nature genetics

Genome-wide association study identifies variants at *CLU* and *CR1* associated with Alzheimer's disease

Jean-Charles Lambert[1–3], Simon Heath[4], Gael Even[1,2], Dominique Campion[5], Kristel Sleegers[6,7], Mikko Hiltunen[8], Onofre Combarros[9], Diana Zelenika[4], Maria J Bullido[10], Béatrice Tavernier[11], Luc Letenneur[12], Karolien Bettens[6,7], Claudine Berr[13], Florence Pasquier[3,14], Nathalie Fiévet[1,2], Pascale Barberger-Gateau[12], Sebastiaan Engelborghs[7,15], Peter De Deyn[7,15], Ignacio Mateo[9], Ana Franck[16], Seppo Helisalmi[8], Elisa Porcellini[17], Olivier Hanon[18], the European Alzheimer's Disease Initiative Investigators[19], Marian M de Pancorbo[20], Corinne Lendon[21], Carole Dufouil[22,23], Céline Jaillard[24], Thierry Leveillard[24], Victoria Alvarez[25], Paolo Bosco[26], Michelangelo Mancuso[27], Francesco Panza[28], Benedetta Nacmias[29], Paola Bossù[30], Paola Piccardi[31], Giorgio Annoni[32], Davide Seripa[33], Daniela Galimberti[34], Didier Hannequin[5], Federico Licastro[17], Hilkka Soininen[8], Karen Ritchie[13], Hélène Blanché[35], Jean-François Dartigues[12], Christophe Tzourio[22,23], Ivo Gut[4], Christine Van Broeckhoven[6,7], Annick Alpérovitch[22,23], Mark Lathrop[4,35] & Philippe Amouyel[1–3,14]

The gene encoding apolipoprotein E (*APOE*) on chromosome 19 is the only confirmed susceptibility locus for late-onset Alzheimer's disease. To identify other risk loci, we conducted a large genome-wide association study of 2,032 individuals from France with Alzheimer's disease (cases) and 5,328 controls. Markers outside *APOE* with suggestive evidence of association ($P < 10^{-5}$) were examined in collections from Belgium, Finland, Italy and Spain totaling 3,978 Alzheimer's disease cases and 3,297 controls. Two loci gave replicated evidence of association: one within *CLU* (also called *APOJ*), encoding clusterin or apolipoprotein J, on chromosome 8 (rs11136000, OR = 0.86, 95% CI 0.81–0.90, $P = 7.5 \times 10^{-9}$ for combined data) and the other within *CR1*, encoding the complement component (3b/4b) receptor 1, on chromosome 1 (rs6656401, OR = 1.21, 95% CI 1.14–1.29, $P = 3.7 \times 10^{-9}$ for combined data). Previous biological studies support roles of *CLU* and *CR1* in the clearance of β amyloid (Aβ) peptide, the principal constituent of amyloid plaques, which are one of the major brain lesions of individuals with Alzheimer's disease.

Alzheimer's disease is a neurological disorder primarily affecting the elderly that manifests through memory disorders, cognitive decline and loss of autonomy. Two principal types of neuropathologic lesions are observed: (i) neurofibrillary degeneration resulting from the intraneuronal accumulation of hyperphosphorylated Tau proteins and (ii) amyloid deposits resulting from the extracellular accumulation of amyloid plaques, which are primarily composed of Aβ peptides. Currently, the processes leading to the formation of these lesions and their combined association with Alzheimer's disease are not adequately understood[1].

Genetic studies have provided significant insights into the molecular basis of Alzheimer's disease. Rare hereditary early-onset forms of the disease have been linked to mutations in three different genes: *APP*, encoding amyloid precursor protein, on chromosome 21; *PS1*, encoding presenilin 1, on chromosome 14; and *PS2*, encoding presenilin 2, on chromosome 1 (ref. 2). These mutations, however, explain less than 1% of all cases of Alzheimer's disease, whereas the vast majority (especially for late-onset forms of the disease) have other, more complex genetic determinants[3].

Genetic studies have led to the consistent identification of the ε4 allele of *APOE* as a susceptibility locus for late-onset Alzheimer's disease[4]. Twin studies suggest that genes may have a role in more than 60% of Alzheimer's disease susceptibility[5] and that *APOE* may account for as much as 50% of this genetic susceptibility[6]. More than 550 other genes have been proposed as candidates for Alzheimer's disease susceptibility, but thus far none have been confirmed to have a role in Alzheimer's disease pathogenesis[7].

As with other multifactorial diseases, this knowledge gap has motivated more comprehensive investigations using genome-wide association studies (GWAS). The first GWAS of case-control Alzheimer's disease data collections have examined a relatively small number of cases (<1,000)[8–12]. Similar to studies done on other multifactorial disorders, these GWAS have shown that, except in the case of *APOE*, larger samples will be necessary to locate common genetic factors of Alzheimer's disease. Here, we report results from a large two-stage GWAS of late-onset Alzheimer's disease.

In the first stage of this study, we undertook a GWA analysis of 537,029 SNPs in 2,032 French Alzheimer's disease cases and 5,328 French controls. Patients with probable Alzheimer's disease were ascertained by neurologists. Individuals without symptoms of

*A full list of authors and affiliations appears at the end of the paper.

Received 12 May; accepted 31 July; published online 6 September 2009; doi:10.1038/ng.439

Table 1 Association of SNPs at the *CLU* locus with Alzheimer's disease in the Stage 1 and Stage 2 samples

	rs2279590						
	N		MAF			Association test	
	Cases	Controls	Cases	Controls	HWE	OR (95% CI)	*P* value
Stage 1	**2,025**	**5,328**	**0.36**	**0.41**	**3.1×10^{-1}**	**0.83 (0.77–0.90)**	**1.0×10^{-6}**
Stage 2	**3,803**	**3,097**	**0.38**	**0.41**	**4.9×10^{-1}**	**0.88 (0.81–0.95)**	**8.2×10^{-4}**
Belgium	1,071	505	0.38	0.41	3.0×10^{-1}	0.82 (0.69–0.98)	3.1×10^{-2}
Finland	587	645	0.40	0.42	4.3×10^{-1}	0.92 (0.78–1.09)	3.5×10^{-1}
Italy	1,410	1,206	0.38	0.41	3.4×10^{-1}	0.87 (0.77–0.97)	1.3×10^{-2}
Spain	738	806	0.38	0.40	5.6×10^{-1}	0.91 (0.79–1.06)	2.4×10^{-1}
Stage 1 & 2	**5,828**	**8,425**	**0.37**	**0.41**	**2.9×10^{-1}**	**0.86 (0.82–0.91)**	**8.9×10^{-9}**
APOE ε4 carriers	3,060	1,714	0.36	0.40	7.5×10^{-1}	0.81 (0.74–0.89)	1.9×10^{-5}
APOE ε4 non carriers	2,727	6,697	0.39	0.41	1.8×10^{-1}	0.92 (0.85–0.98)	1.3×10^{-2}
	rs11136000						
Stage 1	**2,016**	**5,266**	**0.35**	**0.39**	**6.0×10^{-1}**	**0.83 (0.77–0.90)**	**1.5×10^{-6}**
Stage 2	**3,775**	**3,154**	**0.35**	**0.38**	**2.6×10^{-1}**	**0.88 (0.81–0.95)**	**8.8×10^{-4}**
Belgium	987	467	0.35	0.37	5.2×10^{-2}	0.80 (0.66–0.97)	2.2×10^{-2}
Finland	596	640	0.38	0.41	2.5×10^{-1}	0.92 (0.78–1.08)	3.1×10^{-1}
Italy	1,454	1,241	0.35	0.38	6.0×10^{-1}	0.88 (0.80–0.99)	2.8×10^{-2}
Spain	738	806	0.35	0.37	6.1×10^{-1}	0.91 (0.79–1.06)	2.4×10^{-1}
Stage 1 + 2	**5,791**	**8,420**	**0.35**	**0.38**	**2.7×10^{-1}**	**0.86 (0.81–0.90)**	**7.5×10^{-9}**
APOE ε4 carriers	3,053	1,707	0.36	0.41	9.0×10^{-1}	0.81 (0.74–0.88)	2.7×10^{-5}
APOE ε4 non-carriers	2,693	6,699	0.39	0.41	1.9×10^{-1}	0.91 (0.85–0.97)	7.0×10^{-3}
	rs9331888						
Stage 1	**2,025**	**5,328**	**0.31**	**0.28**	**8.9×10^{-1}**	**1.19 (1.11–1.30)**	**1.8×10^{-5}**
Stage 2	**3,862**	**3,180**	**0.31**	**0.29**	**7.9×10^{-1}**	**1.12 (1.04–1.21)**	**2.9×10^{-3}**
Belgium	1,072	501	0.29	0.28	9.3×10^{-1}	1.16 (0.96–1.40)	1.2×10^{-1}
Finland	586	638	0.38	0.37	1.3×10^{-1}	1.09 (0.92–1.29)	3.2×10^{-1}
Italy	1,474	1,241	0.30	0.26	9.9×10^{-1}	1.22 (1.07–1.39)	4.5×10^{-3}
Spain	730	800	0.29	0.28	5.4×10^{-1}	1.06 (0.90–1.24)	4.8×10^{-1}
Stage 1 + 2	**5,887**	**8,508**	**0.30**	**0.28**	**3.3×10^{-1}**	**1.16 (1.10–1.23)**	**9.4×10^{-8}**
APOE ε4 carriers	3,098	1,723	0.32	0.29	8.6×10^{-1}	1.21 (1.10–1.33)	7.8×10^{-5}
APOE ε4 non-carriers	2,748	6,770	0.30	0.28	3.1×10^{-1}	1.09 (1.02–1.18)	1.8×10^{-2}

P values and ORs with the associated 95% CI have been calculated under an additive model using logistic regression models adjusted for age, gender and centers when necessary. MAF, minor allele frequency; HWE, *P* value for the test of Hardy-Weinberg equilibrium in controls.

dementia from French Three-City (3C) prospective population-based cohort were obtained as controls (**Supplementary Note**). Samples were genotyped with Illumina Human 610-Quad BeadChip and subjected to standard quality control procedures. The resulting GWA data were then analyzed with logistic regression taking into account sex and age and using principal components to adjust for possible population stratification. The genomic control parameter was 1.20 before this adjustment but 1.04 afterward. Comparison of the observed and expected χ^2 distributions (**Supplementary Fig. 1**) did not indicate substantial inflation of the test statistics after adjustment. We carried out additional tests to establish the robustness of the statistical results as described in the online methods section.

Several *APOE*-linked SNPs gave strong evidence of disease association (**Supplementary Table 1**). Outside of *APOE*, one marker (rs11136000) within *CLU* on chromosome 8p21-p12 showed significance ($P = 9.0 \times 10^{-8}$) in the association test. This slightly surpassed the criteria for genome-wide significance as evaluated with a conservative Bonferroni correction ($P < 9.3 \times 10^{-8}$). We observed markers in several chromosome regions with suggestive evidence of association ($P < 10^{-5}$) as shown in a Manhattan plot (**Supplementary Fig. 2**). The results of the entire GWAS are available online. Finally, we imputed genotypes using the HapMap CEU samples to increase the number of SNPs examined in these regions. The genotyped and imputed markers from the regions that gave $P < 10^{-5}$ are shown in **Supplementary Table 2**.

In the second stage, we sought replication of association by genotyping a selection of markers at the loci identified in **Supplementary Table 2** in additional collections totaling 3,978 probable Alzheimer's disease cases and 3,297 controls obtained from Belgium, Finland, Italy and Spain (**Supplementary Table 3** and **Supplementary Note**). Genotyping of the second-stage samples was performed using either Taqman or Sequenom assays. Data were analyzed using logistic regression under an additive genetic model taking into account sex and age at diagnosis (cases) or at confirmation of the absence of dementia (controls). Five SNPs at two loci showed association, with *P* values ranging from 1.6×10^{-2} to 8.2×10^{-4}. The first of these loci encompasses *CLU* on 8p21-p12, and the second spans the gene encoding complement component (3b/4b) receptor 1 (*CR1*) on 1q32.

At *CLU*, three markers (rs2279590, rs11136000, rs9331888) showed statistically significant association with Alzheimer's disease in the replication collections, even with a conservative Bonferroni correction for all 11 SNPs tested in stage 2 ($P < 4.5 \times 10^{-3}$, **Table 1**). The odds ratios were statistically homogeneous across the stage 1 and 2 study collections (**Supplementary Table 4**). We found strong evidence for association in the combined GWA and replication datasets taking into account the sample origin in the logistic regression, with two markers exceeding the criterion for genome-wide significance ($P < 9.3 \times 10^{-8}$). For the marker showing the strongest evidence of association (rs11136000), the odds ratio for the minor

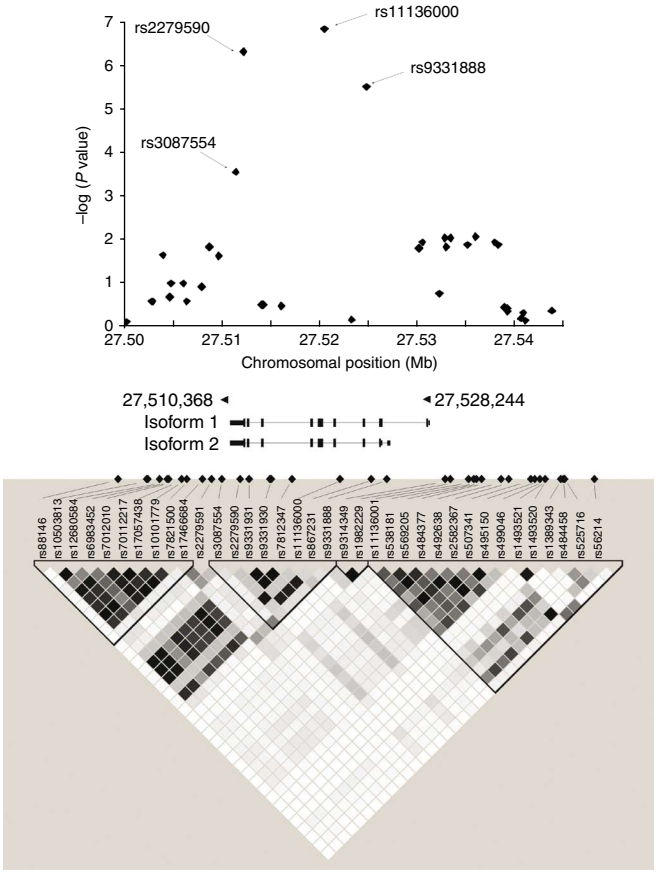

Figure 1 Schematic overview of *CLU* and LD patterns at the *CLU* locus. *P* values for association of SNPs (resulted from imputation and with minor allele frequency (MAF) >5%) encompassing the *CLU* locus with Alzheimer's disease risk under an additive model with adjustment for age and gender are plotted against physical distance. The LD plot is shown for SNPs at the *CLU* locus in controls (stage 1 data in Haploview 4.0, solid spine haplotype block definition, r^2 color scheme).

allele was 0.86 (95% CI 0.81–0.90, $P = 7.5 \times 10^{-9}$). Similar results were obtained using a Mantel-Haenszel statistic for the combined analysis (**Supplementary Table 5**).

We detected a statistical interaction between the *APOE* ε4 status and the *CLU* SNPs (ranging from 3.0×10^{-2} to 5.2×10^{-2} from the SNP tested). For rs11136000, although the association was significant in both ε4 carriers and non-carriers, it was more significant in carriers (**Table 1**).

The three *CLU* locus markers replicated in stage 2 are within a linkage disequilibrium (LD) block that encompasses only the *CLU* gene (**Fig. 1**). They define three common haplotypes (frequency >2%) that together account for 98.2% of the observations in stage 1 controls. Compared to the most frequent haplotype, TTC, the other two most frequent haplotypes were all associated with a statistically significant increased disease risk, with similar odd ratios in the stage 1 and stage 2 collections (**Table 2**). The odds ratio was highest for the CCG haplotype compared to the TTC haplotype (OR = 1.22, $P = 5.6 \times 10^{-10}$ for the combined samples).

A second locus of potential interest lies within an LD block that encompasses *CR1* on 1q32 (**Fig. 2**). We tested two SNPs at this locus in the second stage, and one (rs6656401) showed evidence of association with Alzheimer's disease in the replication collections ($P = 8.2 \times 10^{-4}$, **Table 3**). There was no significant heterogeneity by origin (**Supplementary Table 4**), and the odds ratio and *P* values in the combined data were 1.21 (95% CI 1.14–1.29) and 3.7×10^{-9}. We found similar results when using the Mantel-Haenszel method (**Supplementary Table 5**). At this locus, we also detected a statistical interaction with *APOE* ε4 status and risk of disease ($P = 9.6 \times 10^{-3}$) with significant association in both carriers and non-carriers, but with a more significant association in the former. Although the association of the second marker tested at this locus (rs3818361) only showed a suggestive significance overall in the stage 2 collections (OR = 1.11, 95% CI 1.02–1.22, $P = 1.6 \times 10^{-2}$), there was evidence of association of this SNP in the *APOE* ε4 carriers in stage 1, stage 2 and combined samples (respectively, OR = 1.38, 95% CI 1.19–1.60, $P = 2.3 \times 10^{-5}$; OR = 1.29, 95% CI 1.08–1.56, $P = 5.4 \times 10^{-3}$; OR = 1.34, 95% CI 1.20–1.50, $P = 2.9 \times 10^{-7}$). The genotyped markers define two principal haplotypes that account for 97.8% of the observations at the *CR1* locus, and a third haplotype has an estimated frequency of 1.2% in the combined control population (**Table 4**). The odds ratio was highest for the AA haplotype compared to the GG haplotype (OR = 1.22, $P = 3.1 \times 10^{-10}$ for the combined samples).

The association results for *CLU* are supported by a recent meta-analysis of linkage studies of Alzheimer's disease, in which a region of 15.1 Mb on chromosome 8p that encompasses *CLU* had the strongest evidence genome wide for linkage[13]. The biological functions of the *CLU* markers associated with Alzheimer's disease, and of other markers from the region that may be in LD with them, is presently unknown. However, *CLU* is the only described gene located within the association region, and several lines of evidence suggest that it is a strong candidate for involvement in the disease.

APOE and CLU are the most abundantly expressed apolipoproteins in the central nervous system[14,15], with strong analogies in terms of possible impact on the Alzheimer's disease physiopathological process. Like APOE, CLU is present in amyloid plaques[16,17] and can bind Aβ[18,19]. Consistent with this, several experiments in transgenic PDAPP mice lacking *CLU* have shown that this protein may strongly influence Aβ fibrillogenesis and increase Aβ neurotoxicity *in vivo*[20]. It has also been proposed that CLU may participate in Aβ clearance from the brain across the blood-brain barrier, mainly of the highly pathogenic Aβ42 peptide[21,22], which is a function similar to that of APOE[23–25].

In addition, we observed an association between Alzheimer's disease risk and markers spanning *CR1*, the gene that encodes for the main receptor of the complement C3b protein[26]. The biological function of these CR1 markers is unknown. The LD block containing these markers

Table 2 Association results for haplotypes at the *CLU* locus

	Stage 1				Stage 2				Stage 1 + 2			
Haplotypes	Cases	Controls	OR (95% CI)	*P* value	Cases	Controls	OR (95% CI)	*P* value	Cases	Controls	OR (95% CI)	*P* value
TTC	0.344	0.388	Ref.	–	0.343	0.372	Ref.	–	0.344	0.382	Ref.	–
CCC	0.334	0.329	1.14 (1.04–1.24)	3.0×10^{-3}	0.338	0.327	1.10 (1.01–1.21)	3.3×10^{-2}	0.336	0.328	1.12 (1.06–1.20)	8.0×10^{-5}
CCG	0.302	0.265	1.28 (1.17–1.41)	1.5×10^{-7}	0.277	0.258	1.19 (1.09–1.31)	1.5×10^{-4}	0.286	0.263	1.22 (1.14–1.29)	5.6×10^{-10}

The results have been calculated using the THESIAS program with adjustment for age, gender and center (see Online Methods). The *P* values for the global association were 1.8×10^{-6}, 2.9×10^{-3} and 2.7×10^{-9}, respectively, for the data from stage 1, stage 2 and stage 1 and 2 combined. The markers as ordered from left to right (5′ to 3′) are rs2279590, rs11136000 and rs9331888. Minor alleles are underlined.

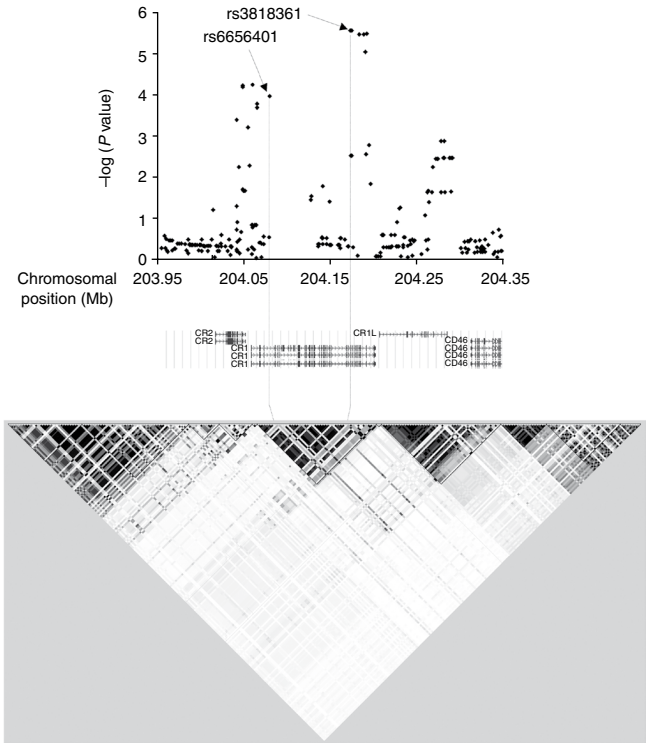

Figure 2 Schematic overview of *CR1* and LD patterns at the *CR1* locus. *P* values for association of SNPs (resulted from imputation and with MAF >5%) encompassing the *CR1* locus under an additive model with adjustment for age and gender are plotted against physical distance. The LD plot shown is for SNPs at the *CR1* locus in controls (stage 1 data in Haploview 4.0, solid spine haplotype block definition, r^2 color scheme).

only includes *CR1*. Several observations suggest that pathways involving C3b and CR1 are involved in the Alzheimer's disease process, particularly in Aβ clearance. Briefly, complement activation leads to the formation of C3-cleaving enzymes, known as C3 convertases, on the surface of the pathogen or protein undergoing complement attack. Cleavage of C3 results in the formation of C3a and C3b fragments. Whereas C3a is involved in the chemotaxis of phagocytes, C3b binds covalently to acceptor molecules and can mediate phagocytosis through CR1[27,28]. Such a mechanism has been proposed to participate in Aβ clearance for several reasons: (i) aggregated Aβ are able to activate and become bound by C3b[29–31]; (ii) in a transgenic mouse model, an increase in C3 expression coincides with a smaller degree of Aβ deposition and neuropathology[32,33]; (iii) conversely, expression of a C3 convertase inhibitor such as the rodent complement receptor 1-related gene/protein y (Crry) results in increased Aβ deposition and neurodegeneration[33]; (iv) using human erythrocytes (the cell type that most abundantly expresses CR1), it has been observed that this cell type is able to sequester $A\beta_{42}$(refs. 31,32) and to favor its clearance via the C3b-mediated adherence to erythrocyte CR1 (ref. 32); (v) finally, this mechanism may be deficient in individuals with Alzheimer's disease[31]. Altogether, these data support a protective role for CR1 via the generation and binding of C3b, which may contribute to Aβ clearance[33].

We calculated the attributable fractions of risk to be 25.5% for *APOE*, 8.9% for *CLU* and 3.8% for *CR1*. As these calculations are based on the combined stage 1 and 2 samples, the estimates are biased upward. Nevertheless, if the estimate that 60–80% of the Alzheimer's disease risk is due to genetic factors[5] is correct, additional genetic susceptibility loci remain to be identified, which is also true for many other diseases in which loci have been successfully mapped by GWA[34,35]. In a recent GWAS, the *PCDH11X* gene was found to be a genetic determinant of Alzheimer's disease[12]. We examined the effect of 17 SNPs in the *PCDH11X* locus in the stage 1 samples. Although none reached our criteria for suggestive level of significance, we did observe $0.01 < P < 0.05$ for several SNPs (**Supplementary Table 6**). Thus, this locus may also contribute to the Alzheimer's disease risk in our dataset, albeit not as significantly as the other loci that we have identified.

Table 3 Association of SNPs at the *CR1* locus with Alzheimer's disease in the Stage 1 and Stage 2 samples

	rs6656401						
	N		MAF			Association test	
	Cases	Controls	Cases	Controls	HWE	OR (95% CI)	*P* value
Stage 1	**2,025**	**5,324**	**0.22**	**0.18**	**9.9×10^{-1}**	**1.27 (1.16–1.39)**	**1.8×10^{-7}**
Stage 2	**3,880**	**3,198**	**0.20**	**0.18**	5.3×10^{-2}	**1.16 (1.06–1.27)**	**8.2×10^{-4}**
Belgium	1,066	500	0.22	0.18	2.2×10^{-1}	1.24 (0.99–1.24)	5.6×10^{-2}
Finland	608	654	0.21	0.17	2.9×10^{-1}	1.38 (1.12–1.70)	2.5×10^{-3}
Italy	1,472	1,243	0.20	0.20	2.2×10^{-1}	1.03 (0.88–1.17)	8.4×10^{-1}
Spain	734	801	0.20	0.16	4.2×10^{-2}	1.23 (1.02–1.47)	2.6×10^{-2}
Stage 1 + 2	**5,905**	**8,526**	**0.21**	**0.19**	1.1×10^{-1}	**1.21 (1.14–1.29)**	**3.5×10^{-9}**
APOE ε4 carriers	2,497	1,632	0.22	0.18	7.5×10^{-1}	1.38 (1.22–1.55)	1.8×10^{-7}
APOE ε4 non-carriers	2,761	6,780	0.21	0.19	4.3×10^{-2}	1.13 (1.04–1.23)	3.7×10^{-3}
	rs3818361						
Stage 1	**2,018**	**5,324**	**0.22**	**0.18**	**8.5×10^{-1}**	**1.28 (1.17–1.40)**	**8.5×10^{-8}**
Stage 2	**3,717**	**3,094**	**0.22**	**0.20**	**1.1×10^{-1}**	**1.11 (1.02–1.22)**	**1.6×10^{-2}**
Belgium	972	436	0.24	0.24	9.9×10^{-1}	1.05 (0.84–1.48)	6.8×10^{-1}
Finland	590	634	0.22	0.19	1.5×10^{-1}	1.26 (1,03–1.57)	2.6×10^{-2}
Italy	1,423	1,232	0.20	0.20	9.2×10^{-1}	1.03 (0.89–1.18)	7.2×10^{-1}
Spain	732	792	0.21	0.18	1.0×10^{-2}	1.21 (1.01–1.44)	3.8×10^{-2}
Stage 1 + 2	**5,735**	**8,418**	**0.22**	**0.19**	**2.2×10^{-1}**	**1.19 (1.11–1.26)**	**8.9×10^{-8}**
APOE ε4 carriers	3,032	1,696	0.22	0.18	8.9×10^{-1}	1.34 (1.20–1.50)	2.9×10^{-7}
APOE ε4 non-carriers	2,661	6,707	0.21	0.19	1.8×10^{-1}	1.12 (1.03–1.22)	6.4×10^{-3}

P values and ORs with their associated 95% CIs have been calculated under an additive model using logistic regression models adjusted for age, gender and centers when necessary. MAF, minor allele frequency; HWE, *P* value for the test of Hardy-Weinberg equilibrium in controls.

Table 4 Association results for haplotypes at the *CR1* locus

	Stage 1				Stage 2				Stage 1 + 2			
Haplotypes	Cases	Controls	OR (95% CI)	*P* value	Cases	Controls	OR (95% CI)	*P* value	Cases	Controls	OR (95% CI)	*P* value
GG	0.772	0.813	Ref.	–	0.778	0.786	Ref.		0.776	0.803	Ref.	
GA	0.011	0.009	1.25 (0.87–1.80)	2.2×10^{-1}	0.016	0.017	1.01 (0.77–1.33)	6.4×10^{-1}	0.014	0.012	1.13 (0.89–1.43)	3.9×10^{-1}
AA	0.207	0.170	1.28 (1.17–1.40)	1.4×10^{-7}	0.198	0.182	1.17 (1.07–1.27)	6.8×10^{-4}	0.202	0.175	1.22 (1.15–1.30)	3.1×10^{-10}

The results have been calculated using the THESIAS program with adjustment for age, gender and center (see Online Methods). The *P* values for the global association were 7.5×10^{-7}, 3.0×10^{-3} and 1.5×10^{-10}, respectively, for the data from stage 1, stage 2 and stage 1 and 2 combined. The markers as ordered from left to right (5′ to 3′) are rs6656401 and rs3818361. Minor alleles are underlined.

In an independent study published in this issue of *Nature Genetics*[36], Williams and colleagues report an independent GWAS of late-onset Alzheimer's disease and also report association with markers at *CLU*. In comparing these two studies, the Belgian samples used in the replication phase of each were found to overlap. To test for the independence of these results, we tested our association results with the entire Belgian collection removed from the replication phase. We find that the association to *CLU* remains significant, with OR = 0.89, 95% CI 0.83–0.97 and $P = 5.3 \times 10^{-3}$ for association at rs11136000 in the stage 2 samples when the Belgian collection is excluded. Although this is marginally less significant than the conservative *P* value that we applied at replication phase, the combined stage 1 and 2 samples gave OR = 0.86, 95% CI 0.81–0.91 and $P = 4.8 \times 10^{-8}$ with the Belgian collection excluded, meeting our criterion for genome-wide significance. Thus these two studies provide strong independent evidence of an association of Alzheimer's disease risk with *CLU* markers. In addition, evidence of association with the *CR1* markers at rs3818361 in the GWA data from Williams and her colleagues[36] exceeded our criteria for replication in their study ($P = 9.2 \times 10^{-6}$), providing additional confirmation of the relationship between *CR1* markers and late-onset Alzheimer's disease (OR = 1.17, 95% CI 1.09–1.25). We also examined *PILCAM* markers on chromosome 11 for which association has been identified and replicated in the second study. In our stage 1 data, we also detected evidence of association for *PILCAM* markers with *P* values between 10^{-2} and 10^{-3} (rs541458 OR = 0.88, 95% CI 0.81–0.96, $P = 2.8 \times 10^{-3}$ and results available online in GWAS results).

In summary, in addition to the previously known *APOE* locus, we have identified loci at *CLU* and *CR1* that are potentially associated with the risk of late-onset Alzheimer's disease. Biological evidence suggests that the genes at these loci, along with *APOE*, are involved in Aβ clearance. These data may indicate that whereas familial early-onset forms of Alzheimer's disease are mainly linked to genes implicated in Aβ overproduction, genetic variants at *APOE* and these newly defined loci may influence susceptibility to late-onset forms of the disease as a result of roles in Aβ clearance.

METHODS

Methods and any associated references are available in the online version of the paper at http://www.nature.com/naturegenetics/.

Note: Supplementary information is available on the Nature Genetics website.

ACKNOWLEDGMENTS

The work was made possible by the generous participation of the control subjects, the subjects with Alzheimer's disease, and their families. We thank A. Boland (Centre National de Génotypage) for her technical help in preparing the DNA samples for analyses. This work was supported by the National Foundation for Alzheimer's Disease and Related Disorders, the Institut Pasteur de Lille and the Centre National de Génotypage.

The Three-City Study was performed as part of a collaboration between the Institut National de la Santé et de la Recherche Médicale (Inserm), the Victor Segalen Bordeaux II University and Sanofi-Synthélabo. The Fondation pour la Recherche Médicale funded the preparation and initiation of the study. The 3C Study was also funded by the Caisse Nationale d'Assurance Maladie des Travailleurs Salariés, Direction Générale de la Santé, Mutuelle Générale de l'Education Nationale, Institut de la Longévité, Agence Française de Sécurité Sanitaire des Produits de Santé, the Aquitaine and Bourgogne Regional Councils, Fondation de France and the joint French Ministry of Research/Inserm 'Cohortes et collections de données biologiques' programme. Lille Génopôle received an unconditional grant from Eisai.

Belgium sample collection: the Antwerp site (CVB) was in part supported by the VIB Genetic Service Facility (http://www.vibgeneticservicefacility.be/) and the Biobank of the Institute Born-Bunge; the Special Research Fund of the University of Antwerp; the Fund for Scientific Research-Flanders (FWO-V); the Foundation for Alzheimer Research (SAO-FRMA); and the Interuniversity Attraction Poles (IAP) program P6/43 of the Belgian Federal Science Policy Office, Belgium; K.S. is a postdoctoral fellow and K.B. a PhD fellow of the FWO-V.

Finnish sample collection: Financial support for this project was provided by the Health Research Council of the Academy of Finland, EVO grant 5772708 of Kuopio University Hospital, and the Nordic Centre of Excellence in Neurodegeneration.

Italian sample collection: the Bologna site (FL) obtained funds from the Italian Ministry of research and University as well as Carimonte Foundation. The Florence site was supported by a grant from the Italian ministry of Health (RFPS-2006-7-334858). The Milan site was supported by a grant from the 'fondazione Monzino'. We appreciate the expert contribution of C. Romano.

Spanish sample collection: the Madrid site (MB) was supported by grants of the Ministerio de Educación y Ciencia and the Ministerio de Sanidad y Consumo (Instituto de Salud Carlos III) and an institutional grant of the Fundación Ramón Areces to the Centro de Biología Molecular Severo Ochoa. We thank I. Sastre and A. Martínez-García for the preparation and control of the DNA collection and P. Gil and P. Coria for their cooperation in the recruitment of cases and controls. We are grateful to the Asociación de Familiares de Alzheimer de Madrid (AFAL) for continuous encouragement and help.

AUTHOR CONTRIBUTIONS

Project conception, design, management. J.-C.L., M.L., P.A. **Phenotype collection, data management.** *France:* D.C., B.T., L.L., C.B., F.Pasquier, N.F., P.B.-G., O.H., C.L., C.D., C.J., T.L., D.H., K.R., H.B., J.-F.D., C.T., A.A. *Belgium*: K.S., K.B., S.E., P.D.D., C.V.B. *Finland:* M.H., S.Helisalmi, H.S. *Italy:* E.P., P.Bosco, M.M., F.Panza, B.N., P.Bossù, P.P., G.A., D.S., D.G., F.L. *Spain*: O.C., M.J.B., I.M., A.F., M.M.d.P., V.A. **Genome-wide, validation genotyping.** S.Heath., D.Z., I.G. **Data analysis:** J.-C.L., S.Heath., G.E., M.L., P.A. **Writing group:** J.-C.L., S.Heath., M.L., P.A.

COMPETING INTERESTS STATEMENT

The authors declare competing financial interests: details accompany the full-text HTML version of the paper at http://www.nature.com/naturegenetics/.

Published online at http://www.nature.com/naturegenetics/.
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/.

1. Small, S.A. & Duff, K. Linking Abeta and tau in late-onset Alzheimer's disease: a dual pathway hypothesis. *Neuron* **60**, 534–542 (2008).
2. Hardy, J. & Selkoe, D.J. The amyloid hypothesis of Alzheimer's disease: progress and problems on the road to therapeutics. *Science* **297**, 353–356 (2002).
3. Campion, D. *et al.* Early-onset autosomal dominant Alzheimer disease: prevalence, genetic heterogeneity, and mutation spectrum. *Am. J. Hum. Genet.* **65**, 664–670 (1999).
4. Farrer, L.A. *et al.* Effects of age, sex, and ethnicity on the association between apolipoprotein E genotype and Alzheimer disease. A meta-analysis. APOE and Alzheimer Disease Meta Analysis Consortium. *J. Am. Med. Assoc.* **278**, 1349–1356 (1997).
5. Gatz, M. *et al.* Role of genes and environments for explaining Alzheimer disease. *Arch. Gen. Psychiatry* **63**, 168–174 (2006).

6. Ashford, J.W. & Mortimer, J.A. Non-familial Alzheimer's disease is mainly due to genetic factors. *J. Alzheimers Dis.* **4**, 169–177 (2002).
7. Bertram, L., McQueen, M.B., Mullin, K., Blacker, D. & Tanzi, R.E. Systematic meta-analyses of Alzheimer disease genetic association studies: the AlzGene database. *Nat. Genet.* **39**, 17–23 (2007).
8. Reiman, E.M. *et al.* GAB2 alleles modify Alzheimer's risk in APOE epsilon4 carriers. *Neuron* **54**, 713–720 (2007).
9. Li, H. *et al.* Candidate single-nucleotide polymorphisms from a genomewide association study of alzheimer disease. *Arch. Neurol.* **65**, 45–53 (2008).
10. Bertram, L. *et al.* Genome-wide association analysis reveals putative Alzheimer's disease susceptibility loci in addition to APOE. *Am. J. Hum. Genet.* **83**, 623–632 (2008).
11. Beecham, G.W. *et al.* Genome-wide association study implicates a chromosome 12 risk locus for late-onset Alzheimer disease. *Am. J. Hum. Genet.* **84**, 35–43 (2009).
12. Carrasquillo, M.M. *et al.* Genetic variation in PCDH11X is associated with susceptibility to late-onset Alzheimer's disease. *Nat. Genet.* **41**, 192–198 (2009).
13. Butler, A.W. *et al.* Meta-analysis of linkage studies for Alzheimer's disease-A web resource. *Neurobiol. Aging* **30**, 1037–1047 (2009).
14. Roheim, P.S., Carey, M., Forte, T. & Vega, G.L. Apolipoproteins in human cerebrospinal fluid. *Proc. Natl. Acad. Sci. USA* **76**, 4646–4649 (1979).
15. May, P.C. & Finch, C.E. Sulfated glycoprotein 2: new relationships of this multifunctional protein to neurodegeneration. *Trends Neurosci.* **15**, 391–396 (1992).
16. May, P.C. *et al.* Dynamics of gene expression for a hippocampal glycoprotein elevated in Alzheimer's disease and in response to experimental lesions in rat. *Neuron* **5**, 831–839 (1990).
17. Calero, M. *et al.* Apolipoprotein J (clusterin) and Alzheimer's disease. *Microsc. Res. Tech.* **50**, 305–315 (2000).
18. Ghiso, J. *et al.* The cerebrospinal-fluid soluble form of Alzheimer's amyloid beta is complexed to SP-40,40 (apolipoprotein J), an inhibitor of the complement membrane-attack complex. *Biochem. J.* **293**, 27–30 (1993).
19. Zlokovic, B.V. *et al.* Glycoprotein 330/megalin: probable role in receptor-mediated transport of apolipoprotein J alone and in a complex with Alzheimer disease amyloid beta at the blood-brain and blood-cerebrospinal fluid barriers. *Proc. Natl. Acad. Sci. USA* **93**, 4229–4234 (1996).
20. DeMattos, R.B. *et al.* Clusterin promotes amyloid plaque formation and is critical for neuritic toxicity in a mouse model of Alzheimer's disease. *Proc. Natl. Acad. Sci. USA* **99**, 10843–10848 (2002).
21. DeMattos, R.B. *et al.* ApoE and clusterin cooperatively suppress Abeta levels and deposition: evidence that ApoE regulates extracellular Abeta metabolism in vivo. *Neuron* **41**, 193–202 (2004).
22. Bell, R.D. *et al.* Transport pathways for clearance of human Alzheimer's amyloid beta-peptide and apolipoproteins E and J in the mouse central nervous system. *J. Cereb. Blood Flow Metab.* **27**, 909–918 (2007).
23. Holtzman, D.M. In vivo effects of ApoE and clusterin on amyloid-beta metabolism and neuropathology. *J. Mol. Neurosci.* **23**, 247–254 (2004).
24. Lambert, J.C. & Amouyel, P. Genetic heterogeneity of Alzheimer's disease: complexity and advances. *Psychoneuroendocrinology* **32**, S62–S70 (2007).
25. Holtzman, D.M. *et al.* Expression of human apolipoprotein E reduces amyloid-beta deposition in a mouse model of Alzheimer's disease. *J. Clin. Invest.* **103**, R15–R21 (1999).
26. Iida, K., Mornaghi, R. & Nussenzweig, V. Complement receptor (CR1) deficiency in erythrocytes from patients with systemic lupus erythematosus. *J. Exp. Med.* **155**, 1427–1438 (1982).
27. Song, W.C., Sarrias, M.R. & Lambris, J.D. Complement and innate immunity. *Immunopharmacology* **49**, 187–198 (2000).
28. Khera, R. & Das, N. Complement receptor 1: disease associations and therapeutic implications. *Mol. Immunol.* **46**, 761–772 (2009).
29. Webster, S., Bradt, B., Rogers, J. & Cooper, N. Aggregation state-dependent activation of the classical complement pathway by the amyloid beta peptide. *J. Neurochem.* **69**, 388–398 (1997).
30. Rogers, J. *et al.* Peripheral clearance of amyloid beta peptide by complement C3-dependent adherence to erythrocytes. *Neurobiol. Aging* **27**, 1733–1739 (2006).
31. Kuo, Y.M. *et al.* Amyloid-beta peptides interact with plasma proteins and erythrocytes: implications for their quantitation in plasma. *Biochem. Biophys. Res. Commun.* **268**, 750–756 (2000).
32. Zhou, J., Fonseca, M.I., Pisalyaput, K. & Tenner, A.J. Complement C3 and C4 expression in C1q sufficient and deficient mouse models of Alzheimer's disease. *J. Neurochem.* **106**, 2080–2092 (2008).
33. Wyss-Coray, T. *et al.* Prominent neurodegeneration and increased plaque formation in complement-inhibited Alzheimer's mice. *Proc. Natl. Acad. Sci. USA* **99**, 10837–10842 (2002).
34. Donnelly, P. Progress and challenges in genome-wide association studies in humans. *Nature* **456**, 728–731 (2008).
35. Maher, B. Personal genomes: The case of the missing heritability. *Nature* **456**, 18–21 (2008).
36. Harold, D. *et al.* Genome-wide association study identifies variants at *CLU* and *PICALM* associated with Alzheimer's disease. *Nat. Genet.* advance online publication, doi:10.1038/ng.440 (6 September 2009).

[1]Inserm U744, Lille, France. [2]Institut Pasteur de Lille, Lille, France. [3]Université de Lille Nord de France, Lille, France. [4]Centre National de Génotypage, Institut Genomique, Commissariat à l'énergie Atomique, Evry, France. [5]Inserm U614, Faculté de Médecine-Pharmacie de Rouen, Rouen, France. [6]Neurodegenerative Brain Diseases group, Department of Molecular Genetics, VIB, Antwerpen, Belgium. [7]Institute Born-Bunge and University of Antwerp, Antwerpen, Belgium. [8]Department of Neurology, Kuopio University and University Hospital, Finland. [9]Neurology Service and Centro Investigación Biomédica en Red para Enfermedades Neurodegenerativas (CIBERNED), 'Marqués de Valdecilla' University Hospital, University of Cantabria, Santander, Spain. [10]Centro de Biologia Molecular Severo Ochoa, Universidad Autónoma de Madrid–Consejo Superior de Investigaciones Científicas and CIBERNED, Universidad Autonoma, Cantoblanco, Madrid, Spain. [11]Departement de Geriatrie, Centre Hospitalier Universitaire de Dijon, Dijon, France. [12]Inserm U897, Victor Segalen University, Bordeaux, France. [13]Inserm U888, Hôpital La Colombière, Montpellier, France. [14]Centre Hospitalier Régional Universitaire de Lille, Lille, France. [15]Memory Clinic and Department of Neurology, Ziekenhuisnetwerk Antwerpen Middelheim, Antwerpen, Belgium. [16]Servicio de Neurologia, Hospital Universitario La Paz, Universidad Autónoma de Madrid and CIBERNED, Madrid, Spain. [17]Department of Experimental Pathology, School of Medicine, University of Bologna, Italy. [18]Inserm U894 Faculté de Médecine, Université Paris Descartes, Paris, France. [19]A full list of members is provided in the **Supplementary Note**. [20]Banco de ADN, Universidad del Pais Vasco(UPV/EHU), Facultad de Farmacia, Vitoria-Gasteiz, Spain. [21]Molecular Psychiatry Laboratory, Queensland Institute of Medical Research, Royal Brisbane Hospital, Queensland, Australia. [22]Inserm U708, Paris, France. [23]Université Pierre et Marie Curie Paris 6 , Paris, France. [24]Department of genetics, Institut de la vision, Inserm, Université Pierre et Marie Curie Paris 6, Paris, France. [25]Genetic Molecular Unit, Hospital Universitario Central de Asturias, Oviedo, Spain. [26] Istituto Di Ricovero e Cura a Carattere Scientifico (IRCCS) Associazione Oasi Maria Santissima, Institute for Research on Mental Retardation and Brain Aging, Troina, Italy. [27]Department of Neuroscience, Neurological Clinic, University of Pisa, Italy. [28]Department of Geriatrics, Center for Aging Brain, Memory Unit, University of Bari, Bari, Italy. [29]Department of Neurological and Psychiatric Sciences, Florence, Italy. [30]Department of Clinical and Behavioral Neurology, IRCCS Santa Lucia Foundation, Roma, Italy. [31]Laboratory of Molecular Genetics, Section of Clinical Pharmacology, Department of Neuroscience, University of Cagliari, Italy. [32]Department of Clinical Medicine and Prevention, University of Milano-Bicocca, Monza, Italy. [33]Geriatric Unit & Gerontology-Geriatric Research Laboratory, Department of Medical Science, IRCCS Casa Sollievo della Sofferenza, San Giovanni Rotondo, Italy. [34]Department of Neurological Sciences, Dino Ferrari Center, University of Milan, IRCCS Ospedale Maggiore Policlinico, Milan, Italy. [35]Fondation Jean Dausset–Centre d'Etudes du Polymorphisme Humain, Paris, France. Correspondence should be addressed to P.A. (philippe.amouyel@pasteur-lille.fr).

ONLINE METHODS

Sample populations. The case-control cohorts are described in the **Supplementary Note**.

Genotyping. DNA samples were transferred to the French Centre National de Génotypage for genotyping. First stage samples that passed DNA quality control were genotyped with Illumina Human610-Quad BeadChips. Genotype data were retained in the study for samples that had been successfully genotyped for >98% of the SNP markers. SNPs with call rates of <98%, with MAF <1% or showing departure from Hardy-Weinberg equilibrium in the control population ($P < 10^{-6}$) were excluded. We removed 134 Alzheimer's disease case and 980 control samples because the individuals who had provided them were found to be first- or second-degree relatives of other study participants or were assessed as being of non-European descent based on genetic analysis using methods described[37]. This led us to retain 537,029 autosomal SNPs genotyped in 2,032 cases and 5,328 controls. Stage 2 genotyping was performed using Taqman (Applied Biosystems) or Sequenom assays. The primer and probe sequences for the genotyping assays are available upon request. To avoid any genotyping bias, cases and controls were randomly mixed when genotyping, and laboratory personnel were blinded to case or control status. The genotyping success rate was at least 95%, and no departure from Hardy-Weinberg equilibrium was observed for the markers included in the second stage.

Statistical analysis. We evaluated the case and control differences using logistic regression, which optionally incorporated principal components that were significantly associated with disease status to account for possible population stratification as described[37,38]. We hypothesized that the relatively high genomic control parameter found in the absence of the principal-components adjustment was due to differences in the representation of various French regions in the case and control series. Therefore, we further explored the robustness of our conclusions by incorporating 6,734 anonymized samples from France and other European countries as additional controls (unpublished data and ref. 37). With the inclusion of the additional samples, the genomic control parameter was 1.04 without principle-components adjustment and 1.03 after the adjustment. Inclusion of the additional controls did not substantially modify the association statistics for markers in regions showing suggestive evidence of association ($P < 10^{-5}$) after correction for population structure in the primary analysis (**Supplementary Tables 1** and **2**). Markers in these regions were analyzed at stage 2. We included imputed markers in these regions and selected markers for genotyping at stage 2 if the marker showed at least moderate association with disease status ($P < 10^{-4}$) (either with or without correction for population structure), the marker was not in strong LD with another marker entering stage 2, and a genotyping assay for the marker could be successfully designed.

We did statistical analyses under an additive genetic model using logistic regression taking age, sex and disease status into account and using SAS software version 9.1 (SAS Institute). Population controls that were not genotyped specifically for this study were excluded, as were any samples with missing age or gender data. This gave a maximum of 2,025 Alzheimer's disease cases and 5,328 controls in stage 1 and 3,359 cases and 2,633 controls in stage 2. Information on age and gender of the cases and controls included in these analyses are shown in **Supplementary Table 3**. Homogeneity of the odds ratios in different collections was tested using the Breslow-Day test[39]. We also used the Mantel-Haenszel method as implemented in Review Manager 5.0 to evaluate odds ratios across collections. Interactions between *CLU* or *CR1* SNPs and *APOE* ε4 polymorphism were tested in logistic regression models adjusted for age, gender and center where the samples were taken. The solid spine haplotype block definition in Haploview 4.0 was used to generate LD blocks of the genomic regions encompassing the *CLU* or *CR1* genes from imputed SNPs (MAF >5%)[40]. Associations of the *CLU* and *CR1* haplotypes were estimated using either logistic regression or a proportional hazards models using THESIAS 3.0, which implements a maximum likelihood model and uses an s.e.m. algorithm[41]. The population attributable risk (PAR) fraction was estimated using the following formula: $\mathrm{PAR} = F(\mathrm{OR} - 1)/(F(\mathrm{OR} - 1) + 1)$, where F is the frequency of the deleterious allele in the sample and OR is the odds ratio of Alzheimer's disease risk associated with the deleterious allele.

The criterion for genome-wide significance was defined in stage 1 with a conservative Bonferroni correction ($P < 9.3 \times 10^{-8}$; 0.05/537,029 SNPs tested). In stage 2, the level for significance was defined at $P < 4.5 \times 10^{-3}$ (0.05/11 SNPs tested).

URLs. Haploview: http://www.broad.mit.edu/mpg/haploview/index.php; Revman: http://www.cc-ims.net/revman/; THESIAS: http://ecgene.net/genecanvas/uploads/THESIAS3.1/; HapMap: http://www.hapmap.org; website for GWAS results: http://www.cng.fr/alzheimer/.

37. Heath, S.C. *et al.* Investigation of the fine structure of European populations with applications to disease association studies. *Eur. J. Hum. Genet.* **16**, 1413–1429 (2008).
38. Price, A.L. *et al.* Principal components analysis corrects for stratification in genome-wide association studies. *Nat. Genet.* **38**, 904–909 (2006).
39. Breslow, N.E., Day, N.E., Halvorsen, K.T., Prentice, R.L. & Sabai, C. Estimation of multiple relative risk functions in matched case-control studies. *Am. J. Epidemiol.* **108**, 299–307 (1978).
40. Barrett, J.C., Fry, B., Maller, J. & Daly, M.J. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* **21**, 263–265 (2005).
41. Tregouet, D.A. & Tiret, L. Cox proportional Hazards survival regression in haplotype-based association analysis using the stochastic-EM algorithm. *Eur. J. Hum. Genet.* **12**, 971–974 (2004).

doi:10.1038/ng.439