Human Molecular Genetics, 2010, Vol. 19, No. 10 **2068–2078**
doi:10.1093/hmg/ddq079
Advance Access published on February 18, 2010

Analysis of lipid pathway genes indicates association of sequence variation near *SREBF1/TOM1L2/ATPAF2* with dementia risk

Chandra A. Reynolds[1], Mun-Gwan Hong[2], Ulrika K. Eriksson[2], Kaj Blennow[3], Fredrik Wiklund[2], Boo Johansson[4], Bo Malmberg[5], Stig Berg[5], Andrey Alexeyenko[2], Henrik Grönberg[2], Margaret Gatz[2,6], Nancy L. Pedersen[2,6] and Jonathan A. Prince[2,*]

[1]Department of Psychology, University of California at Riverside, Riverside, CA 92521, USA, [2]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, 171 77 Stockholm, Sweden, [3]Institute of Neuroscience and Physiology, Sahlgrenska Academy at University of Gothenburg, 431 80 Mölndal, Sweden, [4]Department of Psychology, University of Gothenburg, 405 30 Göteborg, Sweden, [5]Institute of Gerontology, School of Health Sciences, Jönköping University, 551 11 Jönköping, Sweden and [6]Department of Psychology, University of Southern California, Los Angeles, CA 90089-1061, USA

Received November 11, 2009; Revised and Accepted February 17, 2010

We conducted dense linkage disequilibrium (LD) mapping of a series of 25 genes putatively involved in lipid metabolism in 1567 dementia cases [including 1270 with Alzheimer disease (AD)] and 2203 Swedish controls. Across a total of 448 tested genetic markers, the strongest evidence of association was as anticipated for *APOE* (rs429358 at $P \sim 10^{-72}$) followed by a previously reported association of *ABCA1* (rs2230805 at $P \sim 10^{-8}$). In the present study, we report two additional markers near the *SREBF1* locus on chromosome 17p that were also significant after multiple testing correction (best $P = 3.1 \times 10^{-6}$ for marker rs3183702). There was no convincing evidence of association for remaining genes, including candidates highlighted from recent genome-wide association studies of plasma lipids (*CELSR2/PSRC1/SORT1*, *MLXIPL*, *PCSK9*, *GALNT2* and *GCKR*). The associated markers near *SREBF1* reside in a large LD block, extending more than 400 kb across seven candidate genes. Secondary analyses of gene expression levels of candidates spanning the LD region together with an investigation of gene network context highlighted two possible susceptibility genes including *ATPAF2* and *TOM1L2*. Several markers in strong LD ($r^2 > 0.7$) with rs3183702 were found to be significantly associated with AD risk in recent genome-wide association studies with similar effect sizes, providing independent support of the current findings.

INTRODUCTION

The discovery that sequence variants of the gene encoding Apolipoprotein E (*APOE*) contribute to variation in Alzheimer disease (AD) risk provided one of the first strong indications that lipid metabolism might be integrally involved in dementia (1). This has been widely supported by replication studies showing genetic association of *APOE* and plasma/serum levels of cholesterol and phospholipids (2,3). Although multiple lipids are affected by *APOE* polymorphism, the strongest effect is on low-density lipoprotein (LDL) (4,5). *APOE*'s effect on lipid metabolism is presumed to significantly contribute to AD risk and progression (6–8), although the precise molecular mechanisms still remain unclear. Another key finding connecting AD and lipid metabolism is that APOE also appears to strongly affect brain β-amyloid (Aβ) deposition, which is the main component of senile plaques (9). Furthermore, another prominent phenotypic consequence of *APOE* polymorphism is to contribute to a large proportion of variance in cerebrospinal fluid (CSF) levels of the

*To whom correspondence should be addressed at: Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Nobels väg 12A, 171 77 Stockholm, Sweden. Tel: +46 852486008; Fax: +46 8314975; Email: jonathan.prince@ki.se

42 amino acid fragment of β-amyloid ($A\beta_{1-42}$) (10). Other genes that are proven to lead to rare familial forms of AD (*APP*, *PSEN1* and *PSEN2*) (11,12) also have been shown to affect Aβ metabolism in some way (13) and have been implicated in cholesterol metabolism (*APP*) or the processing of lipoprotein receptors such as LRP1 (*PSEN1* and *PSEN2*) (14). Together, these results provide a plausible indication of a pathway basis for AD, with both lipid and Aβ metabolism playing central roles. This suggests that systematically testing additional genes involved in this process may be an appropriate strategy to narrow the search field for genetic risk factors for AD.

Although numerous candidate gene studies have been performed over the years, *APOE* remains the only widely validated common genetic risk factor for AD. From single candidate gene studies (15), to pathway approaches (16), to genome-wide association (17), no additional genes with as convincing statistical support have emerged. By and large, genetic association studies for AD still reflect a shifting field of support and refutation. A summary of replicated and non-replicated findings can be found at the Alzgene database (www.alzgene.org) (18), which echoes the sentiment of only a few modestly significant results. This is also true for the few pathway based approaches that have been conducted that have focused specifically on lipid metabolism and AD (16), although coverage of selected lipid metabolism genes has been relatively sparse in such studies. We sought in the present study to explore a number of genes that, like *APOE*, also have a probable role in lipid metabolism. We selected most of the commonly studied genes in relation to plasma lipid levels, including some novel candidates that have emerged from genome-wide association studies of lipid traits that have not yet been specifically tested in relation to dementia and AD. We were particularly interested in genes that had prior evidence of functional sequence variation and note that the present study provides dense genetic coverage of 18 of 23 previously listed loci with strong evidence of genetic association with plasma lipids (5).

RESULTS

We prioritized a list of 506 genetic markers spanning 25 candidate genes (Table 1) for which genotyping assays were then designed and tested in all 3770 DNA samples. For all markers, genotyping was performed in duplicate in a subset of 100 individuals. We also tested for significant allele-frequency deviation with HapMap data for European populations, for markers overlapping with our study. Genotypes were obtained for a total of 454 markers with call rates above 70% across all samples. Nineteen of these markers had call rates between 70 and 90%, with the remainder being above 90%. For duplicate genotyping, one marker had three discrepancies, 15 had one discrepancy only and the remaining markers had no discrepant genotype calls. Marker loss was not randomly distributed. As an example, only one marker of 10 tested succeeded for *ABCA2*. Also of note, genotyping of *APOE* markers failed for both rs7412 and rs429358 and these were therefore reassessed with alternative chemistry on a different platform. These were the only markers chosen for additional genotyping. There were six cases of what we considered extreme ($P < 10^{-9}$) deviation from Hardy–Weinberg equilibrium (HWE). For all six markers, genotyping success rates were below 90%, and these were thus excluded. There were a further 15 markers that deviated from HWE at a more stringent threshold ($P < 0.001$), but all had call rates above 90% and were therefore included in primary analyses.

The remaining 448 markers were initially tested for association with total dementia using all 3770 individuals under an additive logistic regression model without accounting for family structure. The results of that analysis for the top 20 single markers are shown in Table 2. The top five markers exceeded significance in the context of multiple testing and strict Bonferroni correction based upon number of tests (the threshold given this number is $P < 0.00011$). There were 54 markers that exceeded an uncorrected significance level of $P < 0.05$, which is more than the 22 expected by chance. There were 23 markers that exceeded an uncorrected $P < 0.01$, compared with four expected by chance. Part of this inflation is due to linkage disequilibrium (LD), but we noted that for these last 23 markers in particular, they represented 13 distinct loci. The most significant finding was for rs429358 in *APOE*, followed by rs2230805 in *ABCA1*. The third and fourth most significant markers (rs3183702 and rs9899634) are in relatively strong LD and reside on chromosome 17p in the vicinity of *SREBF1*, which was the specific target gene in the original study design among the genes located in that block. For both rs3183702 and rs9899634 the common alleles were associated with increased risk. Testing specifically rs3183702 across gender indicated similar effect sizes in both women [odds ratio (OR) = 1.23, confidence interval (CI) 1.08–1.40, $P = 9.7 \times 10^{-4}$] and men (OR = 1.27, CI 1.08–1.50, $P = 2.2 \times 10^{-3}$). The common allele of marker rs3183702 was also significantly enriched when only possible and probable AD cases were tested (OR = 1.25, CI 1.12–1.39, $P = 3.7 \times 10^{-5}$). Given the established role of *APOE* in dementia risk, stratification by *APOE* e4 carrier status was also performed in addition to a joint logistic regression model considering both rs429358 and rs3183702 together. Effect size estimates for rs3183702 were similar in e4 carriers (OR 1.39, CI 1.17–1.64, $P = 6.2 \times 10^{-5}$) and e4 non-carriers (OR 1.22, CI 1.05–1.42, $P = 4.3 \times 10^{-3}$). The effect size for rs3183702 improved slightly when rs429358 in *APOE* was included in an adjusted model (OR 1.29, CI 1.16–1.45, $P = 3.1 \times 10^{-6}$).

Marker rs3183702 resides in a tight LD block spanning 400 kb across multiple gene targets, and is specifically located in the 3′-UTR of *TOM1L2*. There are seven genes that reside completely or partially in this LD block, including (from pter–qter along chromosome 17p) *RAI*, *SREBF1*, *TOM1L2*, *LRRC48*, *ATPAF2*, *DRG2* and *MYO15A*. We used information from a Swedish genome-wide data set (Materials and Methods) to establish which markers might be in strong LD with rs3183702 and thus potentially identify better functional candidates for the putative association signal. Across the LD region and a total of 327 markers, there were 180 markers with an LD with rs3183702 in excess of $r^2 = 0.2$, 26 markers with $r^2 > 0.9$ and 8 perfect proxies ($r^2 = 1$). This reiterates the problem of fine-mapping in this interval, in that more than half of all validated markers in the region

Table 1. Genes central to lipid metabolism

Gene symbol	Gene name	Location	Function	Markers
ABCA1	ATP-binding cassette protein 1	9q31	Cellular efflux of cholesterol	53
ABCA2	ATP-binding cassette protein 2	9q34	Cellular efflux of cholesterol	10
ACAT1	acetyl-Coenzyme A acetyltransferase 1	11q22	Esterification of free cholesterol	13
ACAT2	acetyl-Coenzyme A acetyltransferase 2	6q25	Esterification of free cholesterol	8
APOA1[a]	Apolipoprotein A1	11q23	Transport of cholesterol	22
APOB	Apolipoprotein B	2p24	Transport of cholesterol	28
APOE	Apolipoprotein E	19q13	Transport of cholesterol	2
CETP	Cholesteryl-ester transfer protein	16q13	Transfer of insoluble cholesteryl esters	32
CYP46A1	Cytochrome P451 family 46, subfamily A1	14q32	Converts cholesterol to 24S-hydroxycholesterol	17
HMGCR	3-Hydroxy-3-methylglutaryl-coenzyme A reductase	5q13	Rate-limiting enzyme of cholesterol biosynthesis	12
LCAT	Lecithin-cholesterol acetyl transferase	16q22	Esterifies free cholesterol in plasma lipoproteins	7
LIPC	Hepatic lipase	15q21	Hydrolysis of phospholipids, triglycerides and acyl-CoA thioesters	42
LIPG	Endothelial lipase	18q21	Hydrolysis of phospholipids and triglycerides	18
LDLR	Low density lipoprotein receptor	19p13	Binds LDL for transport into cells	22
LPL	Lipoprotein lipase	8p22	Hydrolysis and clearance of circulating triglycerides	29
LRP1	Low density lipoprotein-related protein 1	12q13	Multiligand receptor for lipoproteins	19
PLTP	Phospholipid transfer protein	20q13	Generation and modulation of HDL particles	12
SCARB1	Scavenger receptor class B, member 1	12q24	Cellular flux of free and esterified cholesterol	38
SREBF1	Sterol regulatory element binding transcription factor 1	17p11	Regulates the transcription of genes for sterol biosynthesis	10
SREBF2	Sterol regulatory element binding transcription factor 2	22q13	Regulates the transcription of genes for sterol biosynthesis	15
PCSK9	Proprotein convertase subtilisin	1p32.3	Cholesterol Homeostasis	24
SORT1[b]	Sortilin 1	1p13.3	Sorting of lysosomal proteins	15
GALNT2	Polypeptide *N*-acetylgalactosaminyltransferase 2	1q42.1	Oligosaccharide biosynthesis	30
GCKR	Glucokinase regulatory protein	2p23.3	Inactivates glucokinase	14
MLXIPL	MLX interacting protein-like	7q11.2	Activates triglyceride synthesis genes	14

[a]*APOA1* resides in a cluster of apolipoproteins *APOA5/APOA4/APOC3/APOA1*—we have considered all genes in this block.
[b]*SORT1* resides in a cluster consisting of *CELSR2/PSRC1/SORT1*.
Markers specifies the number of SNPs for which genotyping assays were designed and tested in all samples.

Table 2. Top 20 most associated markers with dementia

Marker	Nearest gene	Logistic p	Logistic OR	Risk allele (SNP type)	$A\beta_{42}$ *P*	ALR *P*	ALR OR
rs429358	*APOE*	1.07E−72	3.59; 3.12–4.14	G (G/A)	2.53E−13	2.91E−68	3.06; 2.70–3.47
rs2230805	*ABCA1*	2.45E−07	1.34; 1.20–1.52	G (G/A)	6.71E−03	4.11E−07	1.35; 1.20–1.52
rs3183702	*SREBF1*	8.50E−06	1.25; 1.13–1.39	G (G/A)	1.33E−01	1.33E−05	1.25; 1.13–1.38
rs9899634	*SREBF1*	2.56E−05	1.24; 1.11–1.37	A (A/T)[a]	1.31E−01	2.16E−05	1.24; 1.12–1.37
rs2230806	*ABCA1*	8.16E−05	1.23; 1.11–1.37	G (G/A)	1.01E−01	1.21E−04	1.23; 1.10–1.37
rs11615630	*SCARB1*	1.93E−04	1.19; 1.08–1.31	G (G/A)	5.17E−01	8.50E−05	1.21; 1.10–1.34
rs2065412	*ABCA1*	5.19E−04	1.16; 1.06–1.28	G (G/A)	2.04E−01	1.07E−03	1.16; 1.05–1.27
rs4379922	*SCARB1*	8.16E−04	1.17; 1.06–1.29	A (G/A)	7.75E−01	2.05E−03	1.16; 1.05–1.28
rs17585355	*SORT1*	1.49E−03	1.36; 1.11–1.66	A (A/C)	6.74E−01	1.49E−03	1.36; 1.11–1.67
rs1800168	*LRP1*	1.75E−03	1.16; 1.05–1.29	G (G/A)	3.94E−01	2.12E−03	1.17; 1.05–1.29
rs10744182	*SCARB1*	1.87E−03	1.16; 1.05–1.28	G (G/A)	4.92E−02	3.17E−03	1.15; 1.04–1.28
rs3751542	*LIPC*	2.26E−03	1.16; 1.05–1.28	G (G/A)	5.54E−01	3.07E−03	1.16; 1.04–1.28
rs12904012	*LIPC*	2.40E−03	1.23; 1.06–1.42	A (G/A)	5.88E−01	3.79E−03	1.22; 1.05–1.40
rs2066716	*ABCA1*	2.48E−03	1.30; 1.08–1.56	G (G/A)	1.27E−01	2.33E−03	1.29; 1.08–1.55
rs1801695	*APOB*	2.89E−03	1.53; 1.13–2.07	G (G/A)	9.41E−01	1.50E−02	1.42; 1.04–1.95
rs1800590	*LPL*	3.17E−03	2.08; 1.23–3.52	A (A/C)	2.03E−01	6.95E−03	2.05; 1.16–3.63
rs2738464	*LDLR*	5.21E−03	1.24; 1.07–1.43	C (C/G)[a]	8.38E−01	5.87E−03	1.22; 1.04–1.42
rs12350560	*ABCA1*	5.66E−03	1.25; 1.07–1.48	G (G/A)	4.07E−01	7.34E−03	1.22; 1.04–1.44
rs5104	*APOA4*	6.95E−03	1.22; 1.06–1.41	G (G/A)	2.61E−01	2.89E−03	1.23; 1.06–1.43
rs2197089	*LPL*	9.50E−03	1.13; 1.03–1.24	A (G/A)	7.87E−01	1.10E−02	1.12; 1.02–1.23

ORs are reported with 95% CIs and have been inverted where ORs were <1.
Association with $A\beta_{42}$ was evaluated with ANOVA in AD cases only ($n = 620$). ALR *P*-values and ORs are shown for comparison.
[a]Ambiguous risk alleles: see Supplementary Material, Table S1 for allele naming conventions.

are in significant LD with rs3183702. Across this region, there were only two validated single nucleotide polymorphisms (SNPs) that potentially affected amino acid sequences of the protein products of their respective genes. This included rs11649804 located in *RAI1* (P165T) and rs4584886 located in *LRRC48* (R191W). We used the F-SNP database tool to consider the potential functional significance of these SNPs (e.g. splicing, transcription, translation, post-translation)

which affirmed that these variants may be considered potentially deleterious changes (19). LD between rs3183702 and rs11649804 was $r^2 = 0.70$ and $r^2 = 0.75$ with rs4584886.

In order to potentially implicate a specific gene in the 17p LD block, we sought a strategy to explore for effects of regulatory sequence variation in the region. Thus, our hypothesis was that if markers can be detected that significantly associate with expression levels of one of the represented genes, this might lend weight to that particular gene over the others. Towards this goal, we downloaded publicly available genetic marker and expression data from previous genome-wide studies that had both sources of information (20,21). Of the genetic markers in the study of Myers *et al.*, four SNPs sufficiently tagged most variation in the region (a proxy was needed for rs3183702 since it was not present) and were tested for association with a single target transcripts from each of the genes in the region using expression data in 193 brain samples using PLINK (22). Of the seven genes in the region, *RAI1*, *SREBF1*, *TOM1L2*, *ATPAF2* and *DRG2* were detected (*LRRC48* and *MYO15A* were not present). None of the tested variants across the above transcripts showed significant association at *P*-values <0.05. However, in investigating association between the expression of the various genes in lymphocytes of 400 individuals and genetic variants (20), we did obtain relatively strong evidence of an effect on *ATPAF2* levels (we note that all seven genes were represented in that study). The best marker, rs4925114 is in nearly complete LD with rs3183702 ($r^2 = 0.95$) with a significance level of $P = 1.9 \times 10^{-4}$ for the *ATPAF2* expression trait. The risk allele implicated in the association with dementia was in LD with the allele associated with increased *ATPAF2* expression.

To further address the question of which of the seven genes on 17p might be a better candidate for dementia, we investigated their functional context using FunCoup (23). Each candidate from 17p region was tested individually for network connectivity, i.e. whether protein interactions might occur between the 17p candidates with the base set of known AD genes (*APOE*, *APP*, *PSEN1* and *PSEN2*). None of the seven candidate genes on 17p was directly linked to the base set. When evaluating indirect connections, *TOM1L2* exhibited a significant enrichment of links (specifically, shared 21 neighbors with the four base genes; $P_\alpha = 0.01$). Figure 1 presents the most important part of this pattern—a group of histones H2B that, according to FunCoup database, are functionally coupled to both *TOM1L2* and presenilins. The results of the full analysis of all candidates are shown in Table 3.

Using the Swedish LD information as a reference, we consulted two recent genome-wide association studies on AD (17,24) to pursue evidence of independent replication of variation in the *SREBF1* region with AD risk. For the first study, there was no perfect proxy for rs3183702 represented in that data set, but in evaluating the region around *SREBF1*, we noted three markers significantly associated with AD at a liberal threshold of $P < 0.05$. LD between these markers and rs3183702 was on the order of $r^2 = 0.75$ in our Swedish sample. We show the results of this in Table 4, where additional markers are highlighted with evidence at $P < 0.1$. In each case, increased risk was associated with the common allele, and that was also in LD with the risk allele for rs3183702. We illustrate the monotonic decay of LD around rs3183702 in Figure 2. There we also highlight the position of the nearest associated marker (rs4459604) from the genome-wide association study of Reiman *et al.* (17). We also indicate the position of the marker that was found to associate with *ATPAF2* expression levels (rs4925114). We note that in the Reiman *et al.* study, there are only 12 typed markers that extend from rs4459604 towards the beginning of the LD block. All markers within the LD block, without exception, have higher MAFs in controls, consistent with the results in the present study. We also examined the top results of the largest genome-wide study of AD risk yet performed (24) (full genotyping results in contrast to the Reiman study were not publicly available). The *SREBF1* region is also implicated there, with the best marker (rs854813) being again in high LD with rs3183702 ($r^2 = 0.67$) at a significance level of 6.3×10^{-4}. Importantly, the common allele of rs854813 is also associated with an increased risk for AD. To facilitate a comparison with the GWAS data sets we used IMPUTE (25) to probabilistically infer additional genotypes from the *SREBF1* region. This resulted in a total of 118 markers spanning the LD region with $>80\%$ genotype calls across individuals at a posterior probability in excess of 80%. The results for all markers with dementia risk in all samples are shown in Figure 3. No additional markers exceeded the statistical significance of rs3183702. The result for rs854813 in our sample was also highly significant, again with the common allele contributing to increased risk (OR 1.23; CI 1.11–1.37, $P = 4.6 \times 10^{-5}$). Meta-analysis for rs854813 was conducted using our imputed data together with the three independent data sets from Harold *et al.* (24) as well as rs4426402 from the Reiman *et al.* study (17) which was a perfect proxy for rs854813. This analysis included a total of 6107 dementia cases and 10263 controls and provided an increase in overall significance (OR 1.15, CI 1.10–1.20, $P = 7.7 \times 10^{-9}$; *P* for homogeneity = 0.3).

Apart from with AD risk, one of the most prominent genetic associations of *APOE* is with CSF $A\beta_{42}$ levels (10), which is consistent with a perturbation of amyloid metabolism being central to disease pathology. Against that background, the top 20 single findings (Table 2) were tested for association with CSF measures of $A\beta_{42}$. Only AD cases were included in this analysis. Across these markers there was as previously reported evidence of association with rs429358 and rs2230805 and CSF $A\beta_{42}$ levels and in each case the risk alleles for AD were associated with lower $A\beta_{42}$ levels. However, although there was one marginal association of rs10744182 in *SCARB1* ($P = 0.049$), none of the other markers were significant and the results for this are summarized in Table 2.

Finally, twin-pair adjusted tests were conducted for rs3183702 to account for dependencies that may impact standard errors and thus significance levels. The total *N* was 3497 with 2740 clustered pairs, where those in the case–control sample were treated as members of incomplete DZ pairs. Similar to the findings that did not take family structure in to account, the alternating logistic regression (ALR) procedure also demonstrated a significant elevated risk of rs3183702 on dementia at $P = 2.67 \times 10^{-5}$ (OR = 1.25, 95% CI = 1.12–1.38). Moreover, adjusting for *APOE*, the effect of

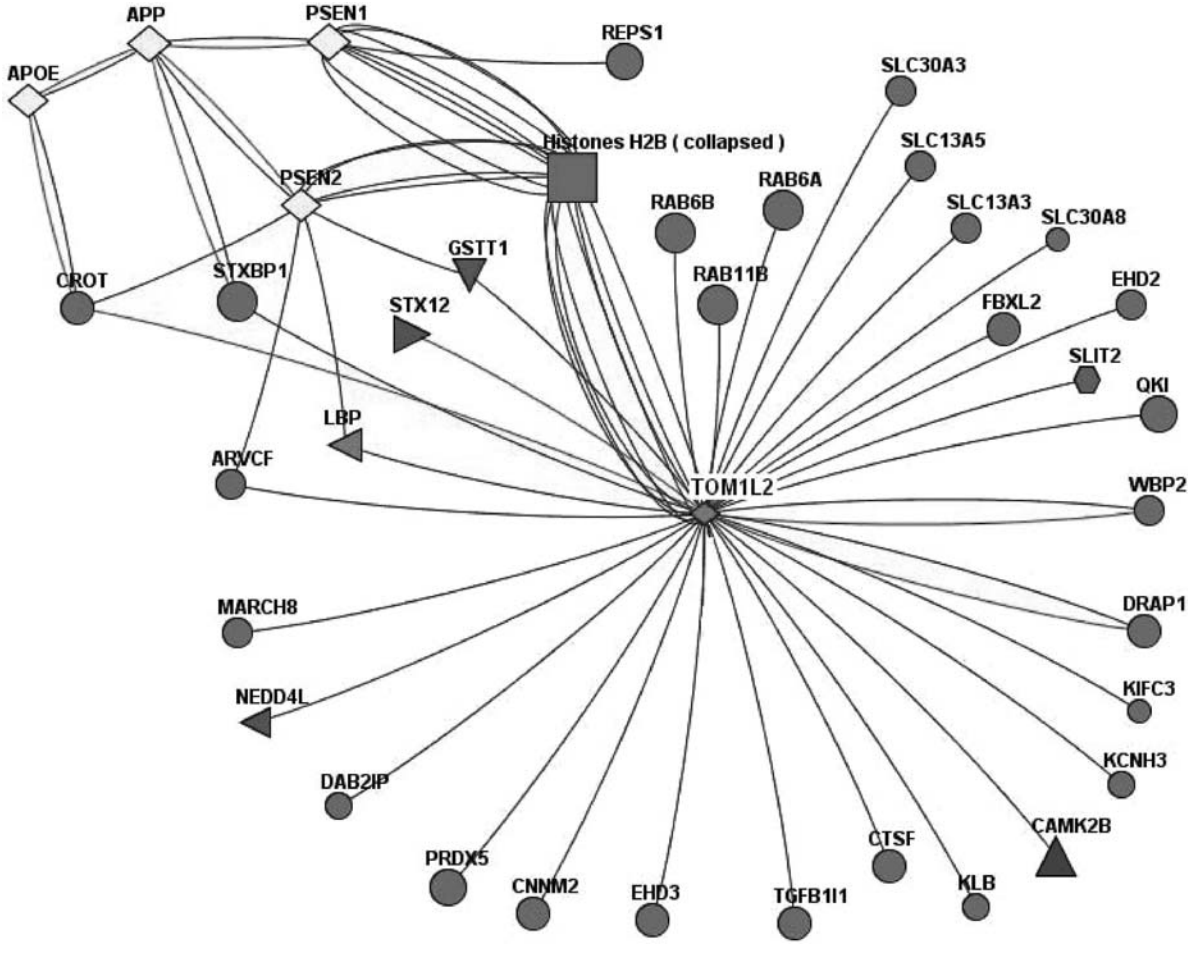

Figure 1. *TOM1L2* and genes associated with Alzheimer's disease in the human interactome. A human network of functional coupling was built by probabilistic integration of different data sets in human and six model eukaryotic species (23) and is available via web network browser at http://funcoup.sbc.su.se (web page http://funcoup.sbc.su.se/algo.html#noislets also describes algorithm 1 employed to retrieve this sub-network). Shapes: diamonds to the top left, the four genes previously associated with Alzheimer's disease and *TOM1L2* is in the center of the figure; square, group of histones (collapsed into single node) that connect presenilins and *TOM1L2*; triangles, genes assigned to a pathway in the KEGG database; circles, other genes. Lines: Evidence of functional coupling between genes with protein–protein interactions and mRNA co-expression.

rs3183702 remained significant at $P = 1.95 \times 10^{-5}$ (OR = 1.28, 95% CI = 1.14–1.43).

DISCUSSION

We performed a survey of candidate genes that have a well-defined role in lipid metabolism for association with dementia and AD risk. This study was designed to provide relatively dense marker coverage of these genes, at a level that exceeds what is obtained on most genome-wide platforms. Notably, most of the candidates have been tested and confirmed to be strongly associated with plasma lipid traits in genome-wide association studies (5,26–28). Although the present study is not the first to employ a lipid pathway approach to explore AD and dementia genetics (16), it is considerably larger, several-fold larger in terms of marker coverage and nearly a third larger in total sample size. In addition, we performed secondary analyses of gene expression levels and considered potential functional coupling of proteins in the cholesterol pathway that were identified in association analyses.

Although we point out significant association of previously reported genes (i.e. *APOE* and *ABCA1*), the focal point of the present study is on the novel finding that sequence variation near the *SREBF1* locus on chromosome 17p associates with both dementia and AD risk. The interval around *SREBF1* contains seven genes, but the extent of LD in the region limits the implication of one particular candidate over the others. We attempted to single-out one of the genes in the interval with a combination of gene expression analysis, together with gene network analyses (23). For the former, there was relatively strong evidence of association with expression levels of transcripts coding for ATPAF2, with a lack of significant signals for the other genes in the region (no other markers at $P < 0.001$). To put this in context, on average each of the detectable ~15 000 transcripts in the genome has around 80 significant markers at $P < 1.9 \times 10^{-4}$ (the significance level for the best associated marker for *ATPAF2* expression). However, for only around 100 of these transcripts does one

Table 3. Analysis of functional relations via common interactors in gene networks

Gene	Links Observed	Links Expected	SD	*Z*-score
PSEN2	443	355.7	5.813	15.0185
PSEN1	362	280.4	5.985	13.6337
APOE	135	117.2	2.098	8.4858
APP	278	247.8	5.75	5.2518
TOM1L2	21	12.1	3.479	2.5586
SREBF1	24	24.6	3.864	−0.1553
LRRC48	0	0.2	0.632	−0.3162
MYO15A	9	10.9	3.143	−0.6045
RAI1	2	7.4	2.171	−2.4879
ATPAF2	11	55.3	8.486	−5.2204
DRG2	103	207.6	18.204	−5.7461

The base set of genes (bold italics) consisted of PSEN2, PSEN1, APOE and APP. Significance was determined for each gene against the base set by permutation (Materials and Methods).

Table 4. Evidence of association of markers near SREBF1 with AD from the TGEN study

TGEN markers	rsID	LD with rs3183702	MAF (affected/unaffected)	*P*-value
SNP_A-2234401	rs4459604	0.75	0.31/0.35	0.027
SNP_A-4241614	rs4368210	0.75	0.30/0.33	0.097
SNP_A-4290890	rs4584886	0.74	0.30/0.33	0.097
SNP_A-4268811	rs4924832	0.74	0.32/0.36	0.037
SNP_A-2220775	rs6502632	0.75	0.30/0.33	0.082
SNP_A-1877154	rs9896837	0.75	0.31/0.34	0.09
SNP_A-1874877	rs2955384	0.75	0.33/0.36	0.097
SNP_A-1897101	rs4643387	0.75	0.30/0.33	0.08
SNP_A-2005092	rs4426402	0.67	0.29/0.35	0.0023

MAF, minor allele frequency; LD is according to r^2. Note that the MAF for rs3183702 was 0.31 in dementia and 0.36 in controls in the present study.

of these significant markers fall within the specific transcripts genomic interval. We note that this number is quite a bit less than that observed in other studies (29), but was nonetheless sufficiently low to merit considering *ATPAF2* a stronger candidate gene than the others in the region, at least in terms of being under control of regulatory polymorphism.

As a complementary attempt to prioritize one (or more) genes in the LD block, a network-based modeling strategy was used to understand the various genes as they relate to the biology of known AD genes (specifically *APOE*, *APP*, *PSEN1* and *PSEN2*). Our guidance for this kind of strategy is the growing appreciation of functional network analyses to understanding the biological role and thus prioritization of candidate genes (30–32). Thus, if any of the candidates in the 17p block have stronger evidence of functional coupling to the aforementioned AD genes, it might provide an important item of evidence implicating that particular gene over the others. The software selected for this analysis (FunCoup) was chosen specifically since it is not biased by text-mining strategies, which are commonly incorporated into network analysis programs (e.g. STRING) (33). Rather, it is driven primarily by experimental data, with the greatest weight being obtained by protein–protein interaction data and gene expression correlations. In testing the various genes, the strongest evidence of functional coupling was obtained for *TOM1L2*. We did note that a deeper investigation revealed that coupling of *TOM1L2* was computationally inferred from orthologs in *Saccharomyces cerevisiae* and *Arabidopsis thaliana*. Hence, it might be premature to conclude that these common neighbors—histones—are involved in the Alzheimer's disease. At the same time, the four base set genes were very well connected with each other and to the previously implicated candidate *ABCA1* (not shown). Lack of significance for any of the other six genes on 17p might result from scarcity of available (high-throughput and text-mining) evidence, as the network confidence scores depend on data abundance and, eventually, focus of the research community.

We used recently published data from genome-wide association studies of AD to explore for significant findings overlapping the results of the present study (17). Although there was no perfect proxy (i.e. a marker in perfect LD) for our primary candidate variant (rs3183702), there were several markers in fairly strong LD with rs3183702 that did show association with AD risk in those studies. Importantly, the effect sizes were similar to our results and the direction of allele effects were equivalent, with common alleles of associated markers appearing to increase risk in both studies. Using imputation in our dataset and conducting meta-analysis together with the other GWAS data sets including over 6000 cases and 10 000 controls, the combined evidence of association is quite strong ($P \sim 10^{-9}$).

We have highlighted two possible candidates in the 17p LD block. The first is *ATPAF2*, which encodes a protein involved in mitochondrial function. Specifically, the protein is an assembly factor for the F1 subunit of the mitochondrial ATP synthase, and is important in preventing the formation of homo-oligomers. Although numerous association studies have been performed, to date there are no strong genetic candidates suggesting that mitochondrial genetics plays a role in dementia. Nonetheless, defects in mitochondrial function remain a central theme in the biological pathway descriptions of dementia (e.g. http://www.genome.jp/kegg/pathway/hsa/hsa05010.html). There is relatively limited literature on *ATPAF2*, with only 11 published articles refer to it, and little more is known about its biology. The second potential candidate is *TOM1L2*, which encodes a protein putatively involved in intracellular protein transport and has a proportionately higher level of connectivity with known AD genes than other genes in the 17p LD block. Like *ATPAF2*, relatively little is known about *TOM1L2* other than that it may to be important for the intracellular recruitment of clathrin onto endosomes (34). Thus, although we remain cautious about which gene in the region is the true culprit, the present results suggest that more attention to *ATPAF2* and *TOM1L2* may be warranted for future replication studies.

There was a tendency towards an enrichment of significant signals across all markers tested in this study. We noted significant findings at an uncorrected $P < 0.05$ for 53 markers, which exceeds the number expected by chance. However, when we considered LD as well as numbers of markers per gene, any significant enrichment of signals disappears. Thus, although the primary goal of the project was to identify new candidates, which we consider successful, there is no strong

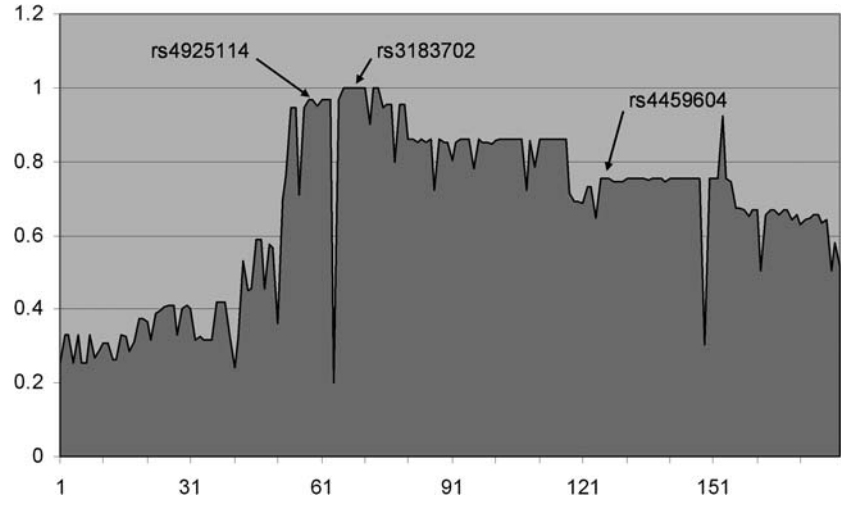

Figure 2. The monotonic decay of LD in Swedes around marker rs3183702 is illustrated, with all markers that exceed $r^2 = 0.2$. This indicates the strength of LD in the region and the difficulty of fine-mapping causative variants. Three key markers are highlighted, including rs3183702 that showed association with dementia in the present study, along with rs4459604 that was associated with AD in a recently published genome-wide association study. Also shown is marker rs4925114 that exhibited evidence of association with expression levels of one of the genes in the LD block (*ATPAF2*).

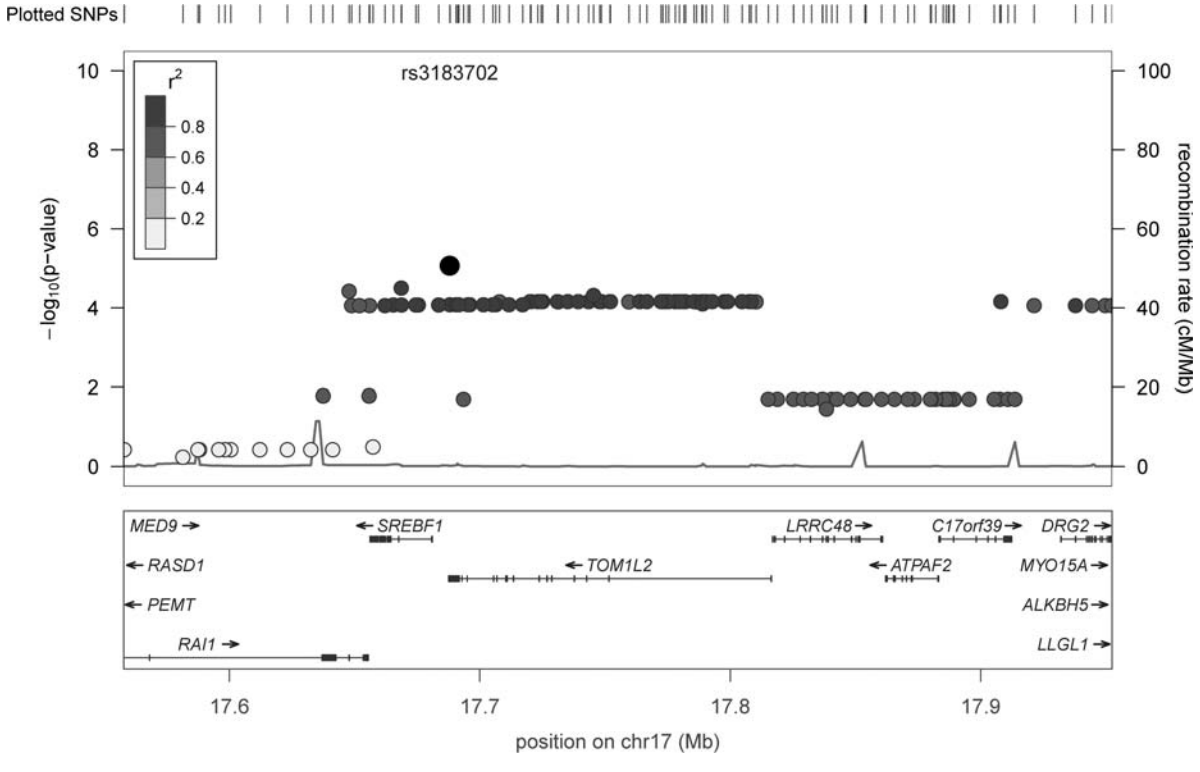

Figure 3. Schematic overview of the *SREBF1* region extending from marker rs11654482 to marker rs854764 on chromosome 17, depicting unadjusted allelic association of $-\log_{10}P$ (*Y*-axis) values for 118 genetic markers in the *X*-axis in relation to the full dementia versus control sample with a maximum for rs3183702. The figure was generated using Locuszoom (http://csg.sph.umich.edu/locuszoom).

evidence that lipid genes in general contain polymorphism that impacts AD to an extent that exceeds any other pathway. This lack of lipid pathway enrichment is also suggested in that the best candidate from the *SREBF1* region may not be a lipid related gene, but instead a gene related to a distinct biological process.

In summary, we have identified a potential association of sequence variation in the vicinity of *SREBF1* on chromosome 17p with dementia and AD. This represents the third strongest association signal that we detect in these particular samples (after *APOE* and *ABCA1*) and the result is supported by findings in recent genome-wide association studies. Further

replication studies in alternative populations are warranted, as are functional studies and a continued application of bioinformatics to refine the more precise location of the association signal.

MATERIALS AND METHODS

Human samples

The present study drew participants from four aging twin studies stemming from the population-based Swedish Twin Registry (35), and an independent non-twin case–control Swedish AD sample (36). Across the samples, DNA was available for 1567 dementia cases and 2203 controls (1275 with possible or probable AD diagnoses). Among dementia cases, 1233 were unrelated. There were 990 men and 1213 women in the control group, and 598 men and 969 women in the dementia group. Average age-at-sampling across cases and controls was 77.7 ± 8.7 (SD) years and age-at-onset for dementia/AD cases was 75.3 ± 8.3 (SD) years. The twin and case–control samples were described in detail recently (36). In brief, the twin samples are from the Swedish Adoption/Twin Study of Aging (SATSA) (37), the Origins of Variance in the Oldest-Old (OCTO-Twin) (38), Sex Differences in Health and Aging Study (GENDER) (39) and the Study of Dementia in Swedish Twins (HARMONY) (40). The case–control sample was comprised of unrelated individuals recruited from three prospective longitudinal studies of patients with dementia from Mölndal, Piteå, and Malmö, Sweden. This total sample affords 80% power for the detection of a genetic effect size of approximately 1.2 (OR) at an alpha of 0.05 and given a risk allele frequency of 0.25.

An additional population was included to establish LD characteristics of markers around *SREBF1/TOM1L2/ATPAF2* in Swedes in order to facilitate comparisons with other genome-wide association data (17). This consisted of 3013 men from the CAPS (Cancer Prostate in Sweden) study with genome-wide SNP data that were also fully imputed based upon HapMap release 22 marker data (25).

Lastly, we conducted secondary analysis of publicly available data of gene expression levels in peripheral blood lymphocytes collected in 400 children (20) and in 193 brain samples from adults who lived to 65 years or older (21).

Gene and marker selection

Our selection of candidate genes involved an exhaustive search of PubMed literature from 2003 to 2008 using the Boolean search string [(cholesterol OR lipid), gene*, polymorphism*]. Titles of papers were screened to establish relevance and abstracts read of papers highlighting gene names and/or cholesterol/lipid pathway approaches. Key papers purporting evidence of association of candidate genes with target phenotypes were read in their entirety. Our specific criteria for candidate gene selection from this search has entailed (i) the established biological role of each gene, in particular if the gene product acts in a rate-limiting manner (ii) the extent to which genetic variation in a gene has been explored in relation to either cholesterol/lipid, CVD or AD (iii) The existence of at least one coding non-synonymous amino acid changing polymorphic site listed in dbSNP (http://www.ncbi.nlm.nih.gov/projects/SNP/) or at least one potential regulatory polymorphism that could affect expression or splicing. This list reflects the most commonly studied genes in relation to lipid homeostasis and related disease phenotypes, in addition to some less studied but intriguing new candidates (e.g. *ABCA2*). We augmented our gene list with candidates highlighted as significantly associated with HDL, LDL or TG in recent genome-wide association studies (5,26). The resultant list is shown in Table 1, where we also highlight the primary function of each of the genes. We note that although several of the gene candidates have been explored in terms of AD risk (16,41), each of the genes has its own history in terms of being prioritized and this may not necessarily reflect the actual contribution of variation of the gene on a population level to any measurable phenotype. Of note, an important feature of the set is the inclusion of the *APOA5/APOA4/APOC3/APOA1* gene cluster on chromosome 11q, which covers a relatively tight genomic interval. This region is strongly implicated in plasma lipid regulation that may harbor multiple interacting trait loci (42). We acknowledge that this list is not yet exhaustive, especially given the emerging detail from high-throughput analyses depicted in pathway and network based descriptions (e.g. KEGG) (43).

Genetic markers for all candidate genes (including 20 kb upstream of the transcription start site and 10 kb downstream of the transcription end site) were selected with previous findings, functional candidature, LD and Illumina SNP design score as criteria (the full final list is shown in Supplementary Material, Table S1). Illumina scores were calculated by an algorithm developed by the company that predicts success of the assay for the marker. On the basis of the genotype data for CEU samples of HapMap Release 22, all polymorphic markers in the dataset were taken into consideration. At first, LD blocks were searched with Haploview 4.0 (44). Prioritization included markers in exons, within 80 bp of exon boundaries, and within 1 kb upstream from the first exon or downstream from the last exon of any predicted gene in UCSC genome browser, and SNPs that can tag the LD blocks. For new candidate genes for lipid traits from genome-wide studies, we ensured that our marker selection either included the previously described best associated marker or a perfect proxy. Among the markers outside of LD blocks, those which could be prioritized by the same scheme were included. After selection, Illumina scores for all markers were calculated and those that did not satisfy the criteria for Illumina probe chemistry were replaced with a SNP in perfect LD if available ($r^2 = 1$) or other tagging SNP. Coverage in terms of total number of markers for each candidate gene is shown in Table 1.

Genotyping

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

Amplification (WGA) using standard kits involving Phi29 DNA polymerase (Amersham).

CSF biomarkers

CSF samples were obtained in the AD case–control study by lumbar puncture in the L3/L4 or L4/L5 inter-space. Further details of CSF collection can be found elsewhere (45). CSF $A\beta_{42}$ was determined using a sandwich enzyme-linked immunosorbent assay (ELISA) [Innotest b-amyloid (1–42), Innogenetics, Ghent, Belgium] constructed to specifically measure $A\beta_{42}$ (46). The microtubule-associated protein tau, a CSF marker of neuronal degeneration, was determined using a sandwich ELISA (Innotest hTAU-Ag, Innogenetics, Ghent, Belgium) constructed to measure total tau, i.e. all isoforms of tau irrespective of phosphorylation state (47).

Statistics

HWE for individual loci was assessed using the Pearson χ^2 statistic. To account for relatedness, secondary analyses were conducted where only one twin was included among MZ pairs, both members of DZ pairs and all non-twin case–control participants. Initial tests of association between individual markers and dementia risk were performed without considering family structure under an additive model using logistic regression. To account for pair dependency, ALR was used that included both members of the pair while accounting for MZ and DZ pair correlation structures (48,49). ALR analyses were performed in SAS 9.1 using the GENMOD procedure (SAS Institute, Inc., Raleigh, NC, USA). Haplotypes were further estimated after LD block definition in individual blocks using Haploview v4.1 (44). Tests of genotypes versus quantitative traits (age-at-onset, CSF tau and CSF $A\beta_{42}$) were conducted using ANOVA in STATA v9.0. Analyses of gene expression levels were performed on log-transformed data as described previously (50) using PLINK (22).

Network analyses

We considered potential functional coupling of proteins in the cholesterol pathway that were identified in association analyses using the FunCoup gene network resource (23) where smaller sub-networks are presented via a web-based network browser. Figure 1 in this article was prepared with this tool. However, estimating significance of functional coupling between a particular gene of interest and a pre-defined gene group (pathway) is not trivial. The result would be strongly biased by the fact that number of network links per gene varies according to power law distribution. To address the problem in estimating significance, we developed custom software that implemented a previously proposed randomization algorithm (51). The randomized network was thus re-wired in such a way that the number of links for each node was preserved, although its network neighbors were shuffled. The real, i.e. FunCoup-predicted, network was randomized 100 times. In FunCoup, each link is characterized by a confidence value termed a *final Bayesian score*—a sum of individual log likelihood ratios of the integrated data sets (51 sets from seven eukaryotes) that informed on functional coupling. For the analysis, we selected network edges with final Bayesian score 4.8 (natural logarithm), that defined a network of 14 899 genes connected with 709 343 links. After every randomization, connections between a gene of interest i and a gene group j were counted. These values were used to calculate the mean $\hat{n}_{ij}$ and standard deviation σ_{ij}. Together with the respective number of links in the real network n_{ij}, these values produced one-sided Z-scores that estimated significance:

$$Z = \frac{n_{ij} - \hat{n}_{ij}}{\sigma_{ij}}$$

Optionally, either direct or indirect (via a shared network neighbor) connections could be analyzed (results of the latter are given in Table 4). If the analyzed single gene belonged to the gene group, it was discarded as the group member.

SUPPLEMENTARY MATERIAL

Supplementary Material is available at *HMG* online.

Conflict of Interest statement. None declared.

FUNDING

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

REFERENCES

1. Strittmatter, W.J., Saunders, A.M., Schmechel, D., Pericak-Vance, M., Enghild, J., Salvesen, G.S. and Roses, A.D. (1993) Apolipoprotein E: high-avidity binding to beta-amyloid and increased frequency of type 4 allele in late-onset familial Alzheimer disease. *Proc. Natl. Acad. Sci. USA*, **90**, 1977–1981.
2. Viiri, L.E., Loimaala, A., Nenonen, A., Islam, S., Vuori, I., Karhunen, P.J. and Lehtimaki, T. (2005) The association of the apolipoprotein E gene promoter polymorphisms and haplotypes with serum lipid and lipoprotein concentrations. *Atherosclerosis*, **179**, 161–167.
3. Yue, P., Isley, W.L., Harris, W.S., Rosipal, S., Akin, C.D. and Schonfeld, G. (2005) Genetic variants of ApoE account for variability of plasma low-density lipoprotein and apolipoprotein B levels in FHBL. *Atherosclerosis*, **178**, 107–113.
4. Bennet, A.M., Di Angelantonio, E., Ye, Z., Wensley, F., Dahlin, A., Ahlbom, A., Keavney, B., Collins, R., Wiman, B., de Faire, U. *et al.* (2007) Association of apolipoprotein E genotypes with lipid levels and coronary risk. *J Am Med Assoc*, **298**, 1300–1311.
5. Kathiresan, S., Melander, O., Guiducci, C., Surti, A., Burtt, N.P., Rieder, M.J., Cooper, G.M., Roos, C., Voight, B.F., Havulinna, A.S. *et al.* (2008) Six new loci associated with blood low-density lipoprotein cholesterol, high-density lipoprotein cholesterol or triglycerides in humans. *Nat. Genet.*, **40**, 189–197.
6. Corder, E.H., Saunders, A.M., Strittmatter, W.J., Schmechel, D.E., Gaskell, P.C., Small, G.W., Roses, A.D., Haines, J.L. and Pericak-Vance, M.A. (1993) Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. *Science*, **261**, 921–923.
7. Poirier, J., Davignon, J., Bouthillier, D., Kogan, S., Bertrand, P. and Gauthier, S. (1993) Apolipoprotein E polymorphism and Alzheimer's disease. *Lancet*, **342**, 697–699.
8. Strittmatter, W.J., Saunders, A.M., Goedert, M., Weisgraber, K.H., Dong, L.M., Jakes, R., Huang, D.Y., Pericak-Vance, M., Schmechel, D. and

Roses, A.D. (1994) Isoform-specific interactions of apolipoprotein E with microtubule-associated protein tau: implications for Alzheimer disease. *Proc. Natl. Acad. Sci. USA*, **91**, 11183–11186.
9. Beffert, U., Aumont, N., Dea, D., Lussier-Cacan, S., Davignon, J. and Poirier, J. (1999) Apolipoprotein E isoform-specific reduction of extracellular amyloid in neuronal cultures. *Brain Res. Mol. Brain Res.*, **68**, 181–185.
10. Prince, J.A., Zetterberg, H., Andreasen, N., Marcusson, J. and Blennow, K. (2004) APOE epsilon4 allele is associated with reduced cerebrospinal fluid levels of Abeta42. *Neurology*, **62**, 2116–2118.
11. Goate, A., Chartier-Harlin, M.-C., Mullan, M., Brown, J., Crawford, F., Fidani, L., Giuffra, L., Haynes, A., Irving, N., James, L. *et al.* (1991) Segregation of a missense mutation in the amyloid precursor protein gene with familial Alzheimer's disease. *Nature*, **349**, 704–706.
12. Sherrington, R., Rogaev, E.I., Liang, Y., Rogaeva, E.A., Levesque, G., Ikeda, M., Chi, H., Lin, C., Li, G., Holman, K. *et al.* (1995) Cloning of a gene bearing missense mutations in early-onset familial Alzheimer's disease. *Nature*, **375**, 754–760.
13. Scheuner, D., Eckman, C., Jensen, M., Song, X., Citron, M., Suzuki, N., Bird, T.D., Hardy, J., Hutton, M., Kukull, W. *et al.* (1996) Secreted amyloid beta-protein similar to that in the senile plaques of Alzheimer's disease is increased in vivo by the presenilin 1 and 2 and APP mutations linked to familial Alzheimer's disease. *Nat. Med.*, **2**, 864–870.
14. Carter, C.J. (2007) Convergence of genes implicated in Alzheimer's disease on the cerebral cholesterol shuttle: APP, cholesterol, lipoproteins, and atherosclerosis. *Neurochem. Int.*, **50**, 12–38.
15. Bertram, L. and Tanzi, R.E. (2008) 30 years of Alzheimer's disease genetics: the implications of systematic meta-analyses. *Nat. Rev. Neurosci.*, **9**, 768–778.
16. Wollmer, M.A., Sleegers, K., Ingelsson, M., Zekanowski, C., Brouwers, N., Maruszak, A., Brunner, F., Huynh, K.D., Kilander, L., Brundin, R.M. *et al.* (2007) Association study of cholesterol-related genes in Alzheimer's disease. *Neurogenetics*, **8**, 179–188.
17. Reiman, E.M., Webster, J.A., Myers, A.J., Hardy, J., Dunckley, T., Zismann, V.L., Joshipura, K.D., Pearson, J.V., Hu-Lince, D., Huentelman, M.J. *et al.* (2007) GAB2 alleles modify Alzheimer's risk in APOE epsilon4 carriers. *Neuron*, **54**, 713–720.
18. Bertram, L., McQueen, M.B., Mullin, K., Blacker, D. and Tanzi, R.E. (2007) Systematic meta-analyses of Alzheimer disease genetic association studies: the AlzGene database. *Nat. Genet.*, **39**, 17–23.
19. Lee, P.H. and Shatkay, H. (2008) F-SNP: computationally predicted functional SNPs for disease association studies. *Nucleic Acids Res.*, **36**, D820–D824.
20. Dixon, A.L., Liang, L., Moffatt, M.F., Chen, W., Heath, S., Wong, K.C., Taylor, J., Burnett, E., Gut, I., Farrall, M. *et al.* (2007) A genome-wide association study of global gene expression. *Nat. Genet.*, **39**, 1202–1207.
21. Myers, A.J., Gibbs, J.R., Webster, J.A., Rohrer, K., Zhao, A., Marlowe, L., Kaleem, M., Leung, D., Bryden, L., Nath, P. *et al.* (2007) A survey of genetic human cortical gene expression. *Nat. Genet.*, **39**, 1494–1499.
22. Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A., Bender, D., Maller, J., Sklar, P., de Bakker, P.I., Daly, M.J. *et al.* (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.*, **81**, 559–575.
23. Alexeyenko, A. and Sonnhammer, E.L. (2009) Global networks of functional coupling in eukaryotes from comprehensive data integration. *Genome Res.*, **19**, 1107–1116.
24. Harold, D., Abraham, R., Hollingworth, P., Sims, R., Gerrish, A., Hamshere, M.L., Pahwa, J.S., Moskvina, V., Dowzell, K., Williams, A. *et al.* (2009) Genome-wide association study identifies variants at CLU and PICALM associated with Alzheimer's disease. *Nat. Genet.*, **41**, 1088–1093.
25. Marchini, J., Howie, B., Myers, S., McVean, G. and Donnelly, P. (2007) A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat. Genet.*, **39**, 906–913.
26. Willer, C.J., Sanna, S., Jackson, A.U., Scuteri, A., Bonnycastle, L.L., Clarke, R., Heath, S.C., Timpson, N.J., Najjar, S.S., Stringham, H.M. *et al.* (2008) Newly identified loci that influence lipid concentrations and risk of coronary artery disease. *Nat. Genet.*, **40**, 161–169.
27. Kathiresan, S., Willer, C.J., Peloso, G.M., Demissie, S., Musunuru, K., Schadt, E.E., Kaplan, L., Bennett, D., Li, Y., Tanaka, T. *et al.* (2009) Common variants at 30 loci contribute to polygenic dyslipidemia. *Nat. Genet.*, **41**, 56–65.
28. Aulchenko, Y.S., Ripatti, S., Lindqvist, I., Boomsma, D., Heid, I.M., Pramstaller, P.P., Penninx, B.W., Janssens, A.C., Wilson, J.F., Spector, T. *et al.* (2009) Loci influencing lipid levels and coronary heart disease risk in 16 European population cohorts. *Nat. Genet.*, **41**, 47–55.
29. Veyrieras, J.B., Kudaravalli, S., Kim, S.Y., Dermitzakis, E.T., Gilad, Y., Stephens, M. and Pritchard, J.K. (2008) High-resolution mapping of expression-QTLs yields insight into human gene regulation. *PLoS Genet.*, **4**, e1000214.
30. Goh, K.I., Cusick, M.E., Valle, D., Childs, B., Vidal, M. and Barabasi, A.L. (2007) The human disease network. *Proc Natl Acad Sci USA*, **104**, 8685–8690.
31. Calvano, S.E., Xiao, W., Richards, D.R., Felciano, R.M., Baker, H.V., Cho, R.J., Chen, R.O., Brownstein, B.H., Cobb, J.P., Tschoeke, S.K. *et al.* (2005) A network-based analysis of systemic inflammation in humans. *Nature*, **437**, 1032–1037.
32. Chen, Y., Zhu, J., Lum, P.Y., Yang, X., Pinto, S., MacNeil, D.J., Zhang, C., Lamb, J., Edwards, S., Sieberts, S.K. *et al.* (2008) Variations in DNA elucidate molecular networks that cause disease. *Nature*, **452**, 429–435.
33. Jensen, L.J., Kuhn, M., Stark, M., Chaffron, S., Creevey, C., Muller, J., Doerks, T., Julien, P., Roth, A., Simonovic, M. *et al.* (2009) STRING 8–a global view on proteins and their functional interactions in 630 organisms. *Nucleic Acids Res.*, **37**, D412–D416.
34. Katoh, Y., Imakagura, H., Futatsumori, M. and Nakayama, K. (2006) Recruitment of clathrin onto endosomes by the Tom1-Tollip complex. *Biochem. Biophys. Res. Commun.*, **341**, 143–149.
35. Lichtenstein, P., De Faire, U., Floderus, B., Svartengren, M., Svedberg, P. and Pedersen, N.L. (2002) The Swedish Twin Registry: a unique resource for clinical, epidemiological and genetic studies. *J. Intern. Med.*, **252**, 184–205.
36. Reynolds, C.A., Hong, M.G., Eriksson, U.K., Blennow, K., Bennet, A.M., Johansson, B., Malmberg, B., Berg, S., Wiklund, F., Gatz, M. *et al.* (2009) A survey of ABCA1 sequence variation confirms association with dementia. *Hum. Mutat.*, **30**, 1348–1354.
37. Pedersen, N.L., Friberg, L., Floderus-Myrhed, B., McClearn, G.E. and Plomin, R. (1984) Swedish early separated twins: identification and characterization. *Acta Genet. Med. Gemellol. (Roma)*, **33**, 243–250.
38. McClearn, G.E., Johansson, B., Berg, S., Pedersen, N.L., Ahern, F., Petrill, S.A. and Plomin, R. (1997) Substantial genetic influence on cognitive abilities in twins 80 or more years old. *Science*, **276**, 1560–1563.
39. Gold, C.H., Malmberg, B., McClearn, G.E., Pedersen, N.L. and Berg, S. (2002) Gender and health: a study of older unlike-sex twins. *J. Gerontol. B Psychol. Sci. Soc. Sci.*, **57**, S168–S176.
40. Gatz, M., Fratiglioni, L., Johansson, B., Berg, S., Mortimer, J.A., Reynolds, C.A., Fiske, A. and Pedersen, N.L. (2005) Complete ascertainment of dementia in the Swedish Twin Registry: the HARMONY study. *Neurobiol. Aging*, **26**, 439–447.
41. Katzov, H., Chalmers, K., Palmgren, J., Andreasen, N., Johansson, B., Cairns, N.J., Gatz, M., Wilcock, G.K., Love, S., Pedersen, N.L. *et al.* (2004) Genetic variants of ABCA1 modify Alzheimer disease risk and quantitative traits related to beta-amyloid metabolism. *Hum. Mutat.*, **23**, 358–367.
42. Hamon, S.C., Kardia, S.L., Boerwinkle, E., Liu, K., Klos, K.L., Clark, A.G. and Sing, C.F. (2006) Evidence for consistent intragenic and intergenic interactions between SNP effects in the APOA1/C3/A4/A5 gene cluster. *Hum. Hered.*, **61**, 87–96.
43. Goto, S., Bono, H., Ogata, H., Fujibuchi, W., Nishioka, T., Sato, K. and Kanehisa, M. (1997) Organizing and computing metabolic pathway data in terms of binary relations. *Pac. Symp. Biocomput.*, **1997**, 175–186.
44. Barrett, J.C., Fry, B., Maller, J. and Daly, M.J. (2005) Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics*, **21**, 263–265.
45. Andreasen, N., Minthon, L., Clarberg, A., Davidsson, P., Gottfries, J., Vanmechelen, E., Vanderstichele, H., Winblad, B. and Blennow, K. (1999) Sensitivity, specificity, and stability of CSF-tau in AD in a community-based patient sample. *Neurology*, **53**, 1488–1494.
46. Vanmechelen, E. and Vanderstichele, H. (1998) Towards an earlier diagnosis of Alzheimer's disease. *J. Biotechnol.*, **66**, 229–231.
47. Blennow, K., Wallin, A., Agren, H., Spenger, C., Siegfried, J. and Vanmechelen, E. (1995) Tau protein in cerebrospinal fluid: a biochemical

marker for axonal degeneration in Alzheimer disease? *Mol. Chem. Neuropathol.*, **26**, 231–245.
48. Carey, V., Zeger, S.L. and Diggle, P. (1993) Modelling multivariate binary data with alternating logistic regressions. *Biometrika*, **80**, 517–526.
49. Jansson, M., Gatz, M., Berg, S., Johansson, B., Malmberg, B., McClearn, G.E., Schalling, M. and Pedersen, N.L. (2003) Association between depressed mood in the elderly and a 5-HTR2A gene variant. *Am. J. Med. Genet. B Neuropsychiatr. Genet.*, **120B**, 79–84.
50. Hong, M.G., Myers, A.J., Magnusson, P.K. and Prince, J.A. (2008) Transcriptome-wide assessment of human brain and lymphocyte senescence. *PLoS One*, **3**, e3024.
51. Maslov, S. and Sneppen, K. (2002) Specificity and stability in topology of protein networks. *Science*, **296**, 910–913.