Neurobiology of Aging 34 (2013) 1309.e1–1309.e7

Contents lists available at SciVerse ScienceDirect

Neurobiology of Aging

journal homepage: www.elsevier.com/locate/neuaging

Discovery by the Epistasis Project of an epistatic interaction between the *GSTM3* gene and the *HHEX/IDE/KIF11* locus in the risk of Alzheimer's disease

James M. Bullock [a], Christopher Medway [a], Mario Cortina-Borja [b], James C. Turton [a], Jonathan A. Prince [c], Carla A. Ibrahim-Verbaas [d,e], Maaike Schuur [d,e], Monique M. Breteler [d], Cornelia M. van Duijn [d], Patrick G. Kehoe [f], Rachel Barber [f], Eliecer Coto [g], Victoria Alvarez [g], Panos Deloukas [h], Naomi Hammond [h], Onofre Combarros [i], Ignacio Mateo [i], Donald R. Warden [j], Michael G. Lehmann [j], Olivia Belbin [a,k], Kristelle Brown [a], Gordon K. Wilcock [l], Reinhard Heun [m,n], Heike Kölsch [m], A. David Smith [j], Donald J. Lehmann [j], Kevin Morgan [a,*]

[a] *Human Genetics, School of Molecular Medical Sciences, Queen's Medical Centre, University of Nottingham, Nottingham, UK*
[b] *Centre for Paediatric Epidemiology and Biostatistics, Institute of Child Health, University College London, London, UK*
[c] *Department of Medical Epidemiology and Biostatistics, Karolinska Institute, Stockholm, Sweden*
[d] *Department of Epidemiology, Erasmus MC University Medical Center, Rotterdam, The Netherlands*
[e] *Department of Psychiatry, Donders Institute for Brain, Cognition and Behavior, and Department of Human Genetics, Radboud University Nijmegen, Medical Centre, Nijmegen, The Netherlands*
[f] *Dementia Research Group, School of Clinical Sciences, University of Bristol, Frenchay Hospital, Frenchay Bristol, UK*
[g] *Genética Molecular, Hospital Central de Asturias, Oviedo, Spain*
[h] *Wellcome Trust Sanger Institute, Hinxton, Cambridge, UK*
[i] *Neurology Service and Centro de Investigación Biomédica en Red sobre Enfermedades Neurodegenerativas (CIBERNED), Marqués de Valdecilla University Hospital (University of Cantabria), Santander, Spain*
[j] *Oxford Project to Investigate Memory and Ageing (OPTIMA), University Department of Pharmacology, Oxford, UK*
[k] *Memory Unit, Alzheimer's Laboratory, Neurology Department, Hospital de la Santa Creu i Sant Pau, Sant Antoni Ma Claret, Barcelona, Spain*
[l] *Nuffield Department of Medicine, University of Oxford, John Radcliffe Hospital, Oxford, UK*
[m] *Department of Psychiatry, University of Bonn, Bonn, Germany*
[n] *Royal Derby Hospital, Derby, UK*

ARTICLE INFO

Article history:
Received 18 April 2012
Received in revised form 13 August 2012
Accepted 17 August 2012
Available online 1 October 2012

Keywords:
Epistasis
Alzheimer's risk
GSTM3
IDE locus

ABSTRACT

Despite recent discoveries in the genetics of sporadic Alzheimer's disease, there remains substantial "hidden heritability." It is thought that some of this missing heritability may be because of gene–gene, i.e., epistatic, interactions. We examined potential epistasis between 110 candidate polymorphisms in 1757 cases of Alzheimer's disease and 6294 control subjects of the Epistasis Project, divided between a discovery and a replication dataset. We found an epistatic interaction, between rs7483 in *GSTM3* and rs1111875 in the *HHEX/IDE/KIF11* gene cluster, with a closely similar, significant result in both datasets. The synergy factor (SF) in the combined dataset was 1.79, 95% confidence interval [CI], 1.35–2.36; $p = 0.00004$. Consistent interaction was also found in 7 out of the 8 additional subsets that we examined post hoc: i.e., it was shown in both North Europe and North Spain, in both men and women, in both those with and without the ε4 allele of apolipoprotein E, and in people older than 75 years (SF, 2.27; 95% CI, 1.60–3.20; $p < 0.00001$), but not in those younger than 75 years (SF, 1.06; 95% CI, 0.59–1.91; $p = 0.84$). The association with Alzheimer's disease was purely epistatic with neither polymorphism showing an independent effect: odds ratio, 1.0; $p \geq 0.7$. Indeed, each factor was associated with protection in the absence of the other factor, but with risk in its presence. In conclusion, this epistatic interaction showed a high degree of consistency when stratifying by sex, the ε4 allele of apolipoprotein E genotype, and geographic region.

1. Introduction

The etiology of sporadic, late-onset Alzheimer's disease (AD) is much more complex than that of the familial, early-onset condition,

* Corresponding author at: Human Genetics, School of Molecular Medical Sciences, Queen's Medical Centre, University of Nottingham, Nottingham NG7 2UH, UK. Tel.: +44(0)1 115 82 30724; fax: +44 115 970 9167.
E-mail address: kevin.morgan@nottingham.ac.uk (K. Morgan).

0197-4580/$ – see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.neurobiolaging.2012.08.010

which displays dominant Mendelian inheritance. The former depends on both genetic and environmental factors. Uncovering those factors is made difficult by the small effect size each exhibits. For many years, only the ε4 allele of apolipoprotein E (*APOEε4)* was known as a susceptibility allele for sporadic AD. Recently, other reproducible gene candidates, such as *PICALM, CLU, CR1*, and *BIN1* (Harold et al., 2009; Lambert et al., 2009; Seshadri et al., 2010), have been identified through genome-wide association studies (GWAS). However, their small effect sizes (odds ratios $\leq$1.5) mean that there still remains much heritability to uncover (Manolio et al., 2009). This is believed to be primarily because of the genetically complex and heterogeneous nature of the disorder, with interactions between multiple genetic mutations and polymorphisms, as well as between those and other, nongenetic, factors (Bertram and Tanzi, 2004).

The term, epistasis, was originally coined approximately 100 years ago by William Bateson to represent the masking of one allelic locus by another (Bateson, 1910). Although it has sometimes been used in a wider sense, we use the term here conventionally, i.e., when an increased risk is only seen in the presence of 2 genetic factors and not seen when they act apart. Such interactions may be one cause of the hidden heritability mentioned above. In such cases, studies that examine single loci individually, such as GWAS, will fail to detect an effect. Examples of epistasis in genetic studies on Alzheimer's disease have been reviewed by Combarros et al. (2009a).

Though GWAS have proven effective in detecting single-locus effects, such an unbiased approach might not be appropriate for the study of epistasis. A typical GWAS may examine perhaps 500,000 loci but the number of potential 2-way interactions between those 500,000 loci is more than 100 billion (10^{11}). In order therefore to reduce the number of potential interactions to a manageable figure, a hypothesis-driven approach might be required.

The approach we adopted is shown in Fig. 1 (Study design). We first carried out a systematic review of claims of epistasis in sporadic AD (Combarros et al., 2009a). From that investigation, we selected 31 genes, involved in 32 interactions, with biological plausibility and previous evidence of association with AD. We have previously replicated several of those interactions (Combarros et al., 2009b, 2010; Heun et al, 2012; Kölsch et al., 2012; Lehmann et al., 2012). In this study, we looked instead for potential binary interactions not previously examined. To do that, we used a discovery dataset of 1366 AD cases and 1184 controls and a replication set of 391 AD cases and 5111 controls, both drawn from the Epistasis Project (Fig. 1).

We discovered and replicated an interaction between two single nucleotide polymorphisms (SNPs), rs7483 and rs1111875. Rs7483 is in the *GSTM3* gene, encoding glutathione S-transferase μ3, involved in the detoxification of products of oxidative stress in the brain (Mannervik and Danielson, 1988). Rs1111875 is in the gene cluster of the hematopoietically expressed homeobox (*HHEX*), the insulin-degrading enzyme (*IDE*), and the kinesin family member 11 (*KIF11*). Associations with AD have previously been reported in this region (Carrasquillo et al., 2010).

2. Methods

2.1. Subjects

The Epistasis Project aims to study interactions between genetic loci that affect the risk of AD. It is a collaboration of 7 AD research groups: Bonn, Bristol, Nottingham, Oviedo, Oxford (OPTIMA), Rotterdam, and Santander. Sample characteristics by geographic region are given in Supplementary Table 1. All AD cases were diagnosed "definite" or "probable" by Consortium to Establish A Registry for Alzheimer's Disease (CERAD) (Mirra et al., 1993) or National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer's Disease and Related Disorders Association (NINCDS-ADRDA) (McKhann et al., 1984) criteria. AD cases were sporadic, i.e., possible autosomal dominant cases were excluded, based on family history. The median ages (interquartile ranges) of cases were 79.0 (73.0–85.2) and of controls were 76.9 (71.3–83.0). Full details of our sample sets and genotyping methods are given elsewhere (Combarros et al., 2009b). Rs7483 and rs1111875 were directly genotyped, not imputed. Research ethics approval was obtained by each of the participating groups (Supplementary Table 2). All participants of the study gave informed written consent.

2.2. Selection and screening of candidate interactions

Fig. 1 describes our study design in detail. Major features included: a systematic literature review of epistasis in sporadic AD

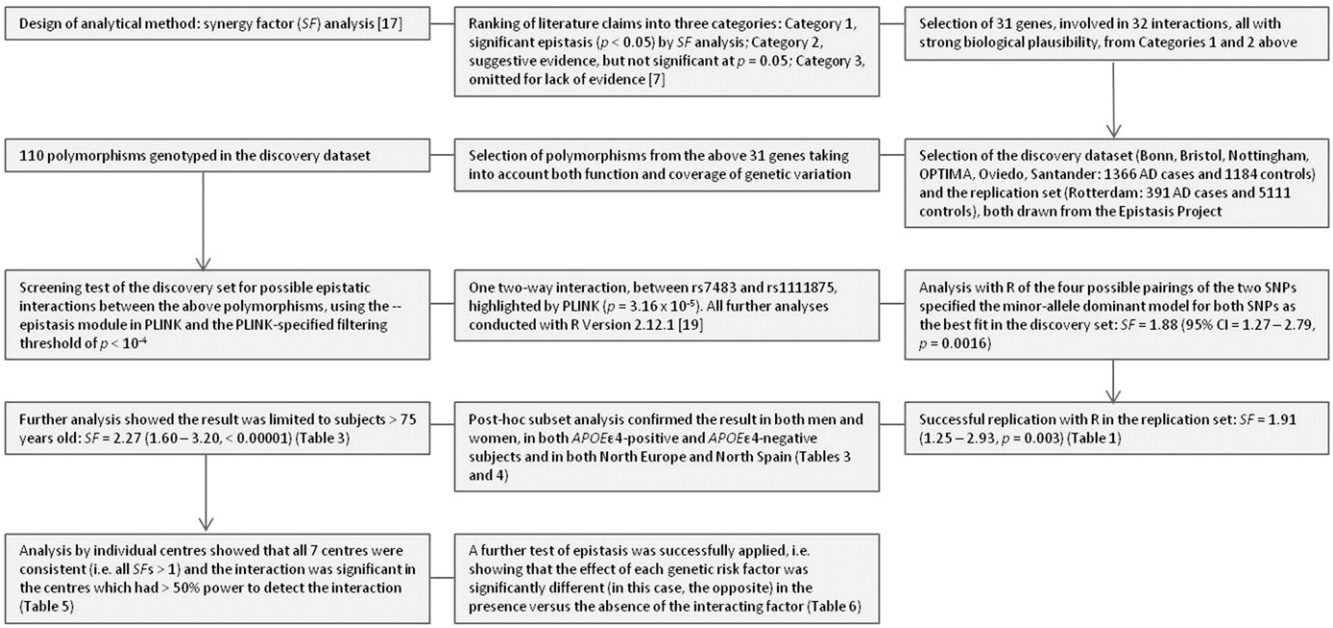

Fig. 1. Study design. Experimental design for selecting genes and single-nucleotide polymorhphisms (SNPs) for study in the Epistasis Project.

(Combarros et al., 2009a); the selection of 110 polymorphisms in 31 genes with previous evidence of interactions; the screening of potential interactions in a discovery dataset, followed by the replication of the highlighted interaction in a separate set, and further subset analysis to confirm the interaction (Fig. 1).

2.3. Statistical analysis

For screening the –epistasis module in the PLINK (version 1.07) whole-genome analysis toolkit was used, with the recommended filtering threshold of $p < 10^{-4}$ (Purcell et al., 2007). For all further analyses (Fig. 1), we used logistic regression models in R Version 2.12.1 (R Development Core Team, 2008), controlling for age, study center, sex, and *APOE*ε4 genotype in all analyses. The best fitting model was selected using Akaike's Information Criterion. Analyses were carried out first in the 6-center discovery set (Fig. 1), then in the Rotterdam replication set. We analyzed each center independently, as well as the 2 geographical regions of North Europe (Bonn, Bristol, Nottingham, Oxford, and Rotterdam) and North Spain (Oviedo and Santander). In addition we also examined the interaction (identified in our screen, see section 3.1.) when stratified by age, sex, or *APOE*ε4 genotype, and the 3-way interactions of these 2 SNPs with age, sex, and *APOE*ε4.

Heterogeneity among centers and overdispersion was controlled for as described elsewhere (Combarros et al., 2010). Epistasis was established using synergy factor (SF) analysis; this approach measures the size and significance of an interaction (Cortina-Borja et al., 2009). It is calculated as the ratio of the observed odds ratio for both factors combined to the predicted odds ratio assuming independent effects of each factor. A brief explanation of SF analysis is given in Box 1 of Combarros et al. (2009a); a full explanation is given in Cortina-Borja et al. (2009). Power calculations were based on the observed SF values, using the method described previously (Cortina-Borja et al., 2009). Allelic frequencies between North Spain and North Europe were compared using Fisher's exact test. All tests of significance and power calculations were 2-sided.

2.4. Exploring functional mutations

In an attempt to understand the functional basis of the interaction, we established the linkage disequilibrium block in which the SNPs were operating using SNAP Proxy (Johnson et al., 2008). SNPnexus and F-SNP were used to prioritize functional proxies within these loci (Chelala et al., 2009; Lee and Shatkay, 2008). We also used publically available expression data for quantitative trait loci to identify polymorphisms regulating the expression of interacting genes and FunCoup to explore any evidence of biological interaction (Alexeyenko and Sonnhammer, 2009; Dixon et al., 2007; Myers et al., 2007).

3. Results

3.1. Screening for interactions

A single 2-way interaction, between rs7483 in *GSTM3* and rs1111875 in the *HHEX/IDE/KIF11* gene cluster, was highlighted by the –epistasis module in PLINK among the 110 candidate polymorphisms in the 6-center discovery cohort as being below the filtering threshold, i.e., 1×10^{-4}; odds ratio, 0.67; $p = 3.16 \times 10^{-5}$. This threshold was used to identify potential interactions that possibly merit further study. The minor allele frequency of rs7483 (A allele) was 29% in North Europe and 26.5% in North Spain, while that of rs1111875 (A allele) was 41% in North Europe and 37% in North Spain (Supplementary Table 3). Genotype distributions in all 7 centers are given in Supplementary Table 4.

Analysis with R allowed for the specification of the interaction model used in the logistic regression calculations, as well as for controlling for covariates. Of the 4 combinations of minor or major allele-dominant pairings for the 2 SNPs, the best fit by Akaike's Information Criterion and the greatest significance was observed with a minor allele-dominant model for both SNPs (SF, 1.88; 95% confidence interval [CI], 1.27–2.79, $p = 0.0016$) in the discovery set (Table 1).

Calculating the interaction in the Rotterdam replication set produced consistent results, with the same model showing the best fit. This was further reinforced when all 7 centers were combined (Table 1). Examining the odds ratio of each factor separately showed that neither SNP was independently associated with AD, i.e., all odds ratios were very close to 1.0 (Table 2).

The 3-way interaction of the 2 SNPs with age as a continuous variable was significant ($p = 0.002$), as was that with age ±75 years ($p = 0.02$), but not that with age ±80 years ($p = 0.06$) (the median ages of cases and controls were 79.0 and 76.9 years, respectively). There were no 3-way interactions with sex or *APOE*ε4 ($p = 0.38$ and 0.86, respectively).

3.2. Post hoc stratification

Because the 2-way interaction between rs7483 in *GSTM3* and rs1111875 in the *HHEX/IDE/KIF11* gene cluster (Table 1) would not survive strict Bonferroni correction for all potential binary interactions ($n = 5995$), we conducted subset analysis to confirm the generality of the result. In the combined dataset (Table 3), significant results were found in men and women, and also in those with at least 1 copy of *APOE*ε4 (*APOE*ε4+) and those without (*APOE*ε4−). In contrast, stratification by age produced a clear difference: highly significant results were seen in people older than the age of 75 years ($p < 0.00001$), but not in those younger than 75 ($p = 0.84$) (Table 3). Further analysis was therefore restricted to people older than 75 years (Tables 4–6).

3.3. Consistent effects between regions and centers

The results from the 2 geographic regions, North Europe and North Spain, were consistent (Table 4), as were those of the 7 centers (Table 5). All 7 SFs pointed in the same direction (SF >1) including the 4 sets with low power (13%–24%) to replicate the interaction (Table 5). Additionally, the 2 sets with >50% power to detect the interaction reached significance (Rotterdam: SF, 2.03; 95% CI, 1.26–3.26; $p = 0.0036$; and Santander: SF, 3.58; 95% CI, 1.39–9.22; $p = 0.009$). There was heterogeneity in the SF values between centers which is unsurprising, given the low power (<25%, Table 5) in more than half of the centers. However, we controlled for this heterogeneity in all combined analyses.

3.4. The epistatic effect

The 2 interacting risk factors were the *GSTM3* rs7483 A allele and the *HHEX/IDE/KIF11* rs1111875 GG genotype. As a further test for

Table 1
Interaction in AD risk between rs7483 AA+AG versus GG and rs1111875 GG versus GA+AA in the hypothesis-generating and replication datasets

Dataset	Numbers		Power[a]	Adjusted[b] SF	95% CI	*p* value
	Control	AD				
Discovery	982	1102	93%	1.88	1.27–2.79	0.0016
Replication	4939	375	81%	1.91	1.25–2.93	0.003
Combined	5921	1477	99%	1.79	1.35–2.36	0.00004

Key: AD, Alzheimer's disease; *APOE*ε4, the ε4 allele of apolipoprotein E; CI, confidence interval; SF, synergy factor.
[a] Power to detect a synergy factor of 1.8 at $p < 0.05$.
[b] Controlling for age, study center, sex, and *APOE*ε4 genotype.

Table 2
Odds ratios of AD for rs7483 AA+AG versus GG and rs1111875 GG versus GA+AA

SNP	Dataset	Adjusted[a] odds ratio	95% CI	*p*-value
rs7483	Discovery	0.96	0.80–1.16	0.68
	Replication	0.97	0.80–1.19	0.80
	Combined	0.98	0.86–1.12	0.81
rs1111875	Discovery	1.01	0.83–1.23	0.92
	Replication	1.02	0.83–1.26	0.85
	Combined	1.00	0.87–1.14	0.97

Key: AD, Alzheimer's disease; *APOE*ε4, the ε4 allele of apolipoprotein E; CI, confidence interval; SNP, single nucleotide polymorphism.
[a] Controlling for age, study center, sex, and *APOE*ε4 genotype.

epistasis, we examined the effect of each factor in the presence versus the absence of the interacting factor. Each factor was associated with an increased risk in the presence of the other factor and was protective in its absence (Table 6). Results obtained in the discovery and replication sets (data not shown) were consistent with those in the combined dataset (Table 6).

4. Discussion

4.1. Interpretation of results

We have identified an interaction associated with AD risk, between rs7483 in *GSTM3* and rs1111875 in the *HHEX/IDE/KIF11* gene cluster, in our discovery and replication datasets (Table 1). Not only have we shown the interaction in our total dataset (1757 AD, 6292 controls), but to demonstrate the generality and robustness of the interaction, we also found it in post hoc analyses of an additional 7 out of 8 subsets. It was shown throughout the European populations we used, both in North Europe and in North Spain (Tables 4 and 5). It was seen in both men and women, and both in those with and without the *APOE*ε4 allele (Table 3). It appeared, however, to be limited to people older than the age of 75 (Tables 3 and 4), consistent with the significant 3-way interaction between the 2 SNPs and age (section 3.4.). This older subset had slightly greater power to detect the interaction (with an SF of 2.3): 99.8% power compared with 89% power in those younger than 75 years. Nevertheless, the marked contrast in results in those older than 75 years (SF, 2.27; 95% CI, 1.60–2.20; $p < 0.00001$), compared with those in individuals younger than 75 years (SF, 1.06; 0.59–1.91; $p = 0.84$) (Table 3), suggests that our findings might indeed be limited to older people. The effect was purely epistatic: neither SNP showed any main effect overall (Table 2); this result would not therefore have been detected by any GWAS reported to date. In epistasis, the effect of each genetic risk factor should be significantly different in the presence versus the absence of the interacting factor, as was the case here (Table 6).

Hong et al. (2009) had found an association with AD of *GSTM3* SNPs, particularly rs1799735, possibly interacting with *APOE*ε4.

Table 3
The interaction between rs7483 AA+AG versus GG and rs1111875 GG versus GA+AA in various subsets

Subset[a]	Adjusted[b] SF	95% CI	*p*-value
Men	1.64	1.01–2.68	0.046
Women	1.95	1.38–2.76	0.00016
*APOE*ε4+	1.63	1.01–2.64	0.046
*APOE*ε4–	1.79	1.27–2.51	0.0009
≥75 y	2.27	1.60–3.20	<0.00001
<75 y	1.06	0.59–1.91	0.84

Key: *APOE*ε4, the ε4 allele of apolipoprotein E; CI, confidence interval; SF, synergy factor.
[a] See Supplementary Table 5 for the numbers in each subset.
[b] Controlling for age, study center, sex, and *APOE*ε4 genotype.

Table 4
The interaction between rs7483 AA+AG vs GG and rs1111875 GG vs GA+AA in people ≥75 years of age

Dataset	Adjusted[a] SF	95% CI	*p*-value
Discovery	2.96	1.72–5.08	0.00009
Replication	2.03	1.26–3.26	0.0036
North Europe	2.19	1.48–3.26	0.0001
North Spain	2.80	1.22–6.41	0.015
Combined	2.27	1.60–3.20	<0.00001

'North Europe' comprises Bonn, Bristol, Nottingham, Oxford, and Rotterdam; 'North Spain' comprises Oviedo and Santander. See Supplementary Table 3 for the numbers from North Europe and North Spain.
Key: *APOE*ε4, the ε4 allele of apolipoprotein E; CI, confidence interval; SF, synergy factor.
[a] Controlling for age, study center, sex, and *APOE*ε4 genotype.

Rs1799735 is in weak linkage disequilibrium (LD) with rs7483 (D' = 1.00; $r^2 = 0.07$). In preliminary analyses in a subset of our data, our results were consistent with those of Hong et al. Thus, that association merits further study.

4.2. HHEX/IDE/KIF11 *rs1111875*

Rs1111875 is located on chromosome 10 in a 58 kb block of tight LD ($r^2 > 0.6$; chr10 coordinates 94,414,402–94,472,056) encompassing *HHEX*. Rs1111875 falls 7.5 kb downstream of *HHEX* in a conserved domain. A GWAS signal for type 2 diabetes has been tracked to 2 linked SNPs in this region (rs1111875 and rs5015480, $r^2 = 0.97$), although the functional mutation is yet to be described (Sladek et al., 2007; Zeggini et al., 2007). Of the 9 SNPs in high LD with rs1111875 ($r^2 > 0.8$), none fall within a gene or a predicted functional region.

Upstream of rs1111875, there is an extended LD block of 299 kb ($r^2 > 0.5$; chr10 coordinates 94,192,885–94,491,751) which contains 2 additional genes (*IDE* and *KIF11*) which might hold functional variants (Fig. 2). However, our attempts to prioritize candidate SNPs using functional prediction (F-SNP and SNPnexus) have offered little insight. Interestingly, rs1111875 is in modest LD ($r^2 = 0.47$) with the intronic *IDE* SNP, rs6583817, which has shown association with AD, increased *IDE* mRNA levels in AD cerebellum and reduced levels of serum β-amyloid 40. Luciferase reporter assays suggested the minor allele increases *IDE* expression in a neuroblastoma cell line (Carrasquillo et al., 2010).

In a further attempt to establish which gene within this locus tracks with rs1111875, we have also used expression data for quantitative trait loci from two previous studies. Unfortunately, one study did not measure the expression of *IDE*, *HHEX*, or *KIF11* (Myers et al., 2007). Although the data of the other study, Dixon et al.,

Table 5
The interaction between rs7483 AA+AG versus GG and rs1111875 GG versus GA+AA in people 75 years of age and older, by center

Centre	Numbers (≥75 y)		Adjusted[a] synergy factor (95% CI, *p*)	Power[b]
	Control	AD		
Bonn	55	84	1.92 (0.29–12.6, 0.50)	24%
Bristol	36	129	5.51 (0.82–36.8, 0.08)	15%
Nottingham	41	44	2.49 (0.27–23.1, 0.43)	13%
OPTIMA	140	148	2.81 (0.91–8.61, 0.07)	45%
Oviedo	27	142	1.05 (0.13–8.14, 0.96)	15%
Rotterdam	2917	351	2.03 (1.26–3.26, 0.004)	95%
Santander	262	166	3.58 (1.39–9.22, 0.009)	58%
Combined	3478	1064	2.27 (1.60–3.20, <0.00001)	99.8%

Key: AD, Alzheimer's disease; *APOE*ε4, the ε4 allele of apolipoprotein E; CI, confidence interval; OPTIMA, Oxford Project To Investigate Memory and Ageing.
[a] Controlling for age, sex, *APOE*ε4 genotype and, in the combined analysis, study center.
[b] Power to detect a synergy factor of 2.27 at $p < 0.05$.

Table 6
Odds ratios of AD for rs7483 AA+AG versus GG and rs1111875 GG versus GA+AA in the presence or absence of each other, in people 75 years of age and older

Odds ratio of AD for:	In the subset:	Numbers		Adjusted[a] odds ratio of AD (95% CI, *p*)
		Controls	AD	
rs7483 AG+AA versus GG	rs1111875 GG	AG+AA: 584 GG: 705	AG+AA: 222 GG: 175	1.70 (1.28–2.27, 0.0003)
	rs1111875 AG+AA	AG+AA:1164 GG: 1180	AG+AA: 315 GG: 429	0.74 (0.61–0.91, 0.004)
rs1111875 GG versus AG+AA	rs7483 AG+AA	GG: 584 AG+AA:1164	GG: 222 AG+AA: 315	1.49 (1.17–1.90, 0.001)
	rs7483 GG	GG: 705 AG+AA:1180	GG: 175 AG+AA: 429	0.66 (0.51–0.84, 0.001)

Key: AD, Alzheimer's disease; *APOEε4*, the ε4 allele of apolipoprotein E; CI, confidence interval.
[a] Controlling for age, study center, sex, and *APOEε4* genotype.

tagged the expression of *IDE*, *HHEX*, and *KIF11*, none of these were significantly associated with SNPs in this locus (Dixon et al., 2007). Attempts to look at functional interaction between *IDE/HHEX/KIF11* and known AD candidate genes (using FunCoup; Alexeyenko and Sonnhammer, 2009) also failed to elevate the candidacy of any 1 gene.

Though *IDE* might seem to be the obvious candidate because of its established role in the degradation of β-amyloid, the other 2 genes merit some attention. *HHEX* has been associated with poor forebrain neurodevelopment in *HHEX* (−/−) mutants (Martinez Barbera et al., 2000). A recent meta-analysis implicated rs1111875, which is 3′ of the *HHEX* gene, in the risk of type 2 diabetes, although the function of this SNP is yet to be assessed (Wang et al., 2011). Type 2 diabetes is itself a risk factor for AD. Members of the kinesin family (KIF) have been implicated in microtubule axonal transport of amyloid precursor protein (APP) from the latter's synthesis sites in the neuronal cell body to the terminus. This activity is afforded by APP binding to the light chain of the KIF protein. Mutant KIF proteins have been studied in mice and a reduced capacity for APP microtubule binding, and thus for transport, was identified (Feuk et al., 2005; Kamal et al., 2000).

Consequently, at present it is not possible to select any particular gene within this extended LD block as the definitive candidate, despite using a number of bioinformatic approaches; at the moment the evidence supports any or all of these genes being implicated when rs1111875 is used as a tag.

4.3. GSTM3 *rs7483*

Rs7483 is located in the *GSTM3* gene and affects transcription of that gene (Maes et al., 2010). The GSTM3 protein has long been shown to be related to detoxification of the effects of oxidative stress (Mannervik and Danielson, 1988), particularly in the brain (Aksenov, 2001). Detoxification of a variety of carcinogens takes place in the cytosol of cells via the activity of glutathione S-transferase (GST) enzymes. Polymorphisms in the genes for these enzymes have been linked to certain carcinomas in a variety of populations (Yengi et al., 1996). One report has identified the presence of GSTM1 on the inner membrane of mitochondria in humans (Gallagher et al., 2006). Based on this study, it has been suggested that GST enzymes might play a role in the detoxification of carcinogens and reactive oxygen species (ROS) in the cytosol and the mitochondria. Consequently, risk alleles at GST loci may also demonstrate altered function and therefore increase the risk of damage from ROS (Datta et al., 2007).

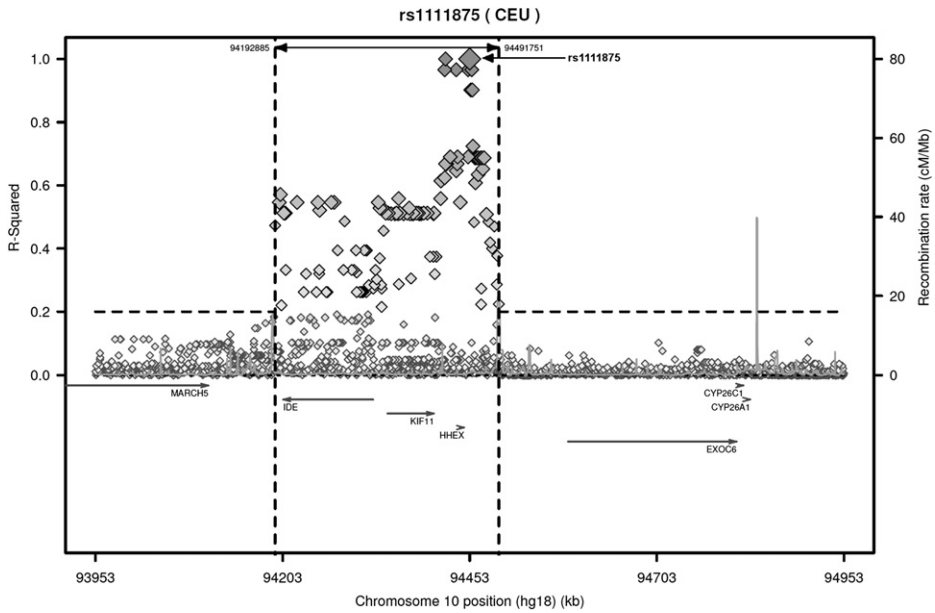

Fig. 2. Linkage disequilibrium (LD) architecture of the locus containing *HHEX*/*KIF11*/*IDE*. Rs1111875, located downstream of *HHEX*, is in high LD ($r^2 > 0.8$) with single-nucleotide polymorhphisms (SNPs) surrounding *HHEX*. Upstream of rs1111875 is a large block of modest linkage ($r^2 > 0.5$) encompassing *KIF11* and *IDE*, after which LD is broken down by high recombination (dashed vertical lines). Image generated with SNAP proxy server (Maes et al., 2010).

Additionally, increased oxidative stress has been shown to occur with the aging process, and to be particularly prominent in the proteins of AD patient brain tissues (Aksenov, 2001). This is consistent with our finding being limited to people older than the age of 75. The oxidative stress hypothesis for AD affords that axonal transport of mitochondria is functionally impaired, resulting in the production of ROS (Pappolla et al., 1992). Finally, it has been observed that GSTM3 is also expressed in distinct patterns in the brain tissue of AD patients, which is particularly noticeable in the plaques and tangles that are diagnostic of AD and in the activated microglia surrounding those plaques (Tchaikovskaya et al., 2005). This last observation suggests that inflammation might also be involved in this interaction.

5. Conclusions

We have confirmed an interaction between the *HHEX/IDE/KIF11* locus (rs1111875) and *GSTM3 (*rs7483) that is associated with AD risk through replication in 2 cohorts and through post hoc stratification by sex, *APOE*ε4, and geographic region. Post hoc analyses also suggested that the effect was mainly or solely in people older than 75 years. Although we were unable to pinpoint 1 gene in particular from the *HHEX/IDE/KIF11* locus as being tagged by rs1111875, it has been shown that IDE is inactivated when subjected to chemically induced oxidative stress (Shinall et al., 2005). Given the importance of GST enzymes in the prevention of oxidative stress in the brain, this suggests a potential functional interaction. However, we cannot rule out other possible mechanisms, e.g., inflammation.

Attempts at further replication of these results in European or other ethnic groups should be limited to large datasets, i.e., with at least 1000 cases and 1000 controls. Such datasets would have at least 91% power to replicate the above interaction, assuming similar allelic frequencies as in this study, and a synergy factor of 1.8, whereas datasets with only 200 cases and 200 controls, for instance, would have only 38% power.

Disclosure statement

The authors declare that there are no conflicts of interest.

Research ethics approval was obtained by each of the participating groups (Supplementary Table 2). All participants of the study gave informed written consent.

Acknowledgments

The authors thank the Moulton Charitable Foundation for a grant to fund the Epistasis Project, to all those who have provided support for the individual clinical studies, and to the Alzheimer's Research UK, and the Thomas Willis Oxford Brain Collection for tissue for DNA extraction. G.K.W. was partly funded by the NIHR Biomedical Research Centre Programme, Oxford. UCL Institute of Child Health receives funding from the Department of Health's NIHR Biomedical Research Centres funding scheme. The Centre for Paediatric Epidemiology and Biostatistics also benefits from funding support from the Medical Research Council in its capacity as the MRC Centre of Epidemiology for Child Health (G0400546). The Rotterdam Study is supported by Erasmus Medical Center and Erasmus University, Rotterdam, The Netherlands Organization for Scientific Research (NWO), Netherlands Organization for the Health Research and Development (ZonMw), the Research Institute for Diseases in the Elderly (RIDE), The Netherlands Genomics Initiative, the Ministry of Education, Culture and Science, the Ministry for Health, Welfare and Sports, the European Commission (DG XII), and the Municipality of Rotterdam. The generation and management of GWAS genotype data for the Rotterdam Study is supported by the Netherlands Organisation of Scientific Research NWO Investments (number 175.010.2005.011, 911-03-012). The contribution of inhabitants, general practitioners, and pharmacists of the Ommoord district to the Rotterdam Study is gratefully acknowledged. The group at Nottingham is funded by Alzheimer's Research UK and the Big Lottery fund. J.M.B. was funded by a Medical Molecular Sciences PhD studentship.

Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.neurobiolaging.2012.08.010.

References

Aksenov, M., 2001. Protein oxidation in the brain in Alzheimer's disease. Neuroscience 103, 373–383.

Alexeyenko, A., Sonnhammer, E.L.L., 2009. Global networks of functional coupling in eukaryotes from comprehensive data integration. Genome Res. 19, 1107–1116.

Bateson, W., 1910. Mendel's principles of heredity. Zeitschrift für Induktive Abstammungs- und Vererbungslehre 3, 108–109.

Bertram, L., Tanzi, R.E., 2004. The current status of Alzheimer's disease genetics: what do we tell the patients? Pharmacol. Res. 50, 385–396.

Carrasquillo, M.M., Belbin, O., Zou, F., Allen, M., Ertekin-Taner, N., Ansari, M., Wilcox, S.L., Kashino, M.R., Ma, L., Younkin, L.H., Younkin, S.G., Younkin, C.S., Dincman, T.A., Howard, M.E., Howell, C.C., Stanton, C.M., Watson, C.M., Crump, M., Vitart, V., Hayward, C., Hastie, N.D., Rudan, I., Campbell, H., Polasek, O., Brown, K., Passmore, P., Craig, D., McGuinness, B., Todd, S., Kehoe, P.G., Mann, D.M., Smith, A.D., Beaumont, H., Warden, D., Holmes, C., Heun, R., Kölsch, H., Kalsheker, N., Pankratz, V.S., Dickson, D.W., Graff-Radford, N.R., Petersen, R.C., Wright, A.F., Younkin, S.G., Morgan, K., 2010. Concordant association of insulin degrading enzyme gene (IDE) variants with IDE mRNA, Abeta, and Alzheimer's disease. PloS One 5, e8764.

Chelala, C., Khan, A., Lemoine, N.R., 2009. SNPnexus: a web database for functional annotation of newly discovered and public domain single nucleotide polymorphisms. Bioinformatics 25, 655–661.

Combarros, O., Cortina-Borja, M., Smith, A.D., Lehmann, D.J., 2009. Epistasis in sporadic Alzheimer's disease. Neurobiol. Aging 30, 1333–1349.

Combarros, O., van Duijn, C.M., Hammond, N., Belbin, O., Arias-Vásquez, A., Cortina-Borja, M., Lehmann, M.G., Aulchenko, Y.S., Schuur, M., Kölsch, H., Heun, R., Wilcock, G.K., Brown, K., Kehoe, P.G., Harrison, R., Coto, E., Alvarez, V., Deloukas, P., Mateo, I., Gwilliam, R., Morgan, K., Warden, D.R., Smith, A.D., Lehmann, D.J., 2009. Replication by the Epistasis Project of the interaction between the genes for IL-6 and IL-10 in the risk of Alzheimer's disease. J. Neuroinflammation 6, 22.

Combarros, O., Warden, D.R., Hammond, N., Cortina-Borja, M., Belbin, O., Lehmann, M.G., Wilcock, G.K., Brown, K., Kehoe, P.G., Barber, R., Coto, E., Alvarez, V., Deloukas, P., Gwilliam, R., Heun, R., Kölsch, H., Mateo, I., Oulhaj, A., Arias-Vásquez, A., Schuur, M., Aulchenko, Y.S., Ikram, M.A., Breteler, M.M., van Duijn, C.M., Morgan, K., Smith, A.D., Lehmann, D.J., 2010. The dopamine beta-hydroxylase -1021C/T polymorphism is associated with the risk of Alzheimer's disease in the Epistasis Project. BMC Med. Genet. 11, 162.

Cortina-Borja, M., Smith, A.D., Combarros, O., Lehmann, D.J., 2009. The synergy factor: a statistic to measure interactions in complex diseases. BMC Res. Notes 2, 105.

Datta, S., Majumder, M., Biswas, N., Sikdar, N., 2007. Increased risk of oral cancer in relation to common Indian mitochondrial polymorphisms and Autosomal GSTP1 locus. Cancer 110, 1991–1999.

Dixon, A.L., Liang, L., Moffatt, M.F., Chen, W., Heath, S., Wong, K.C., Taylor, J., Burnett, E., Gut, I., Farrall, M., Lathrop, G.M., Abecasis, G.R., Cookson, W.O., 2007. A genome-wide association study of global gene expression. Nat. Genet. 39, 1202–1207.

Feuk, L., McCarthy, S., Andersson, B., Prince, J.A., Brookes, A.J., 2005. Mutation screening of a haplotype block around the insulin degrading enzyme gene and association with Alzheimer's disease. Am. J. Med. Genet. B Neuropsychiatr. Genet. 136B, 69–71.

Gallagher, E.P., Gardner, J.L., Barber, D.S., 2006. Several glutathione S-transferase isozymes that protect against oxidative injury are expressed in human liver mitochondria. Biochem. Pharmacol. 71, 1619–1628.

Harold, D., Abraham, R., Hollingworth, P., Sims, R., Gerrish, A., Hamshere, M.L., Pahwa, J.S., Moskvina, V., Dowzell, K., Williams, A., Jones, N., Thomas, C., Stretton, A., Morgan, A.R., Lovestone, S., Powell, J., Proitsi, P., Lupton, M.K., Brayne, C., Rubinsztein, D.C., Gill, M., Lawlor, B., Lynch, A., Morgan, K., Brown, K.S., Passmore, P.A., Craig, D., McGuinness, B., Todd, S., Holmes, C., Mann, D., Smith, A.D., Love, S., Kehoe, P.G., 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, A.M., Kauwe, J.S., Cruchaga, C., Nowotny, P., Morris, J.C., Mayo, K., Sleegers, K., Bettens, K., Engelborghs, S., De Deyn, P.P., Van Broeckhoven, C., Livingston, G., Bass, N.J., Gurling, H., McQuillin, A., Gwilliam, R., Deloukas, P., Al-Chalabi, A., Shaw, C.E., Tsolaki, M., Singleton, A.B., Guerreiro, R.,

Mühleisen, T.W., Nöthen, M.M., Moebus, S., Jöckel, K.H., Klopp, N., Wichmann, H.E., Carrasquillo, M.M., Pankratz, V.S., Younkin, S.G., Holmans, P.A., O'Donovan, M., Owen, M.J., Williams, J., 2009. Genome-wide association study identifies variants at CLU and PICALM associated with Alzheimer's disease. Nat. Genet. 41, 1088–1093.

Heun, R., Kölsch, H., Ibrahim-Verbaas, C.A., Combarros, O., Aulchenko, Y.S., Breteler, M., Schuur, M., van Duijn, C.M., Hammond, N., Belbin, O., Cortina-Borja, M., Wilcock, G.K., Brown, K., Barber, R., Kehoe, P.G., Coto, E., Alvarez, V., Lehmann, M.G., Deloukas, P., Mateo, I., Morgan, K., Warden, D.R., Smith, A.D., Lehmann, D.J., 2012. Interactions between PPAR-α and inflammation-related cytokine genes on the development of Alzheimer's disease, observed by the Epistasis Project. Int. J. Mol. Epidemiol. Genet. 3, 39–47.

Hong, G.-S., Heun, R., Jessen, F., Popp, J., Hentschel, F., Kelemen, P., Schulz, A., Maier, W., Kölsch, H., 2009. Gene variations in GSTM3 are a risk factor for Alzheimer's disease. Neurobiol. Aging 30, 691–696.

Johnson, A.D., Handsaker, R.E., Pulit, S.L., Nizzari, M.M., O'Donnell, C.J., de Bakker, P.I., 2008. SNAP: a web-based tool for identification and annotation of proxy SNPs using HapMap. Bioinformatics 24, 2938–2939.

Kamal, A., Stokin, G.B., Yang, Z., Xia, C.H., Goldstein, L.S., 2000. Axonal transport of amyloid precursor protein is mediated by direct binding to the kinesin light chain subunit of kinesin-I. Neuron 28, 449–459.

Kölsch, H., Lehmann, D.J., Ibrahim-Verbaas, C.A., Combarros, O., van Duijn, C.M., Hammond, N., Belbin, O., Cortina-Borja, M., Lehmann, M.G., Aulchenko, Y.S., Schuur, M., Breteler, M., Wilcock, G.K., Brown, K., Kehoe, P.G., Barber, R., Coto, E., Alvarez, V., Deloukas, P., Mateo, I., Maier, W., Morgan, K., Warden, D.R., Smith, A.D., Heun, R., 2012. Interaction of insulin and PPAR-alpha genes in Alzheimer's disease: the Epistasis Project. J. Neural Transm. 119, 473–479.

Lambert, J.-C., Heath, S., Even, G., Campion, D., Sleegers, K., Hiltunen, M., Combarros, O., Zelenika, D., Bullido, M.J., 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, M.M., 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, J.F., Tzourio, C., Gut, I., Van Broeckhoven, C., Alpérovitch, A., Lathrop, M., Amouyel, P., 2009. Genome-wide association study identifies variants at CLU and CR1 associated with Alzheimer's disease. Nat. Genet. 41, 1094–1099.

Lee, P.H., Shatkay, H., 2008. F-SNP: computationally predicted functional SNPs for disease association studies. Nucleic Acids Res. 36, D820–D824.

Lehmann, D.J., Schuur, M., Warden, D.R., Hammond, N., Belbin, O., Kölsch, H., Lehmann, M.G., Wilcock, G.K., Brown, K., Kehoe, P.G., Morris, C.M., Barker, R., Coto, E., Alvarez, V., Deloukas, P., Mateo, I., Gwilliam, R., Combarros, O., Arias-Vásquez, A., Aulchenko, Y.S., Ikram, M.A., Breteler, M.M., van Duijn, C.M., Oulhaj, A., Heun, R., Cortina-Borja, M., Morgan, K., Robson, K., Smith, A.D., 2012. Transferrin and HFE genes interact in Alzheimer's disease risk: the Epistasis Project. Neurobiol. Aging 202, e1–e13.

Maes, O.C., Schipper, H.M., Chong, G., Chertkow, H.M., Wang, E., 2010. A GSTM3 polymorphism associated with an etiopathogenetic mechanism in Alzheimer disease. Neurobiol. Aging 31, 34–45.

Mannervik, B., Danielson, U.H.,, 1988. Glutathione transferases–structure and catalytic activity. CRC Crit. Rev. Biochem. 23, 283–337.

Manolio, T.A., Collins, F.S., Cox, N.J., Goldstein, D.B., Hindorff, L.A., Manolio, T.A., Collins, F.S., Cox, N.J., Goldstein, D.B., Hindorff, L.A., Hunter, D.J., McCarthy, M.I., Ramos, E.M., Cardon, L.R., Chakravarti, A., Cho, J.H., Guttmacher, A.E., Kong, A., Kruglyak, L., Mardis, E., Rotimi, C.N., Slatkin, M., Valle, D., Whittemore, A.S., Boehnke, M., Clark, A.G., Eichler, E.E., Gibson, G., Haines, J.L., Mackay, T.F., McCarroll, S.A., Visscher, P.M., 2009. Finding the missing heritability of complex diseases. Nature 461, 747–753.

Martinez Barbera, J.P., Clements, M., Thomas, P., Rodriguez, T., Meloy, D., Kioussis, D., Beddington, R.S., 2000. The homeobox gene Hex is required in definitive endodermal tissues for normal forebrain, liver and thyroid formation. Development 127, 2433–2445.

McKhann, G., Drachman, D., Folstein, M., Katzman, R., Price, D., Stadlan, E.M., 1984. Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. Neurology 34, 939–944.

Mirra, S.S., Hart, M.N., Terry, R.D., 1993. Making the diagnosis of Alzheimer's disease. A primer for practicing pathologists. Arch. Pathol. Lab. Med. 117, 132–144.

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

Pappolla, M.A., Omar, R.A., Kim, K.S., Robakis, N.K., 1992. Immunohistochemical evidence of oxidative [corrected] stress in Alzheimer's disease. Am. J. Pathol. 140, 621–628.

Purcell, S., Neale, B., Toddbrown, K., Thomas, L., Ferreira, M., Bender, D., Maller, J., Sklar, P., de Bakker, P.I., Daly, M.J., Sham, P.C., 2007. PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. Am. J. Hum. Genet. 81, 559–575.

R Development Core Team, 2008. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna Austria.

Seshadri, S., Fitzpatrick, A.L., Ikram, M.A., DeStefano, A.L., Gudnason, V., Boada, M., Bis, J.C., Smith, A.V., Carassquillo, M.M., Lambert, J.C., Harold, D., Schrijvers, E.M., Ramirez-Lorca, R., Debette, S., Longstreth Jr., W.T., Janssens, A.C., Pankratz, V.S., Dartigues, J.F., Hollingworth, P., Aspelund, T., Hernandez, I., Beiser, A., Kuller, L.H., Koudstaal, P.J., Dickson, D.W., Tzourio, C., Abraham, R., Antunez, C., Du, Y., Rotter, J.I., Aulchenko, Y.S., Harris, T.B., Petersen, R.C., Berr, C., Owen, M.J., Lopez-Arrieta, J., Varadarajan, B.N., Becker, J.T., Rivadeneira, F., Nalls, M.A., Graff-Radford, N.R., Campion, D., Auerbach, S., Rice, K., Hofman, A., Jonsson, P.V., Schmidt, H., Lathrop, M., Mosley, T.H., Au, R., Psaty, B.M., Uitterlinden, A.G., Farrer, L.A., Lumley, T., Ruiz, A., Williams, J., Amouyel, P., Younkin, S.G., Wolf, P.A., Launer, L.J., Lopez, O.L., van Duijn, C.M., Breteler, M.M., CHARGE Consortium; GERAD1 Consortium; EADI1 Consortium, 2010. Genome-wide analysis of genetic loci associated with Alzheimer disease. JAMA 303, 1832–1840.

Shinall, H., Song, E.S., Hersh, L.B., 2005. Susceptibility of amyloid beta peptide degrading enzymes to oxidative damage: a potential Alzheimer's disease spiral. Biochemistry 44, 15345–15350.

Sladek, R., Rocheleau, G., Rung, J., Dina, C., Shen, L., Serre, D., Boutin, P., Vincent, D., Belisle, A., Hadjadj, S., Balkau, B., Heude, B., Charpentier, G., Hudson, T.J., Montpetit, A., Pshezhetsky, A.V., Prentki, M., Posner, B.I., Balding, D.J., Meyre, D., Polychronakos, C., Froguel, P., 2007. A genome-wide association study identifies novel risk loci for type 2 diabetes. Nature 445, 881–885.

Tchaikovskaya, T., Fraifeld, V., Urphanishvili, T., Andorfer, J.H., Davies, P., Listowsky, I., 2005. Glutathione S-transferase hGSTM3 and ageing-associated neurodegeneration: relationship to Alzheimer's disease. Mech. Ageing Dev. 126, 309–315.

Wang, Y., Qiao, W., Zhao, X., Tao, M., 2011. Quantitative assessment of the influence of hematopoietically expressed homeobox variant (rs1111875) on type 2 diabetes risk. Mol. Genet. Metab. 102, 194–199.

Yengi, L., Inskip, A., Gilford, J., Alldersea, J., Bailey, L., Smith, A., Lear, J.T., Heagerty, A.H., Bowers, B., Hand, P., Hayes, J.D., Jones, P.W., Strange, R.C., Fryer, A.A., 1996. Polymorphism at the glutathione S-transferase locus GSTM3: interactions with cytochrome P450 and glutathione S-transferase genotypes as risk factors for multiple cutaneous basal cell carcinoma. Cancer Res. 56, 1974–1977.

Zeggini, E., Weedon, M.N., Lindgren, C.M., Frayling, T.M., Elliott, K.S., Lango, H., Timpson, N.J., Perry, J.R., Rayner, N.W., Freathy, R.M., Barrett, J.C., Shields, B., Morris, A.P., Ellard, S., Groves, C.J., Harries, L.W., Marchini, J.L., Owen, K.R., Knight, B., Cardon, L.R., Walker, M., Hitman, G.A., Morris, A.D., Doney, A.S., Wellcome Trust Case Control Consortium (WTCCC), McCarthy, M.I., Hattersley, A.T., 2007. Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes. Science 316, 1336–1341.