TITLE

MAPT allele and haplotype frequencies in Nigerian Africans: population distribution and association with Parkinson's disease risk and age at onset.

Commented [OO1]: We may not necessarily need this since it is a co variate. Just link to PD risk would be fine.

Authors

Olaitan Okunoye, Oluwadamilola Ojo, **NPDR list**, Dena Hernandez, Sara Bandres-Ciga, Cornelius Blauwendraat, Andy Singleton, Henry Houlden, John Hardy, Mie Rizig, Njideka Okubadejo

Affiliations

Correspondence

Njideka U. Okubadejo

Neurology Unit, Department of Medicine, Faculty of Clinical Sciences, College of Medicine, University of Lagos, Lagos State, Nigeria

E-mail: nokubadejo@unilag.edu.ng

ORCID: 0000-0003-2975-8803

Key words

MAPT, tau haplotype, Nigeria, black ancestry, Africa, Parkinson's disease, population

Abstract

Background: The microtubule-associated protein tau (*MAPT*) gene is critical because of its putative role in the causal pathway of neurodegenerative diseases including Parkinson's disease (PD). However, there is a lack of clarity regarding the link between the main H1 haplotype and risk of PD. Methodological differences and ethnic genetic variability may underlie the inconsistent study results. Data on *MAPT* haplotype frequencies in the general population and association studies exploring the role of *MAPT* haplotypes in conferring PD risk in black Africans are lacking.

Objectives: To determine the frequencies of *MAPT* haplotypes and to explore whether the H1 haplotype is linked to PD risk and age at onset in Nigerian Africans.

Methods: The allelic and genotypic frequencies of *MAPT* rs1052553) were analysed using PCR-based KASP™ in 907 persons with PD and 1,022 age- and gender-matched healthy controls without PD from the Nigeria Parkinson's Disease Research network cohort. Clinical data related to PD included age at study, age at onset, and disease duration.

Results: The frequency of the main *MAPT* H1 haplotype in this cohort was 98.7% in persons with PD, and 99.1% in healthy controls (p=0.19), whereas the H2 haplotype was present in 1.3% of PD and 0.9% of controls (p=0.24). The most frequent *MAPT* genotype was H1H1 (97.5% of PD and 98.2% of controls). The H1 haplotype was not associated with PD risk after accounting for gender and age at onset (Odds ratio for H1/H1 vs H1/H2 and H2/H2: 0.68 (95%CI:0.39-1.28);p=0.23).

Commented [N2]:
Results: MAPT rs1052553 allelic and genotype frequencies did not differ significantly between relapsing bout onset MS patients and controls, and were unrelated with the age of onset of MS or gender.
Conclusions: These results suggest that MAPT rs1052553 polymorphism is not related with the risk for relapsing bout onset MS.

Commented [N3]:

Conclusions: Our findings support previous studies that report a low frequency of the *MAPT* H2 among Africans. Furthermore, in our cohort of black Africans with PD, the *MAPT* H1 haplotype was not associated with an increased risk or age at onset of PD.

Introduction

The microtubule-associated protein tau gene (*MAPT*) located on chromosome 17q21 encodes the tau protein which is primarily involved in modulating the stability of axonal microtubules within the brain. (Guo T Acta Neuropath 2017ef). *MAPT* has two distinct haplogroups (H1 and H2) resulting from inversion of an approximately 900kb region, and which, while having the same amino acid sequence, demonstrate differences in some SNPs that are capable of altering the expression (including transcription, post-transcriptional modification, and rates of translation). (ref: Zhang et al Mol Neurobiol 2016) H1 (the non-inverted sequence) is predominant, whereas the less common H2 (the inverted haplotype) is reported most frequently in Europeans and South East Asians, with a frequency of 23.5 – 37.5%. (ref Donnelly MP et al AJHG 2010). In general, the lowest rates in published literature are in sub-Saharan Africans (0.7% - 6.3%) and East Asians/Pacific Islanders (almost 0%). (ref Donnelly) Notably, the proportion of African samples included in these reports have been low, typically below 50 within any African ethnic group.

Mutations in *MAPT* are associated with neurodegenerative tauopathies including Alzheimer's disease (AD), progressive supranuclear palsy (PSP), and fronto-temporal dementia (FTD), but have also been implicated in Parkinson's disease (PD) which is predominantly a synucleinopathy. Whereas normal tau is integral to maintaining neuronal functioning, stabilizing and assembling microtubules, regulating axonal transport, promoting neurite growth, and preventing DNA damage and apoptosis, abnormal hyperphosphorylated tau (the product of *MAPT* mutations) promotes neurodegeneration. (ref Zhang CC Role of ___as above). Conformational change, increased tendency to misfold, dissociation from mictrotubules, aggregation into neurofibrillary tangles, and ultimate loss of normal function characterize the main pathological process in tauopathies. Clarifying the role of *MAPT* in neurodegenerative

Commented [N4]: Guo T, Noble W, Hanger DP. Roles of tau protein in health and disease. *Acta Neuropathol.* 2017;133(5):665-704. doi:10.1007/s00401-017-1707-9

Commented [N5]: Zhang CC, Xing A, Tan MS, Tan L, Yu JT. The Role of MAPT in Neurodegenerative Diseases: Genetics, Mechanisms and Therapy. *Mol Neurobiol.* 2016;53(7):4893-4904. doi:10.1007/s12035-015-9415-8

Commented [N6]: SE or SW? check paper

Commented [N7]: Donnelly MP, Paschou P, Grigorenko E, et al. The distribution and most recent common ancestor of the 17q21 inversion in humans. *Am J Hum Genet.* 2010;86(2):161-171. doi:10.1016/j.ajhg.2010.01.007

Commented [N8]: Zhang CC, Zhu JX, Wan Y, Tan L, Wang HF, Yu JT, Tan L. Meta-analysis of the association between variants in MAPT and neurodegenerative diseases. Oncotarget. 2017 Jul 4;8(27):44994-45007. doi: 10.18632/oncotarget.16690. PMID: 28402959; PMCID: PMC5546535.

Commented [N9]: Additional reference needed

Commented [N10]: Insert references

disorders has implications for developing neuroprotective interventions to reduce or inhibit tau phosphorylation, aggregation, and promote microtubule stabilization. The H1 haplotype has been postulated to contribute to an increased risk of PD, whereas H2, correlating with reduced tau protein expression, may be protective. (ref Guo T; Wade-Martins 2012, other refs A) Studies of the association between MAPT and risk of PD have shown different effect sizes (and suggestions of a weak effect), the possibility of interactions with other PD associated susceptibility loci, and ethnic variability even within populations of similar geographical origin [4, 5] (*and refs listed as Other refs A*) A highly significant association between PD and the H1/H1 haplotype was documented in the Genetic Epidemiology of Parkinson's Disease (GEO-PD) consortium study that included 5302 PD cases of Caucasian ancestry[6]. The effect was independent of any interaction with *SNCA,* applied to risk of PD, but with no significant effect on the age at onset of PD. (ref)

The main objectives of this study are to describe the population distribution of *MAPT* main haplotypes in Nigerian Africans, and thereby fill the void regarding the frequencies of the MAPT haplotypes in a large population of black sub-Saharan Africans. Secondly, we aim to determine if there is an association between *MAPT* haplotype variation and the risk of, and age at onset of PD in our population.

Commented [N11]: Wade-Martins R. Genetics: The MAPT locus-a genetic paradigm in disease susceptibility. *Nat Rev Neurol.* 2012;8(9):477-478. doi:10.1038/nrneurol.2012.169

Other refs A
- International Parkinson Disease Genomics Consortium, Nalls MA, Plagnol V, Hernandez DG, Sharma M, Sheerin UM, et al. Imputation of sequence variants for identification of genetic risks for Parkinson's disease: a meta-analysis of genome-wide association studies. Lancet 2011;377:641–9.
-Ezquerra M, Pastor P, Gaig C, Vidal-Taboada JM, Cruchaga C, Muñoz E, et al. Different MAPT haplotypes are associated with Parkinson's disease and progressive supranuclear palsy. Neurobiol Aging 2011;32. 547.e11-6.
- Mata IF, Yearout D, Alvarez V, Coto E, de Mena L, Ribacoba R, et al. Replication of MAPT and SNCA, but not PARK16-18, as susceptibility genes for Parkinson's disease. Mov Disord 2011;26:819–23.
- Elbaz A, Ross OA, Ioannidis JP, Soto-Ortolaza AI, Moisan F, Aasly J, et al. Independent and joint effects of the MAPT and SNCA genes in Parkinson disease. Ann Neurol 2011;69:778–92.

Commented [N12]: Reference (effect size) Vandrovcova J, Pittman AM, Malzer E, Abou-Sleiman PM, Lees AJ, Wood NW, de Silva R. Association of MAPT haplotype-tagging SNPs with sporadic Parkinson's disease. Neurobiol Aging. 2009 Sep;30(9):1477-82. doi: 10.1016/j.neurobiolaging.2007.11.019. Epub 2007 Dec 26. PMID: 18162161.

Commented [N13]: Huang Y, Wang G, Rowe D, Wang Y, Kwok JB, Xiao Q, Mastaglia F, Liu J, Chen SD, Halliday G. SNCA Gene, but Not MAPT, Influences Onset Age of Parkinson's Disease in Chinese and Australians. Biomed Res Int. 2015;2015:135674. doi: 10.1155/2015/135674. Epub 2015 Apr 15. PMID: 25960998; PMCID: PMC4413514.

Commented [N14]: Elbaz A, Ross OA, Ioannidis JP, Soto-Ortolaza AI, Moisan F, Aasly J, Annesi G, Bozi M, Brighina L, Chartier-Harlin MC, Destée A, Ferrarese C, Ferraris A, Gibson JM, Gispert S, Hadjigeorgiou GM, Jasinska-Myga B, Klein C, Krüger R, Lambert JC, Lohmann K, van de Loo S, Loriot MA, Lynch T, Mellick GD, Mutez E, Nilsson C, Opala G, Puschmann A, Quattrone A, Sharma M, Silburn PA, Stefanis L, Uitti RJ, Valente EM, Vilariño-Güell C, Wirdefeldt K, Wszolek ZK, Xiromerisiou G, Maraganore DM, Farrer MJ; Genetic Epidemiology of Parkinson's Disease (GEO-PD) Consortium. Independent and joint effects of the MAPT and SNCA genes in Parkinson disease. Ann Neurol. 2011 May;69(5):778-92. doi: 10.1002/ana.22321. Epub 2011 Mar 9. PMID: 21391235; PMCID: PMC3082599.

Methods

Study design and participant recruitment

Study participants in this cohort study were recruited by participating neurologists in the Nigeria Parkinson's Disease Research (NPDR) network as part of an ongoing cohort study conducted in collaboration with the International Parkinson's Disease Genomic Consortium-Africa (IPDGC-Africa). (ref NPDR and IPDGC Africa papers). The NPDR cohort is nationally representative and includes participating sites in all 6 geopolitical zones of Nigeria. Research ethics approval for the study was obtained from the the institutional health research ethics committees, the National Health Research Ethics Committee (NHREC) in Nigeria and the University College London (UCL) Institutional Review Board. All participants provided written informed consent. Persons with PD fulfilled the United Kingdom Parkinson's Disease Society Brain Bank (UKPDSBB) criteria (the only exception being that we did not exclude those with a first degree relative who also had PD). [8] (ref. Hughes, Daniel, Kilford and Lees, 1992). Controls were otherwise healthy volunteers with no known family history of PD, and otherwise healthy with no neurological condition, from the same population, and matched for age and gender. Baseline demographic characteristics (age at study, gender, age at onset of PD, and duration of PD (years)) were collected.

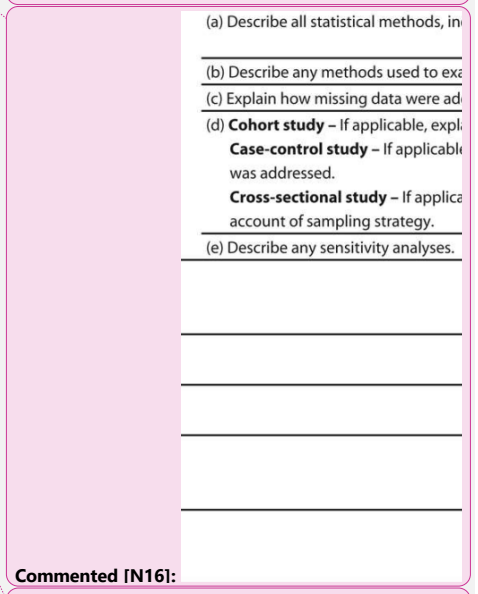
MAPT genotyping

We performed genotyping for the *MAPT* rs1052553 allelic variant on genomic DNA which was extracted from saliva samples collected using DNA Genotek® saliva kits or from venous whole blood samples using standard protocols.

Statistical analysis

Cohort characteristics were expressed as proportions (%) compared between groups (Parkinson's disease cases and controls) using two tailed X2 test for categorical variables. For descriptive purposes, age at onset of PD of PD patients and age at study were categorized by 10-year interval and disease duration by 5-year intervals. SNPs genotypes were assessed for Hardy-Weinberg equilibrium (HWE) using an exact test. In order to characterize MAPT, genotypic and allelic frequencies of the MAPT polymorphism was calculated in PD patients and compared to healthy controls without PD from the same population. We used logistic regression to investigate the association between H1 haplotype and PD risk and age of onset, accounting for gender, age at onset of PD cases and age at recruitment for controls. Data were analysed using Stata/MP version 16.0 statistical software (Stata Corporation, College Station, TX: StataCorp LLC).

Commented [N15]: Please review so it isn't a repetition of the information from the APOE manuscript


(a) Describe all statistical methods, in

(b) Describe any methods used to exa

(c) Explain how missing data were ad

(d) **Cohort study** – If applicable, expl
Case-control study – If applicabl
was addressed.
Cross-sectional study – If applica
account of sampling strategy.

(e) Describe any sensitivity analyses.

Commented [N16]:

Commented [N17]: Olaitan: Shouldn't this be adjusted for age at STUDY for PD and controls since both are available?

Results

Baseline characteristics of study population

In this study, we included 907 Nigerians with PD matched for age and gender with 1,022 healthy controls without PD. There were more male (71.6%) PD participants than females**.** Baseline characteristics of the cohort (sex distribution, age at study, age at onset, duration of PD overall and by gender) are shown in Supplementary Table 1.

MAPT allele and genotypic/haplotype frequencies in Nigerians with PD compared to controls

The *MAPT* rs1052553 allele and genotype frequencies are shown in Table 1, and did not differ significantly between persons with PD and controls either overall or by gender ($p>0.05$). The most frequent allele was the major allele A (H1) present in 1790 (98.7%) of PD and 2026 (99.1%) of controls ($p=0.19$).

Commented [N18]: Please compute and complete

Association between HI haplotype and PD risk and PD age of onset

The mean age at onset of PD did not differ significantly between PD patients with the H1/H1 haplotype (60.0± 10.6) and those with either H1/H2 or H2/H2 (60.0 ± 11.5) ($p=0.98$) as shown in Table 2. There was no association between H1 haplotype and PD even after adjusting for age at onset of PD and gender (Table 3).

DISCUSSION

The biological plausibility of abnormalities in the function of *MAPT* contributing to the pathogenesis of Parkinson's disease make it a target of interest for studies exploring the role of genetic variability in the causation of familial and apparently sporadic PD. Previous studies, including those demonstrating an independent and significant effect have been derived from populations of predominantly Caucasian ancestry, with minimal to no data on persons of black African ancestry. In addition, data provided in our study is important from the general perspective of contributing to data regarding the population genetics of *MAPT* and the distribution of the 17q21 inversion in persons of African ancestry. Donnelly *et al* previously described the distribution of the inverted (H2) haplotype using 21 SNPs including the rs1052553 utilized in our study. The study included populations within Africa, including 3 of the Nigerian ethnic groups included in this study (Ibo, Yoruba and Hausa). The study concluded that the H2 inverted haplotype occurs at low frequencies in Africa and was completely absent in the Nigerian sample. (ref Donnelly et al 2010) Although our findings largely corroborate the low frequency reported in west Africans, we have in fact demonstrated that the H2 haplotype does occur in the Nigerian population, and occurred in 1.1% of our participants. The frequencies for other African populations range between 0.75 to 6.3%, whereas it is reported as occurring in 4% of African Americans. (ref Donnelly) Donnelly and colleagues had reported that within Africa, the H2 haplotype occurs at the highest frequency in North Africa (Mozabites), at low levels in Northwest Africa, Central Africa, and Eastern Africa, and is absent in West Africa (where Nigeria is located) except in the Mandenka. (ref Donnelly) Sleinberg and colleagues made a similar assertion as, in their study that included Nigerian Yorubas, the H2 inverted haplotype was

Commented [N19]: Donnelly MP, Paschou P, Grigorenko E, Gurwitz D, Mehdi SQ, Kajuna SL, Barta C, Kungulilo S, Karoma NJ, Lu RB, Zhukova OV, Kim JJ, Comas D, Siniscalco M, New M, Li P, Li H, Manolopoulos VG, Speed WC, Rajeevan H, Pakstis AJ, Kidd JR, Kidd KK. The distribution and most recent common ancestor of the 17q21 inversion in humans. Am J Hum Genet. 2010 Feb 12;86(2):161-71. doi: 10.1016/j.ajhg.2010.01.007. Epub 2010 Jan 28. PMID: 20116045; PMCID: PMC2820164.

Commented [N20]: Steinberg KM, Antonacci F, Sudmant PH, et al. Structural diversity and African origin of the 17q21.31 inversion polymorphism. *Nat Genet*. 2012;44(8):872-880. Published 2012 Jul 1. doi:10.1038/ng.2335

absent from virtually all Western African individuals except for the Pygmy populations (Bakola, Biaka, and Mbuti). (ref Sleinberg)

Regarding MAPT H1 haplotype and risk of PD, the need to investigate this relationship are linked to previous reports which have showed increased transcriptional activity of H1 in comparison to H2 (REF) implying a possible mechanism for development of PD. Also, the increased proportions of H1 resulting in tau mediated alpha-synuclein fibrillisation has been linked to PD susceptibility and progression (REF). However, our study did not show any association between H1 and PD after correction for gender and age at onset of PD. Although several studies have reported some association between H1 and development of PD, others like our study have failed to validate this link. A proffered explanation for conflicting reports is underpowered studies resulting from small sample sizes. This is unlikely to be the reason in our large cohort (n=907) and another previous study of German origin (n=418) which failed to demonstrate any association between H1 haplotype and PD risk compared to a similar cohort but smaller cohort from Serbia which confirmed an association between H1haplotype and PD risk (REF). Another reason for differing reports is the possibility that the importance of genetic risk factors may vary among diverse ethnic groups. This could be the reason for the divergent result in our study. The role of divergent ethnic background to explain inconsistent results in the relationship between H1 haplotype and PD in white Caucasian populations have been previously reported (REF).

Furthermore, for mutations in monogenic PD, similar population-dependent variations have been reported. The prevalence of the *LRRK2* gene (p.G2019S mutation) for example, ranges from 1-7% among patients of European origin with PD to 20-40% among the North African Arabs and

Commented [N21]: I am going to look at the database to extract the self declared ethnic groups of the 40 plus with H2 haplotype

Commented [OO22]: **Tau haplotypes regulate transcription and are associated with Parkinson's disease**
John B J Kwok [1], Erdahl T Teber, Clement Loy, Marianne Hallupp, Garth Nicholson, George D Mellick, Daniel D Buchanan, Peter A Silburn, Peter R Schofield

Commented [OO23]: **Initiation and synergistic fibrillization of tau and alpha-synuclein**
Benoit I Giasson [1], Mark S Forman, Makoto Higuchi, Lawrence I Golbe, Charles L Graves, Paul T Kotzbauer, John Q Trojanowski, Virginia M-Y Lee

Role of the H1 haplotype of microtubule-associated protein tau (*MAPT*) gene in Greek patients with Parkinson's disease
Nikolaos Refenes*1,2, Juliane Bolbrinker2, Georgios Tagaris3,
Antonio Orlacchio4,5, Nikolaos Drakoulis1 and Reinhold Kreutz2

Commented [OO24]: **Role of ethnicity on the association of MAPT H1 haplotypes and subhaplotypes in Parkinson's disease**
Susan Winkler [1], Inke R König, Katja Lohmann-Hedrich, Peter Vieregge, Vladimir Kostic, Christine Klein

Commented [OO25]: **Role of ethnicity on the association of MAPT H1 haplotypes and subhaplotypes in Parkinson's disease**
Susan Winkler [1], Inke R König, Katja Lohmann-Hedrich, Peter Vieregge, Vladimir Kostic, Christine Klein

Ashkenazi Jews with PD (REF). In addition, ethnic variation is reported to play a key role in the association studies which have examined the possible effect of polymorphisms in *LRRK2.* (REF)

Another example is that of fibroblast growth factor which was linked to PD risk in a diverse US population (REF) but not established in the Finnish and Greek cohorts (REF).

Within our Nigeria PD cohort, in addition to the three main ethnic groups (Yoruba, Igbo and hausa), there are many other ethnic groups resulting in variations suggesting a huge impact of ethnic background on genetic risk factors.

The main limitation of our study is that we did not may be the potential of misclassification of PD patients even though we have used validated diagnostic criteria. Another limitation is that due to the complex nature of the MAPT Locus, we might not have dissected the

The main strength of our study is the large sample size of our cohort within the African continent. To the best of our knowledge this is the first study to evaluate MAPT frequency and provide data on PD risk and H1 haplotype in Africa.

Conclusion: Our data showed no relationship between the *MAPT* H1 haplotype and PD in our cohort of black Africans in line with previous studies suggesting the impact of divergent ethnic populations on genetic risk factors.

Commented [OO26]: Zabetian CP, Morino H, Ujike H et al: Identification and haplotype analysis of LRRK2 G2019S in Japanese patients with Parkinson disease. Neurology 2006; 67: 697–699.

Lesage S, Durr A, Tazir M et al: LRRK2 G2019S as a cause of Parkinson's disease in North African Arabs. N Engl J Med 2006; 354: 422–423.

Ozelius LJ, Senthil G, Saunders-Pullman R et al: LRRK2 G2019S as a cause of Parkinson's disease in Ashkenazi Jews. N Engl J Med 2006; 354: 424–425.

Commented [OO27]: Di Fonzo A, Wu-Chou YH, Lu CS et al: A common missense variant in the LRRK2 gene, Gly2385Arg, associated with Parkinson's disease risk in Taiwan. Neurogenetics 2006; 7: 133–138.

Tan EK, Zhao Y, Skipper L et al: The LRRK2 Gly2385Arg variant is associated with Parkinson's disease: genetic and functional evidence. Hum Genet 2007; 120: 857–863.

Biskup S, Mueller JC, Sharma M et al: Common variants of LRRK2 are not associated with sporadic Parkinson's disease. Ann Neurol 2005; 58: 905–908.

Paisan-Ruiz C, Evans EW, Jain S et al: Testing association between LRRK2 and Parkinson's disease and investigating linkage disequilibrium. J Med Genet 2006; 43: e09.

Commented [OO28]: van der Walt JM, Noureddine MA, Kittappa R et al: Fibroblast growth factor 20 polymorphisms and haplotypes strongly influence risk of Parkinson disease. Am J Hum Genet 2004; 74: 1121–1127.

Commented [OO29]: Clarimon J, Xiromerisiou G, Eerola J et al: Lack of evidence for a genetic association between FGF20 and Parkinson's disease in Finnish and Greek patients. BMC Neurol 2005; 5: 11.

TABLE AND FIGURES

Table 1. *MAPT* rs1052553 allele and genotype distribution in persons with Parkinson's disease and controls

	All participants (n= 1929; 3858 allelles)	Parkinson's disease (n= 907; 1814 allelles)	Control participants (n= 1022; 2044 allelles)	*p* value
MAPT rs1052553 alleles				
A (H1)	3816 (98.9%)	1790 (98.7%)	2026 (99.1%)	0.19
G (H2)	42 (1.1%)	24 (1.3%)	18 (0.9%)	
MAPT rs1052553 genotypes				
A/A (H1/H1)	1888 (97.9%) `	884 (97.5%)	1004 (98.2%)	0.24
A/G (H1/H2)	40 (2.1%)	22 (2.4%)	18 (1.8%)	0.31
G/G (H2/H2)	1 (0.05%)	1 (0.1%)	0 (0.0%)	0.47*
MAPT rs1052553 genotype category				
H1/H2 (A/G) and H2/H2 (G/G)	41(2.2%)	23 (2.5%)	18 (1.8%)	ref
H1/H1 (A/A) only	1888	884 (97.5)	1004 (98.2)	0.23**

Footnote: *p value for Fisher's Exact test. Odds ratio(95% confidence interval) for comparison: 0.68 (0.39 – 1.28). MAPT H1H1 frequency by gender (n vs total) r: PD (females 251/258, 97.3%; males 633/649, 97.5%) and Controls (females 348/351, 99.2%; males 656/671, 97.8%); p>0.05

Commented [N30]: Allele frequency calculated based on total allele counts from genotype, AA = 2 x a, AG = 1 x A and 1 x G, GG = 2 x G. Calclulated for total, PD and controls. (no of counts of allele based on genotypes/no of alleles)

Commented [N31]: Genotype frequency calculated as n with genotype / no of participants in group.

Commented [N32]: Ola please complete

Table 2. ***MAPT*** **rs1052553** genotype and age at onset in Parkinson's disease

	PD with HI/HI			PD without HI/HI			*p-value
	All n=884	Female n=251 (	Male n=633	All n=23	Female n=7	Male n=16	
Mean age at onset ± SD, *years*	60.0± 10.6	60.7±10.6	59.7±10.5	60.0 ± 11.5	55.7± 13.1	61.9±10.6	0.98
Median age at onset (IQR), y*ears*	60 (14.0)	61 (14.0)	60 (14.0)	61 (11.0)	58 (5.0)	64.5 (12.5)	
Age group at onset, *n(%)*							
<45	70 (7.9)	18(24.6)	52(75.4)	2 (8.7)	1(50.0)	1(50.0)	0.94
45-54	178 (20.1)	45(25.3)	133(74.7)	3 (13.0)	0(0.0)	3(100.0)	
55-64	330 (37.3)	89(27.0)	241(73.0)	9 (39.1)	5(55.6)	4(44.4)	
65-74	224 (27.7)	80(32.8)	164(67.2)	7 (30.4)	1(14.3)	6(85.7)	
≥ 75	62 (7.0)	19(30.7)	43(69.4)	2 (8.7)	0(0.0)	2(100.0)	

*chi square. p-value comparing PD with and without H1.H1. All other comparisons p also not significant (>0.05)

Commented [N33]: If we decide to report 95% CI: H1/H1 all (59.3 – 60.7); Without H1/H1 (55.03 – 64.97); H1H1 female/male (59.4 – 62.1/58.8 – 60.5); Without H1H1 female /male (43.6 – 67.8/56.2- 67.5)

Age at onset *n(%)*	**PD with HI/HI**	**PD without HI/HI**	**Odds ratio 95%CI**	***p-value**
<45	70(7.9)	2(8.7)	ref	
45-54	178(20.1)	3(13.0)	1.72 (0.28-10.51)	0.568

Commented [N36]: These tables are same save for the reference category. Which one is preferred? The information needs to be included in the data analysis/statistics section of methodology

55-64	330(37.3)	9(39.1)	1.07 (0.23-5.03)	0.937
65-74	224(27.7)	7(30.4)	1.02 (0.21-5.01)	0.984
≥ 75	62(7.0)	2(8.7)	0.90 (0.12-6.60)	0.919

Age at onset *n(%)*	**PD with HI/HI**	**PD without HI/HI**	**Odds ratio 95%CI**	***p-value**
55-64	330 (37.3)	9(39.1)	ref	
<45	70(7.9)	2(8.7)	0.94(0.20-4.44)	0.937
45-54	178(20.1)	3(13.0)	1.61(0.43-6.05)	0.477
65-74	224(27.7)	7(30.4)	0.95(0.35-2.60)	0.927
≥ 75	62(7.0)	2(8.7)	0.85(0.18-4.01)	0.834

*adjusted for gender

Commented [N34]: Ola – why was the age range 55-64 chosen as the reference? Because it was the most frequent?

Commented [OO35R34]: Yes Prof, we need to decide which to use.

Table 3. ***MAPT*** **rs1052553** genotype and risk of Parkinson's disease

Genotype	**PD cases**	**Controls**	**Odds ratio 95%CI**	***p-value**
H1/H2 and H2/H2	23(2.5)	18(1.8)	ref	
H1/H1	884(97.5)	1,004(98.2)	0.68(0.39-1.28)	0.228

*Adjusted for gender, age at onset for PD cases and age at study for controls

Supplementary Materials

Supplementary Table 1.

Variables	**All participants *n* = 1929**	**PD *n*=907**	**Controls *n*=1,022**	**p-value**
Female	609 (31.6)	258 (28.4)	351 (34.3)	0.005
Male	1320 (68.4)	649 (71.6)	671 (65.7)	
Mean age at study (±SD), *years*	63.8 ±9.8	64.1 ± 10.2	63.4 ± 9.3	0.10
Median age at study (IQR), *years*	64.0 (12)	65.0 (13)	63.0 (12)	
Mean age at study (±SD) (females), *years*	63.6 ±9.5	64.9 ±9.8	62.6 ±9.1	
Mean age at study (±SD) (males), *years*	63.8±9.9	63.9±10.4	63.8±9.4	
Mean age at onset (±SD), *years*	N/A	60.0 ±10.6	N/A	
Median age at onset (IQR), *years*	N/A	60.0 (14.0)	N/A	
Mean age at onset (±SD) (females), *years*	N/A	60.6 ±10.7	N/A	0.25
Mean age at study (±SD) (males), *years*	N/A	59.7 ±10.5	N/A	
Duration of disease (mean ±SD), years	N/A	4.2 ±4.1	N/A	
Median duration of disease (IQR), years	N/A	3.0 (3.0)	N/A	
Duration of disease (mean ±SD) (females), years	N/A	4.3 ±4.5	N/A	0.73
Duration of disease (mean ±SD) (males), years	N/A	4.2 ±3.9	N/A	

Supplementary Table 2. Frequency of G Allele by Nigerian ethnic group

Possible supplementary table: Distribution of the minor G allele in the Nigerian population Ethnicities of cohort with minor allele G

Ethnic group	Minor allele present	Minor allele absent	Total	Frequency of minor allele
Igbo	21	360	381	5.5
Yoruba	8	676	684	1.2
Others	4	198	202	1.98
Fulani	2	23	25	8
Hausa	2	245	247	0.8
Anang	1	8	9	11.1
Efik	1	31	32	3.1
Ibibio	1	53	54	1.9
Idoma	1	42	43	2.3

References

1. Zhang CC, Xing A, Tan MS, Tan L, Yu JT. The Role of MAPT in Neurodegenerative Diseases: Genetics, Mechanisms and Therapy. Mol Neurobiol. 2016;53(7):4893-904.

2. Donnelly MP, Paschou P, Grigorenko E, Gurwitz D, Mehdi SQ, Kajuna SL, et al. The distribution and most recent common ancestor of the 17q21 inversion in humans. Am J Hum Genet. 2010;86(2):161-71.

3. Zhang CC, Zhu JX, Wan Y, Tan L, Wang HF, Yu JT, et al. Meta-analysis of the association between variants in MAPT and neurodegenerative diseases. Oncotarget. 2017;8(27):44994-5007.

4. Vandrovcova J, Pittman AM, Malzer E, Abou-Sleiman PM, Lees AJ, Wood NW, et al. Association of MAPT haplotype-tagging SNPs with sporadic Parkinson's disease. Neurobiol Aging. 2009;30(9):1477-82.

5. Huang Y, Wang G, Rowe D, Wang Y, Kwok JB, Xiao Q, et al. SNCA Gene, but Not MAPT, Influences Onset Age of Parkinson's Disease in Chinese and Australians. Biomed Res Int. 2015;2015:135674.

6. Elbaz A, Ross OA, Ioannidis JP, Soto-Ortolaza AI, Moisan F, Aasly J, et al. Independent and joint effects of the MAPT and SNCA genes in Parkinson disease. Ann Neurol. 2011;69(5):778-92.

7. Rizig M, Okubadejo N, Salama M, Thomas O, Akpalu A, Gouider R. The International Parkinson Disease Genomics Consortium Africa. The Lancet Neurology. 2021;20(5).

8. Hughes AJ, Daniel SE, Kilford L, Lees AJ. Accuracy of clinical diagnosis of idiopathic Parkinson's disease: a clinico-pathological study of 100 cases. Journal of Neurology, Neurosurgery and Psychiatry. 1992;55(3).