COGNITIVE NEUROSCIENCE

The *FAM171A2* gene is a key regulator of progranulin expression and modifies the risk of multiple neurodegenerative diseases

Wei Xu[1,2]*, Si-Da Han[1]*, Can Zhang[3], Jie-Qiong Li[4], Yan-Jiang Wang[5], Chen-Chen Tan[2], Hong-Qi Li[1], Qiang Dong[1], Cui Mei[1], Lan Tan[2], Jin-Tai Yu[1†]

Progranulin (PGRN) is a secreted pleiotropic glycoprotein associated with the development of common neurodegenerative diseases. Understanding the pathophysiological role of PGRN may help uncover biological underpinnings. We performed a genome-wide association study to determine the genetic regulators of cerebrospinal fluid (CSF) PGRN levels. Common variants in region of *FAM171A2* were associated with lower CSF PGRN levels (rs708384, $P = 3.95 \times 10^{-12}$). This was replicated in another independent cohort. The rs708384 was associated with increased risk of Alzheimer's disease, Parkinson's disease, and frontotemporal dementia and could modify the expression of the *FAM171A2* gene. FAM171A2 was considerably expressed in the vascular endothelium and microglia, which are rich in PGRN. The in vitro study further confirmed that the rs708384 mutation up-regulated the expression of FAM171A2, which caused a decrease in the PGRN level. Collectively, genetic, molecular, and bioinformatic findings suggested that *FAM171A2* is a key player in regulating PGRN production.

INTRODUCTION

Progranulin (PGRN) is a secreted pleiotropic glycoprotein expressed in the central nervous system (CNS) and peripheral tissues. Its deficiency has been associated with the development of frontotemporal dementia (FTD) (*1*), Alzheimer's disease (AD) (*2*), and Parkinson's disease (PD) (*3*). Loss-of-function animal models revealed that potential pathogenesis mechanisms may involve neuroinflammation (*4*), autophagy (*5*), and cell signaling. PGRN overexpression provided protection against pathological protein deposition and toxicity and inhibited phenotypic progression in AD- and PD-like disease models (*3*, *6*). Recently, PGRN in cerebrospinal fluid (CSF) is found to be increased during the course of AD (*7*). Together, these lines of evidence indicate that PGRN dysregulation might be involved in the pathogenesis of neurodegenerative diseases and could be a promising therapeutic target. Mounting evidence illustrates that PGRN levels might be modified by genetic variants (*8*, *9*), including the encoding gene *GRN*, which could partially explain variability of the protein (*10*). Previous genome-wide screens have identified multiple loci outside the *GRN* [*SORT1* (*8*) and *PSRC1* (*9*)] as regulators of blood PGRN levels. Nonetheless, to date, no genome-wide association study (GWAS) has been conducted to reveal the genetic modulators of PGRN in CNS. The search for CNS PGRN regulators is of tremendous importance, especially considering that (i) PGRN levels might be regulated differently in the peripheral blood and CNS (*11*) and (ii) increasing PGRN levels as a neuroprotective approach might be hindered by its peripheral effects of promoting carcinogenesis (*12*)

and obesity (*13*). This study aimed to identify genetic modifiers of CSF PGRN levels. We first conducted a GWAS of CSF PGRN levels from 1362 adults without dementia and successfully identified an independent locus within the *FAM171A2* gene region that showed genome-wide significant association with CSF PGRN levels. Then, we used bioinformatics and cellular approaches to annotate the functionality of the identified variant and investigate its potential mechanisms underlying neurodegenerative diseases. On the basis of these findings, we propose that the *FAM171A2* gene is a key regulator of CSF PGRN expression and modifies risk of neurodegenerative diseases.

RESULTS

This study aimed to detect genetic modifiers of CSF PGRN levels and better understand the role of PGRN in biological pathways relevant to neurodegenerative diseases. To achieve this goal, we analyzed genetic data and CSF PGRN levels from 432 individuals without dementia from the Alzheimer's Disease Neuroimaging Initiative (ADNI) in a GWAS. Then, we sought to replicate our findings in an independent cohort composed of 930 Chinese adults without dementia.

CSF PRGN levels were influenced by sex, educational level, and *APOE* ε4

In the GWAS cohort, CSF PGRN levels were significantly influenced by sex ($P = 3.87 \times 10^{-5}$, β = 0.003) and potentially influenced by *APOE* ε4 allele ($P = 0.08$, β = -8.73×10^{-4}) and educational level ($P = 0.12$, β = -1.77×10^{-4}). CSF PGRN levels showed no association with either age ($P = 0.48$) or clinical diagnosis [healthy controls (HC) versus subjects with mild cognitive impairment (MCI) subjects, $P = 0.21$].

Common variants in *FAM171A2* gene were associated with CSF PGRN levels

A total of 432 individuals without dementia (157 HC and 275 subjects with MCI) participated in the GWAS. Age ($P = 1.38 \times 10^{-6}$) and *APOE* ε4 carrier percentage ($P = 1.37 \times 10^{-6}$) differed between

[1]Department of Neurology and Institute of Neurology, Huashan Hospital, State Key Laboratory of Medical Neurobiology and MOE Frontiers Center for Brain Science, Shanghai Medical College, Fudan University, Shanghai, China. [2]Department of Neurology, Qingdao Municipal Hospital, Qingdao University, Qingdao, China. [3]Genetics and Aging Research Unit, McCance Center for Brain Health, Mass General Institute for Neurodegenerative Diseases (MIND), Department of Neurology, Massachusetts General Hospital, Harvard Medical School, Charlestown, MA, USA. [4]Department of Neurology, The Affiliated Hospital of Qingdao University, Qingdao, China. [5]Department of Neurology and Center for Clinical Neuroscience, Daping Hospital, Third Military Medical University, Chongqing, China.
*These authors contributed equally to the present work.
†Corresponding author. Email: jintai_yu@fudan.edu.cn

the two diagnostic groups (Table 1). After adjusting for sex, educational level, *APOE* ε4 allele, and the first three principal components, no inflation due to population stratification was identified (genomic inflation factor λ = 1.012). A total of 18 single-nucleotide polymorphisms (SNPs) (5 in *FAM171A2*, 12 in *ITGA2B*, and rs5848 in *GRN*) on chromosome 17 were significantly associated with CSF PGRN levels, with rs708384 (genotyped) in *FAM171A2* gene yielding the most significant signal ($P = 5.09 \times 10^{-12}$) (Table 2 and Fig. 1A). Located in the transcription factor (TF) binding site (intron 1) of *FAM171A2*, rs708384 also survived the permutation-based empiric corrections for multiple testing (empiric $P = 0.0002$; permutation-based corrected empiric family-wise error rate controlled at 0.0002) (Fig. 1B).

In addition to the loci in *FAM171A2*, 13 SNPs in *ITGA2B* and *GRN* also showed significant associations with CSF PGRN levels (empiric $P = 0.0002$; permutation-based corrected empiric family-wise error rate controlled at <0.005). Notably, eight loci showed high linkage disequilibrium (LD) ($r^2 \geq 0.8$) with rs708384. After controlling for the rs708384 genotype, the association signals for all loci were markedly attenuated ($P > 0.1$), indicating that the relationships in these loci were driven by rs708384 (Table 2 and Fig. 1, C and D). The analysis also identified 149 suggestive loci; 92% were located in *FAM171A2*, *ITGA2B*, and *GPATCH8* genes of chromosome 17. However, they did not survive after permutation-based empiric corrections (table S1). Accordingly, rs708384 in *FAM171A2* gene was selected as the top SNP, for which the minor allele (A allele, MAF = 0.4074) was significantly linked to lower CSF PGRN levels in a dose-dependent manner ($P = 3.95 \times 10^{-12}$) (Fig. 2A).

Subgroup and sensitivity analyses

In the subgroup analysis, the minor allele (A) of rs708384 was associated with lower CSF PGRN levels in different strata according to sex (male, $P = 2.67 \times 10^{-6}$; female, $P = 1.76 \times 10^{-6}$) and baseline diagnosis (HC, $P = 0.017$; MCI, $P = 1.36 \times 10^{-11}$). Notably, the effect size drastically increased in the HC population (beta for HC = 100.9; beta for MCI = 8.77×10^{-5}) (fig. S2A). As CSF PGRN levels were reported to be influenced by age, diagnosis, and rs5848 in *GRN* gene (*11*), sensitivity analysis was performed to further adjust these confounders. Modifying for age and baseline diagnosis did not significantly change the results ($P = 3.51 \times 10^{-12}$). After adjusting for rs5848, a strong, although less significant, association for rs708384 with CSF PGRN levels was found ($P = 8.28 \times 10^{-5}$) (fig. S2B). Excluding those who developed AD within 3 years from baseline did

not change the result, and rs708384 remained the top SNP ($n = 359$, $P = 4.73 \times 10^{-10}$) (fig. S2C). According to Aβ pathology status, subgroup analysis showed the same dose-response relationship for both groups ($P < 5.00 \times 10^{-4}$ for Aβ-negative group, $P = 1.07 \times 10^{-7}$ for Aβ-positive group). Furthermore, we added the rs5848 genotype, which did not pass the GWAS quality control (QC) stage, as a covariate in the regression model. After adjusting for the rs5848 genotype and other covariates mentioned above, the association of PGRN levels with rs708384 remained significant ($n = 337$; $P = 0.0009$).

Specificity of the association with PGRN

We tested whether the observed association may be influenced by other candidate proteins related to neurodegeneration. First, the significance remained unchanged after including CSF levels of Aβ42, T-tau, P-tau, α-synuclein (α-SYN), neurofilament light chain (NFL), and secreted triggering receptor expressed on myeloid cells 2 (sTREM2) as covariates. Next, no correlations with Aβ42 or P-tau biomarkers were identified (r^2 ranging from 0.04 to 0.1) although weak correlations were observed of CSF PGRN with T-tau ($P = 0.04$, $r = 0.1$), α-SYN ($P = 0.0003$, $r = 0.36$), NFL ($P = 0.03$, $r = 0.22$), and sTREM2 ($P = 1.134 \times 10^{-8}$, $r = 0.27$) levels. Furthermore, no associations of rs708384 were revealed with any of the abovementioned proteins when they were separately used as the endophenotype. Together, these results supported the specificity of our results (table S2).

Variability in CSF PGRN levels explained by genetic variants

We used a genome-partitioning analysis to estimate the proportion of variance in CSF PGRN levels explained by chromosomes, significant genes, and loci. All genotyped and imputed variants on chromosome 17 can explain approximately 17.4% of the variability in the CSF PGRN levels ($P < 0.05$) (Fig. 3A). Most significant loci were in high LD with rs708384 ($r^2 \geq 0.8$) (Fig. 3B). The SNPs in the *GRN* region explained 13.3% of the variability in CSF PGRN levels ($P = 8.13 \times 10^{-8}$), suggesting that genetic variants other than the *GRN* gene also played significant roles. Rs708384 was identified to explain 9.1% of the variability with the smallest P value ($P = 2.60 \times 10^{-10}$) (Fig. 3C). Although rs5848 was not genotyped in ADNI, the 1000 Genomes Project data indicated that rs5848 has only low-to-moderate LD ($r^2 = 0.6$), with rs708384 in the CEU (Utah residents with Northern and Western European ancestry) population (Fig. 3D) and no loci showed LD with rs5848. Thus, it can be reasonably implied that the observed associations for other significant loci were more likely to be influenced by rs708384.

Table 1. Characteristics of discovery and replication samples. CV, coefficient of variation.

Baseline diagnosis	ADNI (discovery sample)			CABLE (replication sample)			
	Total	HC	MCI	Total	AA	AC	CC
n	432	157	275	930	93	392	445
Age, years, mean ± SD	72.8 ± 6.8	74.7 ± 5.5	71.8 ± 7.3	62.6 ± 10.4	62.1 ± 9.9	62.9 ± 10.2	62.4 ± 10.7
Male/female, *n*	252/180	85/72	167/108	380/550	46/47	157/235	177/268
Education, years	16.2 ± 2.8	16.4 ± 2.8	16.0 ± 2.8	9.7 ± 5.2	10.6 ± 10.1	9.5 ± 4.5	9.7 ± 4.2
APOE ε4 carrier, %	37.5%	22.3%	46.5%	16.3%	15.05%	17.98%	16.34%
CSF PGRN, mean ± SD, pg/ml	1600.3 ± 612	1551.2 ± 368.6	1628.4 ± 713.9	1739 ± 372	1687 ± 453	1696 ± 335	1787 ± 380
CV of CSF PGRN, mean ± SD, %	2.3 ± 1.7	2.4 ± 1.9	2.2 ± 1.7	2.95 ± 2.69	2.74 ± 2.40	2.90 ± 2.71	3.04 ± 2.72

Table 2. SNPs significantly ($P < 5 \times 10^{-8}$) associated with CSF PGRN levels. CHR, chromosome; BP, base pair; Alt, altered allele; UTR, untranslated region; NA, not available.

CHR	SNP	LD (r^2)	BP (GRCh37)	Alt	Observed MAF	Source	Closest gene	SNP location	Regulome DB score*	P values†			
										P1	P2	P3	P4
17	**rs708384‡**	**1**	**42437682**	**A**	**0.4074**	**Genotyping**	***FAM171A2***	**Intron (TF binding site)**	**5**	**5.09×10^{-12}**	**0.0002**	**0.0002**	**…**
17	rs708383‡	1	42437754	A	0.4074	Imputation	*FAM171A2*	Intron	5	5.09×10^{-12}	0.0002	0.0002	…
17	rs34603233	NA	42455022	T	0.4201	Imputation	*ITGA2B*	Intron	5	6.48×10^{-12}	0.0002	0.0002	0.4561
17	rs850730	0.89	42454463	C	0.4213	Imputation	*ITGA2B*	Intron	5	1.90×10^{-11}	0.0002	0.0002	0.8554
17	rs850729	0.88	42454806	T	0.4213	Imputation	*ITGA2B*	Intron	NA	1.90×10^{-11}	0.0002	0.0002	0.8554
17	rs850736	0.93	42446420	C	0.4225	Imputation	*ITGA2B*	NA	NA	2.47×10^{-11}	0.0002	0.0002	0.9572
17	rs850734	0.91	42449025	G	0.4225	Imputation	*ITGA2B*	NA	5	2.47×10^{-11}	0.0002	0.0002	0.9572
17	rs5910	0.90	42449789	A	0.4225	Genotyping	*ITGA2B*	Synonymous	5	2.47×10^{-11}	0.0002	0.0002	0.9572
17	rs35905006	NA	42451281	C	0.4225	Imputation	*ITGA2B*	Intron	6	2.47×10^{-11}	0.0002	0.0002	0.9572
17	rs850733	0.89	42451305	A	0.4225	Imputation	*ITGA2B*	Intron	NA	2.47×10^{-11}	0.0002	0.0002	0.9572
17	rs5911	0.89	42453065	C	0.4225	Imputation	*ITGA2B*	Missense	2b	2.47×10^{-11}	0.0002	0.0002	0.9572
17	rs850732	0.89	42454243	T	0.4225	Imputation	*ITGA2B*	Intron	NA	2.47×10^{-11}	0.0002	0.0002	0.9572
17	rs708382	0.86	42442344	C	0.4340	Genotyping	*FAM171A2*	NA	6	2.91×10^{-11}	0.0002	0.0002	0.7185
17	rs850737	0.86	42443784	C	0.4340	Imputation	*FAM171A2*	NA	5	2.91×10^{-11}	0.0002	0.0002	0.7185
17	rs850738	0.73	42434630	G	0.4317	Imputation	*FAM171A2*	Intron	NA	4.89×10^{-11}	0.0002	0.0004	0.2513
17	rs11389756	NA	42448210	AG	0.4699	Imputation	*ITGA2B*	NA	6	3.62×10^{-11}	0.0002	0.0022	0.6897
17	rs850731	0.71	42454270	C	0.4699	Imputation	*ITGA2B*	Intron	5	3.62×10^{-11}	0.0002	0.0022	0.6897
17	rs5848	0.60	42430244	T	0.3275	Imputation	*GRN*	3'-UTR	5	1.22×10^{-11}	0.0002	0.0074	0.6925

*Lower scores indicate increasing evidence for a variant to be located in a functional region (www.regulomedb.org).   †P1 = P value for GWAS; P2 = empiric P value; P3 = permutation-based corrected empiric familywise error rate; P4 = empiric P value after controlling for rs708384   ‡rs708383 is in complete LD with rs708384 ($r^2 = 1$) according to the HaploRegv4.1 online database.

Replication in an independent cohort and meta-analysis

To validate our finding that rs708384, as the top signal, was associated with CSF PGRN levels, we measured CSF PGRN levels and genotyped rs708384 in a larger, independent cohort of 930 Northern Han Chinese individuals without dementia (Table 1). In this cohort, the participants' age ranged from 40 to 88 years (mean = 62.6 years, SD = 10.4 years) at the time of CSF extraction. Both sexes were well-represented, including 380 men and 550 women. The mean CSF PGRN levels was 1739 pg/ml (SD = 372 pg/ml), with mean intrabatch coefficient of variation (CV) of 2.95% and mean interbatch CV of 3.82%. Consistent with the result of the discovery cohort, the A allele of rs708384 was significantly associated with lower CSF PGRN levels in a linear regression model accounting for age, sex, educational level, *APOE* ε4 genotype, MMSE (mini-mental state examination) score at baseline, and CV ($P = 7.47 \times 10^{-9}$, Fig. 2B).

To exclude the potential influence of rs5848 in the *GRN* gene, we genotyped and included it as a covariate. Single regression analysis indicated that rs5848 was significantly associated with CSF PGRN levels ($P < 0.005$). However, the significance of rs5848 was no longer present ($P = 0.88$) when rs708384 was added as a covariate. In contrast, the association of rs708384 with the CSF PGRN level still reached borderline significance when controlling for rs5848 ($P = 0.10$). This may be because rs5848 showed relatively high but not complete LD with rs708384 in the Chinese population ($r^2 \approx 0.8$, Fig. 3D). The additive model (AA versus CC) indicated that the association for rs708384 remained significant ($P < 0.05$) after controlling for rs5848

and other covariates (table S3). These results demonstrated that the influence of rs708384 on the CSF PGRN level was not determined by rs5848. The meta-analysis of the two-stage findings further strengthened the association of rs708384 with the CSF PGRN level ($P = 1.74 \times 10^{-18}$). We identified that the loci previously associated with blood PGRN levels had no significant relationships with CSF PGRN level, suggesting that the genetic mechanisms of PGRN regulation may be different in the peripheral system and CNS (table S3).

Functional annotation of the top signal in the *FAM171A2* gene region

Enhancer enrichment analysis showed that the tagged variants by rs708384 were significantly enriched in specific brain regions (e.g., middle hippocampus, inferior temporal lobe, and prefrontal lobe) (table S4). This suggests that the variants might be linked to the regulation of gene expression in these sites. We also found significant differential expression of *FAM171A2* in the temporal cortex ($P = 0.004$) in AD (fig. S3). Moreover, the expression quantitative trait loci (eQTL) analyses showed that rs708384 ($z = 4.74$; $P = 2.5 \times 10^{-6}$) and some of its tagged variants can potentially influence *GRN* expression in both the blood (table S5) and specific brain regions (anterior cingulate cortex and frontal cortex) (table S6). In addition to *GRN*, rs708384 was also associated with the expression of *FAM171A2* and *ITGA2B* in multiple sites, such as the frontal lobe (table S7). Together, the evidence points to the hypothesis that rs708384, a TF binding site, may functionally modulate the expression
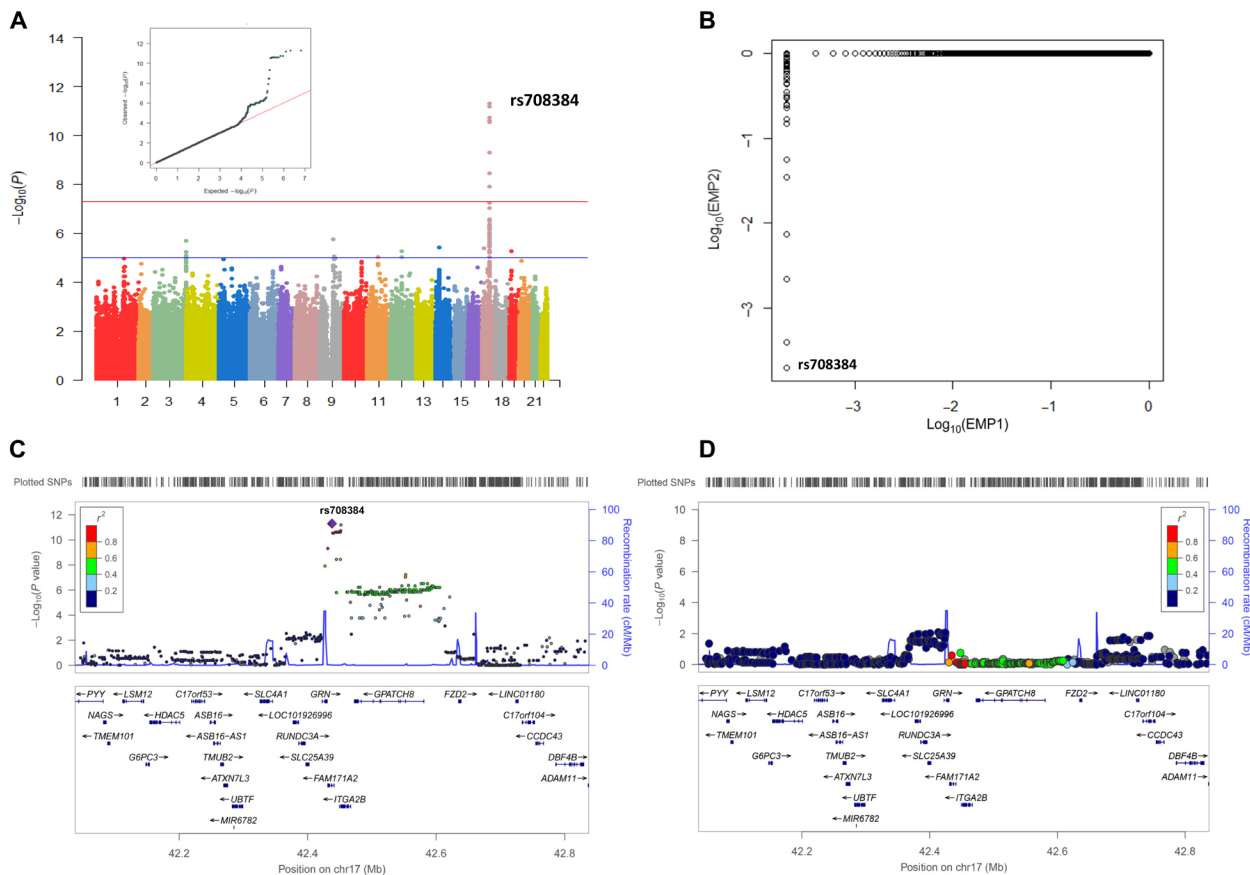
Fig. 1. GWAS results and regional plots for associations with CSF PGRN levels. (**A**) Manhattan plots [showing the $-\log_{10}$ (*P* value) for individual SNP] and q-q plot. (**B**) Association results after permutation test. EMP1, empiric *P* value; EMP2, permutation-based corrected empiric family-wise error rate. (**C**) Regional association results for the *GRN-FAM171A2-ITGA2B* region. (**D**) Regional association results after controlling for rs708384.
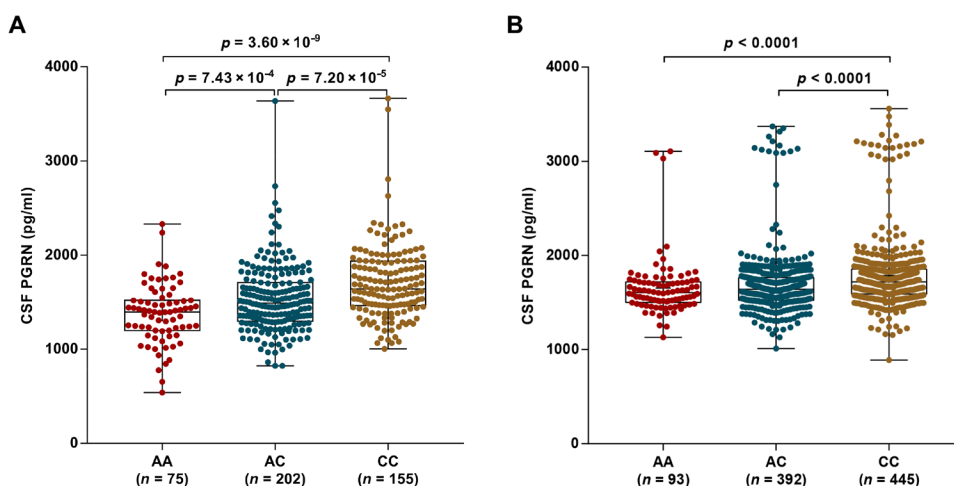

Fig. 2. CSF PGRN levels as a function of rs708384 genotype in two samples. (**A**) The minor allele (A allele, MAF = 0.41) of rs708384 was significantly associated with lower CSF levels of PGRN in a dose-dependent manner. (**B**) CSF PGRN levels were compared across the AA, AC, and CC genotypes of rs708384 in a larger independent cohort of 930 nondemented Chinese participants to validate the initial observed top signal. A significant association of decreasing CSF PGRN levels with increased minor allele (A) dose of rs708384 was observed, independent of age, gender, education, *APOE* ε4 genotype, MMSE score at baseline, CV, and rs5848 genotype.

of *GRN* and its downstream gene *FAM171A2*. We identified that the expression of *GRN* was highly correlated with that of *FAM171A2* ($P = 1.1 \times 10^{-17}$), in either the brain ($P = 5.3 \times 10^{-10}$) or blood ($P = 3.1 \times 10^{-6}$).

FAM171A2 and *GRN* genes were clustered in common pathways

There was a significant excess of enriched categories: 96 categories in the Protein Analysis Through Evolutionary Relationships (PANTHER) analysis and 107 categories in the Consensus PathDB (CPDB) analysis. There were 34 significant categories that overlapped in the two analyses (table S8). These categories were primarily related to regulation of nervous system development, molecular transport, signal transduction, cell-to-cell adhesion, and response to stimuli (Fig. 4A). The gene network analysis revealed a total of 31 significant categories ($P < 0.001$) forming four clusters of genes (Fig. 4B). As expected, the most significant genes (*GRN*, *FAM171A2*, and *ITGA2B*) were clustered together (Fig. 4B). The Gene Ontology
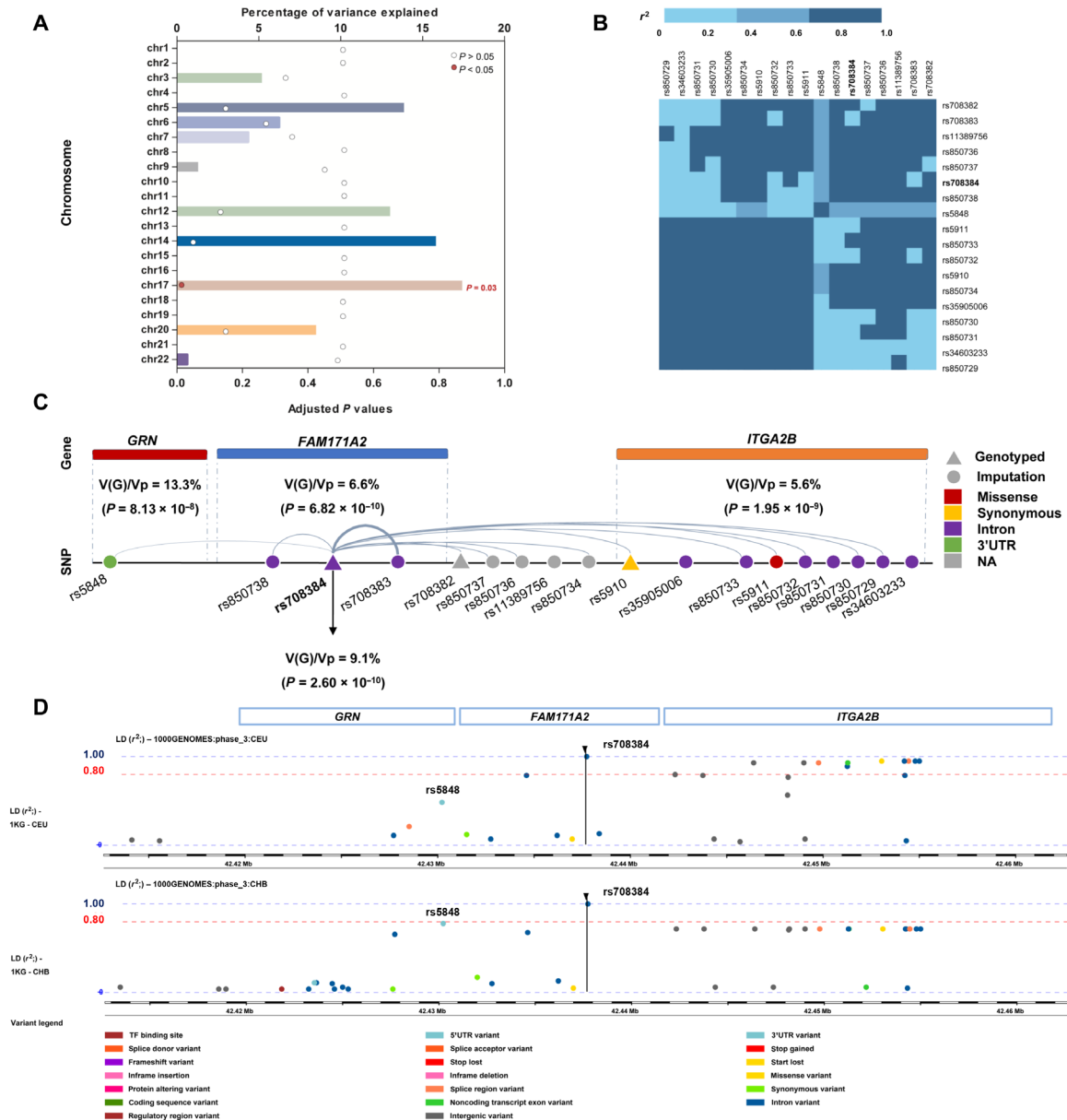
Fig. 3. Variability in CSF PGRN levels explained by genetic variants. (**A**) Chromosome 17 explained approximately 17.4% of the variability in the CSF levels of PGRN. (**B**) Most of the significant loci were in LD with rs708384. (**C**) SNPs in the *GRN* region explained the most but not all of the variability in CSF PGRN. Analysis of *FAM171A2* and *ITGA2B* region showed that these two regions explained 9.1 and 5.6% of the variability in CSF levels of PGRN, respectively. rs708384 was shown to explain 9.1% of the variability. (**D**) rs5848 was only in low-to-moderate LD with rs708384 in CEU population ($r^2 = 0.6$), and no loci were in LD with rs5848. Nonetheless, rs5848 was in high LD with rs708384 in CHB population ($r^2 \approx 0.8$).
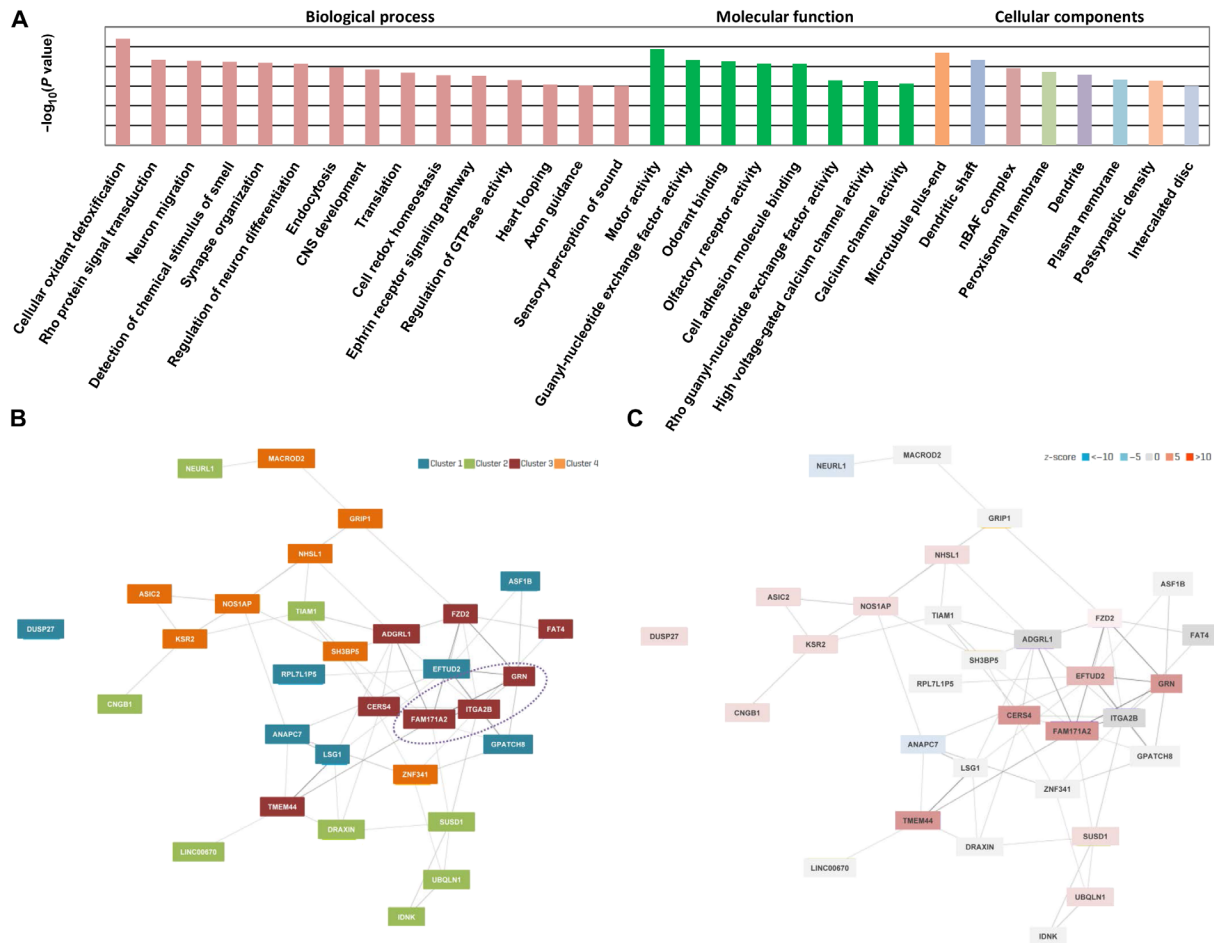
Fig. 4. GO and pathway analysis. (**A**) Functional categories were identified that were significantly enriched ($P < 0.001$), primarily including those involving regulation of nervous system development, molecular transport, signal transduction, cell-cell adhesion, and response to stimuli. GTPase, guanosine triphosphatase. (**B**) Four clusters were identified in the gene network analysis. The most significant genes (*FAM171A2* and *GRN*) were clustered together. (**C**) The highest z-scores were achieved by *FAM171A2* and *GRN*. z-score was used by GeneNetwork Assisted Diagnostic Optimization to prioritize the candidate genes: Gene with a higher z-score is more likely to explain the phenotype.

(GO) items for this specific cluster included all the abovementioned categories ($P < 0.001$) and another two items: regulation of blood pressure ($P = 3.84 \times 10^{-4}$) and interleukin-8 secretion ($P = 4.20 \times 10^{-4}$).

Among 16 pathways [12 from REACTOME and 4 from Kyoto Encyclopedia of Genes and Genomes (KEGG)] that achieved the statistical significance level ($P < 0.005$), the L1 cell adhesion molecule (L1CAM) interaction pathway was the most significant (REACTOME, $P = 2.49 \times 10^{-6}$), for which the highest z-scores were achieved by *FAM171A2* and *GRN* gene (Fig. 4C). Most other significant pathways were also primarily related to cell-to-cell interactions (table S9).

FAM171A2 considerably expressed in the cerebral vascular endothelium and microglia

It was essential to determine the location of FAM171A2 in the brain before investigating its functions. Therefore, we conducted immuno-histochemistry (IHC) and immunofluorescence (IF) staining on brain sections (cortex and hippocampus) of 10-week-old male C57BL/6 mice. Evident DAB (Diaminobenzidine) staining was observed around the vessels (Fig. 5, A1 and B1, tilted arrows) and on cells in a similar form to microglia (Fig. 5, A2 and B2, horizontal arrows). The IF

staining of FAM171A2 colocalized with CD31 and IBA1 on the brain sections (Fig. 5, C1, C2, D1, and D2). Five independent experiments were conducted, and the abovementioned results indicate that FAM171A2 was considerably expressed in the cerebral vascular endothelium and microglia.

Targeting rs708384 regulates expression of *FAM171A2* and *GRN* genes in vitro

First, we conducted dual-luciferase reporter assay in human embryonic kidney 293 (HEK293) cells to evaluate the regulation of rs708384 to FAM171A2 (Fig. 6A). A total of four independent experiments were conducted. We observed that the c > a mutation significantly improved the ratio of firefly/Renilla luciferase reporter gene expression compared to the wild type. This indicates that c > a stimulated the expression of FAM171A2 (Fig. 6B).

Next, we examined the influences of FAM171A2 overexpression on GRN/PGRN levels. Cultured human umbilical vein endothelial cells were used, and five independent experiments were conducted. A significant decrease in the intracellular GRN level (Fig. 6, C and D) and the supernatant PGRN level (Fig. 6E) was observed after FAM171A2 overexpression, which was consistent with the GWAS
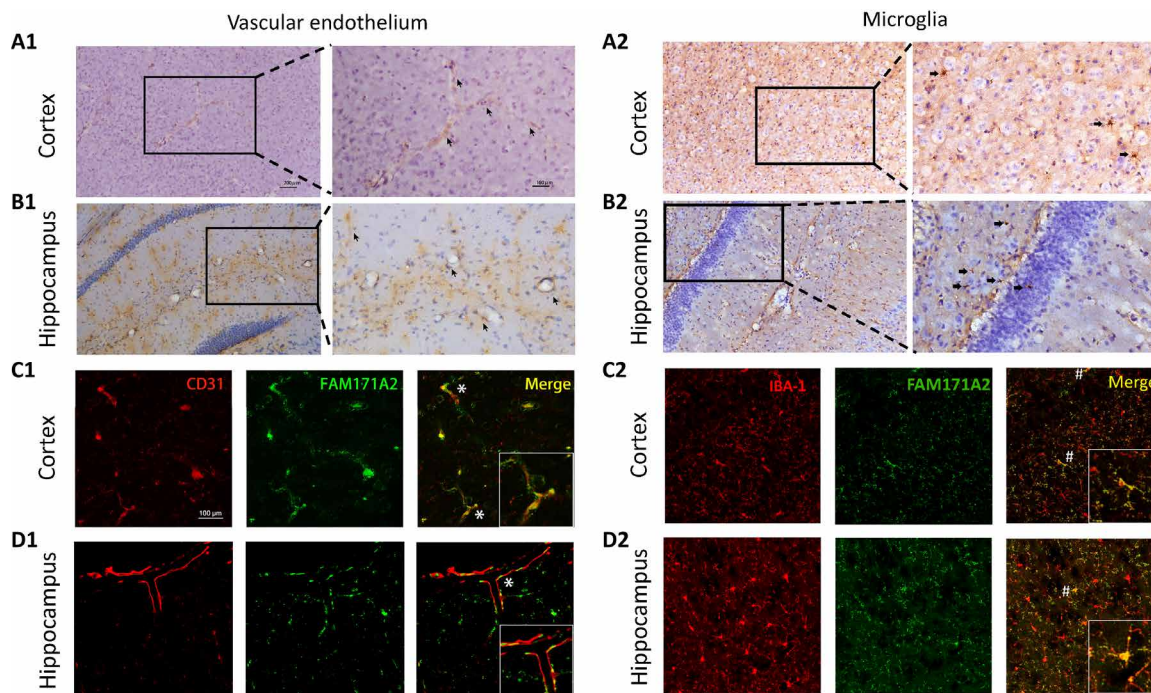
Fig. 5. FAM171A2 high expression on cerebral vascular endothelium and microglia. (**A** and **B**) The IHC staining of FAM171A2 on mouse cortex and hippocampi. The DAB staining along and around the cerebral vascular (A1 and B1) was marked by tilted arrows. The DAB staining on the cells in a similar form of microglia (A2 and B2) was marked by horizontal arrows. (**C** and **D**) The IF staining of FAM171A2 with the CD31 and IBA1 antibodies on mouse cortex and hippocampus. Their colocalization was marked by "*" and "#." $n = 5$ mice in these experiments.
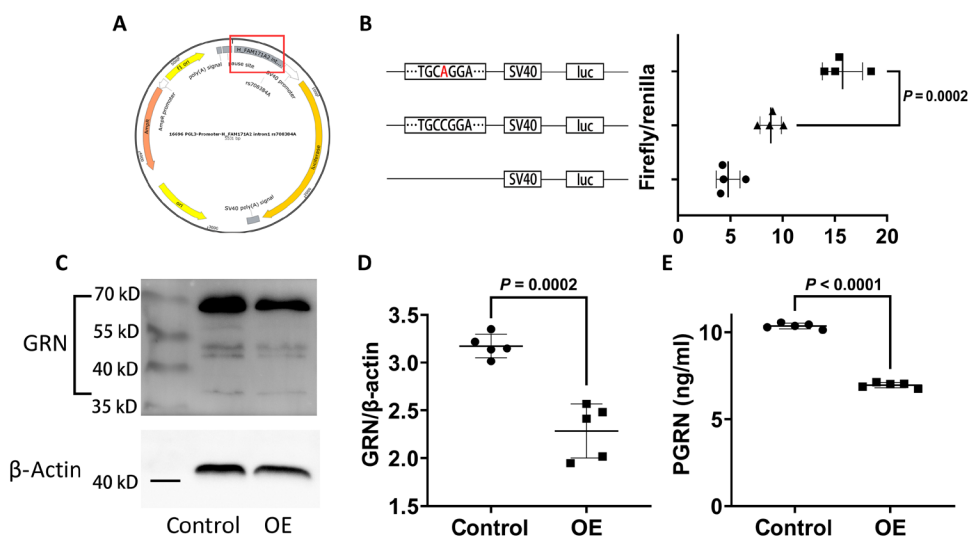
Fig. 6. rs708384 stimulates FAM171A2 expression and subsequently inhibits GRN/PGRN level. (**A**) The structure of firefly luciferase reporter plasmid. The sequence containing rs708384 was labeled by a red square. (**B**) The plasmids with or without rs708384 (c > a), including the empty control, had different expressing levels of firefly luciferase after transfection into the HEK293 cells. The luminous intensity was calibrated by Renilla luciferase, and the result was presented by ratio of firefly/Renilla luciferase. $n = 4$ per group, the data were analyzed by the one-way analysis of variance (ANOVA) ($P < 0.0001$) followed by the Tukey post hoc test ($P = 0.0002$ c > a versus wild type). (**C** and **D**) The change of intracellular GRN after FAM171A2 overexpression. $n = 5$ per group, the data were analyzed by t test ($P = 0.0002$). (**E**) The change of the supernatant PGRN levels by FAM171A2 overexpression (OE). $n = 5$ per group, the data were analyzed by t test ($P < 0.0001$).

data. Meanwhile, FAM171A2 was knocked down by small interfering RNA (siRNA), and we observed a significant increase in the intracellular GRN level (fig. S4, A and B). The efficiency of FAM171A2 siRNA and overexpression were verified by Western blot analysis (fig. S4, C and D). These results showed that FAM171A2 overexpression led to decreased GRN/PGRN levels, while knockdown had the opposite effect.

Impact of rs708384 on phenotypes of neurodegenerative disease

Last, we tested whether rs708384 had relationships with Aβ pathologic features, CSF biomarkers of neuronal injury, cognitive functions, brain structure/metabolism using the ADNI database, and the risk of neurodegenerative diseases [including AD, PD, amyotrophic lateral sclerosis (ALS), and FTD] using previous published GWAS summary statistics. The "AA" genotype of the rs708384 was significantly linked to higher CSF tau levels (fig. S5A) and poorer global cognition (fig. S5, B and C). There were no significant influences of rs708384 genotypes on Aβ pathologic features, CSF P-tau levels, and brain structure/metabolism of the regions of interest. Ten variants were tagged by rs708384 with $r^2 \geq 0.8$. In the large-scale GWAS dataset, these loci were significantly related to AD, PD, and FTD risk, with effect allele (A) of rs708384 associated with increased risks [$P = 0.01$ for AD (14), $P = 1.42 \times 10^{-5}$ for PD (15), and $P = 0.04$ for FTD (16)]. No significant relationships of rs708384 or its tagged loci were found for ALS (table S10).

DISCUSSION

It has previously been shown that loci on genes other than *GRN* could modify blood PGRN levels (8, 9). However, no previous GWAS has been conducted to report the genetic modifiers of PGRN in CSF. In the present study, we conducted a genome-wide analysis of CSF PGRN levels and observed an independent signal in the *FAM171A2* gene. Collectively, evidence from genetic, molecular, and bioinformatic findings suggests that the *FAM171A2* gene is a key regulator of PGRN production.

There were several sources of evidence that support the identified genome-wide locus as a real and specific signal and not an artificial type I error. First, rs708384 showed a highly significant *P* value and survived the permutation-based empiric corrections for multiple testing (Fig. 1). Moreover, the locus was directly genotyped, eliminating the possibility that the signal was a result of an imputation error. Second, the association was successfully replicated in a larger, independent sample. Third, both bioinformatics and in vitro evidence confirmed that rs708384 could modulate the expression of the *GRN* and *FAM171A2* genes. Fourth, FAM171A2 and GRN were closely connected and may functionally act together in the human brain. Last, the genotype of rs708384 was associated with the CSF biomarker and risk of neurodegenerative diseases. The finding was consistent with that of a recent study in which PGRN deficiency was reported to increase AD risk by influencing tau rather than Aβ pathology (17).

It has recently been found that PGRN expression was up-regulated in patients with AD compared to controls (18). Thus, participants with dementia were excluded from the present analysis to preclude the potential selection bias due to case ascertainment. Moreover, sensitivity or subgroup analysis showed that the results were significant in both Aβ-positive and Aβ-negative groups. Subgroup analysis indicated that the influence of rs708384 on CSF PGRN levels was vastly greater in HCs than in patients with MCI.

GRN and *FAM171A2* are clustered in common pathways that are potentially involved in regulating CSF PGRN levels. The most significant is L1CAM interactions. L1CAM is an axonal glycoprotein that is important for neurite outgrowth and neuronal survival (19). Although previous animal studies depicted this molecule to reduce AD pathology (20), it is still unclear how it contributes to the development of neurodegenerative diseases via interaction with PGRN. In addition, our GO analyses included categories related to oxidative stress, transport system of neurons, sensory perception of smell and sound, and endocytosis. Previous studies demonstrated that all these biological processes were associated with neurodegenerative diseases, especially AD, PD, and FTD. Future research is necessary to explore the roles of PGRN in these processes and how these may influence the progression of neurodegenerative diseases.

Several limitations should also be noted. The validation cohort is based on Han Chinese individuals recruited from the hospital, which leads to constrained generalizability. Future studies are required to replicate the association in community-based cohorts of other races. As a critical covariate in the GWAS stage, rs5848, which did not pass the QC filter, was imputed but not genotyped. We provided multiple forms of evidence to lower this risk: (i) rs5848 was in low-to-moderate LD with rs708384 in the CEU population, suggesting that the potential influences were limited; (ii) the consistency test showed a high imputation accuracy ($n = 427$, κ = 0.874) for the rs5848 locus; and (iii) we found that the association of rs708384 with the CSF PGRN level was only slightly influenced by genotyped rs5848, in either ADNI or Chinese Alzheimer's Biomarker and LifestylE (CABLE). The in vitro evidence suggests that rs708384 could influence expression of *FAM171A2*, which showed close relationships with expression of the *GRN* gene. However, the specific mechanisms have yet to be studied. In the following research, we plan to further investigate the specific mechanisms of FAM171A2 to regulate PGRN. We will also focus on the role of FAM171A2 in the expression of PGRN within microglia and various models of neurodegeneration diseases. Further investigation is needed to determine whether the effect size of the top locus on CSF PGRN varies by the diagnostic group, although the present study focused on individuals without dementia.

Therefore, we found evidence to show that *FAM171A2*, downstream of *GRN*, is a novel genetic regulator of PGRN production. Future studies are warranted to fully understand FAM171A2 function in the brain and the mechanisms for which FAM171A2 affects PGRN levels. Targeting FAM171A2 may potentially modify neurodegenerative disease risk by regulating brain PGRN levels.

MATERIALS AND METHODS
Participants for GWAS

In the discovery cohort, 432 (HC = 157, MCI = 275) non-Hispanic white nondemented individuals were included from the Alzheimer's Disease Neuroimaging Initiative GO/2 (ADNI-GO/2). Table 1 showed the summarized characteristics of this sample. Data used in the preparation for GWAS were obtained from the ADNI database (adni.loni.usc.edu). The original cohort with CSF PGRN and GWAS data comprised 508 participants. We confined our analysis to non-Hispanic white individuals ($n = 434$) to constrain the risk of bias from population stratification. We tested unanticipated duplicates and cryptic relatedness among samples using pairwise genome-wide

estimates of proportion identity-by-descent by the PLINK software (beta 6.4) (*21*). The genome-wide complex trait analysis (GCTA) was used to calculate principal component and confirm the ethnicity of the samples (*22*). Two outlier participants were removed, and 432 were finally included (fig. S1). The study was approved by institutional review boards of all participating institutions, and written informed consent was obtained from all participants or authorized representatives.

CSF PGRN measurements

CSF PGRN levels were measured by a previously reported sandwich immunoassay (*7*) using the Meso Scale Discovery platform (*23*). All CSF samples were distributed randomly across plates and measured in duplicate. All the antibodies and plates were from a single lot to exclude variability between batches. Experiments were performed by experienced operators blinded to clinical information. The mean intrabatch CV was 2.2% (with all duplicate measures presented CV < 15%) and the mean interbatch CV was 4.21%. PGRN levels were corrected by interbatch variation, and corrected values were used for analyses.

Genotyping and imputation

The ADNI-GO/2 samples were genotyped by the Human OmniExpress BeadChip (Illumina Inc., San Diego, CA). Before association analysis, all samples and genotypes underwent stringent QC with the following criteria: call rates for SNPs > 95%, call rates for individuals > 95%, minor allele frequencies > 0.2, and Hardy-Weinberg equilibrium test $P > 0.001$. The original dataset included a total of 710,618 genotyped variants, including rs7412 and rs429358, which were genotyped separately by an *APOE* genotyping kit to define the *APOE* ε2/ε3/ε4 isoforms, as previously described (*24*). Following the standard procedure, imputation was conducted using Beagle software (version 5.0) with the HapMap GRCh37 as reference. SNPs with a Beagle $R^2 < 0.8$, a call rate < 95%, a minor allele frequency < 0.2, and Hardy-Weinberg equilibrium test $P < 0.001$ were removed. Last, the filters and QC produced a total of 3,262,988 imputed and genotyped SNPs for analyses.

Statistical methods

Because the CSF PGRN values were skewed (Shapiro-Wilk test $P < 0.05$), data transformation was performed to achieve a normal distribution using "car" package of R. A linear regression model was used to determine the associations of CSF PGRN levels with genetic polymorphisms using the PLINK v1.9 with an additive genetic model. Stepwise linear regression analysis was used to determine whether CSF levels of PGRN were influenced by diagnosis, age, gender, education, *APOE* ε4 allele, and CV of PGRN measurements. Those with $P < 0.2$ were included as covariates. To correct for confounding by genetic ancestry that could lead to population stratification, the first three principal components of a genetic relationship matrix between pairs of individuals were further included as covariates. Given that 30% of nondemented elders met the biomarker criteria for AD (*25*), subgroup analyses were first conducted for Aβ pathological status (A positive versus A negative) according to A-T-N criteria, for which A positive was defined as positive evidence of cerebral Aβ deposition as detected by PET (AV45 > 1.11) or CSF (Aβ < 192 pg/ml) (*26*). We further excluded those who progressed to dementia within 3 years since baseline. We also performed sensitivity analyses including rs5848 (*GRN, imputed)*, age, and diagnosis

as covariates because it was previously shown that levels of PGRN might be influenced by these factors (*11*, *18*). Analyses stratified according to gender and diagnoses were conducted to examine the strata effect.

The specificity of the association with top signal (the SNP with the smallest P value) was explored to examine whether the association was influenced by other CSF proteins linked with neurodegenerative diseases, including Aβ, p-tau181 (P-tau), total tau (T-tau) (*27*), α-SYN (*28*), NFL (*29*, *30*), and sTREM2 (*31*, *32*). The measurement data of all these proteins were downloaded from ADNI: (i) The associations of top signal with these proteins were separately explored via single variable linear regression; (ii) the relationships between normalized levels of CSF PGRN and other CSF proteins were estimated using Pearson's product moment correlation coefficient; (iii) these CSF proteins were respectively included as covariates in the model. Statistical analyses and data visualization were performed in R v3.5.1, PLINK v1.9 (*21*), GCTA v1.91.6beta (*33*), and LocusZoom v1.3 (*34*).

Bonferroni-corrected statistical significance was defined as $P < 5 \times 10^{-8}$, and $P < 1 \times 10^{-5}$ was considered suggestive association. To identify additional independent genetic signals, conditional analyses were performed by adding the top SNP as a covariate and testing all remaining regional SNPs for association. As an additional alternative to exclude possible false-positive results, the PLINK max (T) permutation test with 5000 permutations was used to generate empiric P values for multiple testing correction. Genome-wide associations were visualized with the R package "qqman"; genome partitioning was conducted using the algorithm GCTA to estimate the proportion of phenotypic variance explained (*35*).

Replication cohort

We used an independent replication cohort from nondemented northern Han Chinese elders who were derived from CABLE study. The study has been approved by the Institutional Review Board of Qingdao Municipal Hospital. Since 2017, CABLE is an ongoing large-scale study majorly focused on AD's risk factors and biomarkers in Chinese Han population. CABLE is aimed to determine the genetic and environmental modifiers of AD biomarkers and their utility in early diagnosis. Individuals were recruited at Qingdao Municipal Hospital, Shandong Province, China.

All enrolled participants were Han Chinese aged between 40 and 90 years. The exclusion criteria include the following: (i) CNS infection, head trauma, epilepsy, multiple sclerosis, or other major neurological disorders; (ii) major psychological disorders; (iii) severe systemic diseases (e.g., malignant tumors); (iv) family history of genetic diseases. All participants underwent clinical and neuropsychological assessments, biochemical testing, as well as blood and CSF sample collection. Demographic information and medical history were collected via a structured questionnaire and an electronic medical record system.

CSF PGRN levels in CABLE were determined with the Human PGRN Enzyme-Linked Immunosorbent Assay (ELISA) kit (Biovendor-Laboratornimedicina A.s. Czech Republic) on the microplate reader (Thermo Fisher Scientific Multiskan MK3). Samples were diluted 10-fold and run in duplicate. The association for the most significant locus (top SNP) in the discovery stage was chosen for validation. Several loci were selected for genotyping with restriction fragment length polymorphism technology, including rs708384 (top SNP for validation), rs5848 (important confounder in *GRN*),

two loci related to *APOE* ε4 status (rs7412 and rs429358), and three loci (rs660240, rs4747197, and rs646776) associated with blood PGRN by previous GWAS. Until October 2019, eligible measurements of CSF PGRN and genotyping have been available for 930 non-demented individuals. Data analysis was performed with linear regression via R, adjusting for age, sex, education, *APOE* ε4 status, MMSE at baseline, and CV of PGRN measurements. Sensitivity analysis was conducted by including rs5848 as covariate. Meta-analysis was performed to pool the *P* values of the two-stage findings using "metap" in Stata/SE 12.0 software. In addition, we also examined whether the loci associated with blood PGRN also influenced the levels in CSF.

Bioinformatics analyses

SNP annotations were conducted using the NCBI Database of SNPs (dbSNP, GRCh37/hg19 assembly, 105 release) (www.ncbi.nlm.nih.gov/SNP/) (*36*), SNPnexus (www.snp-nexus.org) (*37*), and SNP and CNV Annotation Database (SCAN) (www.scandb.org/newinterface/index_v1.html). The potential regulatory functions were examined using HaploReg v4.1 (*38*), RegulomeDBv1.1(build 141 of dbSNP) (*39*), and 1000 Genomes Project (www.international-genome.org/). LD analyses were conducted on the basis of data from the 1000 Genomes Project [EUR and Han Chinese in Beijing, China (CHB)]. Variants tagged by the top SNP were used to perform the following enhancer enrichment analysis and expression analysis.

Enhancer enrichment analysis was conducted via HaploReg v4.1 to evaluate in which cell types the tag variants were significantly enriched. Enhancers were defined using Roadmap Epigenomics data with four different methods (including the 15-state core model, the 25-state model incorporating imputed epigenomes, the H3K4me1/H3K4me3 peaks, and the H3K27ac/H3K9ac peaks) (*40*). A binomial test was performed using all 1000 Genomes variants with a MAF > 5% in any population as the background set. A total of 28 blood cells and 13 brain cells were selected for the present analysis.

To validate the result of enhancer enrichment analysis and characterize these associations, expression analyses were performed to determine (i) whether the expression levels of genes where the identified significant loci are located are associated with AD case-control status. The differential expression of identified genes in AD blood (*41*) and brain (*42*) (fig. S3) was investigated by analyzing the gene expression profiles from Gene Expression Omnibus (GEO) via NCBI web tool GEO2R (www.ncbi.nlm.nih.gov/geo/geo2r/) and AlzData web server (www.alzdata.org) (*43*). *P* < 0.05 after adjustment for multiple testing was regarded as statistically significant (*44*); (ii) whether SNPs associated with CSF PGRN also affect the protein-encoding gene (*GRN*) expression levels in brain and blood; (iii) whether the significant SNPs were associated with expression levels of the candidate genes within each locus. To these end, eQTL analyses were conducted using multiple publicly available datasets in human brain tissues [including brain expression GWAS eQTL datasets (*45*) and Genotype-Tissue Expression (GTEx) (*46*)] and the whole blood [including Blood eQTL browser (*47*), Consortium for the Architecture of Gene Expression browser (*48*), NCBI Molecular QTL Browser Search database (*49*), and Framingham Heart Study eQTL database (*49*)]. The eQTL results were also searched from 1000 Genomes Project. (iv) Pair-wise gene expression correlation analysis was performed using public RNA sequencing data of 31,499 samples (https://genenetwork.nl) and GTEx expression data via GEPIA web server (http://gepia.cancer-pku.cn/index.html) (*50*).

GO overrepresentation analyses and pathway analyses

The Association List GO Annotator method (*51*) was used to search for the GO terms and KEGG pathways. SNPs were clumped if they were within 1 Mb and in high LD ($r^2 \geq 0.8$) with the index SNP (defined as $P < 1.0 \times 10^{-4}$). These SNPs were mapped to genes using SNP annotation tools mentioned above. SNPs were mapped to a gene if they were situated within 20 kb of that gene; genes were only counted once irrespective of how many SNPs were mapped to the gene. Last, 34 genes were included.

GO overrepresentation analyses were performed using the PANTHER statistical overrepresentation test v9.0 (www.pantherdb.org/) (*52*) and the CPDB overrepresentation gene set analysis Release 33 (http://cpdb.molgen.mpg.de/) (*53*). Both used data from the GO Consortium (www.geneontology.org) and calculated overrepresentation of candidate genes, relative to different background (18,043 genes for CPDB and 20,814 genes for PANTHER). Of the 34 genes in the analysis, CPDB recognized all and 29 were assigned to at least one GO term by PANTHER. The resulting *P* values are corrected for multiple testing using the false discovery rate method in CPDB (*P* < 0.05). As for PANTHER, Bonferroni-adjusted (six separate tests) *P* values were used ($P < 8.33 \times 10^{-3}$).

In addition, multiple sources of pathway sets were used, including KEGG (1 August 2018, www.genome.jp/kegg/) (*54*), PANTHER v8.1 (*52*), Reactome pathways, v65 (13 June 2018, https://reactome.org/) (*55*), WikiPathways (www.wikipathways.org/index.php/WikiPathways), and Prioritization And Functional Assessment (http://159.226.67.237:8080/pafa) (*56*). Last, the function of gene network was predicted on the basis of the functional enrichment of its coregulatory partners using the GeneNetwork Assisted Diagnostic Optimization tool (www.genenetwork.nl).

In vitro and in vivo experiments
Luciferase reporter plasmid and antibodies
The plasmid was purchased from Genomeditech Co. Ltd. Shanghai, China. Briefly, we cloned an about 500-bp region of the *FAM171A2* intron 1 containing SNP rs708384 or wild-type site and inserted the cloned fragment into a pGL3-promoter vector, which contained the SV40 promoter upstream of the firefly luciferase reporter gene. Empty pGL3-promoter vector was transferred as a systemic control. Antibodies used were as follows: anti-FAM171A2 antibody (1:20 for IHC and IF; ATLAS, HPA019770); anti-FAM171A2 antibody (1:500 for Western blot; Proteintech, 20836-1-AP); anti-CD31 antibody (1:250 for IF; ab24590), anti-granulin antibody (1:1000 for Western bolt; ab108608); anti–β-actin antibody (1:1000 for Western bolt; Affinity, AF7018).

Animals
Eight-week-old C57BL/6 mice were purchased from Beijing Vital River Laboratory Animal Technology Co. Ltd. and housed in 23° ± 1°C, 55 ± 5% humidity and 12-hour light/dark cycles (lights on between 07:00 and 19:00) with libitum food and water. Mice were euthanized for IHC and IF staining (*n* = 5) after adaptation for 2 weeks. The study was performed with the approval of the Institutional Animal Care and Use Committee of Fudan university (approval number: JS-256). The experiments were performed according to the National Institutes of Health Guide for the Care and Use of Laboratory animals.

FAM171A2 knockdown and overexpression
The in vitro FAM171A2 knockdown was achieved by siRNA, purchased from GenePharma Co. Ltd, Shanghai, China. Sequence:

5′-GCAAUGGCACUGGUGUAAUTT, AUUACACCAGUGC-CAUUGCTT-3′. The FAM171A2 overexpression plasmid was purchased from GenePharma. The pEX-3 vector was used for plasmid construction. Both FAM171A2 siRNA and plasmid were transferred into human umbilical vein endothelial cells (HUVECs) by Lipofectamine2000. The effect of knockdown and overexpression was confirmed 48 hours after transfection.

Cell culture
The HUVECs we used were purchased from iCell, Shanghai, China. This cell line was immortalized by lentiviral transformation in 2018 from human primary umbilical vein endothelial cells, which was confirmed by CD31 IF staining. HUVECs was cultured in extracellular matrix + 5% fetal bovine serum (FBS) + 1% endothelial growth factor + 1% penicillin/streptomycin (P/S). The HEK293 cell was purchased from Chinese Academy of Sciences Cell Bank. The cells were cultured in Dulbecco's modified Eagle's medium + 5% FBS + 1% P/S. The incubator maintained an environment of 37ºC and 5% CO_2. Cells were subcultured by 1:3 until a convergence of 80%.

Dual-luciferase reporter assay
The assay was conducted according to the technical manual of Promega Dual-Luciferase Reporter Assay System. The firefly and Renilla luciferase reporter plasmid (5:1) was cotransferred into HEK293 cells by Lipofectamine2000. After 24 hours, cells were homogenized by the passive lysate buffer provided by the manufacture. Luciferase Assay Reagent II and Stop and Glo Reagent were added to the lysates sequentially. Luminous intensity was detected by the Synergy H1 hybrid reader using the luminescence model.

Immunohistochemistry
Mice was injected by pentobarbital sodium into deep anesthesia and perfused with saline solution followed by 4% (w/v) paraformaldehyde. Hippocampi and cortex were frontally sectioned using a freezing microtome, and 30-μm serial sections were collected. Sections were incubated in blocking buffer [0.01% Triton X-100 and 20% normal goat serum in phosphate-buffered saline, (pH 7.4)] for 1 hour and were subsequently incubated at 4°C overnight in rabbit anti-FAM171A2 primary antibody, followed by 2 hours in secondary biotinylated antibody, at room temperature. After three washes, sections were incubated in avidin-biotin complex (ABC Standard, Vector Laboratories) and color developed by using 0.025% 3,3′-diaminobenzidine and 0.1% H_2O_2. Then, the sections were stuck on microslides and immersed into hematite dyeing solution for 1 min followed by deionized water washes. After treated with ethanol and dimethylbenzene, the sections were covered by neutral gum and slips. An Olympus camera (DP72; Olympus) was used to take images.

IF staining
The process was similar to IHC. The brain sections were incubated by the rabbit anti-FAM171A2 antibody together with the mouse anti-CD31 antibody and shaken overnight. After washing, sections were immersed into goat anti-rabbit Alex Fluor 488 and goat anti-mouse Alex Fluor 647 (FAM171A2 and CD31) or donkey anti-rabbit Alex Fluor 488 and donkey anti-goat Alex Fluor 594 (FAM171A2 and IBA1) for 2 hours, followed by washing and imaged using the Olympus FV10 laser scanning confocal microscope.

Western blotting analysis
Thirty-microgram protein for each sample was added to a lane of 10% SDS–polyacrylamide gel electrophoresis gels. After electrophoresis, polyvinylidene difluoride membranes were transferred by constant current of 250 mA for 90 min, and then blocked with 5% bovine serum albumin for 1 hour, and incubated in the rabbit anti-granulin antibody overnight. After washing, membranes were incubated with horseradish peroxidase–conjugated secondary antibodies for 2 hours. The blots were visualized using the Super Signal West Pico Chemiluminescent Substrate (Thermo Fisher Scientific Inc.). The grayscale was measured by ImageJ. All experiments were performed in triplicate. The final data were expressed as a ratio of the relative optical density (OD) of the protein of interest to that of β-actin.

Enzyme-linked immunosorbent assay
We used the R&D PGRN Quantikine ELISA kit to detect the supernatant PGRN content. The experiment was conducted according to the manual of the ELISA kit. Briefly, 50 μl of cell lysate was added per well and incubated for 2 hours at room temperature. A 200-μl human PGRN conjugate was added to each well after washing for four times with 400 μl of wash buffer per well and incubated for 2 hours at room temperature. A 200-μl substrate solution was added to each well subsequently and incubated for 30 min at room temperature. Last, the reaction was ended by 50 μl of stop solution per well. Within 30 min, the OD of each sample was detected at 450 nm and corrected by wavelength at 540 nm. PGRN contents were calculated from OD according to protein standards.

Statistics
The data were presented in the means ± SEM. GraphPad Prism 8 software (version 8.0.2, GraphPad Software Inc., La Jolla, CA, USA) was used to test the difference via the one-way analysis of variance (ANOVA) followed by the Tukey post hoc analysis and unpaired t test. $P < 0.01$ was considered statistically significant.

Associations of the top signal with neurodegenerative diseases
The influences of the top SNP on cognition, brain Aβ deposition, FDG, and volume/thickness of brain regions of interest (hippocampus, parahippocampus, posterior cingulate, precuneus, cuneus, entorhinal cortex, and middle temporal region) were further explored using ADNI. Those who developed AD in the first 3 years were excluded. Linear regression models were conducted to determine the associations with the above indexes at baseline. Subgroup analysis based on the diagnosis (HC versus MCI) was conducted. R packages "lm" was used to perform the above analyses. In addition, previous GWAS for AD (14), PD (15, 57) (www.pdgene.org/), ALS (58), and FTD (16, 59–62) was searched for associations of the top signal with risk of neurodegenerative diseases.

REFERENCES AND NOTES
1. X. Zhou, L. Sun, O. A. Brady, K. A. Murphy, F. Hu, Elevated TMEM106B levels exaggerate lipofuscin accumulation and lysosomal dysfunction in aged mice with progranulin deficiency. *Acta Neuropathol. Commun.* **5**, 9 (2017).
2. S. S. Minami, S.-W. Min, G. Krabbe, C. Wang, Y. Zhou, R. Asgarov, Y. Li, L. H. Martens, L. P. Elia, M. E. Ward, L. Mucke, R. V. Farese Jr., L. Gan, Progranulin protects against amyloid β deposition and toxicity in Alzheimer's disease mouse models. *Nat. Med.* **20**, 1157–1164 (2014).
3. J. M. Van Kampen, D. Baranowski, D. G. Kay, Progranulin gene delivery protects dopaminergic neurons in a mouse model of Parkinson's disease. *PLOS ONE* **9**, e97032 (2014).
4. L. H. Martens, J. Zhang, S. J. Barmada, P. Zhou, S. Kamiya, B. Sun, S.-W. Min, L. Gan, S. Finkbeiner, E. J. Huang, R. V. Farese Jr., Progranulin deficiency promotes neuroinflammation and neuron loss following toxin-induced injury. *J. Clin. Invest.* **122**, 3955–3959 (2012).

5. M. C. Chang, K. Srinivasan, B. A. Friedman, E. Suto, Z. Modrusan, W. P. Lee, J. S. Kaminker, D. V. Hansen, M. Sheng, Progranulin deficiency causes impairment of autophagy and TDP-43 accumulation. *J. Exp. Med.* **214**, 2611–2628 (2017).

6. J. M. Van Kampen, D. G. Kay, Progranulin gene delivery reduces plaque burden and synaptic atrophy in a mouse model of Alzheimer's disease. *PLOS ONE* **12**, e0182896 (2017).

7. M. Suárez-Calvet, A. Capell, M. Á. A. Caballero, E. Morenas-Rodríguez, K. Fellerer, N. Franzmeier, G. Kleinberger, E. Eren, Y. Deming, L. Piccio, C. M. Karch, C. Cruchaga, K. Paumier, R. J. Bateman, A. M. Fagan, J. C. Morris, J. Levin, A. Danek, M. Jucker, C. L. Masters, M. N. Rossor, J. M. Ringman, L. M. Shaw, J. Q. Trojanowski, M. Weiner, M. Ewers, C. Haass; Dominantly Inherited Alzheimer Network; Alzheimer's Disease Neuroimaging Initiative, CSF progranulin increases in the course of Alzheimer's disease and is associated with sTREM2, neurodegeneration and cognitive decline. *EMBO Mol. Med.* **10**, e9712 (2018).

8. M. M. Carrasquillo, A. M. Nicholson, N. C. Finch, J. R. Gibbs, M. Baker, N. J. Rutherford, T. A. Hunter, M. De Jesus-Hernandez, G. D. Bisceglio, I. R. Mackenzie, A. Singleton, M. R. Cookson, J. E. Crook, A. Dillman, D. Hernandez, R. C. Petersen, N. R. Graff-Radford, S. G. Younkin, R. Rademakers, Genome-wide screen identifies rs646776 near sortilin as a regulator of progranulin levels in human plasma. *Am. J. Hum. Genet.* **87**, 890–897 (2010).

9. A. Tönjes, M. Scholz, J. Krüger, K. Krause, D. Schleinitz, H. Kirsten, C. Gebhardt, C. Marzi, H. Grallert, C. Ladenvall, H. Heyne, E. Laurila, J. Kriebel, C. Meisinger, W. Rathmann, C. Gieger, L. Groop, I. Prokopenko, B. Isomaa, F. Beutner, J. Kratzsch, A. Fischer-Rosinsky, A. Pfeiffer, K. Krohn, J. Spranger, J. Thiery, M. Blüher, M. Stumvoll, P. Kovacs, Genome-wide meta-analysis identifies novel determinants of circulating serum progranulin. *Hum. Mol. Genet.* **27**, 546–558 (2018).

10. L. H. H. Meeter, H. Patzke, G. Loewen, E. G. P. Dopper, Y. A. L. Pijnenburg, R. van Minkelen, J. C. van Swieten, Progranulin levels in plasma and cerebrospinal fluid in granulin mutation carriers. *Dement. Geriatr. Cogn. Dis. Extra* **6**, 330–340 (2016).

11. A. M. Nicholson, N. C. A. Finch, C. S. Thomas, A. Wojtas, N. J. Rutherford, M. M. Mielke, R. O. Roberts, B. F. Boeve, D. S. Knopman, R. C. Petersen, R. Rademakers, Progranulin protein levels are differently regulated in plasma and CSF. *Neurology* **82**, 1871–1878 (2014).

12. Y. Lu, L. Zheng, W. Zhang, T. Feng, J. Liu, X. Wang, Y. Yu, M. Qi, W. Zhao, X. Yu, W. Tang, Growth factor progranulin contributes to cervical cancer cell proliferation and transformation *in vivo* and *in vitro*. *Gynecol. Oncol.* **134**, 364–371 (2014).

13. H. Qu, H. Deng, Z. Hu, Plasma progranulin concentrations are increased in patients with type 2 diabetes and obesity and correlated with insulin resistance. *Mediators Inflamm.* **2013**, 360190 (2013).

14. J. C. Lambert, C. A. Ibrahim-Verbaas, D. Harold, A. C. Naj, R. Sims, C. Bellenguez, A. L. De Stefano, J. C. Bis, G. W. Beecham, B. Grenier-Boley, G. Russo, T. A. Thornton-Wells, N. Jones, A. V. Smith, V. Chouraki, C. Thomas, M. A. Ikram, D. Zelenika, B. N. Vardarajan, Y. Kamatani, C. F. Lin, A. Gerrish, H. Schmidt, B. Kunkle, M. L. Dunstan, A. Ruiz, M. T. Bihoreau, S. H. Choi, C. Reitz, F. Pasquier, C. Cruchaga, D. Craig, N. Amin, C. Berr, O. L. Lopez, P. L. De Jager, V. Deramecourt, J. A. Johnston, D. Evans, S. Lovestone, L. Letenneur, F. J. Morón, D. C. Rubinsztein, G. Eiriksdottir, K. Sleegers, A. M. Goate, N. Fiévet, M. W. Huentelman, M. Gill, K. Brown, M. I. Kamboh, L. Keller, P. Barberger-Gateau, B. McGuiness, E. B. Larson, R. Green, A. J. Myers, C. Dufouil, S. Todd, D. Wallon, S. Love, E. Rogaeva, J. Gallacher, P. St George-Hyslop, J. Clarimon, A. Lleo, A. Bayer, D. W. Tsuang, L. Yu, M. Tsolaki, P. Bossù, G. Spalletta, P. Proitsi, J. Collinge, S. Sorbi, F. Sanchez-Garcia, N. C. Fox, J. Hardy, M. C. D. Naranjo, P. Bosco, R. Clarke, C. Brayne, D. Galimberti, M. Mancuso, F. Matthews; European Alzheimer's Disease Initiative (EADI); Genetic, Environmental Risk in Alzheimer's Disease; Alzheimer's Disease Genetic Consortium; Cohorts for Heart, Aging Research in Genomic Epidemiology, S. Moebus, P. Mecocci, M. Del Zompo, W. Maier, H. Hampel, A. Pilotto, M. Bullido, F. Panza, P. Caffarra, B. Nacmias, J. R. Gilbert, M. Mayhaus, L. Lannefelt, H. Hakonarson, S. Pichler, M. M. Carrasquillo, M. Ingelsson, D. Beekly, V. Alvarez, F. Zou, O. Valladares, S. G. Younkin, E. Coto, K. L. Hamilton-Nelson, W. Gu, C. Razquin, P. Pastor, I. Mateo, M. J. Owen, K. M. Faber, P. V. Jonsson, O. Combarros, M. C. O'Donovan, L. B. Cantwell, H. Soininen, D. Blacker, S. Mead, T. H. Mosley Jr., D. A. Bennett, T. B. Harris, L. Fratiglioni, C. Holmes, R. F. de Bruijn, P. Passmore, T. J. Montine, K. Bettens, J. I. Rotter, A. Brice, K. Morgan, T. M. Foroud, W. A. Kukull, D. Hannequin, J. F. Powell, M. A. Nalls, K. Ritchie, K. L. Lunetta, J. S. Kauwe, E. Boerwinkle, M. Riemenschneider, M. Boada, M. Hiltuenen, E. R. Martin, R. Schmidt, D. Rujescu, L. S. Wang, J. F. Dartigues, R. Mayeux, C. Tzourio, A. Hofman, M. M. Nöthen, C. Graff, B. M. Psaty, L. Jones, J. L. Haines, P. A. Holmans, M. Lathrop, M. A. Pericak-Vance, L. J. Launer, L. A. Farrer, C. M. van Duijn, C. Van Broeckhoven, V. Moskvina, S. Seshadri, J. Williams, G. D. Schellenberg, P. Amouyel, Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nat. Genet.* **45**, 1452–1458 (2013).

15. C. M. Lill, C. M. Lill, J. T. Roehr, M. B. McQueen, F. K. Kavvoura, S. Bagade, B.-M. M. Schjeide, L. M. Schjeide, E. Meissner, U. Zauft, N. C. Allen, T. Liu, M. Schilling, K. J. Anderson, G. Beecham, D. Berg, J. M. Biernacka, A. Brice, A. L. De Stefano, C. B. Do, N. Eriksson,

S. A. Factor, M. J. Farrer, T. Foroud, T. Gasser, T. Hamza, J. A. Hardy, P. Heutink, E. M. Hill-Burns, C. Klein, J. C. Latourelle, D. M. Maraganore, E. R. Martin, M. Martinez, R. H. Myers, M. A. Nalls, N. Pankratz, H. Payami, W. Satake, W. K. Scott, M. Sharma, A. B. Singleton, K. Stefansson, T. Toda, J. Y. Tung, J. Vance, N. W. Wood, C. P. Zabetian; 23andMe; The Genetic Epidemiology of Parkinson's Disease (GEO-PD) Consortium; The International Parkinson's Disease Genomics Consortium (IPDGC); The Parkinson's Disease GWAS Consortium; The Wellcome Trust Case Control Consortium (WTCCC), P. Young, R. E. Tanzi, M. J. Khoury, F. Zipp, H. Lehrach, J. P. A. Ioannidis, L. Bertram, Comprehensive research synopsis and systematic meta-analyses in Parkinson's disease genetics: The PDGene database. *PLOS Genet.* **8**, e1002548 (2012).

16. R. Rademakers, J. L. Eriksen, M. Baker, T. Robinson, Z. Ahmed, S. J. Lincoln, N. Finch, N. J. Rutherford, R. J. Crook, K. A. Josephs, B. F. Boeve, D. S. Knopman, R. C. Petersen, J. E. Parisi, R. J. Caselli, Z. K. Wszolek, R. J. Uitti, H. Feldman, I. R. Mackenzie, N. R. Graff-Radford, D. W. Dickson, Common variation in the miR-659 binding-site of GRN is a major risk factor for TDP43-positive frontotemporal dementia. *Hum. Mol. Genet.* **17**, 3631–3642 (2008).

17. H. Takahashi, Z. A. Klein, S. M. Bhagat, A. C. Kaufman, M. A. Kostylev, T. Ikezu, S. M. Strittmatter; Alzheimer's Disease Neuroimaging Initiative, Opposing effects of progranulin deficiency on amyloid and tau pathologies via microglial TYROBP network. *Acta Neuropathol.* **133**, 785–807 (2017).

18. Y. A. Cooper, D. Nachun, D. Dokuru, Z. Yang, A. M. Karydas, G. Serrero, B. Yue; Alzheimer's Disease Neuroimaging Initiative, A. L. Boxer, B. L. Miller, G. Coppola, Progranulin levels in blood in Alzheimer's disease and mild cognitive impairment. *Ann. Clin. Transl. Neurol.* **5**, 616–629 (2018).

19. D. Lutz, G. Loers, R. Kleene, I. Oezen, H. Kataria, N. Katagihallimath, I. Braren, G. Harauz, M. Schachner, Myelin basic protein cleaves cell adhesion molecule L1 and promotes neuritogenesis and cell survival. *J. Biol. Chem.* **289**, 13503–13518 (2014).

20. N. Djogo, I. Jakovcevski, C. Müller, H. J. Lee, J.-C. Xu, M. Jakovcevski, S. Kügler, G. Loers, M. Schachner, Adhesion molecule L1 binds to amyloid beta and reduces Alzheimer's disease pathology in mice. *Neurobiol. Dis.* **56**, 104–115 (2013).

21. C. C. Chang, C. C. Chow, L. C. Tellier, S. Vattikuti, S. M. Purcell, J. J. Lee, Second-generation PLINK: Rising to the challenge of larger and richer datasets. *Gigascience* **4**, 7 (2015).

22. A. L. Price, N. J. Patterson, R. M. Plenge, M. E. Weinblatt, N. A. Shadick, D. Reich, Principal components analysis corrects for stratification in genome-wide association studies. *Nat. Genet.* **38**, 904–909 (2006).

23. A. Capell, S. Liebscher, K. Fellerer, N. Brouwers, M. Willem, S. Lammich, I. Gijselinck, T. Bittner, A. M. Carlson, F. Sasse, B. Kunze, H. Steinmetz, R. Jansen, D. Dormann, K. Sleegers, M. Cruts, J. Herms, C. Van Broeckhoven, C. Haass, Rescue of progranulin deficiency associated with frontotemporal lobar degeneration by alkalizing reagents and inhibition of vacuolar ATPase. *J. Neurosci.* **31**, 1885–1894 (2011).

24. S. Kim, S. Swaminathan, L. Shen, S. L. Risacher, K. Nho, T. Foroud, L. M. Shaw, J. Q. Trojanowski, S. G. Potkin, J. Huentelman, D. W. Craig, B. M. De Chairo, P. S. Aisen, R. C. Petersen, M. W. Weiner, A. J. Saykin; Alzheimer's Disease Neuroimaging Initiative, Genome-wide association study of CSF biomarkers $A\beta_{1-42}$, t-tau, and p-tau$_{181p}$ in the ADNI cohort. *Neurology* **76**, 69–79 (2011).

25. J. A. Schneider, N. T. Aggarwal, L. Barnes, P. Boyle, D. A. Bennett, The neuropathology of older persons with and without dementia from community versus clinic cohorts. *J. Alzheimers Dis.* **18**, 691–701 (2009).

26. L. M. Shaw, H. Vanderstichele, M. Knapik-Czajka, C. M. Clark, P. S. Aisen, R. C. Petersen, K. Blennow, H. Soares, A. Simon, P. Lewczuk, R. Dean, E. Siemers, W. Potter, V. M.-Y. Lee, J. Q. Trojanowski; Alzheimer's Disease Neuroimaging Initiative, Cerebrospinal fluid biomarker signature in Alzheimer's disease neuroimaging initiative subjects. *Ann. Neurol.* **65**, 403–413 (2009).

27. A. Olsson, H. Vanderstichele, N. Andreasen, G. De Meyer, A. Wallin, B. Holmberg, L. Rosengren, E. Vanmechelen, K. Blennow, Simultaneous measurement of β-amyloid$_{(1-42)}$, total tau, and phosphorylated tau (Thr^{181}) in cerebrospinal fluid by the xMAP technology. *Clin. Chem.* **51**, 336–345 (2005).

28. Z. Hong, M. Shi, K. A. Chung, J. F. Quinn, E. R. Peskind, D. Galasko, J. Jankovic, C. P. Zabetian, J. B. Leverenz, G. Baird, T. J. Montine, A. M. Hancock, H. Hwang, C. Pan, J. Bradner, U. J. Kang, P. H. Jensen, J. Zhang, DJ-1 and α-synuclein in human cerebrospinal fluid as biomarkers of Parkinson's disease. *Brain* **133**, 713–726 (2010).

29. A. Nordlund, S. Rolstad, P. Hellström, M. Sjögren, S. Hansen, A. Wallin, The Goteborg MCI study: Mild cognitive impairment is a heterogeneous condition. *J. Neurol. Neurosurg. Psychiatry* **76**, 1485–1490 (2005).

30. H. Zetterberg, T. Skillbäck, N. Mattsson, J. Q. Trojanowski, E. Portelius, L. M. Shaw, M. W. Weiner, K. Blennow; Alzheimer's Disease Neuroimaging Initiative, Association of Cerebrospinal Fluid Neurofilament Light Concentration With Alzheimer Disease Progression. *JAMA Neurol.* **73**, 60–67 (2016).

31. M. Suárez-Calvet, G. Kleinberger, M. Á. A. Caballero, M. Brendel, A. Rominger, D. Alcolea, J. Fortea, A. Lleó, R. Blesa, J. D. Gispert, R. Sánchez-Valle, A. Antonell, L. Rami, J. L. Molinuevo, F. Brosseron, A. Traschütz, M. T. Heneka, H. Struyfs, S. Engelborghs, K. Sleegers, C. Van Broeckhoven, H. Zetterberg, B. Nellgård, K. Blennow, A. Crispin,

M. Ewers, C. Haass, sTREM2 cerebrospinal fluid levels are a potential biomarker for microglia activity in early-stage Alzheimer's disease and associate with neuronal injury markers. *EMBO Mol. Med.* **8**, 466–476 (2016).

32. M. Suárez-Calvet, E. Morenas-Rodríguez, G. Kleinberger, K. Schlepckow, M. Á. A. Caballero, N. Franzmeier, A. Capell, K. Fellerer, B. Nuscher, E. Eren, J. Levin, Y. Deming, L. Piccio, C. M. Karch, C. Cruchaga, L. M. Shaw, J. Q. Trojanowski, M. Weiner, M. Ewers, C. Haass; Alzheimer's Disease Neuroimaging Initiative, Early increase of CSF sTREM2 in Alzheimer's disease is associated with tau related-neurodegeneration but not with amyloid-β pathology. *Mol. Neurodegener.* **14**, 1 (2019).

33. J. Yang, S. H. Lee, M. E. Goddard, P. M. Visscher, GCTA: A tool for genome-wide complex trait analysis. *Am. J. Hum. Genet.* **88**, 76–82 (2011).

34. R. J. Pruim, R. P. Welch, S. Sanna, T. M. Teslovich, P. S. Chines, T. P. Gliedt, M. Boehnke, G. R. Abecasis, C. J. Willer, LocusZoom: Regional visualization of genome-wide association scan results. *Bioinformatics* **26**, 2336–2337 (2010).

35. J. Yang, T. A. Manolio, L. R. Pasquale, E. Boerwinkle, N. Caporaso, J. M. Cunningham, M. de Andrade, B. Feenstra, E. Feingold, M. G. Hayes, W. G. Hill, M. T. Landi, A. Alonso, G. Lettre, P. Lin, H. Ling, W. Lowe, R. A. Mathias, M. Melbye, E. Pugh, M. C. Cornelis, B. S. Weir, M. E. Goddard, P. M. Visscher, Genome partitioning of genetic variation for complex traits using common SNPs. *Nat. Genet.* **43**, 519–525 (2011).

36. S. T. Sherry, M. H. Ward, M. Kholodov, J. Baker, L. Phan, E. M. Smigielski, K. Sirotkin, dbSNP: The NCBI database of genetic variation. *Nucleic Acids Res.* **29**, 308–311 (2001).

37. A. Z. Dayem Ullah, J. Oscanoa, J. Wang, A. Nagano, N. R. Lemoine, C. Chelala, SNPnexus: Assessing the functional relevance of genetic variation to facilitate the promise of precision medicine. *Nucleic Acids Res.* **46**, W109–W113 (2018).

38. L. D. Ward, M. Kellis, HaploReg: A resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. *Nucleic Acids Res.* **40**, D930–D934 (2012).

39. A. P. Boyle, E. L. Hong, M. Hariharan, Y. Cheng, M. A. Schaub, M. Kasowski, K. J. Karczewski, J. Park, B. C. Hitz, S. Weng, J. M. Cherry, M. Snyder, Annotation of functional variation in personal genomes using RegulomeDB. *Genome Res.* **22**, 1790–1797 (2012).

40. L. D. Ward, M. Kellis, HaploReg v4: Systematic mining of putative causal variants, cell types, regulators and target genes for human complex traits and disease. *Nucleic Acids Res.* **44**, D877–D881 (2016).

41. T. Barrett, S. E. Wilhite, P. Ledoux, C. Evangelista, I. F. Kim, M. Tomashevsky, K. A. Marshall, K. H. Phillippy, P. M. Sherman, M. Holko, A. Yefanov, H. Lee, N. Zhang, C. L. Robertson, N. Serova, S. Davis, A. Soboleva, NCBI GEO: Archive for functional genomics data sets--update. *Nucleic Acids Res.* **41**, D991–D995 (2013).

42. A. J. Myers, J. R. Gibbs, J. A. Webster, K. Rohrer, A. Zhao, L. Marlowe, M. Kaleem, D. Leung, L. Bryden, P. Nath, V. L. Zismann, K. Joshipura, M. J. Huentelman, D. Hu-Lince, K. D. Coon, D. W. Craig, J. V. Pearson, P. Holmans, C. B. Heward, E. M. Reiman, D. Stephan, J. Hardy, A survey of genetic human cortical gene expression. *Nat. Genet.* **39**, 1494–1499 (2007).

43. M. Xu, D.-F. Zhang, R. Luo, Y. Wu, H. Zhou, L.-L. Kong, R. Bi, Y.-G. Yao, A systematic integrated analysis of brain expression profiles reveals *YAP1* and other prioritized hub genes as important upstream regulators in Alzheimer's disease. *Alzheimers Dement.* **14**, 215–229 (2018).

44. Y. Benjamini, R. Heller, Screening for partial conjunction hypotheses. *Biometrics* **64**, 1215–1222 (2008).

45. F. Zou, H. S. Chai, C. S. Younkin, M. Allen, J. Crook, V. S. Pankratz, M. M. Carrasquillo, C. N. Rowley, A. A. Nair, S. Middha, S. Maharjan, T. Nguyen, L. Ma, K. G. Malphrus, R. Palusak, S. Lincoln, G. Bisceglio, C. Georgescu, N. Kouri, C. P. Kolbert, J. Jen, J. L. Haines, R. Mayeux, M. A. Pericak-Vance, L. A. Farrer, G. D. Schellenberg; Alzheimer's Disease Genetics Consortium, R. C. Petersen, N. R. Graff-Radford, D. W. Dickson, S. G. Younkin, N. Ertekin-Taner, Brain expression genome-wide association study (eGWAS) identifies human disease-associated variants. *PLOS Genet.* **8**, e1002707 (2012).

46. GTEx Consortium, Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: Multitissue gene regulation in humans. *Science* **348**, 648–660 (2015).

47. H.-J. Westra, M. J. Peters, T. Esko, H. Yaghootkar, C. Schurmann, J. Kettunen, M. W. Christiansen, B. P. Fairfax, K. Schramm, J. E. Powell, A. Zhernakova, D. V. Zhernakova, J. H. Veldink, L. H. Van den Berg, J. Karjalainen, S. Withoff, A. G. Uitterlinden, A. Hofman, F. Rivadeneira, P. A. C. `t Hoen, E. Reinmaa, K. Fischer, M. Nelis, L. Milani, D. Melzer, L. Ferrucci, A. B. Singleton, D. G. Hernandez, M. A. Nalls, G. Homuth, M. Nauck, D. Radke, U. Völker, M. Perola, V. Salomaa, J. Brody, A. Suchy-Dicey, S. A. Gharib, D. A. Enquobahrie, T. Lumley, G. W. Montgomery, S. Makino, H. Prokisch, C. Herder, M. Roden, H. Grallert, T. Meitinger, K. Strauch, Y. Li, R. C. Jansen, P. M. Visscher, J. C. Knight, B. M. Psaty, S. Ripatti, A. Teumer, T. M. Frayling, A. Metspalu, J. B. J. van Meurs, L. Franke, Systematic identification of *trans* eQTLs as putative drivers of known disease associations. *Nat. Genet.* **45**, 1238–1243 (2013).

48. L. R. Lloyd-Jones, A. Holloway, A. M. Rae, J. Yang, K. Small, J. Zhao, B. Zeng, A. Bakshi, A. Metspalu, M. Dermitzakis, G. Gibson, T. Spector, G. Montgomery, T. Esko, P. M. Visscher, J. E. Powell, The genetic architecture of gene expression in peripheral blood. *Am. J. Hum. Genet.* **100**, 228–237 (2017).

49. R. Joehanes, X. Zhang, T. Huan, C. Yao, S.-x. Ying, Q. T. Nguyen, C. Y. Demirkale, M. L. Feolo, N. R. Sharopova, A. Sturcke, A. A. Schäffer, N. Heard-Costa, H. Chen, P.-c. Liu, R. Wang, K. A. Woodhouse, K. Tanriverdi, J. E. Freedman, N. Raghavachari, J. Dupuis, A. D. Johnson, C. J. O'Donnell, D. Levy, P. J. Munson, Integrated genome-wide analysis of expression quantitative trait loci aids interpretation of genomic association studies. *Genome Biol.* **18**, 16 (2017).

50. Z. Tang, C. Li, B. Kang, G. Gao, C. Li, Z. Zhang, GEPIA: A web server for cancer and normal gene expression profiling and interactive analyses. *Nucleic Acids Res.* **45**, W98–W102 (2017).

51. P. Holmans, E. K. Green, J. S. Pahwa, M. A. R. Ferreira, S. M. Purcell, P. Sklar; Wellcome Trust Case-Control Consortium, M. J. Owen, M. C. O'Donovan, N. Craddock, Gene ontology analysis of GWA study data sets provides insights into the biology of bipolar disorder. *Am. J. Hum. Genet.* **85**, 13–24 (2009).

52. H. Mi, A. Muruganujan, P. D. Thomas, PANTHER in 2013: Modeling the evolution of gene function, and other gene attributes, in the context of phylogenetic trees. *Nucleic Acids Res.* **41**, D377–D386 (2013).

53. A. Kamburov, K. Pentchev, H. Galicka, C. Wierling, H. Lehrach, R. Herwig, ConsensusPathDB: Toward a more complete picture of cell biology. *Nucleic Acids Res.* **39**, D712–D717 (2011).

54. M. Kanehisa, S. Goto, Y. Sato, M. Furumichi, M. Tanabe, KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Res.* **40**, D109–D114 (2012).

55. A. Fabregat, K. Sidiropoulos, P. Garapati, M. Gillespie, K. Hausmann, R. Haw, B. Jassal, S. Jupe, F. Korninger, S. M. Kay, L. Matthews, B. May, M. Milacic, K. Rothfels, V. Shamovsky, M. Webber, J. Weiser, M. Williams, G. Wu, L. Stein, H. Hermjakob, P. D'Eustachio, The Reactome pathway knowledgebase. *Nucleic Acids Res.* **42**, D472–D477 (2016).

56. L. Zhou, F. Zhao, Prioritization and functional assessment of noncoding variants associated with complex diseases. *Genome Med.* **10**, 53 (2018).

57. Y. Wang, J. D. McKay, T. Rafnar, Z. Wang, M. N. Timofeeva, P. Broderick, X. Zong, M. Laplana, Y. Wei, Y. Han, A. Lloyd, M. Delahaye-Sourdeix, D. Chubb, V. Gaborieau, W. Wheeler, N. Chatterjee, G. Thorleifsson, P. Sulem, G. Liu, R. Kaaks, M. Henrion, B. Kinnersley, M. Vallée, F. Le Calvez-Kelm, V. L. Stevens, S. M. Gapstur, W. V. Chen, D. Zaridze, N. Szeszenia-Dabrowska, J. Lissowska, P. Rudnai, E. Fabianova, D. Mates, V. Bencko, L. Foretova, V. Janout, H. E. Krokan, M. E. Gabrielsen, F. Skorpen, L. Vatten, I. Njølstad, C. Chen, G. Goodman, S. Benhamou, T. Vooder, K. Välk, M. Nelis, A. Metspalu, M. Lener, J. Lubiński, M. Johansson, P. Vineis, A. Agudo, F. Clavel-Chapelon, H. B. Bueno-de-Mesquita, D. Trichopoulos, K.-T. Khaw, M. Johansson, E. Weiderpass, A. Tjønneland, E. Riboli, M. Lathrop, G. Scelo, D. Albanes, N. E. Caporaso, Y. Ye, J. Gu, X. Wu, M. R. Spitz, H. Dienemann, A. Rosenberger, L. Su, A. Matakidou, T. Eisen, K. Stefansson, A. Risch, S. J. Chanock, D. C. Christiani, R. J. Hung, P. Brennan, M. T. Landi, R. S. Houlston, C. I. Amos, Rare variants of large effect in *BRCA2* and *CHEK2* affect risk of lung cancer. *Nat. Genet.* **46**, 736–741 (2014).

58. C. M. Lill, O. Abel, L. Bertram, A. Al-Chalabi, Keeping up with genetic discoveries in amyotrophic lateral sclerosis: The ALSoD and ALSGene databases. *Amyotroph. Lateral Scler.* **12**, 238–249 (2011).

59. R. Ferrari, M. Grassi, E. Salvi, B. Borroni, F. Palluzzi, D. Pepe, F. D'Avila, A. Padovani, S. Archetti, I. Rainero, E. Rubino, L. Pinessi, L. Benussi, G. Binetti, R. Ghidoni, D. Galimberti, E. Scarpini, M. Serpente, G. Rossi, G. Giaccone, F. Tagliavini, B. Nacmias, I. Piaceri, S. Bagnoli, C. Bruni, R. G. Maletta, L. Bernardi, A. Postiglione, G. Milan, M. Franceschi, A. A. Puca, V. Novelli, C. Barlassina, N. Glorioso, P. Manunta, A. Singleton, D. Cusi, J. Hardy, P. Momeni, A genome-wide screening and SNPs-to-genes approach to identify novel genetic risk factors associated with frontotemporal dementia. *Neurobiol. Aging* **36**, 2904. e13–2904.e2926 (2015).

60. R. Ferrari, D. G. Hernandez, M. A. Nalls, J. D. Rohrer, A. Ramasamy, J. B. J. Kwok, C. Dobson-Stone, P. W. Schofield, G. M. Halliday, J. R. Hodges, O. Piguet, L. Bartley, E. Thompson, E. Haan, I. Hernández, A. Ruiz, M. Boada, B. Borroni, A. Padovani, C. Cruchaga, N. J. Cairns, L. Benussi, G. Binetti, R. Ghidoni, G. Forloni, D. Galimberti, C. Fenoglio, M. Serpente, E. Scarpini, J. Clarimón, A. Lleó, R. Blesa, M. L. Waldö, K. Nilsson, C. Nilsson, I. R. A. Mackenzie, G.-Y. R. Hsiung, D. M. A. Mann, J. Grafman, C. M. Morris, J. Attems, T. D. Griffiths, I. G. McKeith, A. J. Thomas, P. Pietrini, E. D. Huey, E. M. Wassermann, A. Baborie, E. Jaros, M. C. Tierney, P. Pastor, C. Razquin, S. Ortega-Cubero, E. Alonso, R. Perneczky, J. Diehl-Schmid, P. Alexopoulos, A. Kurz, I. Rainero, E. Rubino, L. Pinessi, E. Rogaeva, P. St.George-Hyslop, G. Rossi, F. Tagliavini, G. Giaccone, J. B. Rowe, J. C. M. Schlachetzki, J. Uphill, J. Collinge, S. Mead, A. Danek, V. M. Van Deerlin, M. Grossman, J. Q. Trojanowski, J. van der Zee, W. Deschamps, T. Van Langenhove, M. Cruts, C. Van Broeckhoven, S. F. Cappa, I. L. Ber, D. Hannequin, V. Golfier, M. Vercelletto, A. Brice, B. Nacmias, S. Sorbi, S. Bagnoli, I. Piaceri, J. E. Nielsen, L. E. Hjermind, M. Riemenschneider, M. Mayhaus, B. Ibach, G. Gasparoni, S. Pichler, W. Gu, M. N. Rossor, N. C. Fox, J. D. Warren, M. G. Spillantini, H. R. Morris, P. Rizzu, P. Heutink, J. S. Snowden, S. Rollinson, A. Richardson, A. Gerhard, A. C. Bruni, R. Maletta, F. Frangipane, C. Cupidi, L. Bernardi, M. Anfossi, M. Gallo, M. E. Conidi, N. Smirne, R. Rademakers, M. Baker, D. W. Dickson, N. R. Graff-Radford, R. C. Petersen, D. Knopman, K. A. Josephs, B. F. Boeve,

J. E. Parisi, W. W. Seeley, B. L. Miller, A. M. Karydas, H. Rosen, J. C. van Swieten, E. G. P. Dopper, H. Seelaar, Y. A. L. Pijnenburg, P. Scheltens, G. Logroscino, R. Capozzo, V. Novelli, A. A. Puca, M. Franceschi, A. Postiglione, G. Milan, P. Sorrentino, M. Kristiansen, H.-H. Chiang, C. Graff, F. Pasquier, A. Rollin, V. Deramecourt, F. Lebert, D. Kapogiannis, L. Ferrucci, S. Pickering-Brown, A. B. Singleton, J. Hardy, P. Momeni, Frontotemporal dementia and its subtypes: A genome-wide association study. *Lancet Neurol.* **13**, 686–699 (2014).

61. I. Broce, C. M. Karch, N. Wen, C. C. Fan, Y. Wang, C. H. Tan, N. Kouri, O. A. Ross, G. U. Höglinger, U. Muller, J. Hardy; International FTD-Genomics Consortium, P. Momeni, C. P. Hess, W. P. Dillon, Z. A. Miller, L. W. Bonham, G. D. Rabinovici, H. J. Rosen, G. D. Schellenberg, A. Franke, T. H. Karlsen, J. H. Veldink, R. Ferrari, J. S. Yokoyama, B. L. Miller, O. A. Andreassen, A. M. Dale, R. S. Desikan, L. P. Sugrue, Immune-related genetic enrichment in frontotemporal dementia: An analysis of genome-wide association studies. *PLOS Med.* **15**, e1002487 (2018).

62. A. Mishra, R. Ferrari, P. Heutink, J. Hardy, Y. Pijnenburg, D. Posthuma; International FTD-Genomics Consortium, Gene-based association studies report genetic links for clinical subtypes of frontotemporal dementia. *Brain* **140**, 1437–1446 (2017).

Citation: W. Xu, S.-D. Han, C. Zhang, J.-Q. Li, Y.-J. Wang, C.-C. Tan, H.-Q. Li, Q. Dong, C. Mei, L. Tan, J.-T. Yu, The *FAM171A2* gene is a key regulator of progranulin expression and modifies the risk of multiple neurodegenerative diseases. *Sci. Adv.* **6**, eabb3063 (2020).