Published in final edited form as:
J Alzheimers Dis. 2011 ; 24(3): 475–484. doi:10.3233/JAD-2011-101848.

Genetic Association of sequence variants near *AGER/NOTCH4* and Dementia

Anna M. Bennet[1], **Chandra A. Reynolds**[2], **Ulrika K. Eriksson**[1], **Mun-gwan Hong**[1], **Kaj Blennow**[3], **Margaret Gatz**[4,1], **Andrey Alexeyenko**[1], **Nancy L. Pedersen**[1,4], and **Jonathan A. Prince**[1,*]

[1]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, 171 77 Stockholm, Sweden

[2]Department of Psychology, University of California at Riverside, Riverside, CA 92521, USA

[3]Institute of Neuroscience and Physiology, Sahlgrenska Academy at University of Gothenburg, 431 80 Mölndal, Sweden

[4]Department of Psychology, University of Southern California, Los Angeles, CA 90089-1061, USA

Abstract

We performed a survey of sequence variation in a series of 20 genes involved in inflammation-related pathways for association with dementia risk in twin and unrelated case-control samples consisting in total of 1462 Swedish dementia cases and 1929 controls. For a total of 218 tested genetic markers, strong evidence was obtained implicating a region near *AGER* and *NOTCH4* on chromosome 6p with replication across both samples and maximum combined significance at marker rs1800625 (OR = 1.37, 95% CI 1.19–1.56, $p = 1.36 \times 10^{-6}$). Imputation of the associated genomic interval provided an improved signal at rs8365, near the 3'UTR of *AGER* ($p = 7.34 \times 10^{-7}$). The associated region extends 120kb encompassing 11 candidate genes. While *AGER* encodes a key receptor for β-amyloid protein, an analysis of network context based upon genes now confirmed to contribute to dementia risk (*APP*, *PSEN1*, *PSEN2*, *CR1*, *CLU*, *PICALM*, and *APOE*) suggested strong functional coupling to *NOTCH4*, with no significant coupling to the remaining candidates. The implicated region occurs in the broad HLA locus on chromosome 6p, but associated markers were not in strong LD with known variants that regulate HLA gene function, suggesting that this may represent a signal distinct from immune-system pathways.

Keywords

Dementia; NOTCH4; gene; association; inflammation

Introduction

Dementia entails the progressive loss of cognitive function over the course of many years and is one of the most common disorders of aging with the majority of cases being due to Alzheimer's disease (AD: OMIM 104300). The post-mortem characteristics of AD necessary for a definitive diagnosis include the observation of high levels of β-amyloid (Aβ) deposits

*Correspondence: Dr. Jonathan A. Prince Department of Medical Epidemiology and Biostatistics Karolinska Institutet Nobels väg 12A 171 77 Stockholm, Sweden Phone: +46 (0)8 524 86008 Fax: +46 (0)8 31 49 75 Jonathan.Prince@ki.se.

Authors declare no conflicts of interest.

in specific brain regions, intra-neuronal neuro-fibrillary tangles, neuronal loss, and synaptic damage. The core finding of Aβ deposition is generally considered causative, with the aggregation of the specific $A\beta_{42}$ cleavage product of Aβ aggregating readily into fibrils and contributing to neurotoxicity [1]. Genetic findings also strongly support this basic mechanism, with each of the recognized rare causative genes (*APP*, *PSEN1*, and *PSEN2*) being shown to play a role in Aβ metabolism [2–4]. In addition, the most widely regarded common genetic risk factor for AD (*APOE*) [5] also associates strongly with Aβ levels in the CNS [6–7].

The mechanism by which β-amyloid contributes to neurodegeneration and dementia remains largely unresolved. A central problem is that the majority of normal individuals that die from causes other than dementing illness also have a considerable accumulation of brain Aβ [8]. This indicates that there is a mechanism by which susceptibility to the toxic effects of Aβ varies across individuals. One possible route by which neurons could experience variable vulnerability may be via inflammation, in particular in relation to microglia activation, the complement system, and inflammatory cytokines [9]. In this regard, a link between inflammation and amyloid metabolism has been demonstrated whereby the activation of microglia and subsequent release of inflammatory mediators can be exacerbated by aggregated Aβ [10]. The implication of inflammatory mechanisms to risk for dementia is not new [11–12], but systematic evidence is sparse, especially in comparison to the literature on Aβ. In particular, genetic studies specifically targeting inflammatory genes have been relatively rare, and those that have been conducted have been done so with small sample sizes [13]. Against this background we undertook a pathway approach to comprehensively evaluate common sequence variation in some of the most widely studied genes involved in inflammation in relation to dementia risk.

Materials and Methods

Human Samples

The present study drew participants from four aging twin studies stemming from the population-based Swedish Twin Registry [14], and an independent non-twin case-control Swedish AD sample [15]. This study was approved by the local ethical review board of Stockholm. Across the samples, genotyping was successful in 3429 samples including 1462 dementia cases (1204 with possible or probable AD diagnoses) and 1929 controls. Among the 258 non-AD dementia cases, there were 127 with a vascular dementia diagnosis, 42 with a mixed dementia diagnosis, and the remaining having a non-specific diagnosis. This sample used for analysis derives from a larger sample of 3775 individuals in which one member of MZ twin pairs was removed at random. There were 874 men and 1055 women in the control group, and 551 men and 911 women in the dementia group. Average age-at-sampling across cases and controls was 77.7 ± 8.7 (SD) years and age-at-onset for dementia/AD cases was 75.3 ± 8.3 (SD) years (age-at-onset was available for 1371 dementia cases). Both the twin and case-control samples were described in detail recently [15]. In brief, the twin samples are from the Swedish Adoption/Twin Study of Aging (SATSA) [16], the Origins of Variance in the Oldest-Old (OCTO-Twin) [17], Sex Differences in Health and Aging Study (GENDER) [18], and the Study of Dementia in Swedish Twins (HARMONY) [19]. The case-control sample was comprised of unrelated individuals recruited from three prospective longitudinal studies of patients with dementia from Mölndal, Piteå, and Malmö, Sweden. This total sample affords 80% power for the detection of a genetic effect size of approximately 1.2 (odds ratio) at an alpha of 0.05 and given a risk allele frequency of 0.25. We conducted secondary analysis of publicly available data of peripheral gene expression levels in 193 brain samples from adults who lived to 65 years or older [20].

Gene and Marker Selection

Our selection of candidate genes involved screening of PubMed literature from 2003–2009. Titles of papers were screened to establish relevance and abstracts read of papers highlighting gene names and/or inflammation pathway approaches. Key papers purporting evidence of association of candidate genes with dementia or AD were read in their entirety. The list provided in table 1 reflects some of the most commonly studied genes in relation to inflammation and disease. We acknowledge that this list is far from exhaustive, especially given the emerging detail depicted in pathway and network based descriptions (e.g. KEGG) [21].

Genetic markers for all candidate genes (including 20kbp upstream of the transcription start site and 10kbp downstream of the transcription end site) were selected with previous findings, functional candidature, LD, and Illumina SNP design score as criteria (the full final list is shown in supplementary table 1). Illumina scores were calculated by an algorithm developed by the company that predicts success of the assay for the marker. Based on the genotype data for CEU samples of HapMap Release 22, all polymorphic markers in the dataset were taken into consideration. At first, LD blocks were searched with Haploview 4.0 [22]. Prioritization included markers in exons, within 80bp of exon boundaries, and within 1kb upstream from the 1st exon or downstream from the last exon of any predicted gene in UCSC genome browser, and SNPs that can tag the LD blocks. Among the markers outside of LD blocks, those which could be prioritized by the same scheme were included. After selection, Illumina scores for all markers were calculated and those that did not satisfy the criteria for Illumina probe chemistry were replaced with a SNP in perfect LD if available ($r^2 = 1$) or other tagging SNP. Coverage in terms of total number of markers for each candidate gene is shown in table 1.

Genotyping

Genotyping was performed using the Illumina GoldenGate assay system on Illumina BeadStation 500GX equipment, currently housed and implemented at the Uppsala University SNP Technology Platform (http://www.medsci.uu.se/molmed/snpgenotyping/methods.htm). Prior to use on the Illumina system, all samples were subjected to Whole Genome Amplification (WGA) using standard kits involving Phi29 DNA polymerase (Amersham).

CSF Biomarkers

CSF samples were obtained in the AD case-control study by lumbar puncture in the L3/L4 or L4/L5 inter-space. Further details of CSF collection can be found elsewhere [23]. CSF $A\beta_{42}$ was determined using a sandwich enzyme-linked immunosorbent assay (ELISA) (Innotest b-amyloid (1–42), Innogenetics, Ghent, Belgium) constructed to specifically measure $A\beta_{42}$ [24]. The microtubule-associated protein tau, a CSF marker of neuronal degeneration, was determined using a sandwich ELISA (Innotest hTAU-Ag, Innogenetics, Ghent, Belgium) constructed to measure total tau, i.e., all isoforms of tau irrespective of phosphorylation state [25].

Statistics

Hardy-Weinberg equilibrium for individual loci was assessed using the Pearson χ^2 statistic. To account for intra-class correlation between twin-pairs, the statistical significance of genetic association for each marker with dementia risk was estimated using multilevel mixed-effects logistic regression as implemented as the xtlogit function in STATA v11.0. We also present analyses (supplementary table 1) using alternating logistic regression that likewise account for pair dependency [26–27]. Models were fit by full maximum likelihood.

The sample used in the present study is derived from a larger number of individuals where for MZ twin-pairs, one co-twin was selected at random and removed. Tests of genotypes versus CSF quantitative traits in a maximum of 838 non-twin samples (CSF total tau, phosphorylated tau (p-tau) and CSF $A\beta_{42}$) were conducted using ANOVA in STATA v11.0. For total dementia onset age (AAO), mixed linear regression accounting for twin correlations was performed using xtmixed implemented in STATA.

Network Analyses

We evaluated functional coupling of candidate genes from the genetic association analysis in a gene network predicted using the FunCoup resource [28]. Smaller sub-networks could be investigated via a web-based network browser. To estimate significance of functional coupling between a particular gene of interest and a pre-defined gene group we used custom software that implemented a previously proposed randomization algorithm [29]. The procedure entails the re-wiring of randomized networks in such a way that the number of links for each node in the network is preserved, while network neighbors are shuffled. The real, i.e. FunCoup-predicted, network was randomized 100 times. After every randomization, connections between a gene of interest i and all genes of group j were counted. These counts were used to calculate the mean $\hat{n}_{ij}$ and standard deviation σ_{ij}. Together with the respective number of links in the real network n_{ij}, these values produced one-sided Z-scores that conveyed significance:

$$Z=\frac{n_{ij}-\widehat{n_{ij}}}{\sigma_{ij}} \quad (1)$$

The z-scores were then converted to p-values by a standard procedure for normal distributions. Optionally, either direct or indirect (via a shared network neighbor) connections could be analyzed. For the latter, links were counted in every case where some gene was a network neighbor simultaneously to both the gene of interest i and a gene from group j.

Results

We prioritized a list of 237 genetic markers spanning 20 candidate genes (table 1) for which genotyping assays were then designed and tested in a total of 3430 Swedish DNA samples. For all markers, genotyping was performed in duplicate in a subset of 100 individuals. We also tested for significant allele-frequency from HapMap data for European populations and compared markers overlapping with our study. Genotypes were obtained for a total of 220 markers with call rates above 70% across all samples. Six of these markers had call rates between 70% and 90%, with the remainder being above 90%. For duplicate genotyping, one marker had 2 discrepancies, 8 had 1 discrepancy only, and the remaining markers had no discrepant genotype calls. Marker loss was essentially randomly distributed (i.e. there were no genomic regions with increased missingness). There were 2 cases of what we considered extreme ($p < 10^{-9}$) deviation from Hardy-Weinberg equilibrium (for rs3753827 and rs12438654). For these 2 markers, genotyping success rates were below 90%, and these genotypes were thus excluded from subsequent analyses.

The 218 remaining markers were initially tested for association with dementia in both the main and replication data sets (represented by twin and non-twin samples, respectively) using multilevel mixed-effects logistic regression to account for twin structure within the twin set and standard logistic regression in the non-twin set. The full results in the two independent sets as well as combined evidence of association in the full sample are shown in supplementary table 1. The results of the combined analysis for the top 10 single markers

are shown in table 2. The top two markers exceeded significance in the context of multiple testing and strict Bonferroni correction based upon number of tests (the threshold given this number is $p < 0.00023$). There were 23 markers that exceeded an uncorrected significance level of $p < 0.05$, which is more than the 11 expected by chance. There were 6 markers that exceeded an uncorrected $p < 0.01$, compared to 2 expected by chance. We attributed this inflation to linkage disequilibrium, which can be seen readily in supplementary table 1 where the top 23 markers for the combined analysis occur in 16 distinct loci.

The most significant finding was for rs1800625 which was significant in both independent samples (combined OR 1.37, 95% CI 1.19–1.56, $p = 1.36 \times 10^{-6}$) located near *AGER*, which was the specific candidate gene in that region that was targeted in the present study. Two additional significant markers (rs204995, and rs204993) were in relatively strong LD with rs1800625 ($r^2 = 0.83$ and 0.62, respectively). For all three markers the common alleles were associated with increased risk. Marker rs204995 was also significant in both twin and non-twin samples (see supplementary table 1). Testing rs1800625 across gender in the total sample indicated similar effect sizes in both women (OR 1.38, 95% CI 1.15–1.65, $p = 3.60 \times 10^{-4}$) and men (OR 1.34, 95% CI 1.08–1.64, $p = 3.02 \times 10^{-3}$). The common allele of marker rs1800625 was also significantly enriched when only possible and probable AD cases were compared with controls (OR 1.38, 95% CI 1.19–1.60, $p = 7.82 \times 10^{-6}$). The effect of rs1800625 was also significant when only the non-AD dementia population (258 individuals) was compared with controls (OR 1.38, 95% CI 1.02–1.89, $p = 0.038$). Given the established role of *APOE* in dementia risk, stratification by *APOE* ε4 carrier status was also performed in addition to a joint logistic regression model considering both rs429358 (the marker that defines e4 status) and rs1800625 together. In the ε4 stratified model, the association of dementia and rs1800625 was significant in both groups but the effect size estimate was much higher in ε4 non-carriers (OR 1.45, 95% CI 1.20–1.77, $p = 2.67 \times 10^{-5}$) vs. ε4 carriers (OR 1.23, 95% CI 1.02–1.53, $p = 0.036$). The effect size estimate for rs1800625 remained significant when rs429358 in *APOE* was included as a covariate (OR 1.38, 95% CI 1.19–1.60, $p = 4.43 \times 10^{-6}$).

Marker rs1800625 resides in an LD region spanning at least 120kb across multiple gene targets, and is specifically located between the 3'-UTR of *PBX2* and promoter of *AGER*. There are 11 genes that reside in the local neighborhood of *AGER*, including (from pter-qter along chromosome 6p) ATF6B, FKBPL, PRRT1, PPT2, EGFL8, AGPAT1, RNF5, AGER, PBX2, GPSM3, and *NOTCH4*. To generate a more comprehensive view of LD in our sample, we used IMPUTE [30] to probabilistically infer additional genotypes from the *AGER* region. Linkage disequilibrium in the entire chromosome 6p HLA region is problematic with many markers exhibiting long range correlations and no clear LD block structure. For this reason, we performed imputation in a large 400kb window around rs1800625, but only retained genotypes in which posterior probability estimates exceeded 0.9 with a genotype call rate of greater than 90%. The resultant overview of directly typed and imputed markers tested for association with dementia is shown in figure 1. The window was truncated on either side at points where 3 or more consecutive markers had posterior probabilities of less than 0.9. Within the window, there was one marker (rs8365) that exhibited slightly better significance than rs1800625, but they are perfect proxies ($r^2 = 1$). The posterior probability estimate for rs8365 was 0.971 and the call rate was 91.2%. Using the directly genotyped and imputed marker lists, we searched for validated markers (SNPs with population frequency data from HapMap and average heterozygosities above 0.1) where the alternative alleles code for different amino acids and thus might have a putative functional role for a particular gene. Across the region, there were only 3 such markers that could be imputed with high probability in our data set (from pter to qter; rs2070600, rs422951, and rs520803). The only validated amino-acid changing variant in *AGER* (rs2070600) was in low LD with rs1800625 ($r^2 = 0.007$). The two remaining markers occur

in *NOTCH4*, which is also the longest of the 11 represented genes, and neither was in strong LD with rs1800625 ($r^2 < 0.01$ for both).

In the absence of evidence of association with common protein coding sequence variants, we explored for regulatory sequence variation in the region. For this we utilized a publicly available resource with both genetic marker and expression data from a previous genome-wide study to search for association of markers with expression traits [20]. The sample consists of expression-level measurements of 14,077 transcripts in cortical brain regions for a maximum of 193 individuals. Among the genetic markers tested in the study of Myers and colleagues [20], rs1800625 was not represented, but there were several perfect proxies available. From these proxies, the marker in the closest proximity of rs1800625 (rs3130348) was chosen and tested for association with normalized target transcript levels for each of the genes in the region using PLINK [31]. All eleven genes from the *AGER* region were represented in the expression profiling experiment. From this analysis, there was marginal evidence of association with *AGER* ($p = 0.023$) and *NOTCH4* ($p = 0.019$), with the common allele of rs3130348 being associated with increased *AGER* expression and increased *NOTCH4* expression. We also tested rs3130348 against HLA transcripts within 500kb of *AGER* that were present in the expression experiment (this included *HLA-DRA*, *HLA-DRB5*, *HLA-DRB2*, and *HLA-DQA1*), but there was no evidence of association for any of these genes ($\alpha = 0.05$). We did note that expression levels for several of the genes were highly correlated, the strongest being for *GPSM3* and *NOTCH4* ($r^2 = 0.24$), indicative of common regulation.

In order to further address the question of which of the 11 genes on 6p might be a better candidate for dementia, we investigated their functional context in a gene network using the FunCoup web site for visualization and purpose-built software (see methods) for statistical testing of the same FunCoup gene network [28]. In this analysis, ten of the eleven (no links were available for *GPSM3*) candidates were tested individually for network connectivity with a base set of genes now confirmed to contribute to dementia risk (*APOE*, *APP*, *PSEN1*, *PSEN2, CLU, CR1*, and *PICALM*). *NOTCH4* connected to the base set via direct interactions with *PSEN1* and *PSEN2* (Fig. 2). An alternative test that counted connections indirectly, i.e. via network neighbors, was also used. However again, only *NOTCH4* was significantly enriched in functional connections to the base set (9 shared neighbors with 3.83 expected by chance; $p_0 \sim 0.01$). The results of the network analysis for all the candidates are shown in table 3. We also expanded the network to test for genes that had direct connectivity to the base set with at least 1 direct link to the base set and none in the randomized networks. This resulted in a total of 69 genes, and notably included *NOTCH1*, *NOTCH2*, and *NOTCH3* (see supplementary table 2).

From figure 1 there was an indication that there might be multiple independent association signals in the *AGER* region. This can be seen as markers that are significantly associated with dementia, but not in particularly strong LD with rs1800625. The strongest of these markers, rs1044506 (located near a potential splice-site in *NOTCH4*), was imputed (posterior probability 0.942) with an effect estimate of OR 0.58, 95% CI 0.45–0.76, $p = 8.55 \times 10^{-6}$. We performed a joint logistic regression analysis including both rs1800625 and rs1044506 to explore if they represent independent signals. The independent significance for each marker in the combined test was $p = 1.4 \times 10^{-4}$ for rs1800625 and $p = 1.3 \times 10^{-4}$ for rs1044506. There was no indication that these two variants were interacting, as evidenced by lack of significance of a multiplicative interaction term in the regression model.

Given the focus on *AGER* in the present study and due to its putative involvement in amyloid binding [32] we also sought evidence of genetic association with a measure of CNS beta-amyloid metabolism. For this, CSF $A\beta_{42}$ levels were evaluated in relation to the top ten

results for genetic association with dementia risk and the findings are presented in table 2. We also included two measures of tau protein (total tau and phosphorylated tau), which also accumulates abnormally in dementia, but neither were associated with any of the top 10 markers (not shown). Additionally, there was no convincing evidence of association of the ten variants with age at onset of dementia (table 2).

Finally, among the genes in the full set of results we also searched for evidence of regions where there were 2 or more significant markers associating with dementia but that were in low LD with each other. There was one such case in the region around *TNF* with markers rs1800629 and rs361525 each showing marginal evidence of association but with LD being very low between the two ($r^2 = 0.0041$). As was done for *AGER* above, we tested a joint logistic regression model to demonstrate their independence, which gave $p = 1.5 \times 10^{-3}$ for rs1800629 and $p = 2.6 \times 10^{-3}$ for rs361525. There was no evidence that the two variants were interacting.

Discussion

We conducted an analysis of sequence variation in commonly studied genes involved in inflammation pathways and tested these for genetic association with dementia. From twenty investigated regions, we identified a genomic interval near the *AGER* gene that associates significantly with dementia in two independent samples. We implemented several additional strategies to augment this study in order to single-out a specific gene candidate and these are discussed below.

The issue of isolating a specific gene in a region of association still constitutes one of the greatest challenges in genetic association studies and the present study is no exception. The interval around *AGER* includes 10 additional genes, but the extent of linkage disequilibrium in the region limits the implication of one particular candidate over the others. We attempted to single-out one of the genes in the interval with a combination of gene expression analysis, together with a gene network analysis that we have developed and previously employed [33]. For expression analyses, there was weak evidence of association for the variant most strongly associated with dementia when tested against expression levels of transcripts coding for *AGER* and *NOTCH4*, with a lack of significant signals for the other genes in the region (no other markers at $p < 0.05$). It should be noted that many of the HLA genes near *AGER* are strongly regulated by genetic polymorphism but none of those variants were in LD with our associated markers.

We used a network-based modeling strategy to understand the various genes in the 6p genomic interval as they relate to the biology of known AD genes (specifically *APOE, APP, PSEN1, PSEN2, PICALM, CR1*, and *CLU*). This was based upon the growing appreciation of functional network analyses to understanding the biological role and prioritization of candidate genes [34]. The principle is that if any of the candidates in the 6p region have stronger evidence of functional coupling to the aforementioned AD genes, it might serve as an important indication that the particular gene plays a role in pathways that are common to all or most of the genes. The software selected for this analysis (FunCoup) was chosen specifically since it is driven primarily by experimental data, with the greatest weight being obtained by protein-protein interaction data and gene expression correlations. Interestingly, the seven base genes were highly connected with each other (Fig. 2). In testing the various genes, the only significant evidence of functional coupling was obtained for *NOTCH4*, which had connections to both *PSEN1* and *PSEN2* (figure 2). The literature on the various genes from the associated region favors *AGER* due to its putative role as a receptor for Aβ as well at its involvement in inflammation [35,32]. Although there was weak evidence that *AGER* transcription might be regulated by the polymorphisms that are associated with

dementia, we consider the evidence of network connectivity to provide a more convincing argument for implicating a causative gene. However, FunCoup's sensitivity to true biological links is still incomplete as it heavily depends on the amount of experimental data for a given gene (protein interaction assays, expression profiles etc.). Nonetheless, the significance of the relationship between *NOTCH4* and the base AD gene set was very strong, with few other individual genes in the genome capable of attaining that level of connectivity to this set. There exists only limited literature on *NOTCH4*, and at present nothing that relates it specifically to dementia or AD. NOTCH4, like other NOTCH proteins, functions in the implementation of differentiation, proliferation and apoptotic programs, and may be primarily involved in the development of the vascular system. However, Notch signaling in general has been studied and discussed fairly extensively in relation to dementia [36]. In this regard a common theme is that the gamma-secretase complex influences both Aβ42 production and Notch signaling, though the latter may be more linked to carcinogenesis than dementia [37]. One possible connection with dementia is that mutations in *NOTCH3*, which has strong homology with *NOTCH4*, cause cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy (CADASIL) [38]. With CADASIL, patients progress to subcortical dementia. We considered the possibility that the association near *NOTCH4* might be related to vascular dementia, but noted that the association was equally strong when only pure AD cases were considered.

Three large-scale genome-wide association studies on AD have now been conducted, with at least 3 genes (*CR1, PICALM*, and *CLU*) in addition to *APOE* now believed to contribute to risk [39–41]. Together however, these explain less than 2% of the genetic component of AD. It was interesting that one of the new candidate genes, *CR1*, is intimately involved in the complement system and represents the first convincing genetic finding linking the inflammation system to AD. The *AGER* region also includes several complement genes (*C2, C4A, C4B*), raising the issue that we might be capturing the effects of any of these by virtue of LD with the significant markers. There was however no evidence of association with complement gene expression with our associated variants, nor LD with putative functional protein changing variants in those genes. Neither of the two large-scale genome-wide association studies that have been performed highlights *AGER* specifically but both do indicate association further downstream, closer to HLA genes, at around 32.5Mb. These however most likely represent signals distinct from our study, in that LD is low between our most strongly associated variants and theirs. However we cannot exclude the possibility that LD relationships differ between the various populations which is a strong characteristic of the HLA region, in particular as it relates to north and south gradients in Europe [42] Evseeva et al., 2010]. Our sample consisted primarily of AD cases, but we performed analyses using both a pure AD set, the full dementia sample, and only the non-AD dementia cases. The effect size estimates were equivalent however, suggesting that an effect of sequence variants near *AGER/NOTCH4* may be generalized to multiple forms of dementia.

In summary, we have identified a potential association of sequence variation in the vicinity of *AGER* on chromosome 6p with dementia and AD. Further replication studies in alternative populations are warranted, as are functional studies and a continued application of bioinformatics to refine the more precise location of the association signal.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by the US National Institutes of Health [AG028555, AG08724, AG 04563, AG10175, AG08861]; and the Swedish Research Council [2007-2722].

References

[1]. Yankner BA, Dawes LR, Fisher S, Villa-Komaroff L, Oster-Granite ML, Neve RL. Neurotoxicity of a fragment of the amyloid precursor associated with Alzheimer's disease. Science. 1989; 245:417–20. [PubMed: 2474201]

[2]. Goate A, Chartier-Harlin MC, Mullan M, Brown J, Crawford F, Fidani L, Giuffra L, Haynes A, Irving N, James L, Mant R, Newton P, Rooke K, Roques P, Talbot C, Pericak-Vance M, Roses A, Williamson R, Rossor M, Owen M, Hardy J. Segregation of a missense mutation in the amyloid precursor protein gene with familial Alzheimer's disease. Nature. 1991; 349:704–706. [PubMed: 1671712]

[3]. Rogaev EI, Sherrington R, Rogaeva EA, Levesque G, Ikeda M, Liang Y, Chi H, Lin C, Holman K, Tsuda T, Mar L, Sorbi S, Nacmias B, Piacentini S, Chumakov I, Cohen D, Lannfelt L, Fraser PE, Rommens JM, St. George-Hyslop PH. Familial Alzheimer's disease in kindreds with missense mutations in a gene on chromosome 1 related to the Alzheimer's disease type 3 gene. Nature. 1995; 376:775–778. [PubMed: 7651536]

[4]. Sherrington R, Rogaev EI, Liang Y, Rogaeva EA, Levesque G, Ikeda M, Chi H, Lin C, Li G, Holman K, Tsuda T, Mar L, Foncin J-F, Bruni AC, Montesi MP, Sorbi S, Rainero I, Pinessi L, Nee L, Chumakov I, Pollen D, Brookes A, Sanseau P, Polinsky RJ, Wasco W, Da Silva HAR, Haines JL, Pericak-Vance MA, Tanzi RE, Roses AD, Fraser PE, Rommens JM, St. George-Hyslop PH. Cloning of a gene bearing missense mutations in early-onset familial Alzheimer's disease. Nature. 1995; 375:754–760. [PubMed: 7596406]

[5]. Strittmatter WJ, Saunders AM, Goedert M, Weisgraber KH, Dong LM, Jakes R, Huang DY, Pericak-Vance M, Schmechel D, Roses AD. Isoform-specific interactions of apolipoprotein E with microtubule-associated protein tau: implications for Alzheimer disease. Proc Natl Acad Sci U S A. 1994; 91:11183–11186. [PubMed: 7972031]

[6]. Schmechel DE, Saunders AM, Strittmatter WJ, Crain BJ, Hulette CM, Joo SH, Pericak-Vance MA, Goldgaber D, Roses AD. Increased amyloid beta-peptide deposition in cerebral cortex as a consequence of apolipoprotein E genotype in late-onset Alzheimer disease. Proc Natl Acad Sci U S A. 1993; 90:9649–53. [PubMed: 8415756]

[7]. Prince JA, Zetterberg H, Andreasen N, Marcusson J, Blennow K. APOE epsilon4 allele is associated with reduced cerebrospinal fluid levels of Abeta42. Neurology. 2004; 62:2116–2118. [PubMed: 15184629]

[8]. Dickson DW, Crystal HA, Mattiace LA, Masur DM, Blau AD, Davies P, Yen SH, Aronson MK. Identification of normal and pathological aging in prospectively studied nondemented elderly humans. Neurobiol Aging. 1992; 13:179–89. [PubMed: 1311804]

[9]. McGeer EG, McGeer PL. Neuroinflammation in Alzheimer's disease and mild cognitive impairment: a field in its infancy. J Alzheimers Dis. 2010; 19:355–61. [PubMed: 20061650]

[10]. Cameron B, Landreth GE. Inflammation, microglia, and Alzheimer's disease. Neurobiol Dis. 2010; 37:503–9. [PubMed: 19833208]

[11]. Goedert M. Oskar Fischer and the study of dementia. Brain. 2009; 132:1102–11. [PubMed: 18952676]

[12]. Fischer O. Miliare Nekrosen mit drusigen Wucherungen der Neurofibrillen, eine regelmässige Veränderung der Hirnrinde bei seniler Demenz. Monatsschr Psychiat Neurol. 1907; 22:361–72.

[13]. Bertram L, McQueen MB, Mullin K, Blacker D, Tanzi RE. Systematic meta-analyses of Alzheimer disease genetic association studies: the AlzGene database. Nat Genet. 2007; 39:17–23. [PubMed: 17192785]

[14]. Lichtenstein P, De Faire U, Floderus B, Svartengren M, Svedberg P, Pedersen NL. The Swedish Twin Registry: a unique resource for clinical, epidemiological and genetic studies. J Intern Med. 2002; 252:184–205. [PubMed: 12270000]

[15]. Reynolds CA, Hong MG, Eriksson UK, Blennow K, Wiklund F, Johansson B, Malmberg B, Berg S, Alexeyenko A, Grönberg H, Gatz M, Pedersen NL, Prince JA. Analysis of lipid pathway genes indicates association of sequence variation near SREBF1/TOM1L2/ATPAF2 with dementia risk. Hum Mol Genet. 2010; 19:2068–78. [PubMed: 20167577]

[16]. Pedersen NL, Friberg L, Floderus-Myrhed B, McClearn GE, Plomin R. Swedish early separated twins: identification and characterization. Acta Genet Med Gemellol. 1984; 33:243–250. [PubMed: 6540957]

[17]. McClearn GE, Johansson B, Berg S, Pedersen NL, Ahern F, Petrill SA, Plomin R. Substantial genetic influence on cognitive abilities in twins 80 or more years old. Science. 1997; 276:1560–1563. [PubMed: 9171059]

[18]. Gold CH, Malmberg B, McClearn GE, Pedersen NL, Berg S. Gender and health: a study of older unlike-sex twins. J Gerontol B Psychol Sci Soc Sci. 2002; 57:S168–176. [PubMed: 11983743]

[19]. Gatz M, Fratiglioni L, Johansson B, Berg S, Mortimer JA, Reynolds CA, Fiske A, Pedersen NL. Complete ascertainment of dementia in the Swedish Twin Registry: the HARMONY study. Neurobiol Aging. 2005; 26:439–447. [PubMed: 15653172]

[20]. Myers AJ, Gibbs JR, Webster JA, Rohrer K, Zhao A, Marlowe L, Kaleem M, Leung D, Bryden L, Nath P, Zismann VL, Joshipura K, Huentelman MJ, Hu-Lince D, Coon KD, Craig DW, Pearson JV, Holmans P, Heward CB, Reiman EM, Stephan D, Hardy J. A survey of genetic human cortical gene expression. Nat Genet. 2007; 39:1494–1499. [PubMed: 17982457]

[21]. Goto S, Bono H, Ogata H, Fujibuchi W, Nishioka T, Sato K, Kanehisa M. Organizing and computing metabolic pathway data in terms of binary relations. Pac Symp Biocomput. 1997:175–186. [PubMed: 9390290]

[22]. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics. 2005; 21:263–265. [PubMed: 15297300]

[23]. Andreasen N, Minthon L, Clarberg A, Davidsson P, Gottfries J, Vanmechelen E, Vanderstichele H, Winblad B, Blennow K. Sensitivity, specificity, and stability of CSF-tau in AD in a community-based patient sample. Neurology. 1999; 53:1488–1494. [PubMed: 10534256]

[24]. Vanmechelen E, Vanderstichele H. Towards an earlier diagnosis of Alzheimer's disease. J Biotechnol. 1998; 66:229–231. [PubMed: 9866871]

[25]. Blennow K, Wallin A, Agren H, Spenger C, Siegfried J, Vanmechelen E. Tau protein in cerebrospinal fluid: a biochemical marker for axonal degeneration in Alzheimer disease? Mol Chem Neuropathol. 1995; 26:231–245. [PubMed: 8748926]

[26]. Carey V, Zeger SL, Diggle P. Modelling multivariate binary data with alternating logistic regressions. Biometrika. 1993; 80:517–526.

[27]. Jansson M, Gatz M, Berg S, Johansson B, Malmberg B, McClearn GE, Schalling M, Pedersen NL. Association between depressed mood in the elderly and a 5-HTR2A gene variant. Am J Med Genet B Neuropsychiatr Genet. 2003; 120:79–84. [PubMed: 12815744]

[28]. Alexeyenko A, Sonnhammer EL. Global networks of functional coupling in eukaryotes from comprehensive data integration. Genome Res. 2009; 19:1107–1116. [PubMed: 19246318]

[29]. Maslov S, Sneppen K. Specificity and stability in topology of protein networks. Science. 2002; 296:910–913. [PubMed: 11988575]

[30]. Marchini J, Howie B, Myers S, McVean G, Donnelly P. A new multipoint method for genome-wide association studies by imputation of genotypes. Nat Genet. 2007; 39:906–913. [PubMed: 17572673]

[31]. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007; 81:559–575. [PubMed: 17701901]

[32]. Schmidt AM, Sahagan B, Nelson RB, Selmer J, Rothlein R, Bell JM. The role of RAGE in amyloid-beta peptide-mediated pathology in Alzheimer's disease. Curr Opin Investig Drugs. 2009; 10:672–80.

[33]. Reynolds CA, Hong MG, Eriksson UK, Blennow K, Bennet AM, Johansson B, Malmberg B, Berg S, Wiklund F, Gatz M, Pedersen NL, Prince JA. A survey of ABCA1 sequence variation confirms association with dementia. Hum Mutat. 2009; 30:1348–54. [PubMed: 19606474]

[34]. Goh KI, Cusick ME, Valle D, Childs B, Vidal M, Barabasi AL. The human disease network. Proc Natl Acad Sci U S A. 2007; 104:8685–8690. [PubMed: 17502601]

[35]. Leclerc E, Sturchler E, Vetter SW, Heizmann CW. Crosstalk between calcium, amyloid beta and the receptor for advanced glycation endproducts in Alzheimer's disease. Rev Neurosci. 2009; 20:95–110. [PubMed: 19774788]

[36]. Woo HN, Park JS, Gwon AR, Arumugam TV, Jo DG. Alzheimer's disease and Notch signaling. Biochem Biophys Res Commun. 2009; 390:1093–7. [PubMed: 19853579]

[37]. Li H, Wolfe MS, Selkoe DJ. Toward structural elucidation of the gamma-secretase complex. Structure. 2009; 17:326–34. [PubMed: 19278647]

[38]. Chabriat H, Joutel A, Dichgans M, Tournier-Lasserve E, Bousser MG. Cadasil. Lancet Neurol. 2009; 8:643–53. [PubMed: 19539236]

[39]. Lambert JC, Heath S, Even G, Campion D, Sleegers K, Hiltunen M, Combarros O, Zelenika D, Bullido MJ, Tavernier B, Letenneur L, Bettens K, Berr C, Pasquier F, Fiévet N, Barberger-Gateau P, Engelborghs S, De Deyn P, Mateo I, Franck A, Helisalmi S, Porcellini E, Hanon O, European Alzheimer's Disease Initiative Investigators. de Pancorbo MM, Lendon C, Dufouil C, Jaillard C, Leveillard T, Alvarez V, Bosco P, Mancuso M, Panza F, Nacmias B, Bossù P, Piccardi P, Annoni G, Seripa D, Galimberti D, Hannequin D, Licastro F, Soininen H, Ritchie K, Blanché H, Dartigues JF, Tzourio C, Gut I, Van Broeckhoven C, Alpérovitch A, Lathrop M, Amouyel P. Genome-wide association study identifies variants at CLU and CR1 associated with Alzheimer's disease. Nat Genet. 2009; 41:1094–9. [PubMed: 19734903]

[40]. Harold D, Abraham R, Hollingworth P, Sims R, Gerrish A, Hamshere ML, Pahwa JS, Moskvina V, Dowzell K, Williams A, Jones N, Thomas C, Stretton A, Morgan AR, Lovestone S, Powell J, Proitsi P, Lupton MK, Brayne C, Rubinsztein DC, Gill M, Lawlor B, Lynch A, Morgan K, Brown KS, Passmore PA, Craig D, McGuinness B, Todd S, Holmes C, Mann D, Smith AD, Love S, Kehoe PG, Hardy J, Mead S, Fox N, Rossor M, Collinge J, Maier W, Jessen F, Schürmann B, van den Bussche H, Heuser I, Kornhuber J, Wiltfang J, Dichgans M, Frölich L, Hampel H, Hüll M, Rujescu D, Goate AM, Kauwe JS, Cruchaga C, Nowotny P, Morris JC, Mayo K, Sleegers K, Bettens K, Engelborghs S, De Deyn PP, Van Broeckhoven C, Livingston G, Bass NJ, Gurling H, McQuillin A, Gwilliam R, Deloukas P, Al-Chalabi A, Shaw CE, Tsolaki M, Singleton AB, Guerreiro R, Mühleisen TW, Nöthen MM, Moebus S, Jöckel KH, Klopp N, Wichmann HE, Carrasquillo MM, Pankratz VS, Younkin SG, Holmans PA, O'Donovan M, Owen MJ, Williams J. Genome-wide association study identifies variants at CLU and PICALM associated with Alzheimer's disease. Nat Genet. 2009; 41:1088–93. [PubMed: 19734902]

[41]. Seshadri S, Fitzpatrick AL, Ikram MA, DeStefano AL, Gudnason V, Boada M, Bis JC, Smith AV, Carassquillo MM, Lambert JC, Harold D, Schrijvers EM, Ramirez-Lorca R, Debette S, Longstreth WT Jr, Janssens AC, Pankratz VS, Dartigues JF, Hollingworth P, Aspelund T, Hernandez I, Beiser A, Kuller LH, Koudstaal PJ, Dickson DW, Tzourio C, Abraham R, Antunez C, Du Y, Rotter JI, Aulchenko YS, Harris TB, Petersen RC, Berr C, Owen MJ, Lopez-Arrieta J, Varadarajan BN, Becker JT, Rivadeneira F, Nalls MA, Graff-Radford NR, Campion D, Auerbach S, Rice K, Hofman A, Jonsson PV, Schmidt H, Lathrop M, Mosley TH, Au R, Psaty BM, Uitterlinden AG, Farrer LA, Lumley T, Ruiz A, Williams J, Amouyel P, Younkin SG, Wolf PA, Launer LJ, Lopez OL, van Duijn CM, Breteler MM, CHARGE Consortium; GERAD1 Consortium; EADI1 Consortium. Genome-wide analysis of genetic loci associated with Alzheimer disease. JAMA. 2010; 303:1832–40. [PubMed: 20460622]

[42]. Evseeva I, Nicodemus KK, Bonilla C, Tonks S, Bodmer WF. Linkage disequilibrium and age of HLA region SNPs in relation to classic HLA gene alleles within Europe. Eur J Hum Genet. 2010; 18:924–32. [PubMed: 20354563]

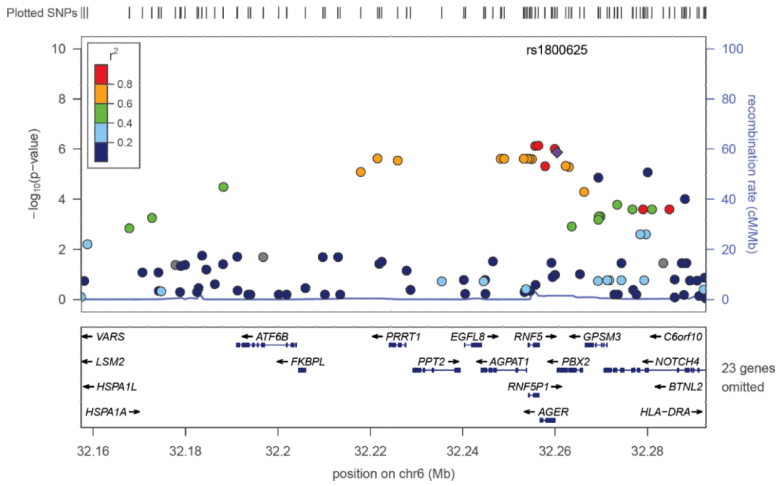

Figure 1.

Schematic overview of the *AGER* region extending from marker rs11654482 to marker rs854764 on chromosome 6p, depicting unadjusted allelic association of $-\log_{10}P$ (Y-axis) values for 110 genetic markers in the X-axis in relation to the full dementia versus control sample with a maximum for rs1800625. The figure was generated using Locuszoom (http://csg.sph.umich.edu/locuszoom).

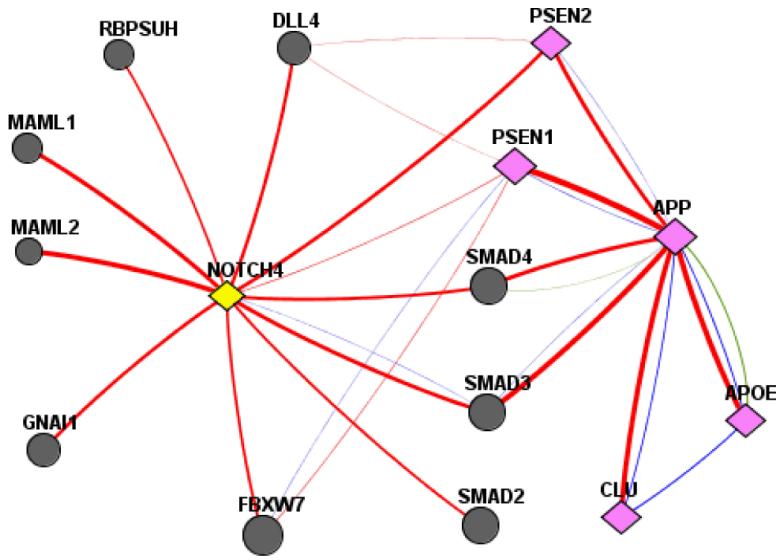

Figure 2.

Sub-network of candidate genes from this study and known Alzheimer's disease genes. The sub-networks were generated with a query to FunCoup web site (http://funcoup.sbc.su.se) so that only a part of the functional links is shown (algorithm #1, selection by centrality and confidence). Most of the indirect links counted in Table 3 are thus hidden. Lines represent the experimental evidence of functional links): Red, known protein-protein interactions; Blue, mRNA co-expression; Green, protein co-expression. Node colors: Yellow, candidate genes from this study; Magenta, the base set of known Alzheimer's disease genes (*CR1* had no links in the FunCoup network); Grey, other genes. Node sizes are proportional to the number of connections for each gene in the whole network

Table 1

Genes with putative involvement in inflammation tested for association with dementia

Gene Symbol	location	SNPs designed	SNPs genotyped
PLA2G2A	Chr1 - 20.1 Mb	18	16
APCS	Chr1 - 157.8 Mb	11	11
CRP	Chr1 - 157.9 Mb	6	6
PTGS2	Chr1 - 184.9 Mb	14	12
IL1A	Chr2 - 113.2 Mb	10	10
IL1B	Chr2 - 113.3 Mb	10	9
REST	Chr4 - 57.4 Mb	2	2
TNF	Chr6 - 31.6 Mb	16	13
AGER	Chr6 - 32.2 Mb	14	12
PLA2G7	Chr6 - 46.7 Mb	21	18
IL6	Chr7 - 22.7 Mb	16	16
SERPINE1	Chr7 - 100.5 Mb	12	11
CHRM2	Chr7 - 136.2 Mb	9	9
LOC729983	Chr9 - 22.1 Mb	2	2
SAA1	Chr11 - 18.2 Mb	21	18
BDNF	Chr11 - 27.6 Mb	11	11
NR1H3	Chr11 - 47.2 Mb	11	11
SERPINA3	Chr14 - 94.1 Mb	17	16
SELS	Chr15 - 99.6 Mb	11	10
NR1H2	Chr19 - 55.5 Mb	5	4

SNPs genotyped refers to markers successfully genotyped with call rates above 90%. Markers were excluded due to complete assay failure, call rates below 90%, or significant deviation from Hardy-Weinberg equilibrium.

Table 2

Top10 most associated markers with dementia

Marker	Nearest Gene	Odds Ratios	p-value	Risk Allele (type)	Aβ342 p	AAO p
rs1800625	PBX2	1.37, 1.19–1.56	1.36E-06	A (G/A)	1.76E-01	9.48E-01
rs204995	PBX2	1.32, 1.17–1.51	4.68E-06	A (G/A)	4.31E-01	7.52E-01
rs10916683	PLA2G2A	1.29, 1.12–1.48	6.61E-04	A (G/A)	1.34E-01	3.60E-02
rs12218	SAA1	1.19, 1.08–1.32	1.31E-03	G (G/A)	2.66E-01	1.06E-01
rs204993	PBX2	1.18, 1.06–1.33	1.47E-03	A (G/A)	7.75E-01	8.63E-01
rs3796530	REST	1.35, 1.11–1.65	3.01E-03	G (G/A)	1.77E-01	7.20E-02
rs1800629	TNF	1.20, 1.04–1.39	3.58E-03	G (G/A)	5.24E-01	5.62E-01
rs361525	TNF	1.48, 1.11–2.00	7.61E-03	G (G/A)	3.62E-01	1.94E-01
rs1446963	APCS	1.18, 1.05–1.33	1.36E-02	G (G/A)	6.14E-01	2.61E-01
rs20417	PTGS2	1.18, 1.03–1.36	2.15E-02	C (C/G)	8.06E-01	5.17E-01

Combined significance for dementia risk in two independent samples was determined using xtlogit (STATA) to correct for twin correlation structure.

Odds ratios (ORs) are reported with 95% confidence intervals.

Association with Aβ342 was evaluated with ANOVA in both cases and controls with a disease class covariate.

Age at onset (AAO) was evaluated using xtmixed (STATA) to account for twin structure.

Table 3

Statistical analysis of candidate genes in a network of functional coupling.

Gene	Direct links to base set*				Indirect links to base set* (via common neighbor)			
	N, in real network	N, mean in 100 permuted networks	Standard deviation	p	N, in real network	N, mean in 100 permuted networks	Standard deviation	p
NOTCH 4	2	0.00	0.000	<1.00e-16	9	3.83	2.38	0.015
EGFL8	0	0.00	0.000	–	9	9.30	3.21	0.413
AGPAT1	0	0.00	0.000	–	0	0.30	0.54	0.713
RNF5	0	0.00	0.000	–	4	5.80	2.12	0.802
AGER	0	0.00	0.000	–	6	10.37	2.92	0.933
FKBPL	0	0.00	0.000	–	10	15.43	3.53	0.938
ATF6B	0	0.10	0.305	0.626	66	93.10	10.45	0.995
PBX2	0	0.00	0.000	–	0	13.53	3.35	0.999
PPT2	0	0.00	0.000	–	0	0.00	0.000	–
PRRT1	0	0.00	0.000	–	0	0.00	0.000	–
GPSM3	NA	NA	NA	NA	NA	NA	NA	NA

*The whole human gene network predicted in FunCoup (version 1.1) was considered, i.e. all 2291931 links between 17151 known genes from evidence of 8 data types in humans and six eukaryotic model organisms. The base gene set consists of *APOE, APP, PSEN1, PSEN2, CR1, CLU,* and *PICALM.*