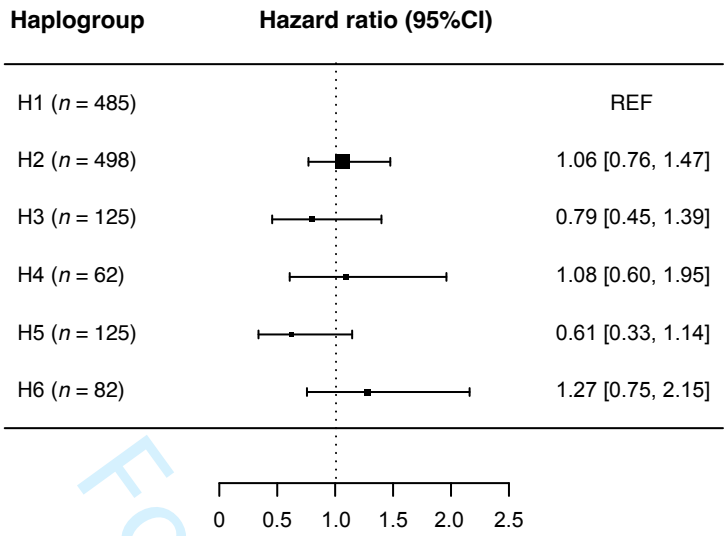
A The haplogroup quality score of 4,491 patients with PD was evaluated from HaploGrep2.0[26] based on Kulczynski measure: (HaplogroupWeight + SampleWeight) × 0.5. The HaploGrep2.0 applied this formula to all haplogroups in Phylotree and returned the overall best hit and the score represented its haplogroup quality. The quality of 0.8 as cutoff was recommended and 4,447 subjects were successfully assigned mitochondrial haplogroup. **B** The donut plot presents the proportion of patients with PD within diverse mitochondrial macro-haplogroups

Supplementary Figure 2 The stacking diagram for distribution of seven macro-haplogroups in patients with PD across 15 cohorts



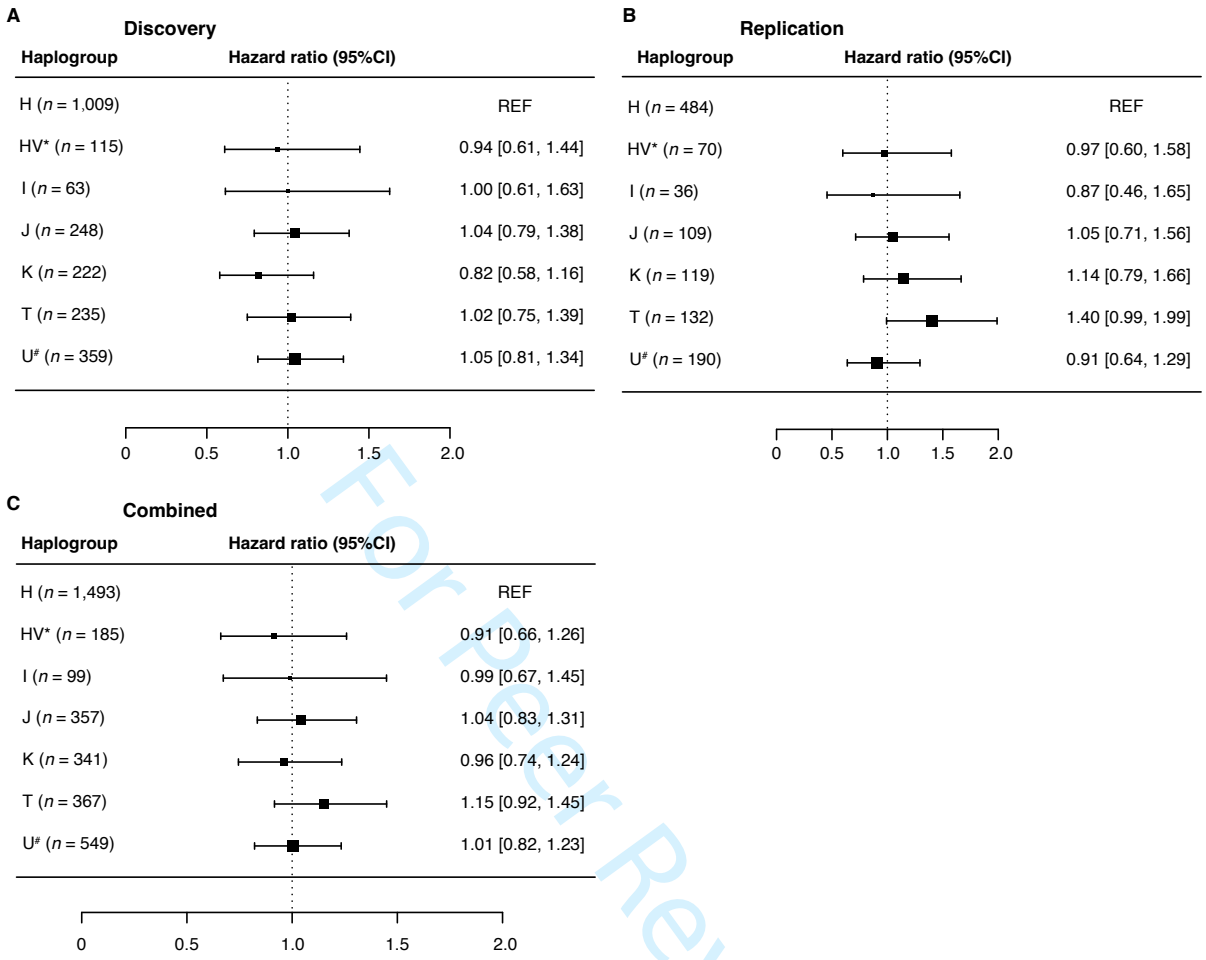
Each vertical bar corresponds to one cohort and consists of 7 sub-bars representing the proportions of the 7 macro-haplogroups H, HV*, I, J, K, T and U# in relevant cohort. There was no any difference in the proportion of seven macro-haplogroups in 15 cohorts ($P \approx 1$, Fisher exact test).

Supplementary Figure 3 Patients with PD with major sub-haplogroups of H have similar risk of progression to global cognitive impairment



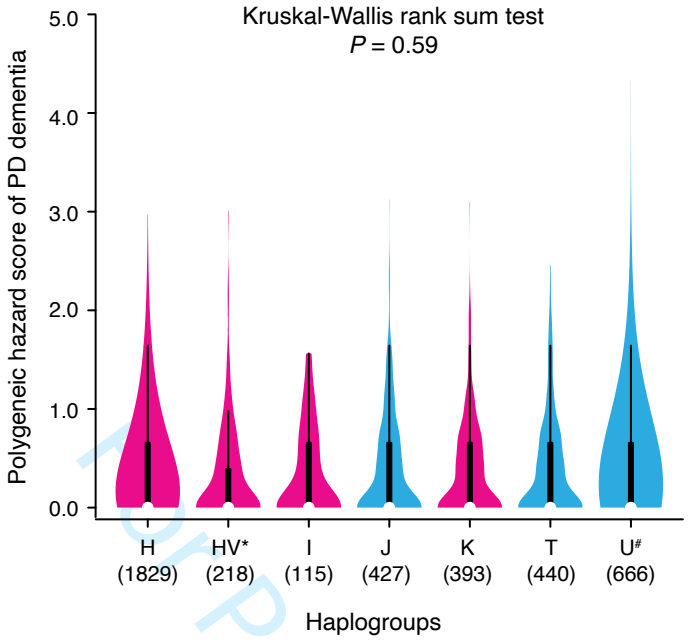
Cox regression analysis did not show any different hazard ratio (HR) to develop global cognitive impairment (MMSE ≤ 25) in combined population, according to the recommendation of the International Parkinson and Movement Disorder Society (MDS) Taks Force[29], among patients with PD in six major sub-haplogroups of H.

Supplementary Figure 4 Patients with PD in seven macro-haplogroups have similar risk of progression to HY3.



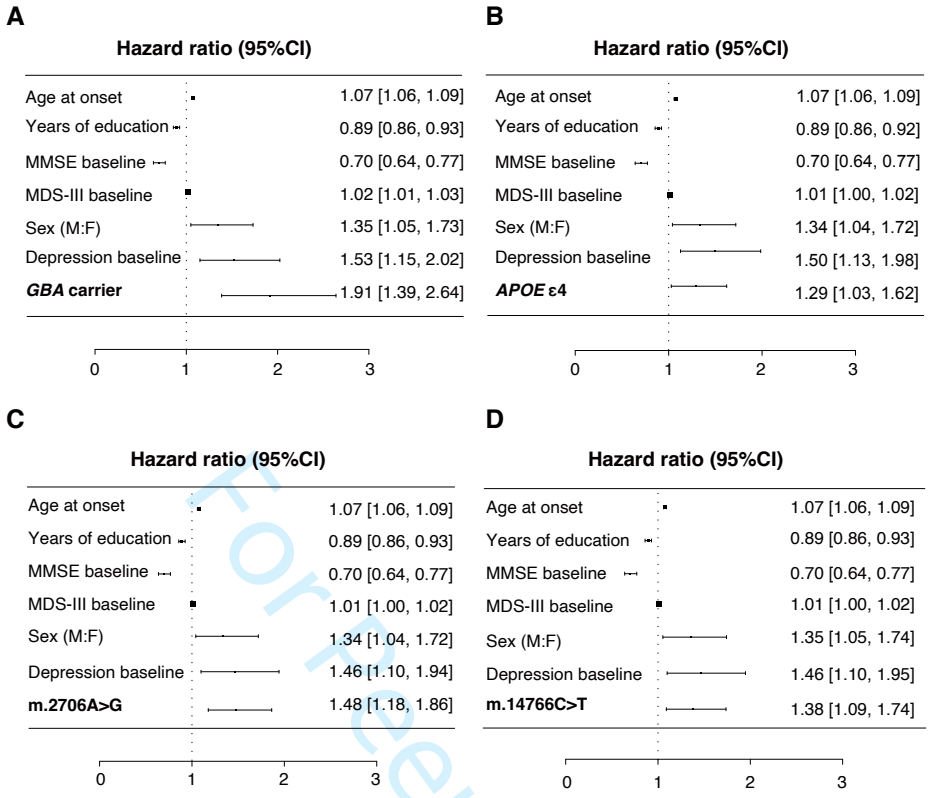
Cox regression analysis did not show any difference in hazard ratio (HR) for development of motor disability with postural instability (Hoehn & Yahr stage 3) during the progression of disease in seven macro-haplogroups from (**A**) discovery, (**B**) replication and (**C**) combined population.

Supplementary Figure 5 Patients with PD in seven mitochondrial macro-haplogroups have similar polygenic hazard scores



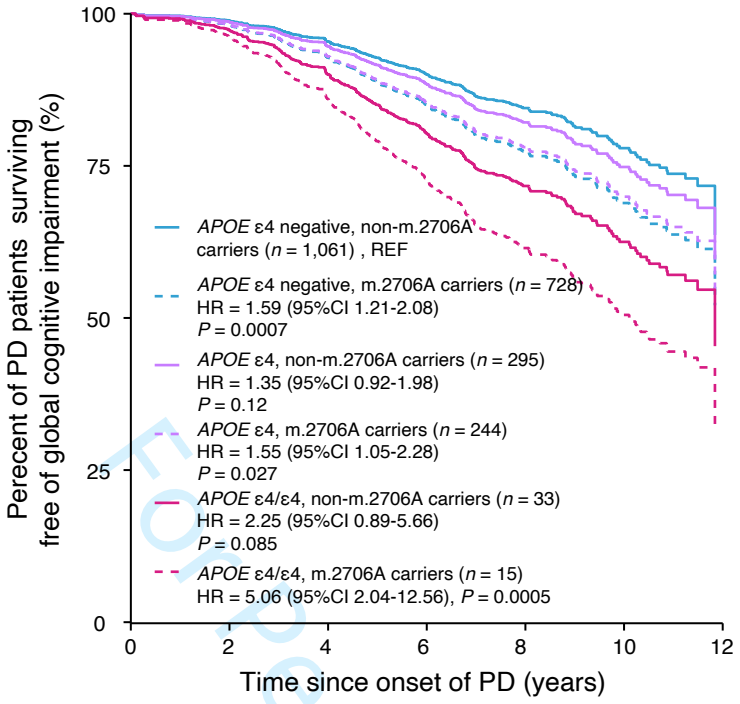
Violin-plot showed no significant difference between Polygenic hazard score to develop PD dementia among seven macro-haplogroups in combined population. Violin plot is a mixed of a box plot and a kernel density plot: the white dot represents the median, and black bar represents the interquartile range of score, the thin black line represents the rest of distribution and each side of the line is a kernel density estimation.

Supplementary Figure 6 The exploratory analysis for global cognitive impairment models with different genetic factors.



The forest plots show hazard ratios (Methods) for global cognitive impairment (GCI) in different genetic models (**A**) *GBA* carrier, (**B**) *APOE* ε4, (**C**) m.2706A>G and (**D**) m.14766C>T with the same six clinical risk factors. The squares represent point estimates, with the height of the square inversely proportional to the standard error of the estimates. The horizontal lines indicate 95% confidence intervals of the estimates.

Supplementary Figure 7 Exploratory analysis for global cognitive impairment models with *APOE* ε4 and m.2706A>G.



Covariate-adjusted survival curves for patients with PD stratified into six subgroups: *APOE* ε4 negative and non-m.2706A carriers (n = 1,061), *APOE* ε4 negative and m.2706A carriers (n = 728), *APOE* ε4 heterozygotes and non-m.2706A carriers (n = 295), *APOE* ε4 heterozygotes and m.2706A carriers (n = 244), *APOE* ε4/ε4 and non-m.2706A carriers (n = 33), *APOE* ε4/ε4 and m.2706A carriers (n = 15).

References

1. Locascio JJ, Eberly S, Liao Z, *et al*. Association between alpha-synuclein blood transcripts and early, neuroimaging-supported Parkinson's disease. *Brain*. Sep 2015;138(Pt 9):2659-71.
2. Writing Group for the NETiPDI, Kieburtz K, Tilley BC, *et al*. Effect of creatine monohydrate on clinical progression in patients with Parkinson disease: a randomized clinical trial. *Jama*. Feb 10 2015;313(6):584-93.
3. Corvol JC, Artaud F, Cormier-Dequaire F, *et al*. Longitudinal analysis of impulse control disorders in Parkinson disease. *Neurology*. Jul 17 2018;91(3):e189-e201.
4. Verbaan D, Marinus J, Visser M, van Rooden SM, Stiggelbout AM, van Hilten JJ. Patient-reported autonomic symptoms in Parkinson disease. *Neurology*. Jul 24 2007;69(4):333-41.
5. Williams-Gray CH, Evans JR, Goris A, *et al*. The distinct cognitive syndromes of Parkinson's disease: 5 year follow-up of the CamPaIGN cohort. *Brain*. Nov 2009;132(Pt 11):2958-69.
6. Williams-Gray CH, Mason SL, Evans JR, *et al*. The CamPaIGN study of Parkinson's disease: 10-year outlook in an incident population-based cohort. *Journal of neurology, neurosurgery, and psychiatry*. Nov 2013;84(11):1258-64.
7. Winder-Rhodes SE, Evans JR, Ban M, *et al*. Glucocerebrosidase mutations influence the natural history of Parkinson's disease in a community-based incident cohort. Research Support, Non-U.S. Gov't. *Brain*. Feb 2013;136(Pt 2):392-9.
8. Breen DP, Evans JR, Farrell K, Brayne C, Barker RA. Determinants of delayed diagnosis in Parkinson's disease. *Journal of neurology*. Aug 2013;260(8):1978-81.
9. Rosenthal LS, Drake D, Alcalay RN, *et al*. The NINDS Parkinson's disease biomarkers program. *Mov Disord*. Jun 2016;31(6):915-23.
10. Beach TG, Adler CH, Sue LI, *et al*. Arizona Study of Aging and Neurodegenerative Disorders and Brain and Body Donation Program. *Neuropathology : official journal of the Japanese Society of Neuropathology*. Aug 2015;35(4):354-89.
11. Alves G, Muller B, Herlofson K, *et al*. Incidence of Parkinson's disease in Norway: the Norwegian ParkWest study. *Journal of neurology, neurosurgery, and psychiatry*. Aug 2009;80(8):851-7.
12. Lucero C, Campbell MC, Flores H, Maiti B, Perlmutter JS, Foster ER. Cognitive reserve and beta-amyloid pathology in Parkinson disease. *Parkinsonism & related disorders*. Aug 2015;21(8):899-904.
13. Jankovic J, McDermott M, Carter J, *et al*. Variable expression of Parkinson's disease: a base-line analysis of the DATATOP cohort. The Parkinson Study Group. *Neurology*. Oct 1990;40(10):1529-34.
14. Ravina B, Tanner C, Dieuliis D, *et al*. A longitudinal program for biomarker development in Parkinson's disease: a feasibility study. *Mov Disord*. Oct 30 2009;24(14):2081-90.
15. Kadastik-Eerme L, Taba N, Asser T, Taba P. The increasing prevalence of Parkinson's disease in Estonia. *Acta Neurol Scand*. Sep 2018;138(3):251-258.
16. Parkinson Progression Marker I. The Parkinson Progression Marker Initiative (PPMI).

Progress in neurobiology. Dec 2011;95(4):629-35.

17. DATATOP: a multicenter controlled clinical trial in early Parkinson's disease. Parkinson Study Group. Clinical Trial
Comparative Study
Multicenter Study
Randomized Controlled Trial
Research Support, Non-U.S. Gov't
Research Support, U.S. Gov't, P.H.S. *Archives of neurology*. Oct 1989;46(10):1052-60.

18. Nalls MA, Pankratz N, Lill CM, *et al*. Large-scale meta-analysis of genome-wide association data identifies six new risk loci for Parkinson's disease. *Nature genetics*. Sep 2014;46(9):989-93.

19. Hughes AJ, Daniel SE, Ben-Shlomo Y, Lees AJ. The accuracy of diagnosis of parkinsonian syndromes in a specialist movement disorder service. *Brain*. Apr 2002;125(Pt 4):861-70.

20. Goetz CG, Poewe W, Rascol O, *et al*. Movement Disorder Society Task Force report on the Hoehn and Yahr staging scale: status and recommendations. *Mov Disord*. Sep 2004;19(9):1020-8.

21. Hoops S, Nazem S, Siderowf AD, *et al*. Validity of the MoCA and MMSE in the detection of MCI and dementia in Parkinson disease. *Neurology*. Nov 24 2009;73(21):1738-45.

22. van Steenoven I, Aarsland D, Hurtig H, *et al*. Conversion between mini-mental state examination, montreal cognitive assessment, and dementia rating scale-2 scores in Parkinson's disease. *Mov Disord*. Dec 2014;29(14):1809-15.

23. Liu G, Peng J, Liao Z, *et al*. Genome-wide survival study identifies a novel synaptic locus and polygenic score for cognitive progression in Parkinson's disease. *Nature genetics*. May 6 2021;53(6):787-793.

24. Genomes Project C, Auton A, Brooks LD, *et al*. A global reference for human genetic variation. *Nature*. Oct 1 2015;526(7571):68-74.

25. Calabrese C, Simone D, Diroma MA, *et al*. MToolBox: a highly automated pipeline for heteroplasmy annotation and prioritization analysis of human mitochondrial variants in high-throughput sequencing. *Bioinformatics*. Nov 1 2014;30(21):3115-7.

26. Weissensteiner H, Pacher D, Kloss-Brandstatter A, *et al*. HaploGrep 2: mitochondrial haplogroup classification in the era of high-throughput sequencing. *Nucleic Acids Res*. Jul 8 2016;44(W1):W58-63.

27. Helgason A, Hickey E, Goodacre S, *et al*. mtDNA and the islands of the North Atlantic: estimating the proportions of Norse and Gaelic ancestry. *Am J Hum Genet*. Mar 2001;68(3):723-37.

28. Stewart JB, Chinnery PF. The dynamics of mitochondrial DNA heteroplasmy: implications for human health and disease. *Nat Rev Genet*. Sep 2015;16(9):530-42.

29. Dubois B, Burn D, Goetz C, *et al*. Diagnostic procedures for Parkinson's disease dementia: recommendations from the movement disorder society task force. *Mov Disord*. Dec 2007;22(16):2314-24.

30. Liu G, Boot B, Locascio JJ, *et al*. Specifically neuropathic Gaucher's mutations accelerate cognitive decline in Parkinson's. *Annals of neurology*. Nov 2016;80(5):674-685.

31. Therneau TM. *A Package for Survival Analysis in S.* 2014.

32. Locascio JJ, Atri A. An overview of longitudinal data analysis methods for neurological research. *Dementia and geriatric cognitive disorders extra*. Jan 2011;1(1):330-57.

33. Liu G, Locascio JJ, Corvol JC, *et al*. Prediction of cognition in Parkinson's disease with a clinical-genetic score: a longitudinal analysis of nine cohorts. *The Lancet Neurology*. Aug 2017;16(8):620-629.

34. Pliss L, Tambets K, Loogvali EL, *et al*. Mitochondrial DNA portrait of Latvians: towards the understanding of the genetic structure of Baltic-speaking populations. *Ann Hum Genet*. Jul 2006;70(Pt 4):439-58.

35. Alvarez JC, Johnson DL, Lorente JA, *et al*. Characterization of human control region sequences for Spanish individuals in a forensic mtDNA data set. *Leg Med (Tokyo)*. Nov 2007;9(6):293-304.

36. Pereira L, Prata MJ, Amorim A. Diversity of mtDNA lineages in Portugal: not a genetic edge of European variation. *Ann Hum Genet*. Nov 2000;64(Pt 6):491-506.

37. Dubut V, Chollet L, Murail P, *et al*. mtDNA polymorphisms in five French groups: importance of regional sampling. *Eur J Hum Genet*. Apr 2004;12(4):293-300.

38. Vidrová V, Tesařová M, Trefilová E, Honzík T, Magner M, Zeman J. Mitochondrial DNA Haplogroups in the Czech Population Compared to Other European Countries. *Human Biology*. 2008;80(6):669-674.

39. Messina F, Scorrano G, Labarga CM, Rolfo MF, Rickards O. Mitochondrial DNA variation in an isolated area of Central Italy. *Ann Hum Biol*. Jun 2010;37(3):385-402.

40. Maruszak A, Canter JA, Styczynska M, Zekanowski C, Barcikowska M. Mitochondrial haplogroup H and Alzheimer's disease--is there a connection? *Neurobiol Aging*. Nov 2009;30(11):1749-55.

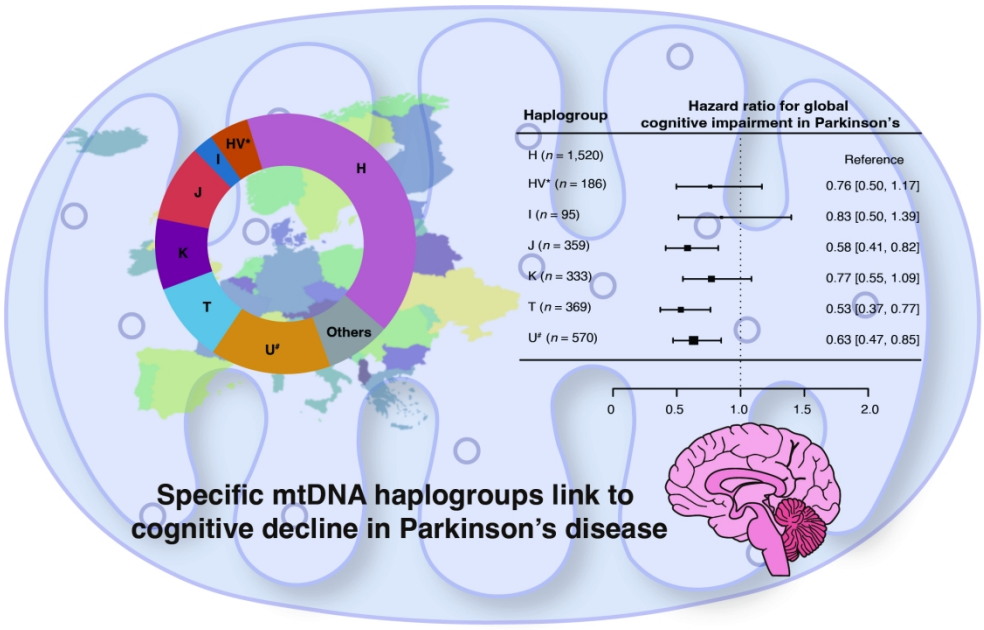
41. Maruszak A, Safranow K, Branicki W, *et al*. The impact of mitochondrial and nuclear DNA variants on late-onset Alzheimer's disease risk. *J Alzheimers Dis*. 2011;27(1):197-210.

42. Carrieri G, Bonafe M, De Luca M, *et al*. Mitochondrial DNA haplogroups and APOE4 allele are non-independent variables in sporadic Alzheimer's disease. *Hum Genet*. Mar 2001;108(3):194-8.

43. Swerdlow RH, Hui D, Chalise P, *et al*. Exploratory analysis of mtDNA haplogroups in two Alzheimer's longitudinal cohorts. *Alzheimer's & Dementia*. 2020;16(8):1164-1172.

44. Ridge PG, Koop A, Maxwell TJ, *et al*. Mitochondrial haplotypes associated with biomarkers for Alzheimer's disease. *PLoS One*. 2013;8(9):e74158.

45. van der Walt JM, Dementieva YA, Martin ER, *et al*. Analysis of European mitochondrial haplogroups with Alzheimer disease risk. *Neuroscience Letters*. 2004;365(1):28-32.



175x113mm (300 x 300 DPI)

STREGA statement: Reporting guidelines checklist for genetic association studies

SECTION	ITEM NUMBER	CHECKLIST ITEM	REPORTED ON PAGE NUMBER:
TITLE AND ABSTRACT			
	1a	Indicate the study's design with a commonly used term in the title or the abstract	Page 1-2
	1b	Provide in the abstract an informative and balanced summary of what was done and what was found	Page 1-2
INTRODUCTION			
Background and objectives	2	Explain the scientific background and rationale for the investigation being reported	Page 5
	3	State specific objectives, including any pre-specified hypotheses. State if the study is the first report of a genetic association, a replication effort, or both	Page 6
METHODS			
Study design	4	Present key elements of study design early in the paper	Page 6-7
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	Page 6-7
Participants	6a	Give information on the criteria and methods for selection of subsets of participants from a larger study, when relevant	Page 7, Supplementary page 7
Variables	7	Clearly define genetic exposures (genetic variants) using a widely-used nomenclature system. Identify variables likely to be associated with population stratification (confounding by ethnic origin)	Page 7
Data sources/measurements	8*	Describe laboratory methods, including source and storage of DNA, genotyping methods and platforms (including the allele calling algorithm used, and its version), error rates and call rates. State the laboratory/centre where genotyping was done. Describe comparability of laboratory methods if there is more than one group. Specify whether genotypes were assigned using all of the data from the study simultaneously or in smaller batches	Page 7, Supplementary page 8
Bias	9	For quantitative outcome variables, specify if any investigation of potential bias resulting from pharmacotherapy was undertaken. If relevant, describe the nature and magnitude of	NA

SECTION	ITEM NUMBER	CHECKLIST ITEM	REPORTED ON PAGE NUMBER:
		the potential bias, and explain what approach was used to deal with this	
Study size	10	Explain how the study size was arrived at	Page 6-7, Supplementary page 7
Quantitative variables	11	If applicable, describe how effects of treatment were dealt with	NA
Statistical methods	12a	Describe all statistical methods, including those used to control for confounding. State software version used and options (or settings) chosen	Page 8, Supplementary page 9-10
	12b	Describe any methods used to examine subgroups and interactions	NA
	12c	Explain how missing data were addressed	Supplementary page 9
	12d	Cohort study—If applicable, explain how loss to follow-up was addressed Case-control study—If applicable, explain how matching of cases and controls was addressed Cross-sectional study—If applicable, describe analytical methods taking account of sampling strategy	Supplementary page 9
	12e	Describe any sensitivity analyses	NA
	12f	State whether Hardy–Weinberg equilibrium was considered and, if so, how	NA
	12g	Describe any methods used for inferring genotypes or haplotypes	Page 7, Supplementary page 8
	12h	Describe any methods used to assess or address population stratification	NA
	12i	Describe any methods used to address multiple comparisons or to control risk of false-positive findings	Supplementary page 10
	12j	Describe any methods used to address and correct for relatedness among subjects	NA
RESULTS			

SECTION	ITEM NUMBER	CHECKLIST ITEM	REPORTED ON PAGE NUMBER:
Participants	13a	Report numbers of individuals in whom genotyping was attempted and numbers of individuals in whom genotyping was successful	Page 9, Supplementary page 8
	13b	Give reasons for non-participation at each stage	Supplementary page 8
	13c	Consider use of a flow diagram	NA
Descriptive Data	14a	Give characteristics of study participants (e.g. demographic, clinical, social) and information on exposures and potential confounders. Consider giving information by genotype	Page 9-10
	14b	Indicate number of participants with missing data for each variable of interest	Supplementary page 8
	14c	Cohort study—Summarise follow-up time (eg, average and total amount)	Supplementary page 9
Outcome Data	15*	Cohort study— Report outcomes (phenotypes) for each genotype category over time Case-control study— Report numbers in each genotype category Cross-sectional study— Report outcomes (phenotypes) for each genotype category	Page 10-11
Main Results	16a	Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (e.g. 95% confidence interval). Make clear which confounders were adjusted for and why they were included	Page 10-11
	16b	Report category boundaries when continuous variables were categorized	Page 11
	16c	If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	NA
	16d	Report results of any adjustments for multiple comparisons	Page 12
Other Analyses	17a	Report other analyses done—e.g. analyses of subgroups and interactions, and sensitivity analyses	Page 13-14
	17b	If numerous genetic exposures (genetic variants) were examined, summarize results from all analyses undertaken	NA
	17c	If detailed results are available elsewhere, state how they can be accessed	NA
DISCUSSION			

SECTION	ITEM NUMBER	CHECKLIST ITEM	REPORTED ON PAGE NUMBER:
Key Results	18	Summarise key results with reference to study objectives	Page 15
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	Page 16-17
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	Page 16-17
Generalisability	21	Discuss the generalisability (external validity) of the study results Other information	Page 17
FUNDING			
	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	Page 18

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

Reference: von Elm E, Altman DG, Egger M, Pocock SJ, Gotzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: guidelines for reporting observational studies. This guideline was published simultaneously in 8 journals. Ann Intern Med. 2007; 147(8):573-577. PMID: 17938396, PLoS Med. 2007;4(10):e296. PMID: 17941714, BMJ. 2007;335(7624):806-808. PMID: 17947786, Prev Med. 2007;45(4):247-251. PMID: 17950122, Epidemiology. 2007;18(6):800-804. PMID: 18049194, Lancet. 2007;370(9596):1453-1457. PMID: 18064739 J Clin Epidemiol. 2008;61(4):344-349. PMID: 18313558, Bull World Health Organ. 2007;85(11):867-872. PMID: 18038077