Neurology Publish Ahead of Print
DOI: 10.1212/WNL.0000000000200699

Genome-wide Association and Meta-analysis of Age-at-Onset in Parkinson Disease: Evidence From COURAGE-PD Consortium

This is an open access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 (CC BY-NC-ND), which permits downloading and sharing the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.

Neurology® Published Ahead of Print articles have been peer reviewed and accepted for publication. This manuscript will be published in its final form after copyediting, page composition, and review of proofs. Errors that could affect the content may be corrected during these processes.
Abstract

Background and Objectives
Considerable heterogeneity exists in the literature concerning genetic determinants of the age of onset (AAO) of Parkinson’s disease (PD), which could be attributed to lack of well-powered replication cohorts. The previous largest GWAS identified SNCA and TMEM175 loci on chromosome (Chr) 4 with a significant influence on AAO of PD, these have not been independently replicated. The present study aims to conduct a meta-analysis of GWAS of PD AAO and validate previously observed findings in worldwide populations.

Methods
A meta-analysis was performed on PD AAO GWAS of 30 populations of predominantly European ancestry from the Comprehensive Unbiased Risk Factor Assessment for Genetics and Environment in Parkinson's Disease (COURAGE-PD) consortium. This was followed up by combining our study with the largest publicly available European ancestry dataset compiled by the International Parkinson disease Genomics Consortium (IPDGC).
Results

The COURAGE-PD included a cohort of 8,535 patients with PD (91.9%: Europeans, 9.1%: East-Asians). The average AAO in the COURAGE-PD dataset was 58.9 years (SD=11.6), with an under-representation of females (40.2%). The heritability estimate for AAO in COURAGE-PD was 0.083 (SE=0.057). None of the loci reached genome-wide significance (P<5x10^-8). Nevertheless, the COURAGE-PD dataset confirmed the role of the previously published TMEM175 variant as genetic determinant of AAO of PD with Bonferroni-corrected nominal levels of significance (P<0.025): (rs34311866: β(SE)COURAGE=0.477(0.203), P_{COURAGE}=0.0185). The subsequent meta-analysis of COURAGE-PD and IPDGC datasets (N_{total}=25,950) led to the identification of two genome-wide significant association signals on Chr 4, including the previously reported SNCA locus (rs983361: β(SE)COURAGE+IPDGC=0.720(0.122), P_{COURAGE+IPDGC}=3.13x10^{-9}) and a novel BST1 locus (rs4698412: β(SE)COURAGE+IPDGC=-0.526(0.096), P_{COURAGE+IPDGC}=4.41x10^{-8}).

Discussion

Our study further refines the genetic architecture of Chr 4 underlying the AAO of the PD phenotype through the identification of BST1 as a novel AAO PD locus. These findings open a new direction for the development of treatments to delay the onset of PD.

Keywords: Burden of disease, Age at onset, Duration of disease, Parkinson’s disease, Genetic heritability

Introduction

In 2019, over 8.51 million individuals (95% uncertainty interval [UI] 7.3–9.8) had PD globally. This disease is one of the fastest-growing neurodegenerative diseases globally, with an estimated 30.9% increase in the number of patients with PD in 2019 compared to 2010. However, the prevalence of a disease depends on both the incidence and duration of disease, making an earlier age at onset of PD an essential contributor to the overall burden of
the disease. While less than 5% of patients with PD harbor pathogenic mutations in known monogenic PD genes, the majority are sporadic with predominantly late age of onset (AAO)\(^2\)\(^3\). A better understanding of genetic factors influencing variability in AAO in sporadic patients could lead to a better understanding of PD pathophysiology.

The emergence of genome-wide association studies (GWAS) has resulted in a rapidly expanding list of loci harboring disease-susceptibility variants for the sporadic form of the disorder\(^4\)\(^6\). To date, genetic variants at 78 loci have been identified for sporadic PD\(^6\). Despite advances in understanding the genetic basis of PD incidence, the heritability underlying PD AAO remains largely unexplained. A recent global effort involving 28,568 sporadic PD patients of European ancestry led to the identification of two loci, \(SNCA\) and \(TMEM175\), as risk factors for an earlier AAO, both of which are also known to play a role in \(\alpha\)-synuclein-linked mechanisms underlying PD pathology\(^1\)\(^7\)\(^8\). More recently, a meta-analysis including 5,166 Chinese patients with PD lead to the identification of another locus \(NDN/PWRN4\)\(^9\). Despite the large disparity in sample size and the genetic loci identified by the two studies, both works estimated similar total heritability of AAO of 10-14%\(^7\)\(^9\). They also showed an inverse correlation between a polygenic risk score (PRS) and AAO based on risk loci for PD on individuals of similar ancestry, suggesting overlap between the pathways underlying disease susceptibility and AAO in PD.

Recent studies have underscored the relevance of inclusion of ethnic diversity in genomic research\(^9\)\(^10\). The COURAGE-PD (COMprehensive Unbiased Risk Factor Assessment for Genetics and Environment in Parkinson’s Disease) is a worldwide collaboration consortium comprising 35 PD study cohorts which aims to address this disparity to some extent in PD research\(^11\). The present study aims to perform an AAO GWAS in COURAGE-PD and to investigate the validity of previously observed loci by conducting one of the largest meta-analysis of PD AAO GWAS to date by combining previous International Parkinson’s Disease Genomics Consortium (IPDGC) AAO GWAS (n=17, 415) with newly generated
COURAGE-PD AAO GWAS (n=8,535), resulting in a combined dataset of 25,950 patients with PD. Lastly, we investigate the influence of a PD PRS on PD AAO to dissect the potential overlapping etiology.

**Methods**

**Study cohorts and participants**

The COURAGE-PD consortium comprises data from 15,849 patients with PD and 11,444 controls of predominantly European ancestry from 35 cohorts with a major contribution from the Genetic Epidemiology of Parkinson’s disease (GEoPD) Consortium (www.geopd.net). Quality control (QC) of genome-wide data was performed in each COURAGE-PD study cohort. See eMethods for more details, including collected phenotypic data. AAO was defined based on the initial manifestation of motor symptoms associated with PD, as described elsewhere. Post-imputation, only sporadic PD patients with data available on AAO and not overlapping with previous IPDGC AAO GWAS were included in the present study, leaving 8,535 samples from 30 cohorts. These comprised 26 European and four East-Asian ancestry cohorts.

**Genotype-phenotype analysis**

**Regression analysis and meta-analysis of study-specific estimates**

Linear regression analysis of imputed dosages with AAO was performed in each study cohort using an additive model, implemented in rvtests, correcting for gender and the first five principal components. The selection of five principal components was based on study cohort specific scree plots. The scree plot flattened out after the third factor for majority of study cohorts, with few exceptions, where five factors explained the highest proportion of the total variance. This was followed by combining study-specific results through an inverse variance weighted (IVW) fixed-effect meta-analyses conducted using METAL. Additionally, only those variants which were successfully genotyped in at least 2/3rd of study cohorts were included for further interpretation. Similarly, the variants with $I^2$ statistic $\geq$50%
were considered to have substantial heterogeneity and were excluded from further interpretation. We also employed additive random-effect meta-analyses using the DerSimonian-Laird estimator to check the influence of heterogeneity on our findings. The quantile-quantile (QQ) plot was generated using R to judge the potential influence of population stratification on the overall significance of the effect estimates. We considered a $P<5\times10^{-8}$ as genome-wide significant and a $P<1\times10^{-6}$ as suggestive evidence for a potential association. We also considered a Bonferroni-corrected $P<0.025$ for reporting replication signals originating from two SNPs (rs356203 (SNCA), rs34311866 (TMEM175)) that reached a genome-wide significance in the previous largest meta-analysis of AAO of PD. The results were visualized using R generated Manhattan and LocusZoom generated regional association plots. We conducted LD score regression with LDSC (using summary-level data) to estimate heritability explained by the PD AAO GWAS. We also performed a meta-analysis of COURAGE-PD AAO (n=8535) with the previous largest AAO meta-analysis comprising IPDGC dataset (n=17,415) to discover potentially new loci and improve heritability estimates.

Correlation between case-control GWAS and AAO GWAS

We used two approaches to assess the correlation between PD case-control GWAS meta-analysis and COURAGE-PD AAO GWAS meta-analysis. Firstly, we computed the genome-wide genetic correlation between PD status and PD AAO in COURAGE-PD dataset using the cross-trait LD score regression method. Secondly, we used effect estimates of significant genetic variants ($P<5\times10^{-8}$) identified by combining of COURAGE-PD case-control GWAS meta-analysis dataset with the IPDGC-PD case-control GWAS meta-analysis dataset to generate individual-specific polygenic risk scores (PRS) in the COURAGE-PD AAO population, using PRSice2. Linear regression analysis of PRS with AAO was performed, correcting for gender and the first five principal components.
Subgroup analysis and power computation

A subgroup analysis was performed to explore the influence of ethnicity and gender on the AAO GWAS as well as the correlation between case-control and AAO GWAS meta-analyses. The power was estimated using QUANTO 1.2.4.19

Expression quantitative trait loci analysis

We further explored the potential influence of novel variants identified in the present study on the expression traits using the gene expression data from the Genotype-Tissue Expression Project using the GTEx portal (gtexportal.org, Data Source: GTEx Analysis Release V8 (dbGaP Accession phs000424.v8.p2) and UK Brain Expression Consortium (UKBEC) using the Braineac portal (braineac.org).20, 21

Standard protocol approvals, registrations, and patient consents

The study was conducted at the University of Tübingen, and the ethical approval was obtained by the local institutional review board (IRB) of respective study sites. All the study participants provided signed informed consent.

Data availability

Summary statistics of COURAGE-PD AAO GWAS used in the meta-analysis are available from the corresponding author upon reasonable request. In addition, IPDGC summary statistics for AAO GWAS was downloaded from the IPDGC website (https://pdgenetics.org/resources). Significant SNPs of risk of PD based on meta-analysis of COURAGE-PD and IPDGC datasets, used in the PRS calculation can be found in the original publication (Grover et al. in preparation). Relevant programming scripts used for the present work are available at the github website of Center of Genetic Epidemiology (CGE) at Tübingen (https://github.com/CGEatTuebingen/Ageatonset_GWAS_Courage-PD).
Results

Main study outcome variable

The final cohort after QC included a total of 8,535 patients with PD, 7,847 of European ancestry (91.9%) and 688 of East-Asian ancestry (9.1%). The average AAO in the COURAGE-PD dataset was 58.9 years (SD=11.6), with an under-representation of females (40.2%) (see eTable 1). We did not observe any major influence of gender or ethnicity on AAO. Furthermore, the average AAO was slightly lower than that reported in the IPDGC dataset (62.1 years; SD=12.1), a difference that was statistically significant (P<0.05).

Genetic heritability of the study outcome

Using summary-level data, the total estimated heritability ($h^2$) in the COURAGE-PD dataset was 0.083 (SE=0.057). Similar heritability estimates were observed in the European sub-cohort ($h^2=0.079$, SE=0.061). However, the heritability estimates in the Asian sub-cohort could not be reliably computed due to an insufficient number of patients. Additionally, we failed to achieve any improvement in heritability estimates, although with improved accuracy by combining COURAGE-PD with IPDGC dataset ($h^2=0.078$, SE=0.018).

Genome-wide meta-analysis

COURAGE-PD

GWAS meta-analysis

The genomic inflation factor $\lambda$ was 1.016 (See eFigure 1 for the QQ plot). None of the loci reached genome-wide significance (Figure 1). We observed one locus reaching suggestive genome-wide significance level, $PDZPH1P$ (chr5) ($\beta$(SE)$_{COURAGE}$=-1.456(0.293), $P_{COURAGE}$=6.91x10$^{-7}$). However, stratifying the analyses by ethnicities, we did not observe any suggestive involvement of $PDZPH1P$ locus in the European sub-cohort. (see eTable 2). Interestingly, despite being a smaller sub-cohort, $SUGCT$ locus on chromosome (Chr) 7 was detected as a suggestive locus in the East-Asian sub-cohort ($\beta$(SE)$_{COURAGE-EASIAN}$=13.681(2.769), $P_{COURAGE-EASIAN}$=7.80x10$^{-7}$). Furthermore, the stratified analysis
provided suggestive evidence of three loci, RHEB (chr8) in males ($\beta$(SE)$_{COURAGE-M}$=-1.112(0.222), $P_{COURAGE-M}$=5.15x10$^{-7}$), and MTHFD1L (chr6) ($\beta$(SE)$_{COURAGE-F}$=-1.995(0.402), $P_{COURAGE-F}$=6.78x10$^{-7}$) and KNH3 (chr12) in females ($\beta$(SE)$_{COURAGE-F}$=2.176(0.432), $P_{COURAGE-F}$=4.59x10$^{-7}$) (see eTable 2).

In the replication of previously reported variants, only the TMEM175 variant (rs34311866: $\beta$(SE)$_{COURAGE}$=0.477(0.203), $P_{COURAGE}$=0.018) reached Bonferroni-corrected nominal levels of significance in the COURAGE-PD dataset. Nevertheless, the SNCA variant also showed a trend towards association (rs356203: $\beta$(SE)$_{COURAGE}$=0.362(0.172), $P_{COURAGE}$=0.035).

**Meta-analysis of COURAGE-PD and IPDGC datasets**

The meta-analysis of COURAGE-PD and IPDGC datasets led to the identification of two loci that reached genome-wide significance (see eTable 2; Figure 2). The SNCA variant, rs983361, was the most strongly associated SNP, with the presence of allele T (frequency=0.204) leading to an average delay in AAO by 0.72 years ($\beta$(SE)$_{COURAGE+IPDGC}$=-0.720(0.122), $P_{COURAGE+IPDGC}$=3.13x10$^{-9}$). This association, however, appeared to be driven by the strong association reported by IPDGC dataset, with negligible effect detected in the COURAGE-PD dataset ($P_{COURAGE/COURAGE-EUR (rs983361)}=0.022$; Not detected in East-Asian sub-population) (see eFigure 2A), which was also reflected in the loss of genome-wide significance, when using an additive random effect model ($P=2.98x10^{-6}$). On the other hand, another independent locus on the same chromosome, BST1 (rs4698412) showed similar effects in COURAGE-PD and IPDGC datasets ($\beta$(SE)$_{COURAGE}$=-0.633(0.175), $P_{COURAGE}$=2.95x10$^{-4}$; $\beta$(SE)$_{IPDGC}$=-0.480(0.115), $P_{IPDGC}$=3.04x10$^{-5}$), and the combination of both estimates resulted in the identification of a novel genome-wide significant BST1 locus for AAO ($\beta$(SE)$_{COURAGE+IPDGC}$=-0.526(0.096), $P_{COURAGE+IPDGC}$=4.41x10$^{-8}$) (see eFigure 2B). The rs4698412 allele A (frequency=0.562) at the locus led to an average earlier AAO of 0.526 years in PD patients. No genetic heterogeneity was detected in the observed association
Furthermore, we did not observe any change in the effect estimates, when using the additive random effect model (P=4.41x10^{-8}).

The previously reported \textit{TMEM175} (rs34311866) showed suggestive association in the combined analysis ($\beta$(SE)$_{\text{COURAGE+IPDGC}}$=0.589(0.114), P$_{\text{COURAGE+IPDGC}}$=2.64 x 10^{-7}) that appeared to be driven by previously reported findings in the IPDGC dataset ($\beta$(SE)$_{\text{IPDGC}}$ =0.642(0.139), P$_{\text{IPDGC}}$=3.72x10^{-6}) (see eFigure 2C). Another locus \textit{AL391867.1/RP11-342F21.1} (rs62582905), a locus of unknown biological significance also crossed the threshold of suggestive association in the same analysis ($\beta$(SE)$_{\text{COURAGE+IPDGC}}$ =-1.456(0.293), P$_{\text{COURAGE}}$=6.62x10^{-7}) (see eTable 2). However, unlike the \textit{TMEM175} association, the association with \textit{AL391867.1/RP11-342F21.1} was observed to be stronger in the COURAGE-PD dataset ($\beta$(SE)$_{\text{COURAGE}}$=-1.925(0.447), P$_{\text{COURAGE}}$=1.64x10^{-5}).

We carried out a sensitivity analysis by excluding the Asian sub-cohort from the COURAGE dataset, followed by combining with the IPDGC dataset. Similar findings were observed for the two genome-wide significant loci (\textit{SNCA} rs983361: P$_{\text{COURAGE(EUR)+IPDGC}}$ =3.13 x 10^{-9}, \textit{BST1} rs4698412: P$_{\text{COURAGE(EUR)+IPDGC}}$=6.27x10^{-8}) (see eTable 2). A similar sensitivity analysis for the previously reported APOE $\varepsilon$4 locus also showed suggestive association with PD AAO (\textit{APOE} rs429358: $\beta$(SE)$_{\text{COURAGE(EUR)+IPDGC}}$ =0.711(0.145), P$_{\text{COURAGE(EUR)+IPDGC}}$=9.33x10^{-7}). However, the association was primarily driven by highly significant findings in the IPDGC dataset ($\beta$(SE)$_{\text{IPDGC}}$=0.754(0.171), P$_{\text{IPDGC}}$=9.86x10^{-6}; $\beta$(SE)$_{\text{COURAGE(EUR)}}$=0.599(0.275), P$_{\text{COURAGE(EUR)}}$=0.029).

**Correlation between genetic risk for PD and PD AAO**

Using complete GWAS summary datasets for COURAGE-PD case-control and COURAGE-PD AAO, we observed a non-significant negative genetic correlation between PD and PD AAO (rg=-0.291, SE=0.224; P=0.186). Furthermore, a slightly stronger genetic correlation was observed when restricting our correlation analysis to European sub-cohorts only (rg=-0.315; SE=0.252; P=0.211).
When using PRS based on the significant loci detected in the meta-analysis of COURAGE-PD and IPDGC European datasets, as reported elsewhere, we observed that each unit increase in SD in PRS lead to a significant decrease in AAO in COURAGE-PD by 0.58 years ($\beta(\text{SE})_{\text{COURAGE}}=-0.581(0.149), P_{\text{COURAGE}}=9.35\times10^{-5}$). Despite the significant findings, the PRS explained only 0.59% of the genetic proportion of PD heritability.

**Expression quantitative trait analysis of novel BST1 locus**

The mining of the Genotype-Tissue Expression (GTEx) portal showed that rs4698412 representing the $BST1$ locus is a highly significant expression quantitative trait locus (eQTL) for $CD38$ in the basal ganglia (caudate, nucleus accumbens and putamen) and cortex (NES=-0.32 – -0.44; $P<1\times10^{-10}$) (Table 1). The expression analysis also showed a strong dosage effect with a consistent lower expression in the presence of AA genotype compared to GG genotype with a higher expression, irrespective of brain tissue type. In addition, we also found that SNP modulates the expression of $BST1$ in whole blood. However, the effect was considerably lower in comparison to that observed on $CD38$ expression levels in brain tissues (NES=-0.071; $P=1.7\times10^{-6}$). The follow-up of association of rs4698412 with expression in brain tissues in the UKBEC database further confirmed the role of basal ganglia with $CD38$ as the most significantly associated expressed gene in the putamen ($P=7.1\times10^{-6}$) (Table 1).

**Discussion**

The identification of genetic determinants that modify the disease progression will not only help to increase our understanding of PD etiopathogenesis, but also enables the development of strategies that could be used for therapeutic intervention for at-risk carriers. Our study not only validates previously reported AAO PD loci in the COURAGE-PD dataset, but our meta-analysis with IPDGC data also provides the first genome-wide significant evidence that the known $BST1$ PD risk locus affects AAO. Interestingly, the variant, rs4698412, representing the $BST1$ locus, showed a similar large effect in COURAGE-PD and
IPDGC providing strong evidence that this is a bona-fide genetic locus for PD AAO. Lastly, using significant SNPs from the meta-analysis of COURAGE-PD and IPDGC case-control datasets, we demonstrate an inverse association between a PD PRS and AAO of PD.

Numerous genetic loci for familial and sporadic PD have been well characterized. The existence of overlapping loci between familial and sporadic PD suggests a complex but interconnected relationship between PD and age. Several meta-analyses of candidate genes and GWAS have previously recognized the BST1 locus as a locus that could influence the development of sporadic late-onset PD. Notably, the BST1 locus has been demonstrated to play a role in both Asian and European PD populations. The genome-wide significant BST1 variant, rs4698412 observed in our AAO meta-analysis, is also identical to the top BST1 variant reported in the latest PD GWAS meta-analysis. Interestingly, regional plots showed that the genome-wide significant variant, rs4694812 was neither the top genetic variant in the BST1 locus in IPDGC nor COURAGE AAO PD datasets. While, rs4694819 (r^2 with rs4694812<0.6) was the most significant variant in the COURAGE AAO dataset, rs11724635 (r^2 with rs4698412=1.0) was the most significant variant in the IPDGC AAO dataset (Figure 2B).

BST1 was first identified as a gene encoding a cell surface receptor on bone marrow stromal cells (bone marrow stromal cell antigen 1) with a role in promoting the growth of hematopoietic stromal cells. In addition to its role as a receptor, it also exhibits ADP-ribosyl cyclase activity, leading to the generation of cyclic ADP-ribose (cADPR), with a role in intrinsic Ca^{2+} regulation. The dual functional protein, a highly conserved glycosylphosphatidylinositol (GPI)-anchored glycoprotein (also known as CD157), is now known to be expressed in a wide variety of tissues, including vascular endothelium and follicular dendritic cells, with an ability to perform a wide variety of immune system and inflammation-related cellular functions. The initial identification of BST1/CD157 as a potential risk locus for sporadic late-onset PD in a GWAS in the Japanese population by
Satake *et al.* 2009, led to several functional studies aimed at deciphering its potential neuronal role in influencing PD phenotype\(^{22}\). Several knockout mouse model studies have shown that BST1 can influence social behaviour. However, the studies failed to demonstrate any influence on motor functioning, the cardinal feature which is impaired in patients with PD\(^{28,29}\). The eQTL analysis demonstrated a highly significant effect of the BST1 locus, rs4694812, on gene expression, with the A allele resulting in a decreased expression of CD38, a paralog of CD157, in a dose-dependent manner. CD38 and CD157 are contiguous gene duplicates, which belong to same gene family with a similar role of dual functional protein and an ability to modulate social behavior\(^{30,31}\). Interestingly, unlike CD157, CD38 knock out mice have been shown to have higher locomotor activity\(^{32}\). Furthermore, the highly significant increased expression of CD38 was mainly observed in the striatum, a region directly implicated in motor dysfunction in PD. Interestingly, a statistically underpowered brain imaging study in humans suggested that allele A of BST1 SNP rs4698412 leads to deficits in the right lingual gyrus region in the brain during the progression of PD\(^{33}\). This brain region is known to play a role in spatial orientation and visuospatial information processing. However, specific molecular and neuronal pathways influenced by altered CD38 expression in basal ganglia, with a potential role in triggering earlier AAO in sporadic PD, remain unclear.

SNCA is one of the most consistently observed significant loci in both early and late-onset PD and has been suggested to play a critical role in the age-related hierarchy of disease onset. While monogenic PD, often with relatively early onset, is attributed to rare point mutations and multiple copies of the SNCA gene, susceptibility to late-onset PD is attributed to common variants\(^ {6,10,34-36}\). In addition to being a leading locus in the largest GWAS of sporadic PD to date, the locus was also recently reported to be a top locus in influencing AAO in Europeans in a meta-analyses comprising IPDGC and 23andMe datasets (n=28,568)\(^7\). An SNP present towards the 3’ end (rs356203) of the SNCA gene was observed as the strongest genome-wide significant variant originating from the region (P=1.9x10\(^{-12}\)). Based on the
conditional analysis, the study also identified an independent signal at the 5’ end of the gene, rs983361 (P=6.8x10^-6). A recent GWAS of AAO in 5166 East Asian (Chinese) PD patients further reported a slightly weaker signal originating from another independent SNCA variant, rs3775458 (P=9.92x10^-7) \(^9\). Using the 1000 genome phase 3 dataset, we failed to detect any LD among the three variants in both European and East Asian populations (data not shown here). Upon screening of the SNCA locus in the COURAGE-PD dataset, we observed nominal significance of all the three variants (P\(_{\text{COURAGE (rs356203)}}=0.035\), P\(_{\text{COURAGE (rs3775458)}}=0.005\), P\(_{\text{COURAGE (rs983361)}}=0.022\), possibly suggesting a consistence influence of different loci around the SNCA region in determining AAO in different worldwide PD populations. The combining of our dataset with IPDGC further showed an independent genome-wide significant signal originating from the 3’ end of the SNCA gene (rs983361), as shown in the results section above. Notably, we also observed an independent signal at the 5’ end (rs356203). However, the variant was excluded for further interpretation due to high heterogeneity observed when combining IPDGC and COURAGE datasets (\(\beta(\text{SE})_{\text{COURAGE+IPDGC}}=-0.591(0.097)\), P\(_{\text{COURAGE+IPDGC}}=9.28 \times 10^{-10}\); \(I^2=61.9\%\)).

Another PD locus, TMEM175, was previously shown to reach genome-wide significance in an AAO study \(^37\). Similar to SNCA, our study also demonstrated replication of the TMEM175 locus in the COURAGE-PD AAO dataset with a nominal level of significance (P=0.018). The subsequent combining of the non-synonymous coding variant, rs3431186 (p.M393T), representing the genome-wide significant locus, in the IPDGC dataset with the COURAGE-PD, resulted in suggestive level of association without any underlying heterogeneity (P\(_{\text{COURAGE+IPDGC}}=2.64 \times 10^{-7}\); \(I^2=0.0\)). On the contrary, a recent East-Asian GWAS failed to observe any signal originating from the locus, possibly suggesting contribution of the locus mainly in the European populations.\(^9\) A previous study also reported a borderline significant association of the variant rs429358 representing APOE \(\varepsilon4\) locus with PD-AAO (P=5.69 x 10^-8) in a combined dataset (n=28568) comprising IPDGC and 23andMe.
The study, however, suggested that the association at the locus could be an age-related effect with a highly significant association with the age of controls ($P=1.49 \times 10^{-5}$). The variant also resulted in a suggestive association upon merging of COURAGE-PD European dataset only with the IPDGC dataset ($P=9.3 \times 10^{-7}$). These findings are consistent with the failure to detect association of APOE ε4 locus with PD-AAO in the recently reported East-Asian GWAS. However, being a longevity marker, the suggestive finding of APOE ε4 locus in Europeans must be interpreted with caution.

Our study has several strengths and limitations. Our study provides the largest independent dataset for testing the reliability of previously discovered AAO loci in a highly diverse and predominantly European population. Another strength of our study was the availability of data on AAO on all the study participants as opposed to age of diagnosis, often used as a proxy for AAO. One of the significant limitations of our findings was the lack of ready access to the recently published East-Asian AAO GWAS dataset that prevented us from drawing any conclusion on the validity of the novel BST1 locus in the East-Asian population. Likewise, the unavailability of the 23andMe dataset to us has precluded us from making an unequivocal claim on our BST1 findings. Hopefully, the inclusion of other datasets, such as 23andMe and East-Asian GWAS datasets, will help further to refine the signals originating from the BST1 locus. We also suggest that loci identified through meta-analysis in the COURAGE-PD dataset (PDZPH1P) and subsequent stratification by gender (RHEB, MTHFD1L, KNH3) and ethnicity (MOAP1/TMEM251, SUGCT) be meta-analysed with these unavailable datasets. Another limitation was our inability to conduct gene-gene interaction due to the limited sample size in the present study. The possibility of complex interactions among various loci on Chr 4 in modulating AAO cannot be ruled out. A recent study showed the association of several genome-wide significant loci on the X Chr with PD. It is also possible that some of these variants may also modulate AAO. However, due to potential analytic challenges from calling and imputation of X Chr genotypes, to model
uncertainty associated with random X Chr inactivation, we excluded the X Chr variants from the present analysis\textsuperscript{39}. And lastly, it is hoped that in the future, the availability of a larger dataset would enable us to integrate additional layers of genetic data, including rare and copy-number variants \textsuperscript{40}.

Our findings clearly highlight the importance of combining GWAS from diverse populations representative of worldwide populations to refine the genetic architecture underlying a complex trait like AAO. Our COURAGE-PD dataset suggests a role for additional pathways in addition to $\alpha$-synuclein mechanisms of modulating PD pathogenesis and influencing AAO in worldwide PD populations.

References

Figure 1. Manhattan plot of COURAGE-PD age at onset GWAS.

Figure 2. Manhattan plot of meta-analysis of COURAGE-PD and IPDGC age at onset GWAS.
### Table 1. eQTL lookup of BST1 SNP rs4698412 from GTEx in brain tissue

<table>
<thead>
<tr>
<th>Database</th>
<th>Gene Symbol</th>
<th>P-Value</th>
<th>NES</th>
<th>Tissue</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>GTEx</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ENSG00000004468</td>
<td>CD38</td>
<td>3.3e-16</td>
<td>-0.44</td>
<td>Caudate (basal ganglia)</td>
</tr>
<tr>
<td>ENSG00000004468</td>
<td>CD38</td>
<td>5.5e-15</td>
<td>-0.39</td>
<td>Cortex</td>
</tr>
<tr>
<td>ENSG00000004468</td>
<td>CD38</td>
<td>1.4e-13</td>
<td>-0.39</td>
<td>Nucleus accumbens (basal ganglia)</td>
</tr>
<tr>
<td>ENSG00000004468</td>
<td>CD38</td>
<td>1.4e-11</td>
<td>-0.32</td>
<td>Putamen (basal ganglia)</td>
</tr>
<tr>
<td>ENSG00000004468</td>
<td>CD38</td>
<td>6.6e-6</td>
<td>-0.21</td>
<td>Frontal Cortex (BA9)</td>
</tr>
<tr>
<td>ENSG000000237765</td>
<td>FAM200B</td>
<td>1.0e-5</td>
<td>0.23</td>
<td>Cerebellar Hemisphere</td>
</tr>
<tr>
<td>ENSG00000004468</td>
<td>CD38</td>
<td>1.2e-5</td>
<td>-0.26</td>
<td>Anterior cingulate cortex (BA24)</td>
</tr>
<tr>
<td>ENSG000000237765</td>
<td>FAM200B</td>
<td>1.4e-5</td>
<td>0.23</td>
<td>Cortex</td>
</tr>
<tr>
<td>ENSG00000004468</td>
<td>CD38</td>
<td>1.4e-5</td>
<td>-0.22</td>
<td>Hypothalamus</td>
</tr>
<tr>
<td><strong>UKBEC</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ENSG000000118564</td>
<td>FBXL5</td>
<td>5.1e-7</td>
<td>NA</td>
<td>Occipital Cortex</td>
</tr>
<tr>
<td>ENSG000000004468</td>
<td>CD38</td>
<td>7.1e-6</td>
<td>NA</td>
<td>Putamen (basal ganglia)</td>
</tr>
<tr>
<td>ENSG000000004468</td>
<td>CD38</td>
<td>2.1e-5</td>
<td>NA</td>
<td>Hippocampus</td>
</tr>
<tr>
<td>ENSG000000137449</td>
<td>CPEB2</td>
<td>2.4e-5</td>
<td>NA</td>
<td>Medulla</td>
</tr>
</tbody>
</table>

GTEx: Genotype-Tissue Expression Project, NA: Not available, UKBEC: UK Brain Expression Consortium
### Genome-wide Association and Meta-analysis of Age-at-Onset in Parkinson Disease:
**Evidence From COURAGE-PD Consortium**
*Neurology* published online May 26, 2022
DOI 10.1212/WNL.0000000000200699

This information is current as of May 26, 2022

<table>
<thead>
<tr>
<th>Updated Information &amp; Services</th>
<th>including high resolution figures, can be found at:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><a href="http://n.neurology.org/content/early/2022/05/26/WNL.0000000000200699.full">http://n.neurology.org/content/early/2022/05/26/WNL.0000000000200699.full</a></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Subspecialty Collections</th>
<th>This article, along with others on similar topics, appears in the following collection(s):</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>Association studies in genetics</strong>&lt;br&gt;<a href="http://n.neurology.org/cgi/collection/association_studies_in_genetics">http://n.neurology.org/cgi/collection/association_studies_in_genetics</a></td>
</tr>
<tr>
<td></td>
<td><strong>Basal ganglia</strong>&lt;br&gt;<a href="http://n.neurology.org/cgi/collection/basal_ganglia">http://n.neurology.org/cgi/collection/basal_ganglia</a></td>
</tr>
<tr>
<td></td>
<td><strong>Burden of disease</strong>&lt;br&gt;<a href="http://n.neurology.org/cgi/collection/burden_of_disease">http://n.neurology.org/cgi/collection/burden_of_disease</a></td>
</tr>
<tr>
<td></td>
<td><strong>Case control studies</strong>&lt;br&gt;<a href="http://n.neurology.org/cgi/collection/case_control_studies">http://n.neurology.org/cgi/collection/case_control_studies</a></td>
</tr>
<tr>
<td></td>
<td><strong>Parkinson's disease/Parkinsonism</strong>&lt;br&gt;<a href="http://n.neurology.org/cgi/collection/parkinsons_disease_parkinsonism">http://n.neurology.org/cgi/collection/parkinsons_disease_parkinsonism</a></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Permissions &amp; Licensing</th>
<th>Information about reproducing this article in parts (figures, tables) or in its entirety can be found online at:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><a href="http://www.neurology.org/about/about_the_journal#permissions">http://www.neurology.org/about/about_the_journal#permissions</a></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Reprints</th>
<th>Information about ordering reprints can be found online:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><a href="http://n.neurology.org/subscribers/advertise">http://n.neurology.org/subscribers/advertise</a></td>
</tr>
</tbody>
</table>