NIH Public Access
Author Manuscript
Nature. Author manuscript; available in PMC 2014 July 23.

Published in final edited form as:
Nature. 2014 January 23; 505(7484): 550–554. doi:10.1038/nature12825.

Rare coding variants in *Phospholipase D3 (PLD3)* confer risk for Alzheimer's disease

A full list of authors and affiliations appears at the end of the article.
[#] These authors contributed equally to this work.

Abstract

Genome-wide association studies (GWAS) have identified several risk variants for late-onset Alzheimer's disease (LOAD)[1,2]. These common variants have replicable but small effects on LOAD risk and generally do not have obvious functional effects. Low-frequency coding variants, not detected by GWAS, are predicted to include functional variants with larger effects on risk. To identify low frequency coding variants with large effects on LOAD risk, we performed whole exome-sequencing (WES) in 14 large LOAD families and follow-up analyses of the candidate variants in several large case-control datasets. A rare variant in *PLD3* (phospholipase-D family, member 3, rs145999145; V232M) segregated with disease status in two independent families and doubled risk for AD in seven independent case-control series (V232M meta-analysis; OR= 2.10, CI=1.47-2.99; p= 2.93×10^{-5}, 11,354 cases and controls of European-descent). Gene-based burden analyses in 4,387 cases and controls of European-descent and 302 African American cases and controls, with complete sequence data for *PLD3*, indicate that several variants in this gene increase risk for AD in both populations (EA: OR= 2.75, CI=2.05-3.68; p=1.44×10^{-11}, AA: OR= 5.48, CI=1.77-16.92; p=1.40×10^{-3}). *PLD3* is highly expressed in brain regions vulnerable to AD pathology, including hippocampus and cortex, and is expressed at lower levels in neurons from AD brains compared to control brains (p=8.10×10^{-10}). Over-expression of PLD3 leads to a significant decrease in intracellular APP and extracellular Aβ42 and Aβ40, while knock-down of PLD3 leads to a significant increase in extracellular Aβ42 and Aβ40. Together, our genetic and functional data indicate that carriers of *PLD3* coding variants have a two-fold increased risk for

[*]To whom correspondence should be addressed at: Department of Psychiatry, Washington University School of Medicine, 660 South Euclid Avenue B8134, St. Louis, MO 63110. ccruchaga@wustl.edu. tel. 314-286-0546, fax. 314-747-2983.

AUTHOR CONTRIBUTION:
All the authors read and approved the manuscript. C.C. conceived and designed the experiments, supervised research. Wrote the manuscript. Performed the family and sample selection for exome-sequencing, analysis of the segregation data, statistical analysis, and the alternative splicing experiments and analysis. C.M.K., S.H, J.C and A.T.J. performed all the cell-based analysis, and the PLD3 total gene expression experiments. S.C.J. performed PLD3 pool-sequencing experiments. B.A.B performed the genotyping of V232M and A442A in the Knight-ADRC and NIA-LOAD datasets. Analyzed public gene-expression databases and bioinformatic analysis of the effect of some variants on splicing. OH, S.B and Y.C. performed statistical and bioinformatic analyses. J.N. and D.L. recruited and assessed the NIA-LOAD families with the *PLD3* variants. J.B. T.S, D.C and B.C Performed Sequenom genotyping. R.G., C.S., J.B., M.K.L., J.P., J.R.G., A.S.,J.H. P.F., P.G.R., C.D.C., J.T.T., M.C.N, R.G.M., C.S., M.L., J.S.K.K., F.Y.D., M.N.B., X.W., O.L., M.G., M.I.K., C.M., J.T., J.L., A.B., I.B., K.B., K.M, O.L., P.P., Z.B., E.S., E.T., E.R., and P.S.G., provided genotype data for the NIA-UK and NIMH datasets, Cache-County dataset, U. Pittsburgh dataset, U. Nottingham dataset, NIA-LOAD, the Wellderly dataset and in the Toronto dataset. M.R. and D.G.H. performed the co-regulation pathway analysis. N.G. performed the neuropathological examination of the PLD3 V232M carriers. J.C.M. supervised recruitment and clinical assessment of the Knight-ADRC subjects, and A.M.G. supervised the functional and genetic experiments and critically reviewed all data and data analysis.

Author Information
Exome-sequencing data is available on NIAGADs (https://www.niagads.org, accession number: NG00033). The authors declare competing financial interests: detailsare available in the online version of the paper.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

LOAD and that *PLD3* influences APP processing. This study provides an example of how densely affected families may be used to identify rare variants with large effects on risk for disease or other complex traits.

The identification of pathogenic mutations in amyloid-beta precursor protein (APP), presenilin (PSEN1) and presenilin 2 (PSEN2) and the association of *apolipoprotein E* (*APOE)* genotype with disease risk led to a better understanding of the pathobiology of Alzheimer's disease (AD), and the development of novel animal models and therapies for AD[3]. Recent studies using next-generation sequencing have also identified a protective variant in *APP*[4], and a low frequency variant in *TREM2* associated with AD risk[5-8] with Odds Ratio (OR) close to that of one *APOE4* allele. In contrast to the loci identified through GWAS[1,2], these studies have led to the identification of functional variants with large effects on AD pathogenesis. Low-frequency coding variants not detected by GWAS may be a source of functional variants with a large effect on LOAD risk[5-8]; however, the identification of such variants remains challenging because most study-designs require WES in very large datasets. One potential solution is to perform WES or whole-genome-sequencing in a highly selected population at increased risk for disease followed by a combination of genotyping and deep resequencing of the variant/gene of interest in large numbers of cases and controls.

We previously reported that families with a clinical history of LOAD in four or more individuals are enriched for genetic risk variants in known AD and frontotemporal dementia (FTD) genes, but some of these families do not carry pathogenic mutations in the known AD or FTD genes[9,10], suggesting that additional genes may contribute to LOAD risk. We ranked 868 LOAD families from the NIA-LOAD study based on number of affected individuals, number of generations affected, the number of affected and unaffected individuals with DNA available, the number of individuals with a definite or probable diagnosis of AD, early age at onset (AAO) and *APOE* genotype (discarding families in which *APOE*4 segregates with disease status). In the 14 selected families, there were at least four affected individuals per family and DNA was available for at least three of these individuals. We sequenced at least two affected individuals per family, prioritizing distantly related affected individuals with the earliest AAO. We also sequenced one unaffected individual in nine families and two unaffected individuals in one family. In total, we performed WES, on 29 affected individuals and eleven unaffected individuals from 14 families of European American ancestry (Table S1-S2).

All variants shared by affected individuals but absent in unaffected individuals within a family, with a minor allele frequency (MAF) lower than 0.5% in the Exome Variant Server (EVS: http://evs.gs.washington.edu/EVS/) were selected and genotyped in the remaining family members to determine segregation with disease (Supplementary results). Next, we examined whether individual variants or variants in the same gene segregated with disease in more than one family. A single variant, rs145999145 (V232M, *PLD3*, chr. 19q13.2), segregated with disease in two independent families (Figure 1 and S1). Next, we determine whether this variant was associated with increased risk for sporadic AD in seven independent datasets (4,998 AD cases and 6,356 controls of European-descent from the

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Knight-ADRC, NIA-LOAD, NIA-UK dataset, Cache-County study, the Universities of Toronto, Nottingham, and Pittsburgh, the NIMH AD series, and the Wellderly study[7,11-14]; Extended Data Table 1) . *PLD3*-V232M was associated with both risk ($p=2.93\times10^{-05}$; OR=2.10, 95%CI=1.47-2.99, Table 1) and AAO ($p=3\times10^{-3}$, Extended Data Figure 1). The frequency of *PLD3*-V232M was higher in AD cases compared to controls in each age-gender-ethnicity matched dataset, with a similar estimated OR for each dataset (Extended Data Table 1 and Extended Data Figure 2) suggesting that the association is unlikely to be a false positive due to population stratification. This was confirmed when population principal components derived from GWAS data were included (Supplementary results, and Figures S2-S3). The association of the V232M variant with AD risk was also independent of *APOE* genotype (Supplementary results, Table S3 and Figure S4).

LOAD risk variants, such as *APOE4*, are most common in AD cases with a family history of disease and least common in elderly controls without disease[8,9]. We examined the frequency of V232M in three non-demented groups stratified by age (>65yrs, >70yrs and >80yrs, table 1) and compare them with in sporadic vs. familial AD cases. As predicted for an AD risk allele, V232M showed age-dependent differences in frequency among controls with the lowest frequency in the Wellderly dataset, a series composed of healthy non-demented individuals older than 80 years (carrier frequency 0.27%). Similarly, no V232M carriers were found among the 303 non-demented individuals with normal cerebrospinal fluid Aβ42 and tau profiles; suggesting that the calculated OR for the V232M variant when compared to all controls may be an underestimation (supplementary results, Table S4). As hypothesized the frequency of V232M was higher in familial cases than in sporadic cases (2.62% in familial vs. 1.36% in sporadic cases).

Several risk variants have been observed in *APP*, *PSEN1-2* and *APOE* supporting the role of these genes in AD risk[3,4]. In order to identify additional risk variants in *PLD3*, we sequenced the *PLD3* coding region in 2,363 cases and 2,024 controls of European-descent (Extended Data Table 2-3). Fourteen variants were observed more frequently in cases than in controls, including nine variants that were unique to cases (Figure 2A, supplementary results). The gene-based burden analysis resulted in a genome-wide significant association of carriers of *PLD3* coding variants among AD cases (7.99%) compared to controls (3.06%; $p=1.44\times10^{-11}$; OR=2.75, 95%CI=2.05-3.68). When the V232M variant was excluded, to avoid the “winner course”, the association remained highly significant, still passing genome-wide multiple test correction ($p=1.58\times10^{-8}$; OR=2.58, 95%CI=1.87-3.57, Extended Data Table 3), indicating that there are additional variants in *PLD3* that increase risk for AD independent of V232M. There were two additional highly conserved variants (Figure S5), that were nominally associated with LOAD risk: M6R (p=0.02; OR=7.73, 95%CI=1.09-61), and A442A ($p=3.78\times10^{-7}$; OR=2.12, 95%CI=1.58-2.83). The A442A variant showed an association with LOAD risk in four independent series (Extended Data Table 4). The A442A variant was included in the gene-based analysis because our bioinformatic and functional analyses indicate that this variant affects splicing and gene expression (see below).

If the association of *PLD3* with AD risk is real, we hypothesized that rare coding variants in *PLD3* in other populations will also increase risk for AD. We therefore, sequenced *PLD3* in

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

302 African-American AD-cases and controls. Both, the V232M and the A442A variants were found in AD cases but not controls, and the A442A variant showed a significant association with AD risk (p=0.03). There was also a significant association with LOAD risk at the gene level ($p=1.4\times10^{-3}$; OR=5.48, 95%CI=1.77-16.92; Figure 1, Extended Data Table 5, supplementary results). This consistent evidence of association with AD risk, at the SNP and gene-level in two different populations strongly supports *PLD3* as an AD risk gene.

To begin to understand the link between *PLD3* and AD, we analyzed *PLD3* expression in AD case and-control brains. In human brain tissue from cognitively normal individuals, *PLD3* showed high levels of expression in the frontal, temporal, and occipital cortices and hippocampus[15] (Figure S6). Using data from gene-expression in laser-captured neurons from AD cases and controls, *PLD3* gene expression was significantly lower in AD cases compared to controls ($p=8.10\times10^{-10}$; Figure 2B). This result was replicated in three additional independent datasets (Supplementary results, Extended Data Figure 3). Bioinformatic analyses predicted that the A442A variant affects alternative splicing (Figure S7, supplementary results). We found that A442A is associated with lower levels of total *PLD3* mRNA (Figure 2D) and lower levels of transcripts containing exon-11 (Figure 2C, and Figure S8), supporting the functional effect of this variant.

PLD3 is a non-classical, poorly characterized member of the PLD superfamily of phospholipases. PLD1 and PLD2 have been previously implicated in APP trafficking and AD[16-18]. To determine whether PLD3 also affects APP processing, wild-type (WT) human PLD3 was overexpressed in mouse neuroblastoma (N2A) cells that stably express wild-type human APP695 (termed N2A-695). In this system extracellular $A\beta_{42}$ and $A\beta_{40}$ were decreased by 48% and 58%, respectively, compared to the empty vector ($P<0.0001$; Figure 3A). Conversely, knockdown of endogenous PLD3 expression by shRNA in N2A-695 cells resulted in higher levels of extracellular $A\beta_{42}$ and $A\beta_{40}$ than in cells transfected with scrambled shRNA. (Figure 3B). To determine if the observed effects on APP processing were unique to PLD3 or common among the phospholipase D protein family, we co-expressed APP695-WT with PLD1, PLD2, and PLD3 in human embryonic kidney (HEK293T) cells. Overexpression of PLD3, but not empty vector, PLD1 or PLD2, resulted in a substantial decrease in full-length APP levels (Figure 3C). Extracellular $A\beta_{42}$ and $A\beta_{40}$ levels were significantly reduced in cells overexpressing PLD1, PLD2, and PLD3 compared to control (Figure 3C). Interestingly, overexpression of catalytically-inactive PLD1 and PLD2 variants (PLD1-K898R and PLD2-K758R) restored extracellular $A\beta_{42}$ and $A\beta_{40}$ levels to control values, demonstrating that this is in part a phospholipase activity-dependent effect (Figure 3C). Overexpression of a PLD3 dominant-negative variant (PLD3-K418R) that inhibits myotube formation[19] failed to restore full-length APP and $A\beta_{42}$ and $A\beta_{40}$ levels to normal levels (Figure 3C). Furthermore, PLD3 can be co-immunoprecipitated with APP in cultured cells (Extended Data Figure 4). Together, these studies demonstrate that PLD3 plays a role in APP processing that is functionally distinct from PLD1 and PLD2. These findings are consistent with the human genetic and brain expression data presented above, whereby lower PLD3 expression/function is correlated with higher APP and Aβ levels and with more extensive AD-specific pathology (Table S4).

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Here, we provide compelling genetic evidence that *PLD3* is an AD risk gene: genome-wide significant evidence that rare variants in *PLD3* increase risk for AD in multiple datasets and two populations. In addition, our functional studies confirm that PLD3 affects APP processing, in a manner consistent with increased risk for AD[3,20]. This work also provides a second example of a novel gene containing rare variants that influence risk for AD[5,7,8]. While these variants have low population attributable fraction and diagnostic utility due to their rarity they provide important and novel insights into AD pathogenesis. Our success in identifying multiple families carrying this rare variant and the enrichment of this variant in LOAD families compared to sporadic AD cases demonstrates the power of using a highly-selected sample of multiplex LOAD families for variant discovery. The studies on *TREM2*[5-8] and this report, suggest that next-generation sequencing project will identify additional low frequency and rare variants associated with Alzheimer's Disease.

METHODS

Participants and Study Design

The Institutional Review Board (IRB) at Washington University School of Medicine in Saint Louis approved the study. Written informed consent was obtained from participants and their family members by the Clinical Core of the Knight ADRC. The approval number for the Knight ADRC Genetics Core is 93-0006.

Knight-ADRC samples

The Knight ADRC sample includes 1,114 late-onset AD (LOAD) cases and 913 cognitively normal controls (377 older than 70yrs), matched for age, gender and ethnicity, European descent and 302 African-American AD cases and controls. These individuals were evaluated by Clinical Core personnel of the Knight ADRC at Washington University. Cases received a clinical diagnosis of AD dementia in accordance with standard criteria, dementia severity was determined using the Clinical Dementia Rating (CDR)[28]. 2,027 individuals were of European descent, and 302 were African-American.

Cerebrospinal fluid levels (CSF) dataset: A subset (n=528) of the Knight-ADRC samples had CSF tau and Aβ42 levels. Of these 528, 303 were non-demented (CDR=0) elderly (Age>65) individuals with high CSF Aβ42 levels (>500 pg/ml). A description of the CSF dataset used in this study can be found in Cruchaga et al[11]. CSF collection and $A\beta_{42}$, tau and ptau181 measurements were performed as described previously[29].

NIA-LOAD

Participants from the National Institute of Aging Late Onset Alzheimer Disease Family Study (NIA-LOAD Family Study) included a single demented individual from each of 868 families with at least three AD-affected individuals and 881 unrelated nondemented elderly controls (545 older than 70). All AD cases were diagnosed with dementia of the Alzheimer's type (DAT) using criteria equivalent to the National Institute of Neurological and Communication Disorders and Stroke-Alzheimer's Disease and Related Disorders Association (NINCDS-ADRDA) for probable AD[30]. NIALOAD families were ascertained based on the following criteria: Probands were required to have a diagnosis of definite or

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

probable late-onset AD (onset >60yrs) and a sibling with definite, probable or possible late-onset AD with a similar age at onset. A third biologically-related family member (first, second or third degree) was also required, regardless of affection status. This individual had to be ≥60 years of age if unaffected, or ≥50 years of age if diagnosed with LOAD or mild cognitive impairment[12]. Within each pedigree, we selected a single individual for the case control series by identifying the youngest affected family member with the most definitive diagnosis (i.e. individuals with autopsy confirmation were chosen over those with clinical diagnosis only). Unrelated nondemented controls used for the NIALOAD case control series had no family history of AD and were matched to the cases as previously described[12]. Only individuals of European descent based on the PCs were included. Written informed consent was obtained from all participants, and the study was approved by local IRB committees.

Wellderly Study

The Scripps Translational Science Institute's Wellderly study has recruited more than 1000 healthy-aged participants. Inclusion criteria specify informed consent, age > 80 years, blood or saliva donation, compliance with protocol-specified procedures, and no/mild aging-related medical conditions. Exclusion criteria includes self-reported cancer (excluding basal and squamous cell skin cancer), coronary artery disease/myocardial infarction, stroke/TIA, DVT/PE, CRF/hemodialysis, Alzheimer's/Parkinson's disease, diabetes, aortic/cerebral aneurysm, or the use of oral chemotherapeutic agents, anti-platelet agents (excluding aspirin), cholinesterase inhibitors for Alzheimer's disease, or insulin. All genotyped individuals were of European descent.

Cache-County study

The Cache County Study was initiated in 1994 to investigate the association of *APOE* genotype and environmental exposures on cognitive function and dementia. A cohort comprised of 5,092 Cache County, Utah, residents (90% of those aged 65 or older) has been followed continually for over fifteen years, completing four triennial waves of data collection including clinical assessments (Breitner et al 1999). Genotypes were obtained for 255 demented individuals and 2471 elderly cognitively normal individuals[13]. All individuals genotyped were of European descent.

UK-NIA dataset

A description of the UK-NIA dataset can be found in Guerreiro et al., 2013[7]. Briefly, this dataset includes WES from 143 AD cases and 183 elderly nondemented controls. All subjects were of European descent.

University of Pittsburgh dataset

The PLD3 V232M variant was genotyped in 2,211 subjects including 1,253 AD cases (62.6% females) and 958 elderly nondemented controls (64.3% females). A complete description of the dataset can be found in Kamboh et al.[14] All samples were of European descent.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Toronto Dataset

The Toronto dataset was composed of 269 unrelated AD cases (53% females) and 250 unrelated non-demented controls (56% females) of European descent. The mean (SD) age at onset of AD was 73 (+/-8) years, and the mean (SD) age at last examination of the controls was 73 (+/-10) years. The study was approved by the Institutional Review Boards of the University of Toronto.

Exome sequencing

Enrichment of coding exons and flanking intronic regions was performed using a solution hybrid selection method with the SureSelect® human all exon 50Mb kit (Agilent Technologies, Santa Clara, California) following the manufacturer's standard protocol. This step was performed by the Genome Technology Access Center at Washington University in St Louis. The captured DNA was sequenced by paired-end reads on the HiSeq 2000 sequencer (Illumina, San Diego, California). Raw sequence reads were aligned to the reference genome hg19 using Novoalign (Novocraft Technologies, Selangor, Malaysia). Base/SNP calling was performed by SNP Samtools. SNP annotation was carried out using version 5.07 of SeattleSeq Annotation server (see URL)[21].

On average, 95% of the exome had >eight-fold coverage. SNP calls were made using SAM tools [31]. SNPs identified with a quality score lower than 20 and a depth of coverage lower than 5 were removed. More than 2,500 novel variants in the coding region were found per individual. We identified all variants shared by the affected individuals in a family. Variants not present in 1,000 genome project or the Exome Variant Server (EVS: http://evs.gs.washington.edu/EVS/) or with a frequency lower than 0.5% in the EVS were selected. On average, 80 coding variants were selected for each family. The selected variants were then genotyped in the remaining sampled family members. We validated >98% of the selected variants, confirming the high specificity of our exome-sequencing method and analysis. On average, we genotyped a total of 13 family members (7 cases and 6 controls) per family.

SNP Genotyping

SNPs were genotyped using the Illumina Golden Gate, Sequenom, Kaspar and/or Taqman genotyping technologies. Only SNPs with a genotyping call rate higher than 98% and in Hardy-Weinberg equilibrium were used in the analyses. The principle of the MassARRAY system is PCR-based where different size products are analyzed by SEQUENOM MALDI-TOF mass spectrometry [22,32]. The KBioscience Competitive Allele-Specific PCR genotyping system (KASP) is FRET-based endpoint-genotyping technology, v4.0 SNP (KBioscience) [22,32]. Genotype call rates were >98%.

PLD3 sequencing

PLD3 was sequenced in 2,363 cases and 2,027 controls of European origin, and 130 cases and 172 controls of African-American descent using a pooled-DNA sequencing design as described previously[9,24,33]. Briefly, equimolar amounts of individual DNA samples were pooled together following quantification using the Quant-iT™ PicoGreen reagent. Pools contained 100 ng of DNA/individual from 94 individuals. The coding exons and flanking

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

regions (a minimum of 50 bp each side) were individually PCR amplified using specific primers and Pfu Ultra high-fidelity polymerase (Stratagene). An average of 20 diploid genomes (approximately 0.14 ng DNA) per individual were used as input. PCR products were cleaned using QIAquick PCR purification kits, quantified using Quant-iT PicoGreen reagent and ligated in equimolar amounts using T4 Ligase and T4 Polynucleotide Kinase. After ligation, concatenated PCR products were randomly sheared by sonication and prepared for sequencing on an Illumina HighSeq2000 according to the manufacturer's specifications. pCMV6-XL5 amplicon (1908 base pairs) was included in the reaction as a negative control. As positive controls, ten different constructs (p53 gene) with synthetically engineered mutations at a relative frequency of one mutated copy per 188 normal copies was amplified and pooled with the PCR products.

Paired-end reads (101 bp) were aligned to the human genome reference assembly build 36.1 (hg19) using SPLINTER[33]. SPLINTER uses the positive control to estimate sensitivity and specificity for variant calling. The wild type: mutant ratio in the positive control is similar to the relative frequency expected for a single mutation in one pool (1 chromosome mutated in 94 samples= 1/188). SPLINTER uses the negative control (first 900bp) to model the errors across the 101-bp Illumina reads and to create an error model from each sequencing run. Based on the error model SPLINTER calculates a p-value for the probability that a predicted variant is a true positive. A p-value at which all mutants in the positive controls were identified was defined as the cut-off value for the best sensitivity and specificity. All mutants included as part of the amplified positive control vector were found upon achieving >30-fold coverage at mutated sites (sensitivity = 100%) and only ~80 sites in the 1908 bp negative control vector were predicted to be polymorphic (specificity = ~95%). The variants with a p-value below this cut-off value were considered for follow-up genotyping confirmation. All rare missense or splice site variants were then validated by Sequenom and KASPar genotyping in each individual included in the pools. In order to avoid any batch/plate effects, cases and controls were included in each genotyping plate and all genotyping was performed in a single experiment. Finally, in order to confirm all of the heterozygous calls, we created a custom DNA plate including all of the heterozygotes (cases and controls) for all of the variants, and then genotyped them again by Sequenom, creating a new Sequenom set.

Gene-expression and Alternative splicing analyses

Total RNA was extracted using the RNeasy mini kit (Qiagen) following the manufacturer's protocol from 82 AD cases and 39 non-demented individuals. Extracted RNA was treated with DNase1 to remove any potential DNA contamination. cDNAs were prepared from the total RNA, using the High-Capacity cDNA Archive kit (ABI). Gene expression levels were analyzed by real-time PCR, using an ABI-7900 real-time PCR system. The PLD3-A442A variant was genotyped in DNA extracted from parietal lobe of 82 AD cases and 39 non-demented individuals by Kaspar as explained below. A total of eight carriers for the A442A variant were identified.

Total PLD3 expression—Gene expression was analyzed by real-time PCR, using an ABI-7500 real-time PCR system. TaqMan assays were used to quantify *PLD3* mRNA

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

levels. Primers and TaqMan probe for the reference gene, *GAPDH*, were designed over exon-exon boundaries, using Primer Express software, Version 3 (ABI) (sequences available on request). Cyclophilin A (ABI: 4326316E) was also used as a reference gene. Each real-time PCR run included within-plate triplicates and each experiment was performed, at least twice for each sample.

Alternative splicing—We selected eight A442A carriers as well as eight CDR, age, *APOE* and PMI matched individuals to analyze the expression level of exon 11 containing transcripts, the exon in which the A442A variant is located. Real-time PCR assays were used to quantify PLD3 exon 7 (forward primer: GCAGCTCCATCCCATCAACT; reverse: CTTGGTTGTAGCGGGTGTCA), exon 8 (forward primer: CTCAACGTGGTGGACAATGC; reverse: AGTGGGCAGGTAGTTCATGACA), 9 (forward primer: ACGAGCGTGGCGTCAAG; reverse: CATGGATGGCTCCGAGTGT), 10 (forward primer: GGTCCCCGCGGATGA; reverse: GGTTGACACGGGCATATGG) and 11 (first pair of primers: forward primer: CCAGCTGGAGGCCATTTTC; reverse: TGTCAAGGTCATGGCTGTAAGG; second pair forward primer: GCTGCTGGTGACGCAGAAT; reverse: AGTCCCAGTCCCTCAGGAAAA). Two pairs of primers were designed for exon 11 as an internal control. Sybr-green primers were designed using Primer Express software, Version 3 (ABI). Each real-time PCR run included within-plate duplicates and each experiment was performed, at least twice for each sample. Real-time data were analyzed using the comparative Ct method. Only samples with a standard error of <0.15% were analyzed. The Ct values for exon 11 were normalized with the Ct value for the exons 7-10. The relative exon 11 levels for the A442A carriers vs. the non-carriers were compared using a t-test.

PLD3 gene expression in public databases

We also used the GEO datasets GSE15222[34] and GSE5281[27] to analyze the association of *PLD3* gene expression and case-control status. In the GSE15222 dataset, there are genotype and expression data from 486 late onset Alzheimer's Disease cases and 279 neuropathologically clean non-demented individuals. In the GSE5281 dataset, samples were laser-captured from cortical regions of 16 normal elderly humans (10 males and 4 females) and from 33 AD cases (15 males and 18 females). Mean age of cases and controls was 80 years. All samples were run on the Affymetrix U133 Plus 2.0 array. RNA data were re-normalized to an average expression of 8 units on a log2 scale. As potential covariates we analyzed brain region, gender, and age for each sample. Stepwise discriminant analysis was used to identify the potential covariates to be included in the ANCOVA. For this dataset we also extracted the gene expression levels for *APP* (probe 211277_x_at), *PSEN1* (1559206_at), and *PSEN2* (203460_s_at) to examine the correlation between *PLD3* and *APP*, *PSEN1*, and *PSEN2* using the Pearson correlation method.

Human brain samples and analysis of Affymetrix Human Exon 1.0 ST array

Quantification and analysis of *PLD3* gene expression in brains was performed as previously described[35]. Briefly, the human data used here were provided by the UK Human Brain Expression Consortium[35] and consisted of 101 control post-mortem brains. All samples originated from individuals with no significant neurological history or neuropathological

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

abnormality and were collected by the MRC Edinburgh Brain Bank[36] ensuring a consistent dissection protocol and sample handling procedure. A summary of the available demographic details of these samples including a thorough analysis of their effects on array quality is provided by Trabzuni et al[37]. All samples had fully informed consent for retrieval and were authorized for ethically approved scientific investigation (Research Ethics Committee number 10/H0716/3). Total RNA was isolated from human post-mortem brain tissues using the miRNeasy 96 well kit (Qiagen, UK). The quality of total RNA was evaluated by the 2100 Bioanalyzer (Agilent, UK) and RNA 6000 Nano Kit (Agilent, UK) before processing with the Ambion® WT Expression Kit and Affymetrix GeneChip Whole Transcript Sense Target Labeling Assay and hybridization to the Affymetrix Exon 1.0 ST. All arrays were pre-processed using Robust Multi-array Average using Partek Genomics Suite v6.6 (Partek Incorporated, USA). The resulting expression data was corrected for individual effects (within which are nested postmortem interval, brain pH, sex, age at death and cause of death) and experimental batch effects (date of hybridization). Transcript-level expression was calculated for 26,993 genes using Winsorized means (Winsorizing the data below 10% and above 90%).

RNA-Pathway analysis

To evaluate the biological and functional relevance of co-expressed genes within the *PLD3*-containing modules, we used Weighted Gene Co-expression Network Analysis (WGCNA)[15] and DAVID v6.7 (http://david.abcc.ncifcrf.gov/), the database for annotation, visualization and integrated discovery[38]. We restricted WGCNA to 15,409 transcripts that passed the Detection Above Background (DABG) criteria (p-value < 0.001 in at least 50% of samples in at least one brain region), had a coefficient of variation > 5% and expression values exceeding 5 in all samples in at least one brain region. We followed a step-by-step network construction and module detection. In short, for each brain region, the Pearson correlations between all genes across all relevant samples were derived. We then calculated a signed-weighted co-expression adjacency matrix, allowing us to keep track of the direction of the correlation. A power 12, the default soft threshold parameter for constructing a signed-weighted Network[39] was used in all brain regions, after checking that it recapitulated scale-free topology[40]. Topological overlap (TO), a more biologically meaningful measure of node interconnectedness (similarity)[9,24], was subsequently calculated and genes were hierarchically clustered using 1 − TO as the distance measure. Finally modules were determined by using a dynamic tree-cutting algorithm. WGCNA led to the identification of several co-expression modules, ranging in number and size between the ten brain regions. We examined the over-representation (i.e. enrichment) of the three Gene Ontology (GO) categories (biological processes, cellular components, and molecular function) and KEGG pathways for each list of co-expressed genes with *PLD3* for each tissue by comparing numbers of significant genes annotated with this biological category with chance.

Statistical Analyses

All of the single SNP analyses were performed using a Fisher's exact test, with no covariates included. Allelic association with risk for AD was tested using “proc logistic” in SAS including *APOE genotype*, age, PCs and study as covariates when available. Odds ratios with 95% confidence intervals and relative risks (RR) were calculated for the alternative

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

allele compared to the most common allele using SAS. Association with age at onset (AAO) was carried out using the Kaplan-Meier method and tested for significant differences, using a proportional hazards model (proc PHREG, SAS) including gender and study as covariates. Non-demented controls were included in the analyses as censored data. The inclusion of these samples did not change the association. Gene-based analyses were performed using the optimal SNP-set (Sequence) Kernel Association Test (SKAT-O)[26].

Population Attributable risk (PAR)

We calculated the PAR using the RR obtained in the study and the MAF from the EVS database (http://evs.gs.washington.edu/EVS/) and in the Cache-County dataset, which is a population-based dataset, using the equation:

$$PAR = \frac{P_e\,(RR_e - 1)}{[1 + P_e\,(RR_e - 1)]}$$

Where P_e is the carrier frequency in the population and RR_e is the relative risk for the different variants.

Neuropathological studies

All study procedures were approved by Washington University's Human Research Protection Office. At autopsy, brain tissue was obtained from participants according to the protocol of the Knight-ADRC. AD neuropathologic change was assessed according to the criteria of the National Institute on Aging-Alzheimer's Association (NIA-AA)[41]. Dementia with Lewy bodies was assessed using the criteria of McKeith et al. (2005)[42].

Cell-based studies

Plasmids and site-directed mutagenesis—The following plasmids were used in this study: pCMV6-XL5 human PLD3 wild-type (Origene), pCS2-myc human APP695 wild-type[43], pCGN-PLD1b WT[44] and K758R[45], pCGN-PLD2 WT[46] and K898R[45], pGFL GFP[47], pGFP-V-RS-PLD3-shRNA-GI548821 (Origene), and pGFP-V-RSScr-shRNA-TR30013 (Origene). A dominant negative mutation (K418R)[19] was introduced into the pCMV6-XL5 human PLD3 wild-type vector by site-directed mutagenesis using the QuikChangeII Site-Directed Mutagenesis kit (Agilent). All constructs were verified by Sanger sequencing.

Cell Culture Assays—Human embryonic kidney (HEK293-T) cells were cultured in Dulbecco's modified eagle medium (DMEM) supplemented with 10% fetal bovine serum (FBS), 1% L-glutamine, and penicillin/streptomycin. HEK293T cells were grown in 6-well lysine coated plates. Mouse neuroblastoma cells (N2A) stably expressing human APP695 wild-type were cultured in DMEM and Opti-MEM (50:50) supplemented with 5% FBS, 1% L-glutamine, penicillin/streptomycin, and 500 ug/mL G418. Upon reaching confluency, cells were transiently transfected with Lipofectamine 2000 (Invitrogen). Culture media were replaced after 24 hours, and cells were incubated for another 24 hours. Conditioned media were collected, treated with protease inhibitor cocktail and centrifuged at 3000xg at 4°C for 10 minutes to remove cell debris. Cell pellets were extracted on ice in lysis buffer (50mM

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Tris, pH 7.6, 2mM EDTA, 150mM NaCl, 1%NP40, 0.5% Triton 100x, protease inhibitor cocktail) and centrifuged at 14,000xg. Protein concentration was measured by the BCA method as described by the manufacturer (Pierce-Thermo).

RT-PCR and qPCR—To confirm effective knockdown of endogenous mouse PLD3 in mouse N2A-695 cells, RNA was extracted from cell lysates with an RNeasy kit (Qiagen) according to the manufacture's protocol. Extracted RNA (10ug) was converted to cDNA by PCR using a High-Capacity cDNA Reverse Transcriptase kit (ABI). Gene expression was analyzed by qPCR using an ABI-7900 Real-Time PCR system. Taqman Real-Time PCR assays were utilized to quantify expression for mouse *PLD3* (ABI:Mm01171272_m1) and *GAPDH* (ABI: Hs02758991_g1). Samples were run in triplicate. To avoid amplication interference, expression asays were run in separate wells from the housekeeping gene *GAPDH*. Real-time data were analyzed by the comparative C_T method. Average C_T values for each sample were normalized to the average C_T values for the housekeeping gene *GAPDH*. The resulting value was corrected for assay efficiency. Samples with a standard error of 20% or less were analyzed.

Immunoblotting—Standard SDS-PAGE was performed in 4-20% Criterion Tris-HCl gels (Bio-Rad). Samples were boiled for 5 min in Laemmli sample buffer prior to electrophoresis[48]. Immunoblots were probed with antibodies: PLD3 (Sigma), 9E10 (Sigma), and β-tubulin (Sigma).

Enzyme-linked immunosorbent assay—The levels of Aβ40 and Aβ42 were measured in cell culture media by sandwich ELISA as described by the manufacturer (Invitrogen). ELISA values were obtained (pg/mL) and corrected for total intracellular protein (ug/mL) based on BCA measurements.

Immunoprecipitation—Cell lysates were pre-cleared with Protein G beads (Thermo Scientific). Pre-cleared supernatants were incubated overnight at 4°C with the antibodies indicated. Supernatant-antibody complexes were then incubated with Protein G beads at room temperature for 2 h. After washing, proteins were dissociated from the Protein G beads by incubating the beads in Laemmli sample buffer[48] supplemented with 5% β-mercaptoethanol at 95°C for 10 min.

Bioinformatics analysis

SIFT (http://sift.jcvi.org/www/SIFT_BLink_submit.html) and Polyphen (http://genetics.bwh.harvard.edu/pph2/) algorithms were used to predict the functional effect of the identified variants. To determine the effect of the A442A variant on splicing we used the ESEfinder (http://rulai.cshl.edu/tools/ESE). Multiple Sequence Alignment was performed by ClustalW2, and the PLD3 orthologues were downloaded from Ensembl (http://www.ensembl.org/).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Authors

Carlos Cruchaga[1,2,*], Celeste M. Karch[#1,2], Sheng Chih Jin[#1], Bruno A. Benitez[1], Yefei Cai[1], Rita Guerreiro[7,8], Oscar Harari[1], Joanne Norton[1], John Budde[1], Sarah Bertelsen[1], Amanda T. Jeng[1], Breanna Cooper[1], Tara Skorupa[1], David Carrell[1], Denise Levitch[1], Simon Hsu[1], Jiyoon Choi[1], Mina Ryten[7,9], Celeste Sassi[7,8], Jose Bras[7], Raphael J. Gibbs[7,8], Dena G. Hernandez[7,8], Michelle K. Lupton[10,30], John Powell[10], Paola Forabosco[11], Perry G. Ridge[12], Christopher D. Corcoran[13,14], JoAnn T. Tschanz[14,15], Maria C. Norton[14,15,16], Ronald G. Munger[16,17], Cameron Schmutz[12], Maegan Leary[12], F. Yesim Demirci[19], Mikhil N. Bamne[19], Xingbin Wang[19], Oscar L. Lopez[20,22], Mary Ganguli[21], Christopher Medway[23], James Turton[23], Jenny Lord[23], Anne Braae[23], Imelda Barber[23], Kristelle Brown[23], The Alzheimer's Research UK (ARUK) Consortium, Pau Pastor[27,28,29], Oswaldo Lorenzo-Betancor[27], Zoran Brkanac[26], Erick Scott[18], Eric Topol[18], Kevin Morgan[23], Ekaterina Rogaeva[24], Andy Singleton[8], John Hardy[7], M. Ilyas Kamboh[20,20,21], Peter St George-Hyslop[24,25], Nigel Cairns[2,3], John C. Morris[3,4,5], John S.K. Kauwe[12], and Alison M. Goate[1,2,4,5,6]

Affiliations

[1] Department of Psychiatry, Washington University, St. Louis, MO, USA

[2] Hope Center Program on Protein Aggregation and Neurodegeneration, Washington University St. Louis, MO, USA

[3] Pathology and Immunology, Washington University, St. Louis, MO, USA

[4] Department of Neurology, Washington University, St. Louis, MO, USA

[5] Knight ADRC, Washington University, St. Louis, MO, USA

[6] Department of Genetics, Washington University, St. Louis, MO, USA

[7] Department of Molecular Neuroscience, UCL Institute of Neurology, London WC1N 3BG, UK

[8] Laboratory of Neurogenetics, National Institute on Aging, National Institutes of Health, Bethesda, Maryland, United States of America

[9] on behalf of UKBEC (UK Brain Expression Consortium)

[10] Institute of Psychiatry, King's College London, London, UK

[11] Istituto di Genetica delle Popolazioni – CNR, Sassari, Italy

[12] Department of Biology, Brigham Young University, Provo, UT, 84602

[13] Department of Mathematics and Statistics, Utah State University, Logan, UT

[14] Center for Epidemiologic Studies, Utah State University, Logan, UT

[15] Department of Psychology, Utah State University, Logan, UT

[16] Department of Family Consumer and Human Development, Utah State University, Logan, UT

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

[17] Department of Nutrition, Dietetics, and Food Sciences, Utah State University, Logan, UT

[18] The Scripps Research Institute, La Jolla, CA, US.

[19] Department of Human Genetics, University of Pittsburgh, Pittsburgh, PA

[20] Alzheimer's Disease Research Center, University of Pittsburgh, Pittsburgh, PA

[21] Department of Psychiatry, University of Pittsburgh, Pittsburgh, PA

[22] Department of Neurology, University of Pittsburgh, Pittsburgh, PA

[23] Human Genetics, School of Molecular Medical Sciences, University of Nottingham, Nottingham, NG7 2UH, UK

[24] Tanz Centre for Research in Neurodegenerative Diseases, University of Toronto

[25] Cambridge Institute for Medical Research, and the Department of Clinical Neurosciences, University of Cambridge

[26] University of Washington. Seattle, WA

[27] Neurogenetics Laboratory, Division of Neurosciences, Center for Applied Medical Research, University of Navarra, Pamplona, Spain

[28] Department of Neurology, Clínica Universidad de Navarra, School of Medicine, University of Navarra, Pamplona, Spain

[29] CIBERNED, Centro de Investigación Biomédica en Red de Enfermedades Neurodegenerativas, Instituto de Salud Carlos III, Spain

[30] Neuroimaging Genetics, QIMR Berghofer Medical Research Institute, Brisbane, Australia

Acknowledgments

We thank Dr. Michael Frohman for providing us with PLD1 and PLD2-WT constructs as well as constructs for the inactive mutations in these genes. This work was supported by grants from the National Institutes of Health (P30-NS069329, R01-AG044546 and R01-AG035083), the Alzheimer Association (NIRG-11-200110) and Barnes Jewish Foundation. This research was conducted while CC was a recipient of a New Investigator Award in Alzheimer's Disease from the American Federation for Aging Research. CC is a recipient of a BrightFocus Foundation Alzheimer's Disease Research Grant (A2013359S). Sequencing of some of the families included in this study was supported by Genentech and Pfizer.

The recruitment and clinical characterization of research participants at Washington University were supported by NIH P50 AG05681, P01 AG03991, and P01 AG026276. This work was supported in part by the Intramural Research Program of the National Institute on Aging, National Institutes of Health, Department of Health and Human Services; project ZO1 AG000950-11. Samples from the National Cell Repository for Alzheimer's Disease (NCRAD), and NIA-LOAD which receives government support under a cooperative agreement (U24 AG21886; U24: 5U24AG026395 and 1R01AG041797) were used in this study. We thank our contributors, including the Alzheimer's Disease Centers, that collected samples used in this study, as well as participants and their families, whose help and participation made this work possible. The Cache County Study is supported by National Institutes of Health, RO1-AG11380, RO1-AG18712, RO1-AG21136. Genotyping and analysis conducted at Brigham Young University was funded by grants from the National Institutes of Health R01-AG042611 and the Alzheimer's Association (MNIRG-11-205368) to Dr. Kauwe. The sequencing at University of Washington was supported by NIH R01-039700. The sequencing for the NIA-UK samples was supported by the Alzheimer's Research UK (ARUK), by an anonymous donor, by the Wellcome Trust/MRC Joint Call in Neurodegeneration award (WT089698) to the UK Parkinson's Disease Consortium (UKPDC) whose members are from the UCL/Institute of Neurology, the University of Sheffield and the MRC Protein Phosphorylation Unit at the University of Dundee, by

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

the Big Lottery (to Dr. Morgan) and by a fellowship from ARUK to Dr. Guerreiro. It was also supported in part by the Intramural Research Programs of the National Institute on Aging and the National Institute of Neurological Disease and Stroke, National Institutes of Health, Department Of Health and Human Services Project number ZO1 AG000950-10. Some samples and pathological diagnoses were provided by the MRC London Neurodegenerative Diseases Brain Bank and the Manchester Brain Bank from Brains for Dementia Research, jointly funded from ARUK and AS via ABBUK Ltd. Supported in part by the NIHR Queen Square Dementia BRU and BRC NIHR grant mechanisms. The sample recruitment and genetic studies at University of Pittsburgh are funded by NIH grants: AG041718, AG030653, AG005133, AG07562, and AG023652. The Toronto sample studies are funded by Canadian Institutes of Health Research, Wellcome Trust, Medical Research Council, National Institute of Health, National Institute of Health Research, Ontario Research Fund and Alzheimer Society of Ontario (to Dr. St George-Hyslop). The Nottingham Lab (KM) is funded by ARUK and Big Lottery. The ARUK Consortium consists of 8 AD research groups from the Universities of Belfast, Bristol, Bonn, Leeds, Manchester, Nottingham, Oxford and Southampton. All AD cases met criteria for either probable (NINCDS-ADRDA, DSM-IV) or definite (CERAD) AD. All controls were either screened for dementia using the MMSE or ADAS-cog or were determined to be free from AD pathology at neuropathological examination. The Alzheimer's Research UK (ARUK) Consortium: Peter Passmore, David Craig, Janet Johnston, Bernadette McGuinness, Stephen Todd, Queen's University Belfast, UK; Reinhard Heun, Royal Derby Hospital, UK; Heike Kölsch, University of Bonn, Germany; Patrick G. Kehoe, University of Bristol, UK; Nigel M. Hooper, University of Leeds, UK; Emma R.L.C. Vardy, University of Newcastle, UK; David M. Mann, Stuart Pickering-Brown, University of Manchester, UK; Kristelle Brown, Noor Kalsheker, James Lowe, Kevin Morgan, University of Nottingham, UK; A. David Smith, Gordon Wilcock, Donald Warden, University of Oxford (OPTIMA),UK, Clive Holmes, University of Southampton, UK. The UK Brain Expression Consortium: John Hardy, Mina Ryten, Daniah Trabzuni, Department of Molecular Neuroscience, UCL Institute of Neurology; Michael E. Weale, Adaikalavan Ramasamy, Department of Medical and Molecular Genetics, King's College London; Colin Smith, MRC Sudden Death Brain Bank Project, University of Edinburgh. This consortium is supported by the UK Medical Research Council through the MRC Sudden Death Brain Bank (C.S.) and by a Project Grant (G0901254 to J.H. and M.W.) and Training Fellowship (G0802462 to M.R.). The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript. This work was supported by grants to P. Pastor from the Department of Health of the Government of Navarra, Spain (refs.13085 and 3/2008) and from the UTE project FIMA, Spain to P.P. J.T.T receives funds from NIA (R01AG21136).

REFERENCES

1. Bertram, L.; McQueen, M.; Mullin, K.; Blacker, D.; Tanzi, R. [1/26/2013] The AlzGene Database. Alzheimer Research Forum. Available at: http://www.alzgene.org
2. Lambert JC, et al. Extended meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. Nat. Genetics. 2013 In Press.
3. Goate A, Hardy J. Twenty years of Alzheimer's disease-causing mutations. J Neurochem. 2012; 120(Suppl 1):3–8. doi:10.1111/j.1471-4159.2011.07575.x. [PubMed: 22122678]
4. Jonsson T, et al. A mutation in APP protects against Alzheimer's disease and age-related cognitive decline. Nature. 2012; 488:96–99. doi:10.1038/nature11283. [PubMed: 22801501]
5. Benitez BA, et al. TREM2 is associated with the risk of Alzheimer's disease in Spanish population. Neurobiol Aging. 2013; 34:1711, e1715–1717. doi:10.1016/j.neurobiolaging.2012.12.018. [PubMed: 23391427]
6. Benitez BA, Cruchaga C. TREM2 and neurodegenerative disease. N Engl J Med. 2013; 369:1567–1568. doi:10.1056/NEJMc1306509#SA4. [PubMed: 24131187]
7. Guerreiro R, et al. TREM2 variants in Alzheimer's disease. N Engl J Med. 2013; 368:117–127. doi: 10.1056/NEJMoa1211851. [PubMed: 23150934]
8. Jonsson T, et al. Variant of TREM2 Associated with the Risk of Alzheimer's Disease. N Engl J Med. 2012 doi:10.1056/NEJMoa1211103.
9. Cruchaga C, et al. Rare variants in APP, PSEN1 and PSEN2 increase risk for AD in late-onset Alzheimer's disease families. PLoS One. 2012; 7:e31039. doi:10.1371/journal.pone.0031039. [PubMed: 22312439]
10. Harms M, et al. C9orf72 Hexanucleotide Repeat Expansions in Clinical Alzheimer Disease. JAMA neurology. 2013:1–6. doi:10.1001/2013.jamaneurol.537.
11. Cruchaga C, et al. GWAS of Cerebrospinal Fluid Tau Levels Identifies Risk Variants for Alzheimer's Disease. Neuron. 2013 doi:10.1016/j.neuron.2013.02.026.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

12. Wijsman EM, et al. Genome-wide association of familial late-onset Alzheimer's disease replicates BIN1 and CLU and nominates CUGBP2 in interaction with APOE. PLoS genetics. 2011; 7:e1001308. doi:10.1371/journal.pgen.1001308. [PubMed: 21379329]
13. Breitner JC, et al. APOE-epsilon4 count predicts age when prevalence of AD increases, then declines: the Cache County Study. Neurology. 1999; 53:321–331. [PubMed: 10430421]
14. Kamboh MI, et al. Genome-wide association study of Alzheimer's disease. Translational psychiatry. 2012; 2:e117. doi:10.1038/tp.2012.45. [PubMed: 22832961]
15. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics. 2008; 9:559. doi:10.1186/1471-2105-9-559. [PubMed: 19114008]
16. Cai D, et al. Phospholipase D1 corrects impaired betaAPP trafficking and neurite outgrowth in familial Alzheimer's disease-linked presenilin-1 mutant neurons. Proc Natl Acad Sci U S A. 2006; 103:1936–1940. doi:10.1073/pnas.0510710103. [PubMed: 16449385]
17. Cai D, et al. Presenilin-1 uses phospholipase D1 as a negative regulator of beta-amyloid formation. Proc Natl Acad Sci U S A. 2006; 103:1941–1946. doi:10.1073/pnas.0510708103. [PubMed: 16449386]
18. Oliveira TG, et al. Phospholipase d2 ablation ameliorates Alzheimer's disease-linked synaptic dysfunction and cognitive deficits. J Neurosci. 2010; 30:16419–16428. doi:10.1523/JNEUROSCI.3317-10.2010. [PubMed: 21147981]
19. Osisami M, Ali W, Frohman MA. A role for phospholipase D3 in myotube formation. PLoS One. 2012; 7:e33341. doi:10.1371/journal.pone.0033341. [PubMed: 22428023]
20. Hardy J, Allsop D. Amyloid deposition as the central event in the aetiology of Alzheimer's disease. Trends Pharmacol Sci. 1991; 12:383–388. [PubMed: 1763432]
21. Benitez BA, et al. Exome-sequencing confirms DNAJC5 mutations as cause of adult neuronal ceroid-lipofuscinosis. PLoS One. 2011; 6:e26741. doi:10.1371/journal.pone.0026741. [PubMed: 22073189]
22. Cruchaga C, et al. Association of TMEM106B gene polymorphism with age at onset in granulin mutation carriers and plasma granulin protein levels. Arch Neurol. 2011; 68:581–586. doi:10.1001/archneurol.2010.350. [PubMed: 21220649]
23. Cruchaga C, et al. Association and Expression Analyses With Single-Nucleotide Polymorphisms in TOMM40 in Alzheimer Disease. Arch Neurol-Chicago. 2011; 68:1013–1019. doi:10.1001/archneurol.2011.155. [PubMed: 21825236]
24. Jin SC, et al. Pooled-DNA sequencing identifies novel causative variants in PSEN1, GRN and MAPT in a clinical early-onset and familial Alzheimer's disease Ibero-American cohort. Alzheimers Res Ther. 2012; 4:34. doi:10.1186/alzrt137. [PubMed: 22906081]
25. Benitez BA, et al. The PSEN1, p.E318G Variant Increases the Risk of Alzheimer's Disease in APOE-epsilon4 Carriers. PLoS Genet. 2013; 9:e1003685. doi:10.1371/journal.pgen.1003685. [PubMed: 23990795]
26. Wu MC, et al. Rare-variant association testing for sequencing data with the sequence kernel association test. Am J Hum Genet. 2011; 89:82–93. doi:10.1016/j.ajhg.2011.05.029. [PubMed: 21737059]
27. Liang WS, et al. Alzheimer's disease is associated with reduced expression of energy metabolism genes in posterior cingulate neurons. Proc Natl Acad Sci U S A. 2008; 105:4441–4446. [PubMed: 18332434]
28. Morris JC. The Clinical Dementia Rating (CDR): current version and scoring rules. Neurology. 1993; 43:2412–2414. [PubMed: 8232972]
29. Fagan AM, et al. Inverse relation between in vivo amyloid imaging load and cerebrospinal fluid Abeta42 in humans. Ann Neurol. 2006; 59:512–519. [PubMed: 16372280]
30. McKhann G, et al. Clinical diagnosis of Alzheimer's disease: report of the NINCDSADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. Neurology. 1984; 34:939–944. [PubMed: 6610841]
31. Li H, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009; 25:2078–2079. doi:btp352 [pii] 10.1093/bioinformatics/btp352. [PubMed: 19505943]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

32. Cruchaga C, et al. SNPs associated with cerebrospinal fluid phospho-tau levels influence rate of decline in Alzheimer's disease. PLoS Genet. 2010; 6 pii: e1001101, doi:10.1371/journal.pgen.1001101.
33. Vallania FL, et al. High-throughput discovery of rare insertions and deletions in large cohorts. Genome Res. 2010; 20:1711–1718. doi:gr.109157.110 [pii] 10.1101/gr.109157.110. [PubMed: 21041413]
34. Myers AJ, et al. A survey of genetic human cortical gene expression. Nat Genet. 2007; 39:1494–1499. [PubMed: 17982457]
35. Forabosco P, Ramasamy A, Hardy J, Ryten M. Insights into TREM2biology using whole genome network analysis of gene expression data from post-mortem human brain tissue. 2013
36. Millar T, et al. Tissue and organ donation for research in forensic pathology: the MRC Sudden Death Brain and Tissue Bank. J Pathol. 2007; 213:369–375. doi:10.1002/path.2247. [PubMed: 17990279]
37. Trabzuni D, et al. Quality control parameters on a large dataset of regionally dissected human control brains for whole genome expression studies. J Neurochem. 2011; 119:275–282. doi:10.1111/j.1471-4159.2011.07432.x. [PubMed: 21848658]
38. Huang da W, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res. 2009; 37:1–13. doi:10.1093/nar/gkn923. [PubMed: 19033363]
39. Mason MJ, Fan G, Plath K, Zhou Q, Horvath S. Signed weighted gene co-expression network analysis of transcriptional regulation in murine embryonic stem cells. BMC genomics. 2009; 10:327. doi:10.1186/1471-2164-10-327. [PubMed: 19619308]
40. Zhang B, Horvath S. A general framework for weighted gene co-expression network analysis. Statistical applications in genetics and molecular biology. 2005; 4 Article17, doi:10.2202/1544-6115.1128.
41. Montine TJ, et al. National Institute on Aging-Alzheimer's Association guidelines for the neuropathologic assessment of Alzheimer's disease: a practical approach. Acta Neuropathol. 2012; 123:1–11. doi:10.1007/s00401-011-0910-3. [PubMed: 22101365]
42. McKeith IG, et al. Diagnosis and management of dementia with Lewy bodies: third report of the DLB Consortium. Neurology. 2005; 65:1863–1872. doi:10.1212/01.wnl.0000187889.17253.b1. [PubMed: 16237129]
43. Schroeter EH, et al. A presenilin dimer at the core of the gamma-secretase enzyme: insights from parallel analysis of Notch 1 and APP proteolysis. Proc Natl Acad Sci U S A. 2003; 100:13075–13080. [PubMed: 14566063]
44. Hammond SM, et al. Characterization of two alternately spliced forms of phospholipase D1. Activation of the purified enzymes by phosphatidylinositol 4,5-bisphosphate, ADP-ribosylation factor, and Rho family monomeric GTP-binding proteins and protein kinase C-alpha. J Biol Chem. 1997; 272:3860–3868. [PubMed: 9013646]
45. Sung TC, et al. Mutagenesis of phospholipase D defines a superfamily including a trans-Golgi viral protein required for poxvirus pathogenicity. Embo J. 1997; 16:4519–4530. doi:10.1093/emboj/16.15.4519. [PubMed: 9303296]
46. Colley WC, et al. Phospholipase D2, a distinct phospholipase D isoform with novel regulatory properties that provokes cytoskeletal reorganization. Curr Biol. 1997; 7:191–201. [PubMed: 9395408]
47. Kauwe JSK, et al. Neurogenetics. 2009; 10:13–17. [PubMed: 18813964]
48. Cleveland DW, Fischer SG, Kirschner MW, Laemmli UK. Peptide mapping by limited proteolysis in sodium dodecyl sulfate and analysis by gel electrophoresis. J Biol Chem. 1977; 252:1102–1106. [PubMed: 320200]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

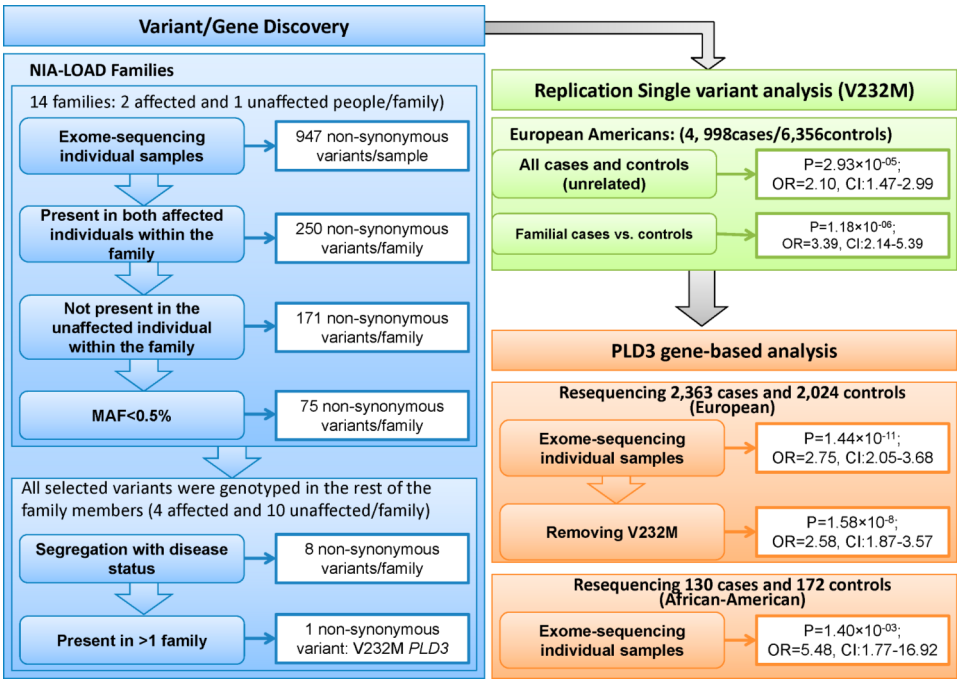


Figure 1. Summary of the main genetic findings

The diagram represents the different steps to filter the variants identified by exome-sequencing, which lead to the identification of the *PLD3*-V232M variant. The diagram also shows the subsequent genetic analyses in large case-control datasets that validated the association of the V232M variant and *PLD3* with risk for AD.

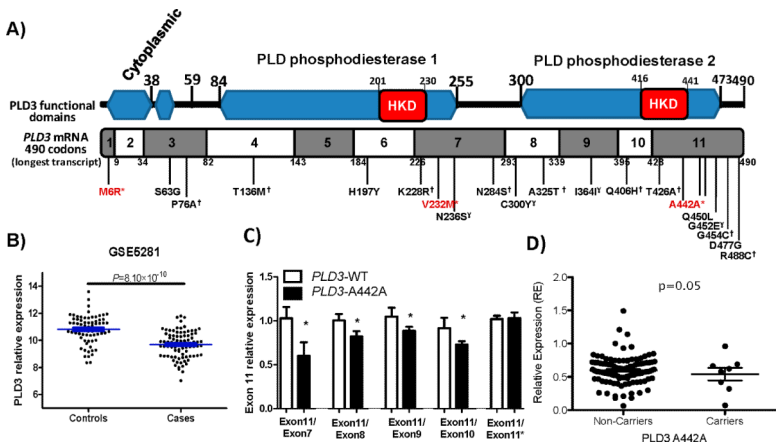


Figure 2.

A) Schematic representation of PLD3 and the relative position of the *PLD3* variants. PLD3 has two PLD phosphodiesterase domains, which contain an HKD signature motif (H-x-K-x(4)-D-x(6)-G-T-x-N, where x represents any amino acid residue). The scheme also shows the exon composition of the longest PLD3 mRNA and the position of the variants found in this study. Variants highlighted in red and noted with an "*" are significantly associated with AD risk. Variants noted with a "†" were found only in AD cases. Variants noted with a "ɣ" are more frequent in AD cases compared to controls. B) PLD3 neuronal gene expression is significantly lower in AD cases compared to controls. We used the GEO dataset GSE5281[27], in which neurons were laser-captured to analyze whether *PLD3 mRNA* expression levels are different between AD cases and cognitively normal elderly individuals.. C-D) The *PLD3 A442A* variant is associated with lower total *PLD3 mRNA* expression and lower levels of exon11 containing transcripts. C) Primers specific to exons 7, to 11 (two pairs of primers) were designed with PrimerExpress. cDNA from eight *PLD3 A442A* carriers and ten age, gender, *APOE*, CDR and PMI-matched individuals were obtained from parietal lobe. Relative expression of exon 11 compared to the other exons was calculated by the ΔCt method. Exon 11 containing transcripts in relation to exon 7-10 containing transcripts were 20% lower in A442A carriers ($P<0.05$). Graphs represent the mean±SEM. **D). Real-time PCR** was used to quantify total *PLD3 mRNA and* standardized using *GADPH mRNA* as a reference. P-value is for the gene-expression levels of major allele carriers vs. minor allele carriers after correcting for dementia severity.

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

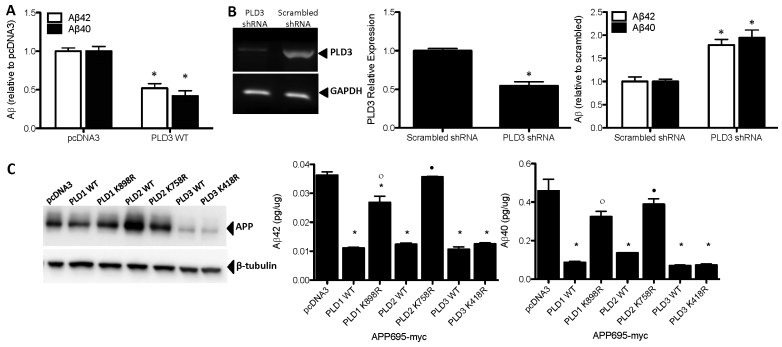
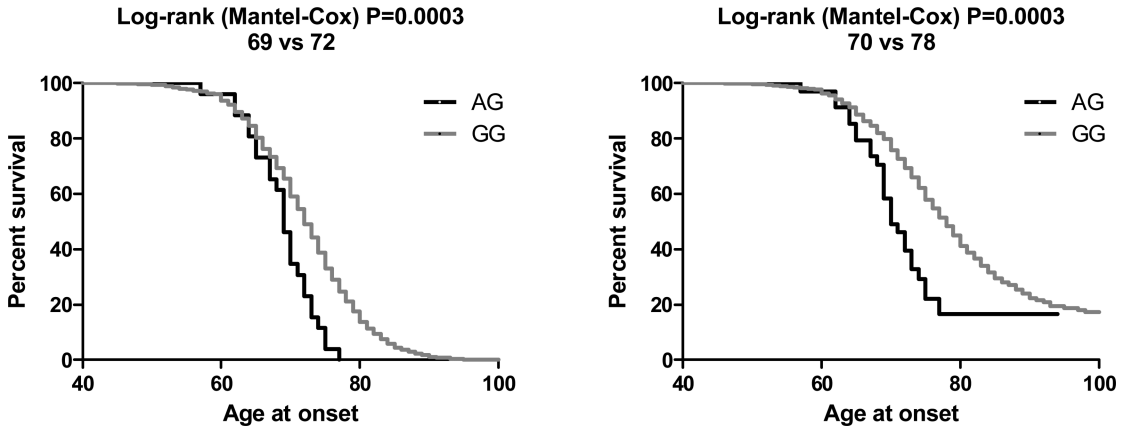


Figure 3. PLD3 affects APP processing

A-B. Overexpression and knockdown of PLD3 produce opposing effects on extracellular Aβ levels. N2A cells stably expressing hAPP695-WT were transiently transfected with vectors containing no insert (pcDNA3), human PLD3-WT, scrambled shRNA (Origene), or mouse PLD3 shRNA (Origene) for 48 hours. Cell media were analyzed with Aβ40 and Aβ42 ELISAs and corrected for total intracellular protein. Aβ levels were then expressed relative to pcDNA3. Graphs represent the mean±SEM. **A)** Overexpression of human PLD3 produces significantly less extracellular Aβ42 and Aβ40. "*", $p<0.0001$. **B).** Knockdown of endogenous PLD3 cells produces significantly more extracellular Aβ42 and Aβ40. "*", $p<0.002$. **C)** Members of the PLD protein family have different effects on APP processing. HEK293T cells were transiently transfected with vectors containing hAPP-WT and an empty vector (pcDNA3), PLD1, PLD2, or PLD3-WT or PLD1, PLD2, PLD3 carrying a dominant negative mutation.. Left panel, PLD3 affects full-length APP levels. Cell lysates were extracted in non-ionic detergent, analyzed by SDS-PAGE and immunoblotting with antibodies to the myc-tag on APP (9E10) or β-tubulin. Middle (Aβ42) and right (Aβ40) panel, cell media were analyzed with Aβ40 and Aβ42 ELISAs and corrected for total intracellular protein. Graphs represent the mean±SEM. "*", $p<0.01$, different from pcDNA3; "o", $p=0.002$, different from PLD1-WT; "•", $p<0.0001$, different from PLD2-WT. Images are representative of at least three replicate experiments.



Extended Data Figure 1. PLD3 V232M is associated with age at onset for AD

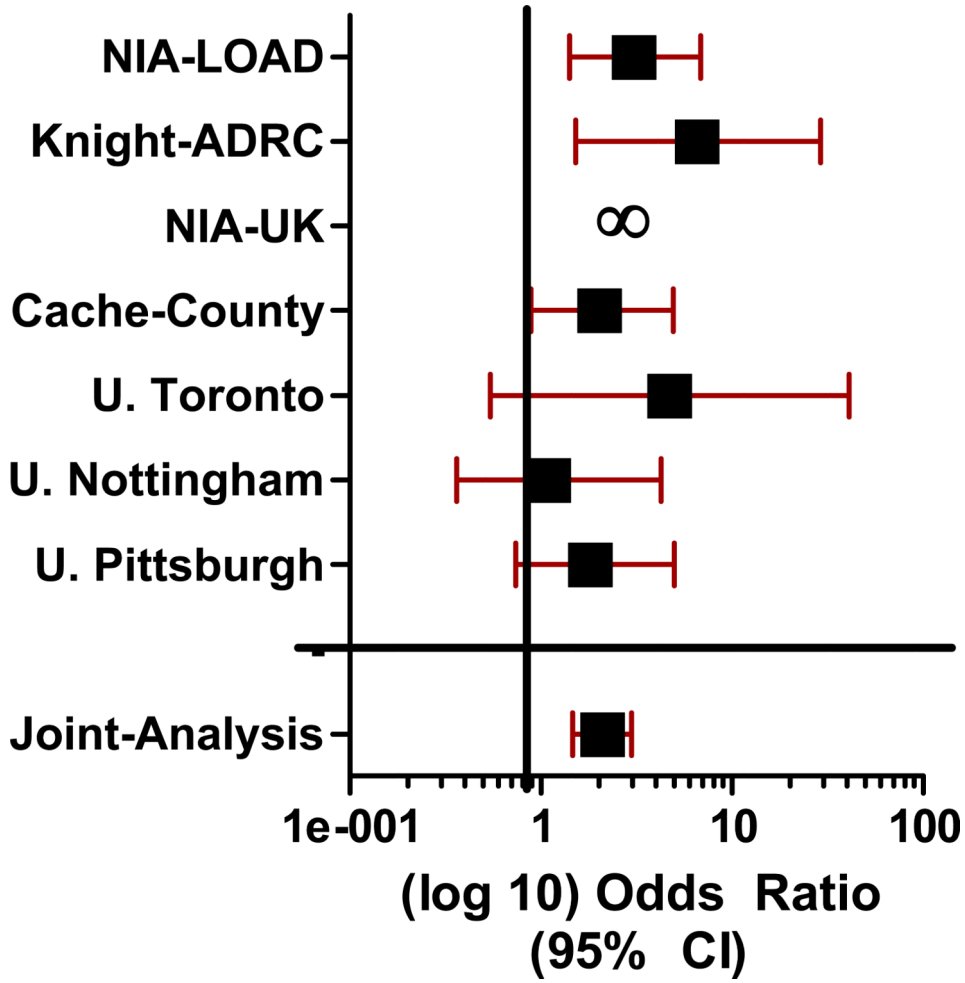
. Age at onset was analyzed for association with the *PLD3* V232M variant in 2,220 cases and 1,841 controls from the Knight-ADRC and NIA-LOAD, by the Kaplan-Meier method and tested for significant differences using the Log-rank test. **A)** Case only analysis. The carriers of the minor allele (AG) have an AAO 3 years lower than the non-carriers (69 vs 73; $p=3\times10^{-3}$). **B)** Controls were included as censored data. The carriers of the minor allele (AG) have an AAO 8 years lower than the non-carriers (70 vs 78; $p=3\times10^{-3}$).

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

NIH-PA Author Manuscript

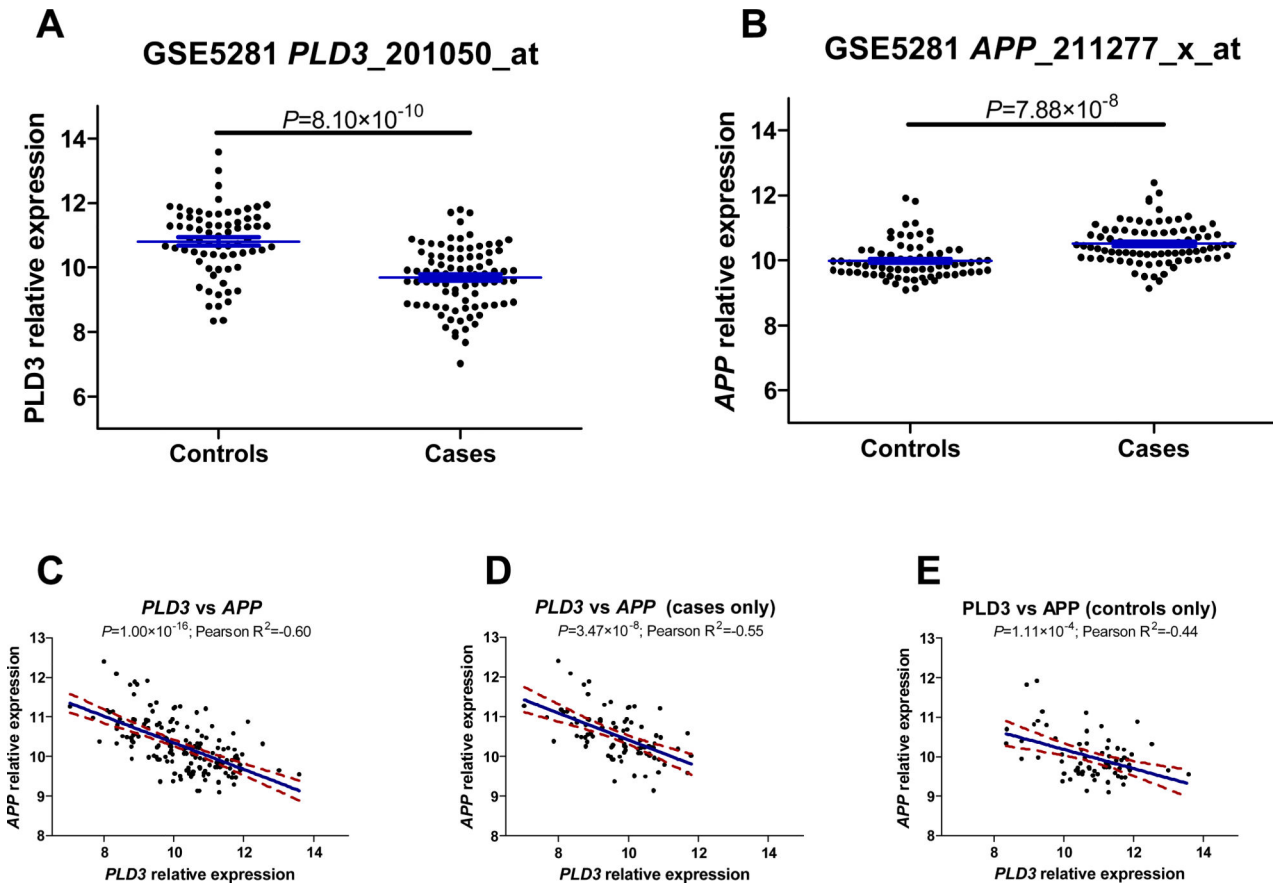
NIH-PA Author Manuscript

NIH-PA Author Manuscript



Extended Data Figure 2.
Forest plot for each case-control series for the V232M variant.

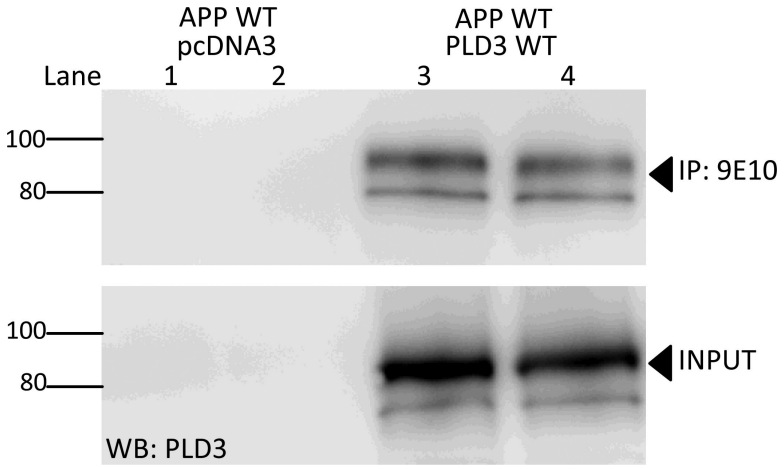
NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript



Extended Data Figure 3. PLD3 and APP mRNA expression are inversely correlated

PLD3 (probe 201050_at) and APP (probe 211277_x_at) expression levels were extracted from the GSE5281 dataset. PLD3 mRNA levels are significantly lower in AD cases compared to controls ($p=8.10\times10^{-10}$), but APP is higher in AD cases ($p=7.88\times10^{-8}$). *PLD3* mRNA levels are inversely correlated with *APP* mRNA expression levels ($p=1.00\times10^{-16}$). The correlation is stronger in AD cases (Person correlation coefficient = -0.55), than in controls (Person correlation coefficient = -0.44), but in both scenarios the correlation is highly significant.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript



Extended Data Figure 4. PLD3 interacts with APP

HEK293T cells were transiently transfected with vectors containing APP WT and an empty vector (pcDNA3) or PLD3 WT for 48 hours. Cell lysates were extracted in non-ionic detergent, pre-cleared with Protein A beads, and immunoprecipitated with an antibody to the myc-tag on APP (9E10). Immunoblots were probed with an antibody specific to human PLD3. PLD1 and PLD2 reportedly do not inmunoprecipitate with APP[17,16].

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Extended Data Table 1

Association of the PLD3-V232M variant in seven independent case-control datasets

Dataset	Cases	Carrier Freq %	Control	Carrier Freq %	OR (95%CI)	p-value
NIA-LOAD	29/1,077	2.62	8/920	0.86	3.09 (1.41-6.81)	4.00×10^{-03}
Knight-ADRC	16/1,098	1.44	2/911	0.22	6.63 (1.52-28.9)	3.40×10^{-03}
NIA-UK	1/142	0.70	0/183	0.00	∞ (NA)	0.438
Cache-County	6/249	2.35	29/2,442	1.17	2.03 (0.83-4.93)	0.131
U. Toronto	5/260	1.89	1/245	0.41	4.71 (0.54-40.7)	0.212
U. Nottingham	6/519	1.14	3/271	1.09	1.05 (0.26-4.25)	1.000
U. Pittsburgh	15/1,253	1.18	6/958	0.62	1.82 (0.74-5.00)*	0.191
NIMH	4/318	1.24	-	-	N/A	
Wellderly	-	-	1/376	0.27	N/A	
Total	**82/4,916**	**1.64**	**50/6,306**	**0.79**	**2.10 (1.47-2.99)**	**2.93×10^{-05}**

The table shows the counts for Carriers and non-carriers. P-values were calculated by Fisher's exact-test.

*For the U. Pittsburgh, age, gender, APOE genotype and principal component factors for population stratification were available. Association of the V232M with AD risk was performed by logistic regression including age, sex, APOE genotype and the first four principal component factors as covariates.

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

Extended Data Table 2

Sequence variants found in *PLD3* in the NIA-LOAD, Knight-ADRC and NIA-UK datasets.

Chr. position	AA		NIA LOAD	Knight ADRC	NIA-UK	total	MAF %	p-value	OR (95% CI)	EVS MAF%	SIFT	Polyphen
40872407	**M6R**	**CA**	0	8	1	9	0.19	**0.02**	7.73 (1.09-61)	NP	tolerated	deleterious
		CO	0	1	0	1	0.02					
40872764	**S63G**	**CA**	3	1	0	4	0.08	0.74	0.68 (0.18-2.55)	0.16	tolerated	neutral
		CO	5	0	0	5	0.12					
40872803	**P76A**	**CA**	3	1	0	4	0.08	0.12	NA	0.03	tolerated	benign
		CO	0	0	0	0	0.00					
40873764	**T136M**	**CA**	0	1	0	1	0.02	0.54	NA	NP	tolerated	deleterious
		CO	0	0	0	0	0.00					
40876055	**H197Y**	**CA**	0	1	0	1	0.02	0.49	0.85 (0.05-13-7)	NP	damaging	benign
		CO	0	1	0	1	0.02					
40877584	**K228R**	**CA**	1	1	1	3	0.06	0.25	NA	NP	damaging	deleterious
		CO	0	0	0	0	0.00					
40877595	**V232M**	**CA**	29	16	1	46	0.99	**1.05×10^{-05}**	3.99 (2.01-7.94)	0.48	damaging	deleterious
		CO	8	2	0	10	0.25					
40877608	**N236S**	**CA**	0	2	0	2	0.04	0.40	1.71 (0.15-18.91)	0.01	damaging	deleterious
		CO	0	1	0	1	0.02					
40877752	**N284S**	**CA**	0	1	0	1	0.02	0.54	NA	NP	tolerated	deleterious
		CO	0	0	0	0	0.00					
40880407	**C300Y**	**CA**	2	3	0	5	0.10	0.46	2.14 (0.41-11.06)	0.09	tolerated	deleterious
		CO	1	0	1	2	0.04					
40880481	**A325T**	**CA**	0	1	0	1	0.02	0.54	NA	NP	damaging	deleterious
		CO	0	0	0	0	0.00					
40883725	**Q406H**	**CA**	1	0	0	1	0.02	0.54	NA	NP	tolerated	neutral
		CO	0	0	0	0	0.00					
40883783	**T426A**	**CA**	1	0	0	1	0.02	0.54	NA	NP	tolerated	neutral
		CO	0	0	0	0	0.00					
40883911	**G435V**	**CA**	0	0	0	0	0.00	0.46	NA	0.02	damaging	deleterious
		CO	1	0	0	1	0.02					

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

Chr. position	AA		NIA LOAD	Knight ADRC	NIA-UK	total	MAF %	p-value	OR (95% CI)	EVS MAF%	SIFT	Polyphen
40883933	**A442A**	**CA**	48	35	12	95	2.09	**1.08×10^{-05}**	2.31 (1.56-3.41)	1.59	-	-
		CO	17	12	7	36	0.90					
40883956	**Q450L**	**CA**	0	0	0	0	0.00	0.46	NA	NP	tolerated	neutral
		CO	0	0	1	1	0.02					
40883962	**G452E**	**CA**	4	6	0	10	0.21	0.16	2.86 (0.78-10.4)	0.09	tolerated	deleterious
		CO	0	2	1	3	0.07					
40883967	**G454C**	**CA**	0	1	0	1	0.02	0.54	NA	NP	damaging	deleterious
		CO	0	0	0	0	0.00					
40884037	**D477G**	**CA**	0	1	0	1	0.02	0.49	0.42 (0.04- 4.72)	0.02	damaging	deleterious
		CO	0	1	0	1	0.02					
40884069	**R488C**	**CA**	0	3	0	3	0.06	0.25	NA	0.02	damaging	deleterious
		CO	0	0	0	0	0.00					
total		**CA**	1106	1114	143	2363						
total		**CO**	928	913	183	2024						

The coding region of *PLD3* was sequenced in 2,363 AD cases and 2,024 controls (see materials and methods) from the Knight-ADRC, NIA-LOAD and the NIA-UK datasets. The table shows the coding variants identified as well as the number of carriers in each dataset. The minor allele frequency (MAF) in cases and in controls, the p-value and the OR for the association with case-control status is shown. The MAF of the identified variants in the Exome Variant Server (EVS) is shown. We also used SIFT and Polyphen to predict the impact of the non-synonymous changes on protein function. NA: not available. NP: not present

Nature. Author manuscript; available in PMC 2014 July 23.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Extended Data Table 3

Comparison of the gene-based analysis including all coding variants or only variants predicted to be deleterious

	Benign + deleterious		Only deleterious	
	p-value	**OR (CI)**	**p-value**	**OR (CI)**
All variants	1.44×10^{-11}	2.75 (2.05-3.68)	2.52×10^{-12}	2.86 (2.10-3.88)
Excluding V232M	1.58×10^{-8}	2.58 (1.87-3.57)	2.95×10^{-8}	2.54 (1.81-3.57)
Excluding A442 and V232M	1.61×10^{-3}	2.86 (1.62-5.06)	5.88×10^{-5}	3.20 (1.59-6.45)

Gene-based analyses were performed by SKAT-O. Variants that were predicted to be benign by both SIFT and Polyphen were removed for the second analysis

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Extended Data Table 4

Association analysis for *PLD3* A442A in four European-descent datasets

	CA	CO	p-value	OR (95% CI)
NIA-LOAD	48/1058	17/911	1.40×10^{-03}	2.43 (1.38-4.25)
Knight-ADRC	35/1079	12/901	7.10×10^{-03}	2.43 (1.25-4.71)
NIA-UK	12/131	7/176	9.76×10^{-02}	2.30 (0.88- 6.0)
Cache-County	9/246	50/2421	1.15×10^{-01}	1.77 (0.86-3.65)
Total	104/2514	86/4409	3.78×10^{-07}	2.12 (1.58-2.83)

The table shows the counts for carriers and non-carriers. P-values were calculated using the Fisher's Exact test.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Extended Data Table 5

PLD3 is associated with risk for AD in African-Americans

	Cases (n=130)		Controls (n=172)			
Variant	**carriers**	**Carrier Freq %**	**carriers**	**Carrier Freq %**	**p-value**	**OR (95% CI)**
G63S	1	0.77%	0	0.00%	0.43	NA
K228R	1	0.77%	0	0.00%	0.43	NA
V232M	3	2.31%	0	0.00%	**0.07**	NA
I364I	6	4.62%	4	2.33%	0.33	2.02 (0.56-7.29)
A442A	4	3.08%	0	0.00%	**0.03**	NA
Total	15	11.54%	4	2.33%	**1.4×10^{-03}**	5.48 (1.77-16.92)

PLD3 was sequenced in a total of 302 African-Americans. Table shows the counts for single SNPs and the gene-based analysis for *PLD3* in 130 AA cases and 172 controls. P-values were calculated using the Fisher's exact test.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

Table 1

Association between *PLD3*-V232M (rs145999145) and Alzheimer's Disease risk in individuals of European-descent.

	count	freq (%)	Odds Ratio (95% CI)	p value
Control Group				
All controls	50/6,306	0.79		
Non-demented >65yrs	9/1,690	0.52		
Non-demented >70yrs	5/1,248	0.39		
Non-demented >80yrs	1/375	0.26		
Cases group				
All AD cases	**82/4,916**	**1.64**	**A 2.10 (1.47-2.99)**	**2.93×10^{-5}**
			B 3.13 (1.57-6.24)	3.54×10^{-4}
			C 4.16 (1.68-10.29)	2.34×10^{-4}
Index cases (Families)	29/1,077	2.62	A 3.39 (2.14-5.39)	1.18×10^{-6}
			B 5.05 (2.38-10.41)	5.14×10^{-6}
			C 6.72 (2.59-17.52)	5.23×10^{-6}
Sporadic AD cases	53/3,839	1.36	A 1.74 (1.18-2.57)	5.70×10^{-3}
			B 2.59 (1.27-5.26)	5.20×10^{-3}
			C 3.44 (1.37-8.63)	3.20×10^{-3}

The table shows the counts for minor allele carriers and non-carriers. P-values were calculated using Fisher's exact test. Only individuals of European-descent were included in this analysis.

The carrier frequency for the V232M variant in the Exome Variant Server (EVS) is 0.99%

A OR and p-value in comparison with all controls

B OR and p-value in comparison with non-demented individuals >65yrs.

C OR and p-value in comparison with non-demented individuals >70yrs.