Journal of Human Genetics (2012) 57, 18–25

www.nature.com/jhg

ORIGINAL ARTICLE

Functional analysis of *APOE* locus genetic variation implicates regional enhancers in the regulation of both *TOMM40* and *APOE*

Lynn M Bekris[1,2], Franziska Lutz[1] and Chang-En Yu[1,2]

Genetic variation within the apolipoprotein E gene (*APOE*) locus is associated with late-onset Alzheimer's disease risk and quantitative traits as well as apoE expression in multiple tissues. The aim of this investigation was to explore the influence of *APOE* locus *cis*-regulatory element enhancer region genetic variation on regional gene promoter activity. Luciferase reporter constructs containing haplotypes of *APOE* locus gene promoters; *APOE*, *APOC1* and *TOMM40*, and regional putative enhancers; *TOMM40* intervening sequence (IVS)2-4, *TOMM40* IVS6 poly-T, as well as previously described enhancers; multienhancer 1 (ME1), or brain control region (BCR), were evaluated for their effects on luciferase activity in three human cell lines. Results of this investigation demonstrate that in SHSY5Y cells, the *APOE* promoter is significantly influenced by the *TOMM40* IVS2-4 and ME1, and the *TOMM40* promoter is significantly influenced by the *TOMM40* IVS6 poly-T, ME1 and BCR. In HepG2 cells, the *TOMM40* promoter is significantly influenced by all four enhancers, whereas the *APOE* promoter is not influenced by any of the enhancers. The main novel finding of this investigation was that multiple *APOE* locus *cis*-elements influence both *APOE* and *TOMM40* promoter activity according to haplotype and cell type, suggesting that a complex transcriptional regulatory structure modulates regional gene expression.
Journal of Human Genetics (2012) **57,** 18–25; doi:10.1038/jhg.2011.123; published online 17 November 2011

Keywords: *APOC1*; *APOE*; BCR; enhancer; luciferase reporter assay; ME1; regulation; *TOMM40*

INTRODUCTION

Multiple genetic variations within the apolipoprotein E gene (*APOE*) locus are strongly associated with late-onset Alzheimer's disease (LOAD) risk.[1] The ε4 allele of *APOE* is the strongest known genetic risk factor for LOAD.[2–4] However, inheritance of the *APOE* ε4 is neither necessary nor sufficient to cause the disease,[5,6] and the mechanism by which the ε4 allele functionally influences the risk and progression of LOAD remains unknown.

Other genetic elements in the *APOE* locus, such as *APOE* promoter polymorphisms, have been reported to be associated with LOAD, including single-nucleotide polymorphisms (SNPs) in position −491, −427, −219 (Th1/E47cs) and +113 as well as gene expression.[7–11] Thus, it has been postulated that expression levels of the *APOE* gene product can also contribute to LOAD risk. In addition to the *APOE* promoter, other *cis*-regulatory elements of *APOE* have been characterized, including multienhancer 1 (ME1), which influences *APOE* regulation in macrophages and adipocytes,[12,13] and a brain control region (BCR) that can modulate *APOE* expression in neurons and microglial.[14]

SNPs within the *TOMM40* gene have robustly shown association with LOAD in multiple genome-wide association studies.[15–18] The *TOMM40* SNP, rs2075650, located within the intervening sequence (IVS) 2, has been associated with both LOAD risk[17,19] and quantitative traits,[20–23] such as age-at-onset and Aβ42 levels, suggesting that the *TOMM40* region contributes to LOAD phenotypes. Our recent study has reported that IVS 2 SNPs within the *TOMM40* gene, as well as other SNPs both proximal and distal to *APOE,* are associated with cerebrospinal fluid apoE levels[24] and post-mortem brain apoE expression in Alzheimer's disease (AD) hippocampus.[25] Recently, it has been reported that a poly-T polymorphism (rs10524523) within IVS 6 of *TOMM40* is associated with LOAD age-at-onset.[26]

A report on *APOE* locus linkage disequilibrium (LD) patterns suggests that strong LD with *APOE* ε4 exists both proximal (in a region partially spanning the *TOMM40* gene) and distal to *APOE* (in the ME1 region).[4] Interestingly, the association between *TOMM40* and LOAD risk is not fully explained by LD between *TOMM40* SNPs and the *APOE* SNP (rs429358) that defines ε4 status, suggesting that other *APOE* locus SNPs contribute to this association with LOAD.[4]

[1]Geriatric Research, Education and Clinical Center, VA Puget Sound Health Care System, Seattle, WA, USA and [2]Department of Medicine, University of Washington, Seattle, WA, USA
Correspondence: Dr LM Bekris, Department of Medicine, University of Washington, Division of Gerontology and Geriatric Medicine, VA Puget Sound Health Care System, Box 358280 VAPSHCS GRECC S182, 1660 South Columbian Way, Seattle, WA 98108, USA.
E-mail: lbekris@u.washington.edu
Received 11 July 2011; revised 4 October 2011; accepted 6 October 2011; published online 17 November 2011

Given that multiple SNPs spanning a large region at the *APOE* locus are associated with LOAD risk, LOAD-specific quantitative traits, age-at-onset, LD and apoE expression levels, we hypothesized that a unique haplotype structure functionally influences the expression of multiple genes at the *APOE* locus. Thus, the aim of this investigation was to demonstrate that promoter activity of three *APOE* locus genes (*TOMM40, APOE* and *APOC1*) is influenced by previously described regulatory element enhancers (ME1, BCR) and putative regulatory element enhancers (*TOMM40* IVS2-4, *TOMM40* IVS6 Poly-T) differentially according to haplotype.

The main novel finding of this investigation was that genomic regions within *TOMM40* influence both *TOMM40* and *APOE*, but not *APOC1* promoter activity, suggesting that these two genes may be co-regulated. Furthermore, *APOE* locus enhancers influence both *TOMM40* and *APOE* promoters according to haplotype and cell type, implicating a complex system of gene regulation that is specific to genetic content and cellular microenvironment.

MATERIALS AND METHODS

APOE locus LD structure and *trans*-acting factor site prediction

The pattern of pairwise LD was measured for the *APOE* locus by D' metrics of chromosome 19: 50 080–50 150 kb. Haplotypes were constructed using the algorithm implemented in Haploview (version 4.2) for SNPs available with a minor allele frequency $\geqslant 0.01$ for the Haploview Caucasian of Northern and Western European descent sample (CEU)[27] (Figure 1).

Trans-acting factor sites were predicted for *APOE* locus promoters and enhancers using UCSC human genome browser ENCODE data for histone marker sites and DNase I hypersensitivity sites[28,29] (Figure 2).

Generation of regulatory haplotype reporters

Genomic DNA was obtained from the University of Washington (UW) Alzheimer's Disease Research Center after approval by the human subject Institutional Review Boards of UW and Veterans Affairs Puget Health Care System. Our regions of interest were PCR amplified and DNA sequenced from genomic DNA of 32 Caucasian subjects. See supplement for list of primer sequences. Two to three haplotypes for each regulatory genetic region of interest were chosen according to sequenced SNP content and frequency (Table 1). Variants of *TOMM40*, *APOE* or *APOC1* promoter core region haplotypes were inserted 5′ to the luciferase gene of the pGL4.10[luc2] vector (Promega, Madison, WI, USA) to produce promoter-only constructs. To produce promoter–enhancer constructs, haplotype variants of *TOMM40* IVS2-4, *TOMM40* IVS6 poly-T, ME1 or BCR genomic regions were inserted into the promoter constructs 3′ to the luciferase gene. The In-Fusion PCR Cloning System (Clontech, Mountain View, CA, USA) was used for all the cloning procedures. After propagating the recombinant DNA in *Escherichia coli* host cell, the reporter constructs were isolated and purified by ion-exchange column (Qiagen, Valencia, CA, USA). Inserts of all constructs were fully DNA sequenced to validate the correct genetic contents.

Cell culture

Human neuroblastoma SHSY5Y cells (ATCC, Manassas, VA, USA) were grown in 44.5% Dulbecco's modified Eagle's medium (Gibco, Grand Island, NY, USA) with 44.5% F12 (Gibco), 10% fetal bovine serum (Gibco) and 1% penicillin/streptomycin (100 μU ml^{-1} penicillin and 100 μg ml^{-1} streptomycin) at 37 °C in a 5% CO_2 atmosphere. Human hepatocytoma HepG2 cells (ATCC) and U118 astrocytoma cells (ATCC) were grown in 89% Dulbecco's modified Eagle's medium (Gibco), 10% fetal bovine serum (Gibco) and 1% penicillin/streptomycin (100 μU ml^{-1} penicillin and 100 μg ml^{-1} streptomycin) at 37 °C in a 5% CO_2 atmosphere. Cells were passaged at 5×10^4 per well into 96-well tissue culture plates 48 h before transfection.

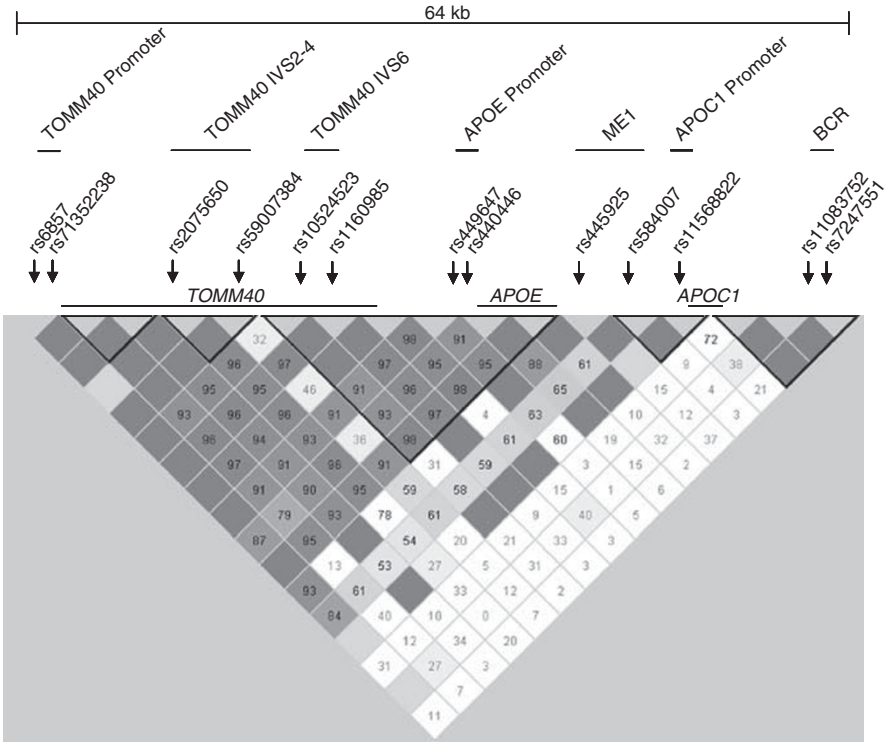


Figure 1 *APOE* locus linkage disequilibrium plot. *APOE* locus linkage disequilibrium (LD) plot demonstrates the strong LD present between *TOMM40* and *APOE* as well as ME1 and *APOC1* but not BCR. Dark gray squares represent strong LD calculated using D′ of Caucasian of Northern and Western European decent (CEU: Haploview: http://www.broadinstitute.org/haploview) where higher numbers represent stronger LD. Dark gray squares without numbers represent a D′ of 100. Bolded regions represent strong haplotype blocks. SNPs were chosen from CEU *APOE* locus available SNPs in Haploview. *APOE* locus size, *APOE* locus genes and haplotype regions inserted into luciferase reporter constructs are noted.

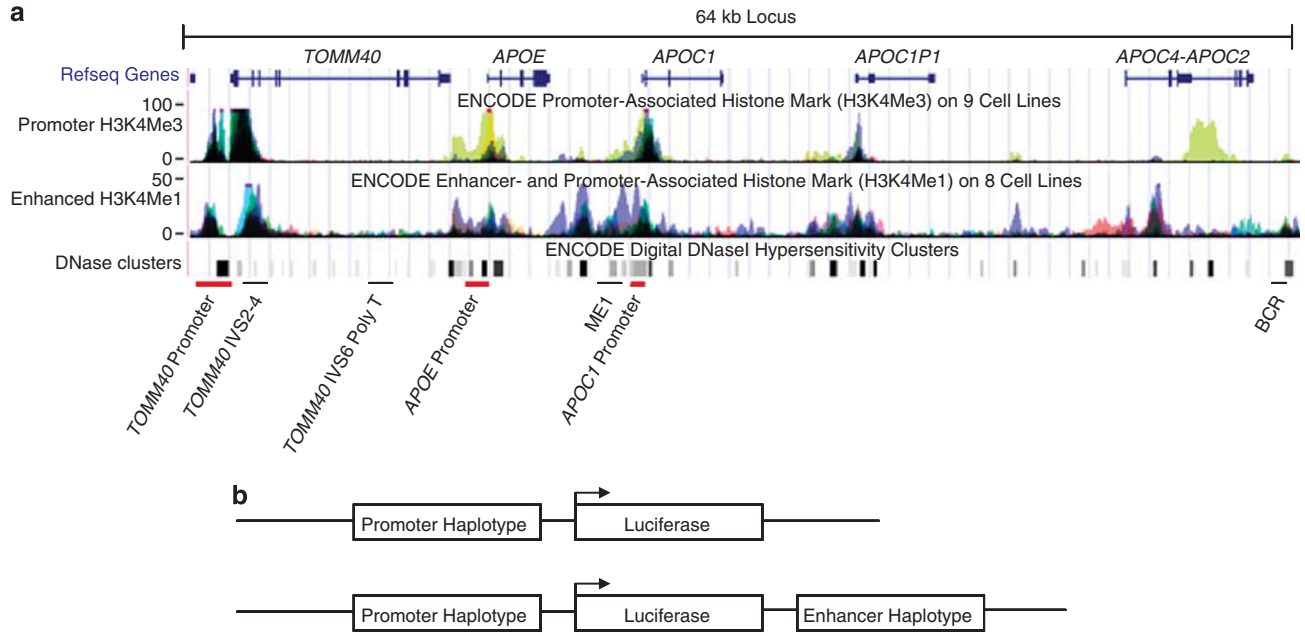


Figure 2 *APOE* locus putative promoter–enhancer regions. Regional genes include:*TOMM40*, *APOE*, *APOC1 APOC4*, and *APOC2*. *TOMM40* promoter, *APOE* promoter, *APOC1* promoter haplotype regions are shown as red bars. Putative enhancer haplotype regions; *TOMM40* IVS2-4, *TOMM40* IVS6 Poly-T and previously described enhancer haplotype regions; ME1 and BCR are shown as black bars. The *APOE* locus containing regulatory haplotype regions tested is located at chr19:50,083,898-50,148,827 as indicated on the UCSC Genome Browser (created by the Genome Bioinformatics Group of UC Santa Cruz, assembly March 2006 (NCBI 36/hg18): http://genome.ucsc.edu/index.html. ENCODE promoter (H3K4Me3) and enhancer (H3K4Me1) associated histone marks from multiple cell lines (H3K4Me3) overlap with regulatory haplotype reporter construct sites. ENCODE DNase1 hypersensitivity cluster results overlap with regulatory haplotype region reporter construct sites (**a**). Regulatory haplotype luciferase reporter construct maps include promoter haplotype only constructs and promoter–enhancer haplotype construct map (**b**).

Luciferase reporter construct transfection and assay

SHSY5Y and HepG2 cells were transiently transfected for 48 h with promoter–enhancer reporter constructs using Lipofectamine 2000 (Invitrogen, Carlsbad, CA, USA), according to manufacturer instructions. All transfection assays included a *TOMM40*, *APOE* and *APOC1* promoter-only (no enhancer) luciferase pGL4.10[luc2] haplotype reporter constructs as basal expression controls and promoter-less luciferase pGL4.10[luc2] as a negative control. In addition, the renilla pGL4.75[hRluc/CMV] was co-transfected as an internal control. All transfection assays included duplicate construct transfections, and transfection assays were performed in triplicate. Failed transfection reactions were duplicated, leading to 6–8 replicate transfections for each haplotype reporter construct.

SHSY5Y, HepG2 and U118 transiently transfected cells were harvested after 48 h using the Dual Glo Luciferase Assay System (Promega), which allows for high throughput analysis of firefly luciferase (pGL4.10[luc2]) constructs in the first step. In the second step, after quenching the firefly luminescence and activating the renilla luciferase (pGL4.75[hRluc/CMV]), luminescence from the internal control was analyzed. Luciferase luminescent counts per second (CPS) were measured using a Wallac Victor2 1420 Multilabel Counter (Wallac, Waltham, MA, USA).

Relative quantitation of haplotypes expression levels and statistical analysis

To measure the enhancer effect on promoter activity, luciferase activities of promoter-only constructs served as baseline controls (set at 1.0) and were compared with promoter–enhancer constructs. First, a ratio of F/R CPS was calculated as firefly (F) luciferase construct CPS per renilla (R) luciferase internal control CPS for each transfected construct, per manufacturers instructions (Dual Glo Luciferase Assay System, Promega). Haplotype expression levels relative to promoter was calculated for each promoter–enhancer reporter construct (that is, F/R *APOE*-promoter haplotype 1-ME1 haplotype 2/F/R *APOE*-promoter haplotype 1). This data were then analyzed using an analysis of variance to compare haplotypes (SPSS Version 13, Armonk, NY, USA). All statistical tests within specific promoter–enhancer haplotype groups were computed by Bonferroni correction for multiple comparisons.

RESULTS

APOE locus genomic signature

To evaluate the LD structure of the 64-kb *APOE* locus, SNPs spanning from the *TOMM40* 5′ region to the BCR were analyzed using Haploview.[27] The promoter and enhancer regions of interest in this investigation,: *TOMM40*, *APOE* and *APOC1* promoter regions; ME1and BCR regulatory element enhancers; *TOMM40* IVS2-4, *TOMM40* IVS6 Poly-T putative regulatory element enhancers, are either in strong or moderate LD with each other or in the case of the BCR, are not in LD with other selected genetic elements (Figure 1). Except for the *TOMM40* IVS6 poly-T, all the promoter and enhancer regions of interest are consistently overlapped with the ENCODE tracks in the UCSC human genome browser.[28,29] These ENCODE tracks include DNase I hypersensitivity sites, histone mark H3K4Me3 sites for promoter and H3K4Me1 sites for enhancers (Figure 2a).

Regulatory haplotype reporter constructs

Putative regulatory haplotype content was identified by SNP genotyping of human subjects and amplified from subject genomic DNA. Putative regulatory haplotype variants were then verified by sequencing, and cloned into the luciferase reporter constructs as shown in Figure 2b. Thus, a large panel of luciferase reporter clones were generated that included eight promoter-only constructs (three promoters with a total of eight variants, Table 1) and 80 promoter–enhancer haplotype constructs (8 promoter haplotypes plus 10 enhancer haplotypes; Table 1). Luciferase gene activities of reporter

Table 1 ***APOE locus haplotype description***

SNP	*Location*	*Alleles*	*MAF*			
				Promoter haplotypes		
TOMM40 promoter				1	2	3
rs6857	17660472	G/a	0.109	G	A	G
rs71352238	17662554	T/c	0.087	C	C	T
APOE promoter				1	2	3
rs449647 (−491)	17676782	T/a	0.146	A	T	A
rs769446 (−427)	17676846	T/c	0.038	C	T	T
rs405509 (−219)	17677054	C/t	0.492	T	C	C
rs440446 (113)	17677385	G/c	0.380	C	G	G
APOC1 promoter				1	2	
rs11568822 *Hpa*I site	17685858	CGTT	0.166	−CGTT	+CGTT	
				Enhancer haplotypes		
TOMM40 IVS2-4				1	2	
rs2075650	17663837	A/g	0.130	G	A	
rs157581	17663932	A/g	0.285	G	A	
rs34095326	17664062	G/a	0.039	A	G	
rs34404554	17664127	C/g	0.114	G	C	
rs11556505	17664362	C/T	0.123	T	C	
rs157582	17664437	G/a	0.276	A	G	
rs59007384	17664883	G/t	0.223	T	G	
TOMM40 IVS6				1	2	3
rs10524523	17671267	Poly-T 16−35	—	16T	24T	35T
rs417357	17671337	C/t	0.061	C	C	C
S_17671434 novel	17671434	C/t	—	C	T	C
rs1160985	17671630	C/t	0.460	T	C	C
ME1				1	2	3
rs445925	17683858	G/a	0.138	G	G	G
rs10414043	17683931	G/a	0.131	A	G	G
rs7256200	17684153	G/t	0.131	T	G	G
rs483082	17684396	G/t	0.258	T	G	G
rs59325138	17684509	C/t	0.304	C	T	C
rs584007	17684696	A/g	0.387	G	G	A
BCR				1	2	
rs11083752	17722581	G/a	0.448	G	A	
rs7248162	17722904	T/c	0.451	T	C	
rs7247227	17722977	A/g	0.460	A	G	
rs7247551	17722984	G/a	0.365	G	A	

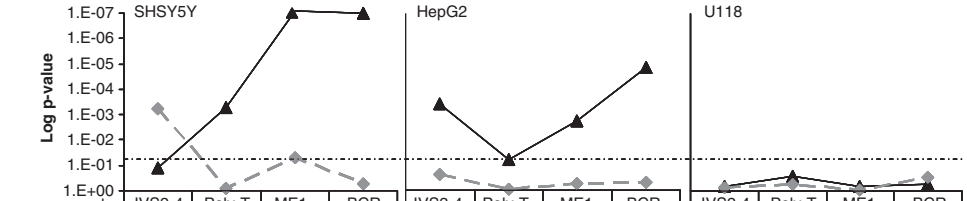
Abbreviations: BCR, brain control region; MAF, minor allele frequency; ME1, multienhancer 1; SNP, single-nucleotide polymorphism.
Allele composition of each haplotype genomic DNA fragment inserted into luciferase reporter constructs. Location in contig: NT_0.11109.16.

constructs were measured in the three human cell lines that included a neuronal cell line (SHSY5Y), a hepatocyte cell line (HepG2) and an astrocyte cell line (U118). No statistically significant differences were observed among *APOC1* promoter–enhancer haplotypes (data not shown). In contrast, statistically significant differences in luciferase gene expression were observed for *TOMM40* and *APOE* promoter–enhancer haplotypes in both SHSY5Y and HepG2 cells, but not U118 cells (Figure 3).

Haplotype specific expression level in SHSY5Y cells

Significant differences in luciferase gene expression were found among two *APOE* promoter–enhancer haplotypes and three *TOMM40* promoter–enhancer haplotypes in the SHSY5Y human neuronal cell line (Figures 3 and 4).

A total of six *APOE* promoter-*TOMM40* IVS2-4 haplotype variants were tested and compared with their promoter-only counterparts. Significant differences in luciferase gene expression were found



p-value	SHSY5Y IVS2-4	SHSY5Y Poly-T	SHSY5Y ME1	SHSY5Y BCR	HepG2 IVS2-4	HepG2 Poly-T	HepG2 ME1	HepG2 BCR	U118 IVS2-4	U118 Poly-T	U118 ME1	U118 BCR
TOMM40	0.130	0.001	0.000	0.000	0.000	0.060	0.002	0.000	0.692	0.268	0.653	0.557
APOE	0.001	0.856	0.048	0.559	0.250	0.894	0.537	0.483	0.722	0.520	0.927	0.281

Figure 3 Cell type (SHSY5Y, HepG2, U118) specific expression of *TOMM40* and *APOE* regulatory haplotypes. *ANOVA P*-values (includes Bonferroni multiple comparison correction) representing the difference between haplotypes are presented for transfection replicates of at least $n=6$. The hatched line represents significant *P*-value cut-off where all values above the line are significant ($P<0.05$). No significant differences in expression were found among *APOC1* promoter–enhancer haplotypes (data not shown).

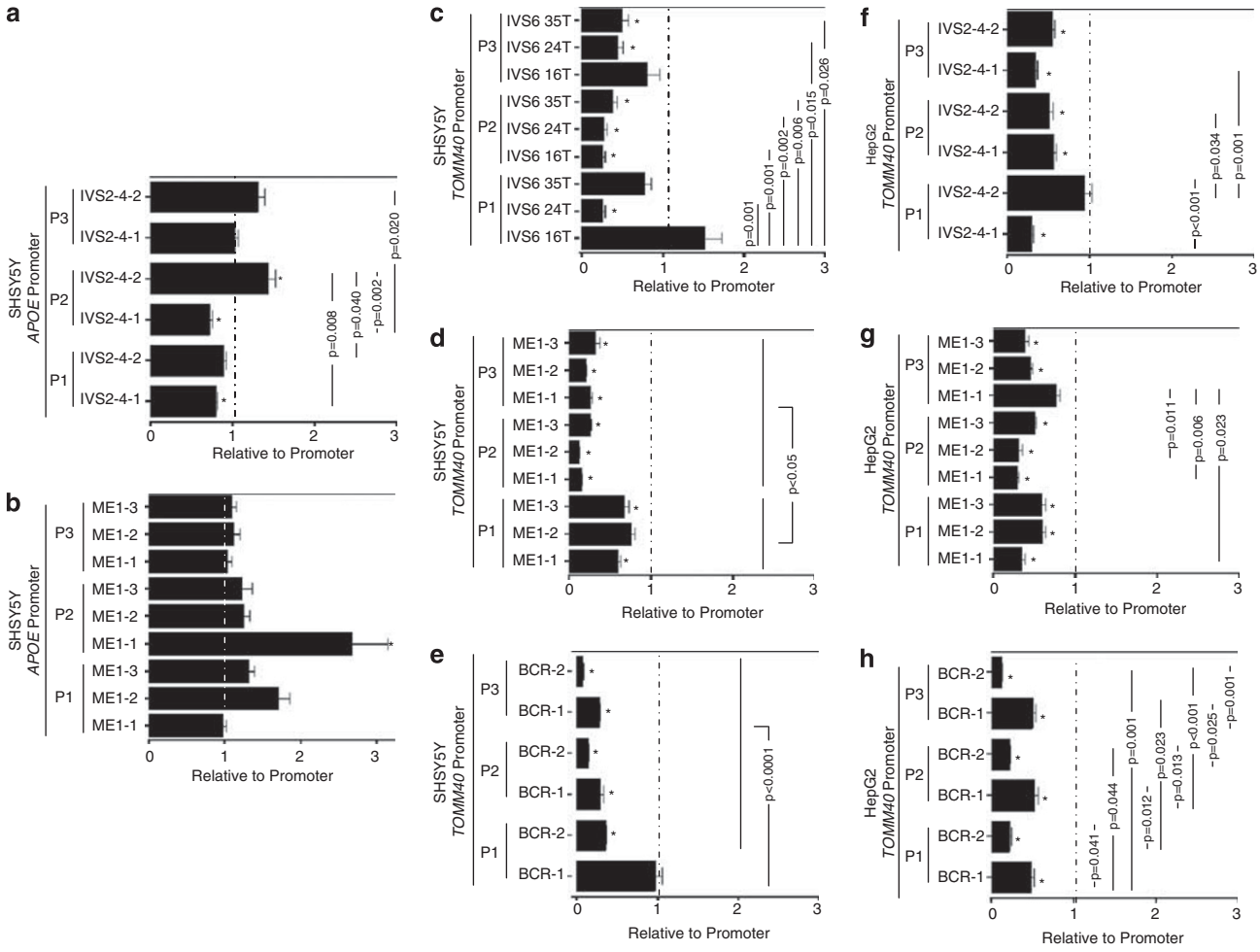


Figure 4 Effect of *APOE* and *TOMM40* Promoter Haplotype on Promoter–Enhancer Regulatory Haplotype Expression. *P*-values represent a significant difference between haplotypes. *APOE* promoter–enhancer haplotypes that showed a significant difference between haplotypes in SHSY5Y cells are *APOE* promoter–*TOMM40* IVS2-4 (**a**) and *APOE* promoter–ME1 (**b**). *TOMM40* promoter–enhancer haplotypes that showed a significant difference between haplotypes in SHSY5Y cells are *TOMM40* promoter–*TOMM40* IVS6 Poly T (**c**), *TOMM40* promoter–ME1 (**d**) and *TOMM40* promoter–BCR (**e**). *TOMM40* promoter–enhancer haplotypes that showed a significant difference between haplotypes in HepG2 cells are *TOMM40* promoter–*TOMM40* IVS2-4 (**f**), *TOMM40* promoter–ME1 (**g**) and *TOMM40* promoter–BCR (**h**). Each bar represents promoter-enhancer levels relative to promoter levels only (for transfection replicates of at least $n=6$). *P*-values represent significant differences between haplotypes and are Bonferroni multiple comparison corrected. Promoter activity was set at 1 (dotted line). Asterisks represent haplotypes with significantly different expression levels relative to the promoter haplotype alone.

between four haplotype pairs (Figure 4a; significant levels relative to 1.0 as noted with asterisk). These differences were not observed for similar *TOMM40* promoter haplotypes expressed in SHSY5Y cells, indicating that an enhancer/silencer effect of this *TOMM40* IVS2-4 region is specific to the *APOE* promoter in SHSY5Y cells (Figure 4).

A total of nine *APOE* promoter-ME1 haplotypes were tested. Expression levels of *APOE* promoter-ME1 haplotypes were higher than the *APOE* promoter-only construct, where one haplotype (*APOE* promoter 2-ME1-1) showed significantly higher expression (Figure 4b; significant levels above 1.0 noted with asterisk). We also tested nine *TOMM40* promoter-ME1 haplotype variants. Significant differences in luciferase gene expression were observed in this setting where the *TOMM40* promoter 1-related haplotypes have higher expression levels than the other haplotypes (Figure 4d). All *TOMM40* promoter-ME1 haplotypes levels were lower than the *TOMM40* promoter-only construct (levels below 1.0; significantly lower levels are noted with asterisks). These results confirm the previously described enhancer activity of ME1 on *APOE* promoter activity[12,13] and suggest an opposite silencer effect of ME1 on *TOMM40* promoter activity.

A total of nine *TOMM40* promoter-IVS6 poly-T haplotype variants were tested. The expression level of *TOMM40* promoter 1-IVS6 16T haplotype was significantly higher than other haplotypes (Figure 4c). Six *TOMM40* IVS6 poly-T haplotypes had significantly lower expression levels compared with the *TOMM40* promoter-only construct (asterisk). These results suggest that the *TOMM40* IVS6 poly-T (or rs10524523) locus has measurable enhancer/silencer activity and that the direction of its effect depends on the specific haplotype content of the *TOMM40* promoter (Figure 4c).

A total of six *TOMM40* promoter-BCR enhancer haplotype variants were tested. Significant differences were found between the *TOMM40* promoter 1-BCR-1 haplotype and the other haplotypes (Figure 4e). The majority of *TOMM40* promoter-BCR haplotypes showed significantly lower expression levels compared with the *TOMM40* promoter-only construct (significant levels below one noted with asterisk), suggesting a silencer effect of BCR on the *TOMM40* promoter in SHSY5Y cells (Figure 4e).

Haplotype specific expression level in HepG2 cells

Expression levels of the same eight promoter-only constructs and 80 promoter–enhancer haplotype constructs were also tested in the HepG2 human neuronal cell line (Table 1). Significant differences were found among three separate *TOMM40* promoter–enhancer haplotypes, differences that were not observed among *APOE* promoter–enhancer haplotypes (Figures 3 and 4).

Among the six haplotype variants of *TOMM40* promoter-*TOMM40* IVS2-4, significant expression differences in HepG2 cells were found between the *TOMM40* promoter 1-*TOMM40* IVS2-4-2 haplotype and two other *TOMM40* promoter-*TOMM40* IVS2-4 haplotypes. All *TOMM40* promoter-*TOMM40* IVS2-4 haplotypes, except *TOMM40* promoter 1-*TOMM40* IVS2-4-2, have significantly lower expression levels relative to*TOMM40* promoter-only construct (asterisk), suggesting that *TOMM40* IVS2-4 has a silencer effect in HepG2 cells (Figure 4f).

Significant differences in expression levels in HepG2 cells were also found between *TOMM40* promoter-ME1 haplotype variants. All *TOMM40* promoter 1-ME1 haplotype levels were lower than the *TOMM40* promoter-only counterpart in HepG2 cells (asterisks), suggesting a silencer effect that is significantly lower for the *TOMM40* promoter 1-ME1-1 and *TOMM40* promoter 2-ME1-1 haplotypes compared with the *TOMM40* promoter 3-ME1-1 haplotype (Figure 4g).

Significant differences in expression levels in HepG2 cells were found between multiple*TOMM40* promoter-BCR haplotypes (Figure 4h). All *TOMM40* promoter-BCR haplotypes have significantly lower expression compared with the *TOMM40* promoter-only construct in HepG2 cells (asterisk), suggesting a silencer effect HepG2 cells that is significantly lower for the BCR-2 haplotype compared with the BCR-1 haplotype (Figure 4h).

DISCUSSION

The core region of *APOE* locus consists of a large haplotype block that contains several genes, including *TOMM40, APOE* and *APOC1* (Figure 1).[4] Genetic markers in all three genes have consistently been shown to be associated with LOAD in multiple studies.[2–4] Current consent is that the ε4 allele of *APOE* is the only true effector of LOAD, and positive association signals of other markers are merely the effect of LD with the ε4 allele. The*TOMM40* gene contains a poly-T repeat within intron 6 (*TOMM40* IVS6 poly-T rs10524523) that has been recently reported to be associated with AD age-at-onset,[26] and SNPs in the *TOMM40* intron 2–4 region (*TOMM40* IVS2-4) are associated with LOAD risk, quantitative trait loci,[20–23] and apoE expression levels.[24,25] In addition, the association between *TOMM40* SNPs and LOAD cannot be fully explained by LD between *TOMM40* SNPs and *APOE* ε4 alone[4] and *APOE* locus enhancers, ME1 and BCR are associated with expression levels.[12–14] However, the functional influence of genetic variation within *APOE* locus enhancers has not been previously characterized. In this investigation, we hypothesized that promoter activity of three *APOE* locus genes (*TOMM40, APOE* and *APOC1*) are functionally influenced by genetic variation within previously described regulatory elements (ME1, BCR) and putative regulatory elements (*TOMM40* IVS2-4, *TOMM40* IVS6 Poly-T) differentially according to haplotype. The main novel finding was that enhancers within *TOMM40* influence both *TOMM40* and *APOE* promoter activity in a haplotype and cell-type-specific manner.

Significant differences in reporter assay expression levels were found in three *TOMM40* promoter–enhancer and two *APOE* promoter–enhancer haplotype variants in SHSY5Y neuronal cells (Figures 3 and 4). Significant differences were also found in three different *TOMM40* promoter–enhancer haplotypes in HepG2 hepatocyte cell line (Figures 3 and 4). However, no significant differences were found in all *APOC1* promoter–enhancer haplotypes and all promoter–enhancer haplotypes transfected into U118 astrocyte cells (Figure 3). Specifically, the *TOMM40* promoter 1-*TOMM40* IVS6 16T haplotype showed increased expression in the SHSY5Y neuronal cell line compared with six other haplotype variant counterparts (Figure 4c). Interestingly, six *TOMM40* IVS6 poly-T haplotypes (asterisks) had significantly lower expression levels compared with the *TOMM40* promoter haplotype alone (levels below one) implicating *TOMM40* promoter activity inhibition by all *TOMM40* IVS6 poly T haplotypes except the *TOMM40* promoter 1-*TOMM40* IVS6 poly-T 16T haplotype (Figure 4c). There was no significant difference between *APOE* promoter-*TOMM40* IVS6 poly-T or *APOC1* promoter-*TOMM40* IVS6 poly-T haplotypes in any of the cell types tested, which suggests that the*TOMM40* IVS6 poly-T 16T haplotype impacts *TOMM40* expression, but not *APOE* or *APOC1* expression. However, as others have reported promoter activity outside of the *APOE* core promoter sequence used in the present study,[8,30] the possibility remains that the *TOMM40* IVS6 poly-T haplotype influences the *APOE* promoter outside of the core promoter region tested in this study.

The *TOMM40* IVS6 poly-T haplotype-specific expression was only significant in the neuronal SHSY5Y cell line, not HepG2 or U118 cell lines, suggesting that this haplotype may be particularly important in neurons. Interestingly, the *TOMM40* IVS6 poly-T 16T haplotype contains a short version of poly-T repeat that has recently been associated with a later AD age-at-onset.[26] This LOAD association, together with our result of higher haplotype expression, suggests that the short version (16T) of *TOMM40* IVS6 poly-T combined with a

specific *TOMM40* promoter haplotype may be a modifier of AD risk through increased *TOMM40* expression not *APOE* expression.

All three of the *TOMM40* promoter 1-ME1 haplotypes showed significant differences compared with other *TOMM40* promoter-ME1 haplotypes in SHSY5Y cells (Figure 4d). In contrast, this *TOMM40* promoter 1 pattern was not seen in HepG2 cells where the most significant increase in expression was seen between *TOMM40* promoter 3-ME1-1 and the other *TOMM40* promoter haplotypes, indicating a cell-type-specific expression pattern for ME1 that is also haplotype specific (Figure 4g). In addition, the *APOE* promoter 2-ME1-1 haplotype showed a marginally significant increase in expression (Figures 3 and 4b). This is consistent with previous reports of increased expression in the presence of a ME1 transgene[12,13,31] and is a novel finding in that genetic variation within ME1 can further modulate expression levels. There was no increased expression of the ME1 haplotypes in the U118 astrocyte cell line, which is in contrast to a previous report that shows apoE expression is increased in transgenic *APOE*-ME1 mouse astrocytes in the brain.[31] It can be speculated that this inconsistency is due to either human versus mouse or *in vitro* versus *in vivo* differences between studies. But, as none of the reporter constructs showed significant differences in the U118 astrocyte cell line, it may represent a lack of *APOE* locus regulatory element activity in this cell line. The *APOE* promoter-ME1 showed marginally significant differences between haplotypes in SHSY5Y cells, but differences did not remain significant after correction of multiple comparisons (Figures 3 and 4b). However, most *APOE* promoter-ME1 haplotype expression was increased above the *APOE* promoter-only baseline of 1.0, significantly *APOE* promoter 2-ME1-1, suggesting that ME1 has a general enhancing effect on *APOE* promoter activity. Notably, all *TOMM40* promoter 1-ME1 haplotype levels were lower than the *TOMM40* promoter-only haplotypes in both SHSY5Y neuronal cells and HepG2 cells (Figures 4d and g; levels below 1.0; significantly lower levels are noted with asterisks). The *TOMM40* promoter-ME1 results suggest that ME1 inhibits *TOMM40* expression in contrast to enhancing *APOE* expression.

The *TOMM40* promoter is influenced by the silencer effect of BCR in both SHSY5Y and HepG2 cells (Figures 4e and h). All *TOMM40* promoter-BCR haplotypes have significantly lower expression compared with the *TOMM40* promoter-only haplotypes (significant levels below 1 noted with asterisk) in both HepG2 cells (Figure 4h) and SHSY5Y cells (Figure 4e) except the *TOMM40* promoter 1-BCR-1 haplotype in SHSY5Y cells (Figure 4e). Moreover, in HepG2 cells the difference in expression between the BCR-1 and BCR-2 haplotypes is significant regardless of the *TOMM40* promoter haplotype, suggesting that genetic variation in the BCR may influence *TOMM40* expression according to cell type independent of *TOMM40* promoter haplotype. Furthermore, the lack of influence by the BCR on the *APOE* promoter haplotypes in HepG2 cells is consistent with a previous report describing apoE expression in mice expressing apoE and BCR transgenes as constrained to the brain and not present in the liver.[14] Why BCR reportedly influences brain apoE levels[14] and not the *APOE* promoter haplotypes in SHSY5Y neuronal cells remains to be determined.

The *APOE* promoter 2 haplotype is significantly influenced by the *TOMM40* IVS2-4-2 haplotype with increased expression compared with the other *TOMM40* IVS2-4 haplotypes in the SHSY5Y cells (Figure 4a). These results are consistent with results from our previous studies where similar genetic variation within *TOMM40* IVS2-4 (rs59007384 G allele) is associated with higher cerebrospinal fluid and post-mortem brain apoE expression.[24,25] The *TOMM40* promoter-1 haplotype is also influenced by the *TOMM40* IVS2-4-2 haplotype, showing higher expression compared with other haplotypes in HepG2 cells (Figure 4f). These results suggest that genetic variation within the *TOMM40* IVS2-4 region impacts expression of *APOE* locus genes in a cell-type-specific manner. Interestingly, a SNP (rs2075650) within the *TOMM40* IVS2-4 region is significantly associated with AD quantitative traits, such as Aβ42, in genome-wide association studies[20–23,32] as well as survival into old age, further implicating the *TOMM40* IVS2-4 region in a biological effect.

Evidence presented here implicates *APOE* locus genetic variation within regional enhancers as contributors to regional promoter activity. These results are further supported by ENCODE data that show promoter (H3K4Me3) and enhancer (H3K4Me1) associated histone marks overlap with the regulatory haplotype reporter construct sites presented here (Figure 2).[28,29] In addition, ENCODE DNase1 hypersensitivity cluster results overlap with regulatory haplotype region reporter construct sites, suggesting *trans*-acting factor binding at these *APOE* locus regulatory haplotype regions (Figure 2).[28,29] Thus, it may be further speculated that both *TOMM40* and *APOE* expression are influenced by regional regulatory haplotypes that in turn biologically impact AD pathogenesis.

In conclusion, functional characterization of *APOE* locus haplotype expression suggests that a complex transcriptional regulatory structure is modulated by distinct haplotypes composed of multiple and distantly located SNPs. The main novel finding of this investigation is that regions within *TOMM40* influence both *TOMM40* and *APOE* promoter activity *in vitro* depending on haplotype and cell type. These *TOMM40* regions include a recently described poly-T in intron 6 that is associated with AD age-at-onset,[26] and an intron 2–4 region that has been associated with AD risk[19,33] and AD quantitative traits in multiple genome-wide association studies.[20–23] This functional study suggests that non-*APOE* ε4 allele SNPs contribute to a promoter–enhancer haplotype structure that influences both *TOMM40* and *APOE* gene regulation, suggesting that the genetic association between AD and *TOMM40*, found in multiple genetic studies, may be due to an AD modifier effect by *TOMM40* that is haplotype and cell type specific.

ACKNOWLEDGEMENTS

This work was supported in part by the US Department of Veterans Affairs, Office of Research and Development Biomedical Laboratory Research Program and the National Institutes of Health/National Institutes on Aging Grant 1K99AG034214-01A1. We thank J Anthony Gill for the editing of the manuscript.

1 Bekris, L. M., Yu, C. E., Bird, T. D. & Tsuang, D. W. Genetics of Alzheimer disease. *J. Geriatr. Psychiatry Neurol.* **23,** 213–227 (2010).
2 Mahley, R. W., Weisgraber, K. H. & Huang, Y. Apolipoprotein E4: a causative factor and therapeutic target in neuropathology, including Alzheimer's disease. *Proc. Natl Acad. Sci. USA* **103,** 5644–5651 (2006).
3 Coon, K. D., Myers, A. J., Craig, D. W., Webster, J. A., Pearson, J. V., Lince, D. H. *et al.* A high-density whole-genome association study reveals that APOE is the major susceptibility gene for sporadic late-onset Alzheimer's disease. *J. Clin. Psychiatry* **68,** 613–618 (2007).
4 Yu, C. E., Seltman, H., Peskind, E. R., Galloway, N., Zhou, P. X., Rosenthal, E. *et al.* Comprehensive analysis of APOE and selected proximate markers for late-onset Alzheimer's disease: patterns of linkage disequilibrium and disease/marker association. *Genomics* **89,** 655–665 (2007).
5 Hyman, B. T., Gomez-Isla, T., Rebeck, G. W., Briggs, M., Chung, H., West, H. L. *et al.* Epidemiological, clinical, and neuropathological study of apolipoprotein E genotype in Alzheimer's disease. *Ann. N Y Acad. Sci.* **802,** 1–5 (1996).
6 Huang, W., Qiu, C., von Strauss, E., Winblad, B. & Fratiglioni, L. APOE genotype, family history of dementia, and Alzheimer disease risk: a 6-year follow-up study. *Arch. Neurol.* **61,** 1930–1934 (2004).

7 Bullido, M. J., Artiga, M. J., Recuero, M., Sastre, I., Garcia, M. A., Aldudo, J. *et al.* A polymorphism in the regulatory region of APOE associated with risk for Alzheimer's dementia. *Nat. Genet.* **18,** 69–71 (1998).
8 Artiga, M. J., Bullido, M. J., Sastre, I., Recuero, M., Garcia, M. A., Aldudo, J. *et al.* Allelic polymorphisms in the transcriptional regulatory region of apolipoprotein E gene. *FEBS Lett.* **421,** 105–108 (1998).
9 Ramos, M. C., Matias, S., Artiga, M. J., Pozueta, J., Sastre, I., Valdivieso, F. *et al.* Neuronal specific regulatory elements in apolipoprotein E gene proximal promoter. *Neuroreport* **16,** 1027–1030 (2005).
10 Town, T., Paris, D., Fallin, D., Duara, R., Barker, W., Gold, M. *et al.* The -491A/T apolipoprotein E promoter polymorphism association with Alzheimer's disease: independent risk and linkage disequilibrium with the known APOE polymorphism. *Neurosci. Lett.* **252,** 95–98 (1998).
11 Parker, G. R., Cathcart, H. M., Huang, R., Lanham, I. S., Corder, E. H. & Poduslo, S. E. Apolipoprotein gene E4 allele promoter polymorphisms as risk factors for Alzheimer's disease. *Psychiatr. Genet.* **15,** 271–275 (2005).
12 Shih, S. J., Allan, C., Grehan, S., Tse, E., Moran, C. & Taylor, J. M. Duplicated downstream enhancers control expression of the human apolipoprotein E gene in macrophages and adipose tissue. *J. Biol. Chem.* **275,** 31567–31572 (2000).
13 Mak, P. A., Laffitte, B. A., Desrumaux, C., Joseph, S. B., Curtiss, L. K., Mangelsdorf, D. J. *et al.* Regulated expression of the apolipoprotein E/C-I/C-IV/C-II gene cluster in murine and human macrophages. A critical role for nuclear liver X receptors alpha and beta. *J. Biol. Chem.* **277,** 31900–31908 (2002).
14 Zheng, P., Pennacchio, L. A., Le Goff, W., Rubin, E. M. & Smith, J. D. Identification of a novel enhancer of brain expression near the apoE gene cluster by comparative genomics. *Biochim. Biophys. Acta.* **1676,** 41–50 (2004).
15 Grupe, A., Abraham, R., Li, Y., Rowland, C., Hollingworth, P., Morgan, A. *et al.* Evidence for novel susceptibility genes for late-onset Alzheimer's disease from a genome-wide association study of putative functional variants. *Hum. Mol. Genet.* **16,** 865–873 (2007).
16 Carrasquillo, M. M., Zou, F., Pankratz, V. S., Wilcox, S. L., Ma, L., Walker, L. P. *et al.* Genetic variation in PCDH11X is associated with susceptibility to late-onset Alzheimer's disease. *Nat. Genet.* **41,** 192–198 (2009).
17 Harold, D., Abraham, R., Hollingworth, P., Sims, R., Gerrish, A., Hamshere, M. L. *et al.* Genome-wide association study identifies variants at CLU and PICALM associated with Alzheimer's disease. *Nat. Genet.* **41,** 1088–1093 (2009).
18 Feulner, T. M., Laws, S. M., Friedrich, P., Wagenpfeil, S., Wurst, S. H., Riehle, C. *et al.* Examination of the current top candidate genes for AD in a genome-wide association study. *Mol. Psychiatry* **15,** 756–766 (2010).
19 Seshadri, S., Fitzpatrick, A. L., Ikram, M. A., DeStefano, A. L., Gudnason, V., Boada, M. *et al.* Genome-wide analysis of genetic loci associated with Alzheimer disease. *JAMA* **303,** 1832–1840 (2010).
20 Potkin, S. G., Guffanti, G., Lakatos, A., Turner, J. A., Kruggel, F., Fallon, J. H. *et al.* Hippocampal atrophy as a quantitative trait in a genome-wide association study identifying novel susceptibility genes for Alzheimer's disease. *PLoS ONE* **4,** e6501 (2009).
21 Shen, L., Kim, S., Risacher, S. L., Nho, K., Swaminathan, S., West, J. D. *et al.* Whole genome association study of brain-wide imaging phenotypes for identifying quantitative trait loci in MCI and AD: a study of the ADNI cohort. *Neuroimage* **53,** 1051–1063 (2010).
22 Han, M. R., Schellenberg, G. D. & Wang, L. S. Genome-wide association reveals genetic effects on human Abeta42 and tau protein levels in cerebrospinal fluids: a case control study. *BMC Neurol.* **10,** 90 (2010).
23 Kim, S., Swaminathan, S., Shen, L., Risacher, S. L., Nho, K., Foroud, T. *et al.* Genome-wide association study of CSF biomarkers Abeta1-42, t-tau, and p-tau181p in the ADNI cohort. *Neurology* **76,** 69–79 (2011).
24 Bekris, L. M., Millard, S. P., Galloway, N. M., Vuletic, S., Albers, J. J., Li, G. *et al.* Multiple SNPs within and surrounding the apolipoprotein E gene influence cerebrospinal fluid apolipoprotein E protein levels. *J. Alzheimers Dis.* **13,** 255–266 (2008).
25 Bekris, L. M., Galloway, N. M., Montine, T. J., Schellenberg, G. D. & Yu, C. E. APOE mRNA and protein expression in postmortem brain are modulated by an extended haplotype structure. *Am. J. Med. Genet. B Neuropsychiatr. Genet.* **153B,** 409–417 (2010).
26 Roses, A. D. An inherited variable poly-T repeat genotype in TOMM40 in Alzheimer disease. *Arch. Neurol.* **67,** 536–541 (2010).
27 Barrett, J. C., Fry, B., Maller, J. & Daly, M. J. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* **21,** 263–265 (2005).
28 Kent, W. J., Sugnet, C. W., Furey, T. S., Roskin, K. M., Pringle, T. H., Zahler, A. M. *et al.* The human genome browser at UCSC. *Genome Res.* **12,** 996–1006 (2002).
29 Rosenbloom, K. R., Dreszer, T. R., Pheasant, M., Barber, G. P., Meyer, L. R., Pohl, A. *et al.* ENCODE whole-genome data in the UCSC genome browser. *Nucleic Acids Res.* **38,** D620–625 (2010).
30 Paik, Y. K., Chang, D. J., Reardon, C. A., Walker, M. D., Taxman, E. & Taylor, J. M. Identification and characterization of transcriptional regulatory regions associated with expression of the human apolipoprotein E gene. *J. Biol. Chem.* **263,** 13340–13349 (1988).
31 Grehan, S., Tse, E. & Taylor, J. M. Two distal downstream enhancers direct expression of the human apolipoprotein E gene to astrocytes in the brain. *J. Neurosci.* **21,** 812–822 (2001).
32 Deelen, J., Beekman, M., Uh, H. W., Helmer, Q., Kuningas, M., Christiansen, L. *et al.* Genome-wide association study identifies a single major locus contributing to survival into old age; the APOE locus revisited. *Aging Cell* **10,** 686–98 (2011).
33 Bertram, L. & Tanzi, R. E. Genome-wide association studies in Alzheimer's disease. *Hum. Mol. Genet.* **18,** R137–145 (2009).