Coding Mutations in *SORL1* and Alzheimer Disease

Badri N. Vardarajan, PhD,[1,2] Yalun Zhang, PhD,[3] Joseph H. Lee, DPH,[1,2,4] Rong Cheng, PhD,[1,2] Christopher Bohm, PhD,[3] Mahdi Ghani, PhD,[3] Christiane Reitz, MD, PhD,[1,2,5] Dolly Reyes-Dumeyer, BA,[1,2] Yufeng Shen, PhD,[6] Ekaterina Rogaeva, PhD,[3] Peter St George-Hyslop, MD,[3,4] and Richard Mayeux, MD[1,2,5,7,8]

Objective: Common single nucleotide polymorphisms in the *SORL1* gene have been associated with late onset Alzheimer disease (LOAD), but causal variants have not been fully characterized nor has the mechanism been established. The study was undertaken to identify functional *SORL1* mutations in patients with LOAD.
Methods: This was a family- and cohort-based genetic association study. Caribbean Hispanics with familial and sporadic LOAD and similarly aged controls were recruited from the United States and the Dominican Republic, and patients with sporadic disease of Northern European origin were recruited from Canada. Prioritized coding variants in *SORL1* were detected by targeted resequencing and validated by genotyping in additional family members and unrelated healthy controls. Variants transfected into human embryonic kidney 293 cell lines were tested for Aβ40 and Aβ42 secretion, and the amount of the amyloid precursor protein (APP) secreted at the cell surface was determined.
Results: Seventeen coding exonic variants were significantly associated with disease. Two rare variants (rs117260922-E270K and rs143571823-T947M) with minor allele frequency (MAF) $< 1\%$ and 1 common variant (rs2298813-A528T) with MAF $= 14.9\%$ segregated within families and were deemed deleterious to the coding protein. Transfected cell lines showed increased Aβ40 and Aβ42 secretion for the rare variants (E270K and T947M) and increased Aβ42 secretion for the common variant (A528T). All mutants increased the amount of APP at the cell surface, although in slightly different ways, thereby failing to direct full-length APP into the retromer-recycling endosome pathway.
Interpretation: Common and rare variants in *SORL1* elevate the risk of LOAD by directly affecting APP processing, which in turn can result in increased Aβ40 and Aβ42 secretion.

ANN NEUROL 2015;77:215–227

SORL1 (sortilin-related receptor, L[DLR class] A-type repeats containing) is a member of the vacuolar protein sorting-10 domain–containing receptor family, and participates in the intracellular vesicular sorting of amyloid precursor protein (APP) after reinternalization from the cell surface.[1,2] *SORL1* determines whether APP is sorted in the retromer recycling–endosome pathway or allowed to drift into the endosome–lysosome pathway where it is cleaved to generate Aβ. Variants in the *SORL1* gene might alter this activity, leading to an increase in Aβ that, in turn, contributes to the pathogenesis of late onset Alzheimer disease (LOAD).[3] To date, despite compelling evidence from case–control, family-based, and genome-wide association studies (GWAS),[3–11] clearly pathogenic variants have not been identified, making it difficult to investigate the functional consequences of specific *SORL1* mutations.

Subjects and Methods

Targeted Resequencing and Analysis Methods

SAMPLE SELECTION AND PREPARATION. We sequenced 1 affected individual with LOAD, usually the proband, from 151

View this article online at wileyonlinelibrary.com. DOI: 10.1002/ana.24305

Received May 19, 2014, and in revised form Oct 10, 2014. Accepted for publication Oct 17, 2014.

Address correspondence to Dr. Mayeux, Taub Institute for Research of Alzheimer's Disease and the Aging Brain, Columbia University, 622 West 168th Street, New York, NY 10032

From the [1]Taub Institute for Research on Alzheimer's Disease and the Aging Brain and [2]Gertrude H. Sergievsky Center, Columbia University, New York, NY; [3]Tanz Centre for Research in Neurodegenerative Diseases and Department of Medicine, University of Toronto, Toronto, Ontario, Canada; [4]Cambridge Institute for Medical Research, Department of Clinical Neurosciences, University of Cambridge, Cambridge, United Kingdom; Departments of [5]Neurology, [6]Systems Biology, and [7]Psychiatry, College of Physicians and Surgeons, Columbia University, New York, NY; and [8]Department of Epidemiology, School of Public Health, Columbia University, New York, NY.

families with multiple affected family members. The mean age at onset for affected subjects was 77.03 years (standard deviation [SD] = 8.93), ranging from 45 to 98 years; 69.5% of the family members were women, and mean education was 4.3 years (SD = 4.61). We extracted genomic DNA from whole blood with 0.16% samples from saliva. Blood samples were extracted using the Qiagen method, and saliva samples were extracted using the Oragene method. The DNA was then quantified using the PicoGreen detection method, following the manufacturer specifications (Invitrogen, Carlsbad, CA).

We validated the prioritized variants by genotyping the sequenced probands and their 464 relatives, of whom 350 were affected and 114 were unaffected. For the sequencing experiment, we pooled DNA samples using 235 samples across 24 pools, with each pool comprising 10 unrelated samples (5 samples failed sequencing).

TARGETED RESEQUENCING. We performed RainDance (http://raindancetech.com/targeted-dna-sequencing) for capture and then followed with pooled sequencing using the Illumina GAII platform (http://www.illumina.com). In total, we sequenced 201,510 base pairs (bp), including both exons and introns of the *SORL1* gene as well as the flanking region, covering from 121,312,961bp to 121,514,471bp.

VARIANT CALLING AND POSTPROCESSING. We aligned the reads obtained from the pooled sequencing to the human reference genome build 37 using the Burrows–Wheeler Aligner[12] (http://bio-bwa.sourceforge.net/). Quality control of the sequencing data was done using established methods, including base alignment quality calibration and refinement of local alignment around putative indels using the Genome Analysis Toolkit.[13] We used SAMTOOLS[14] mpileup to call variants in the pooled data set and validated calls by an independent calling algorithm called CRISP (Comprehensive Read analysis for Identification of Single Nucleotide Polymorphisms [SNPs] from Pooled sequencing).[15] Variant calls were filtered using mpileup filters for base quality (baseQ bias), mapping quality (mapQ bias), strand bias, tail distance bias, and number of nonreference reads to obtain high-quality variants. Reliably called variants were annotated by ANNOVAR,[16] including in silico functional prediction using POLYPHEN[17] software and extent of cross-species conservation using PHYLOP.[18]

GENOTYPING. To validate novel variants discovered in probands, we genotyped the probands and their family members. To investigate whether the allele frequencies for novel variants differed from unaffected persons in the general Caribbean Hispanic population, we genotyped 498 unaffected persons who were unrelated to any of the family members. These 498 individuals underwent the same phenotypic and diagnostic protocols. Genotyping was conducted on the Sequenom platform. When the Sequenom platform failed to generate genotype due to difficulties with primers, we performed Sanger sequencing.

STATISTICAL ANALYSIS. To assess whether a set of rare and common variants in *SORL1* increases the risk of LOAD, we performed a gene-wise analysis using in the SNP-set Kernel Association Test (SKAT)[19] for heterozygous variants in exons and introns with and without adjustments for covariates such as age, sex, and *APOE* genotype. We also used statically estimated haplotypes coupled with generalized estimating equations (GEE) to establish joint burden of 17 single nucleotide variants (SNVs) by accurately adjusting for the correlation between samples. To assess the individual effects of SNPs, we performed joint linkage and association analysis with PSEUDOMARKER[20] using all family members and unrelated controls. This analytical method allows us to analyze family data, unrelated subjects, or both to determine whether a variant is associated with disease. For constructing haplotypes, we used the R-based haplo.stats package[21] (http://mayoresearch.mayo.edu/mayo/research/schaid_lab/software.cfm).

Functional Studies

SITE-DIRECTED MUTAGENESIS. *SORL1* E270K, A528T, and T947M mutations were generated by site-directed mutagenesis using human SORL1-MYC pcDNA3.1 as a backbone according to the manufacturer's instructions.[1,2] All mutant constructs were verified by sequencing.

CELL CULTURE AND TRANSFECTION. Human embryonic kidney 293 (HEK293) cells stably expressing the Swedish APP mutant (APPsw)[22] were maintained in Dulbecco modified Eagle medium (Gibco, Grand Island, NY) with 10% fetal bovine serum and Geneticin (200μg/ml). Wild-type SORL1-MYC pcDNA3.1 and 3 generated SORL1 mutant constructs (SORL1 E270K-MYC pcDNA3.1, SORL1 A528T-MYC pcDNA3.1, SORL1 T947M-MYC pcDNA3.1) were transfected transiently into HEK293 APPsw cells using Lipofectamine 2000 (Invitrogen). Stable clones were selected using Hygromycin (200μg/ml) and Geneticin (400μg/ml) to generate stable cell lines overexpressing either wild-type or mutant *SORL1*.

Aβ ASSAYS. Measurement of secreted Aβ40, Aβ42, and sAPPβ from culture medium in HEK293 APPsw cells,[23] and wild-type *SORL1* and mutant *SORL1* stable HEK293 APPsw cells was made by sandwich enzyme-linked immunosorbent assay according to manufacturer's protocol.

ANTIBODIES AND WESTERN BLOT. Antibodies were used as follows: rabbit antibody to the C-terminus of SORL1 (S9200; Sigma, St Louis, MO), rabbit polyclonal antibody to PS1 N-terminal fragment (NTF; A4, from our laboratory), mouse monoclonal anti–c-MYC (Invitrogen), rabbit polyclonal antibody to the C terminus of APP (Ab365, Sigma), and mouse monoclonal anti-Aβ (6E10; Covance, Princeton, NJ).

Culture medium from HEK293 APPsw cells, and wild-type SORL1 and mutant SORL1 stable cell lines were harvested and subjected to immunoblotting. Secreted sAPPα levels were analyzed by Western blot using anti-Aβ (6E10); samples were normalized to the protein concentration of the collected cell lysates, which were measured by BCA protein assay (Bio-Rad Laboratories, Hercules, CA). The cell lysates were analyzed in a Western blot with full-length APP (FL-APP), PS1-NTF (presenilin 1), and APP C-terminal fragments

TABLE 1. List of Variants Prioritized for Follow-up Genotyping and Probability Values Generated by Linkage and Association Test Implemented in PSEUDOMARKER

SNP	BP, HG19	*p*, LD + Link	A1[a]	A2	Caribbean Hispanic Frequency		ADNI Omni Chip Frequency		ADNI WGS Frequency		Function[b]	AA Change	ESP Frequency	Polyphen[c]	Conservation[d]
					Controls, n = 498	Affected, n = 87 Families, 462	AD + MCI, n = 531	Control, n = 82	AD + MCI, n = 531	Control, n = 281					
rs117260922	121367627	7.68E-07	1	3	0.010040	0.01166					NS	E270K	0.007529	D	C
rs2298813	121393684	6.09E-07	1	3	0.091370	0.15598	0.05849	0.047	0.05039	0.04029	NS	A528T	0.061536	P	C
11–121428111	121428111	1.49E-03	3	1	0.001004	0.00146					NS	E887G		B	C
rs143571823	121429476	7.00E-06	4	2	0.007028	0.00729			0.00000	0.00183	NS	T947M	0.004927	P	C
11–121437722	121437722	5.00E-05	2	3	0.002008	0.00583					NS	R1041S		B	C
rs1699107	121437819	1.41E-10	2	3	0.019080	0.02770			0.00194	0.00183	S	Q1074Q			
rs146903951	121440937	6.00E-06	2	4	0.007042	0.01020			0.01453	0.01648	NS	F1099L	0.007436	B	C
rs62617129	121444958	2.00E-03	3	1	0.003012	0.00583			0.00484	0.00366	NS	I1116V	0.006228	B	N
rs114830255	121454206	4.00E-06	1	3	0.001004	0.00875			0.00097	0.00183	NS	R1207Q	0.005298	B	N
11–121458818	121458818	1.16E-04	0[e]	4	0.000000	0.00292					FSD	C1302fs			
rs146353234	121461799	8.96E-07	4	1	0.002008	0.00583			0.00000	0.00366	NS	T1435S	0.000744	B	C
11–121474914	121474914	2.70E-03	4	2	0.002008	0.00292					NS	T1511I		B	C
11–121476260	121476260	8.86E-04	0	2	0.000000	0.00437					FSD	T1643fs			
rs62622819	121485599	1.50E-04	1	4	0.003012	0.00875			0.00969	0.00916	NS	H1813Q	0.006042	P	N
rs1792120	121491782	7.42E-10	3	1	0.013050	0.03353			0.00194	0.00183	NS	V1967I	0.0002	B	N
rs74811057	121495870	1.74E-03	3	1	0.014060	0.00583			0.00000	0.00183	NS	K2083R	0.013664	B	C
11–121498387	121498387	9.82E-04	0	3	0.000000	0.00437					FSD	R2163fs			

[a]Minor allele.
[b]SNV function: FSD = frame-shift deletion; NS = nonsynonymous SNV; S = synonymous SNV.
[c]Polyphen prediction: B = benign; D = damaging; P = possibly damaging.
[d]Phylop conservation prediction: C = conserved, N = not conserved.
[e]Homozygous wild type.
AA = Amino Acid; AD = Alzheimer disease; ADNI = Alzheimer's Disease Neuroimaging Initiative; BP = Base Pair; ESP = The NHLBI GO Exome Sequencing Project; HG19 = UCSC version of the Human Genome-hg19; LD = Linkage Disequilibrium; MCI = mild cognitive impairment; SNP = single nucleotide polymorphism; SNV = single nucleotide variant; WGS = whole genome sequencing.

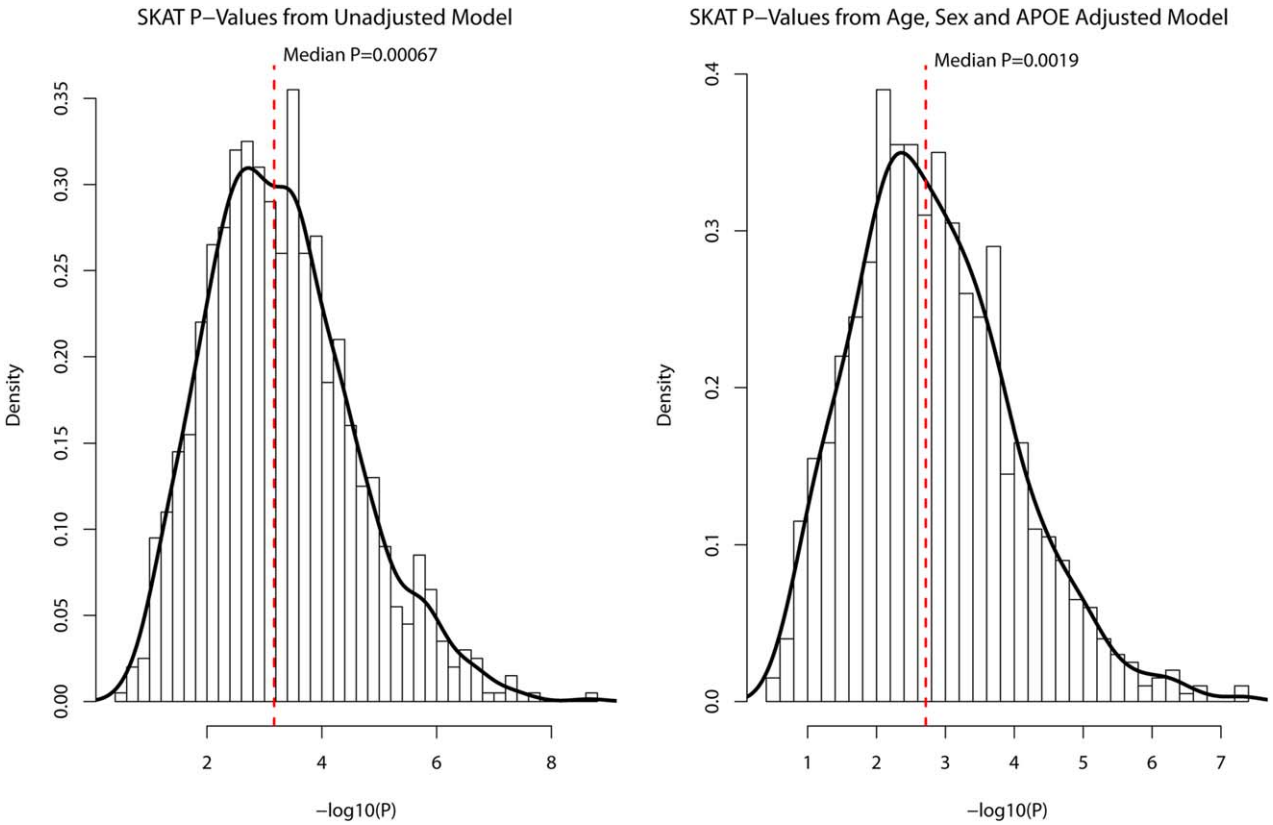

FIGURE 1: Histogram of −log10 of the probability values obtained from SNP-set Kernel Association Test (SKAT) analysis of 1,000 data sets created by randomly choosing 1 subject from each of the 87 families and 498 controls. The SKAT analysis was conducted assuming for the unadjusted model: Alzheimer disease (AD) ∼ single nucleotide polymorphism (SNP) burden; and for the model with age, sex, and apolipoprotein E (APOE) ε4 status as covariates: AD ∼ SNP burden + age + sex + APOE ε4 yes/no. [Color figure can be viewed in the online issue, which is available at www.annalsofneurology.org.]

(CTF; APP-β-CTF[C83] and APP-α-CTF[C99]). Band intensities were quantified using NIH ImageJ software, and relative expression levels of FL-APP, total APP-CTFs, and PS1 were normalized to β-actin. Bar graphs were normalized to wild-type SORL1 control.

CELL SURFACE BIOTINYLATION. Cells were washed with buffer A (phosphate-buffered saline with 1mM $MgCl_2$, pH 8.0) and incubated with 1mg/ml Sulfo-NHS-LC-Biotin (Sigma) in buffer A for 20 minutes at 4°C to prevent internalization. Cells were then washed with ice-cold 20mM glycine in buffer A and lysed, and biotinylated proteins were precipitated with NeutrAvidin beads (Thermo Scientific, Waltham, MA).

Protein lysates were immunoblotted with anti–C-terminal APP antibody (Ab365, Sigma) and anti–C-terminal SORL1 antibody (S9200, Sigma). Immunoprecipitated cell surface APP was normalized to total APP (input). Western blot band intensities were measured with NIH ImageJ software. Bar graphs were normalized to wild-type control.

COIMMUNOPRECIPITATION. Cells were lysed in 1% CHAPSO buffer,[24] immunoprecipitated using G Plus beads with 2μg mouse monoclonal anti–c-MYC antibody (for immunoprecipitation of SORL1-myc), and immunoblotted with anti–C-terminal APP antibody (Ab365), and anti–C-terminal SORL1 (S9200). FL-APP coprecipitated with c-MYC antibody was quantified and normalized to the amount of immunoprecipitated SORL1.

STATISTICAL ANALYSIS. Prism 5 statistical software (GraphPad, San Diego, CA) was used to generate bar charts, and analysis of variance with *t* test was used to analyze statistical difference, followed by Bonferroni correction.

Results

Genetic Analyses

Analysis of the sequence data allowed prioritization of 17 exonic coding variants, including 13 nonsynonymous mutations, 3 frame-shift deletions, and 1 synonymous mutation (Table 1). We validated variant calls by Sequenom genotyping in the sequenced probands, additional family members from 87 families that contained at least 1 heterozygous carrier (464 total familial subjects: 350 affected, 114 unaffected), and 498 unrelated, age-matched Caribbean Hispanic controls. The combined gene burden SKAT[19] analysis confirmed that the joint burden of 17 heterozygous variants was significantly associated with

TABLE 2. Haplotype Analysis SORL1 Coding Mutations: Frequency of Haplotype Combinations of the 17 SNPs (Haplo Stats)

Haplotype	SNP1	SNP2	SNP3	SNP4	SNP5	SNP6	SNP7	SNP8	SNP9	SNP10	SNP11	SNP12	SNP13	SNP14	SNP15	SNP16	SNP17	HF
1	1	3	1	2	3	2	4	1	3	4	1	2	2	4	1	1	3	0.00028
2	1	3	1	2	3	3	4	1	3	4	1	2	2	4	1	1	3	0.01065
3	3	1	1	2	2	3	4	1	3	4	1	2	2	4	1	1	3	0
4	3	1	1	2	3	2	4	1	3	4	1	2	2	4	1	1	3	0.00161
5	3	1	1	2	3	3	4	1	3	4	1	2	2	4	1	1	1	0.00156
6	3	1	1	2	3	3	4	1	3	4	1	2	2	4	1	1	3	0.11336
7	3	1	1	2	3	3	4	1	3	4	1	2	2	4	3	1	3	0.00001
8	3	1	1	4	3	3	4	1	3	4	1	2	2	4	1	1	3	0.00156
9	3	3	1	2	2	3	4	1	3	4	1	2	2	4	1	1	3	0.00416
10	3	3	1	2	3	2	4	1	3	4	1	2	1	4	1	1	3	0.00156
11	3	3	1	2	3	2	4	1	3	4	1	2	2	1	1	1	3	0.00092
12	3	3	1	2	3	2	4	1	3	4	1	2	2	4	1	1	3	0.01759
13	3	3	1	2	3	2	4	1	3	4	1	2	2	4	1	3	3	0.00093
14	3	3	1	2	3	3	2	1	3	4	1	2	2	4	1	1	3	0.00938
15	3	3	1	2	3	3	4	1	1	4	1	2	2	4	1	1	3	0.00469
16	3	3	1	2	3	3	4	1	1	4	1	4	2	4	1	1	3	0.00103
17	3	3	1	2	3	3	4	1	3	2	1	2	2	4	1	1	3	0.00104
18	3	3	1	2	3	3	4	1	3	4	1	2	2	1	1	1	3	0.00533
19	3	3	1	2	3	3	4	1	3	4	1	2	2	4	1	1	3	0.78103
20	3	3	1	2	3	3	4	1	3	4	1	2	2	4	1	3	3	0.00481
21	3	3	1	2	3	3	4	1	3	4	1	2	2	4	3	1	3	0.02184
22	3	3	1	2	3	3	4	1	3	4	1	4	2	4	1	1	3	0.00157
23	3	3	1	2	3	3	4	1	3	4	4	2	2	4	1	1	3	0.00364
24	3	3	1	2	3	3	4	3	3	4	1	2	2	4	1	1	3	0.00468

TABLE 2. Continued

Haplotype	SNP1	SNP2	SNP3	SNP4	SNP5	SNP6	SNP7	SNP8	SNP9	SNP10	SNP11	SNP12	SNP13	SNP14	SNP15	SNP16	SNP17	HF
25	3	3	1	4	3	2	4	1	3	4	1	2	2	4	1	1	3	0.00052
26	3	3	1	4	3	3	4	1	3	4	1	2	2	4	1	3	3	0.00416
27	3	3	1	4	3	3	4	1	3	4	1	2	2	4	3	1	3	0.00052
28	3	3	3	2	3	3	4	1	3	4	1	2	2	4	1	1	3	0.00105
29	3	3	3	2	3	3	4	1	3	4	1	2	2	4	1	3	3	0.00051

Subjects were coded based on haplotype copies of haplotype 19 (0 = 2 copies of 19, 1 = 1 copy of 19, 2 = 0 copies of 19). Subjects with haplotype pairs estimated with posterior probability = 1 were used for the association analysis.
HF = haplotype frequency; SNP = single nucleotide polymorphism.

LOAD ($p_{\text{unadjusted}} = 0.0009$; $p_{\text{adjusted for age and gender covariates}} = 0.0079$). The SKAT test assumes independence of observations but does not adjust for familial correlation. Thus, we conducted SKAT analysis on unrelated subjects creating a data set by randomly selecting 1 member from each of the 87 families, and combined them with the 498 controls to create a case–control set. We repeated this process 1,000 times to create 1,000 case–control data sets and conducted SKAT analysis using unadjusted and age-, sex-, and *APOE*-adjusted models. A total of 961 of 1,000 (96%) unadjusted model data sets and 909 of 1,000 (91%) adjusted model data sets produced significant probability values ($p < 0.05$). We observed median probability values of $p = 0.00067$ for the unadjusted model and $p = 0.002$ for the adjusted model, respectively (Fig 1). These observations are consistent with the SKAT analysis using all family members. In case of a null association, we would have expected 5% of the data sets to produce nominally significant probability values. The significant deviation from the expectation provides further evidence of the joint burden of 17 SNVs in modifying LOAD risk.

Because of lack of appropriate methods for gene- or region-based burden methods for dichotomous traits that adjust for familial correlations, we performed additional haplotype analyses to assess the joint association of the 17 SNPs with LOAD and related traits. Defining the major allele as the most frequent haplotype observed in 78% of the samples (Table 2) and combining the remaining haplotypes into the minor allele, we computed association with LOAD using GEE. We included 933 (of 962) subjects in the association analysis with haplotype pairs estimated at a posterior probability of $p = 1$. The rare haplotypes increased disease risk and were strongly associated with LOAD (odds ratio = 1.9, $p = 6.9\text{E-}05$; Table 3). This observation is consistent with the increased frequency of the minor alleles of several of the 17 SNPs in LOAD versus controls (see Table 1).

To assess individual significance of the SNVs, we conducted joint linkage and association of the 17 variants with LOAD in the subset of 87 families and the unrelated controls. The analysis revealed that all 17 SNVs were significantly associated with disease at a Bonferroni-corrected probability value of $p < 0.0029$. However, 3 of the variants showed significant segregation with disease under a dominant affected subjects only model: rs2298813 (A528T, $p = 6.09\text{E-}7$), rs117260922 (E270K, $p = 7.68\text{E-}7$), and rs143571823 (T947M, $p = 7.0\text{E-}6$). Variant rs2298813 was most frequent, being present in 54 families, in contrast to variant rs117260922, which was detected in 7 families, and variant rs143571823, which was detected in 4 families.

TABLE 3. Haplotype Analysis SORL1 Coding Mutations: Association Test of Haplotype 19 Using Generalized Estimation Equations

	Beta	SE	*z*	*p*
h19haplotype	0.643541	0.16	15.83	6.91E-05

Subjects with haplotype pairs estimated with posterior probability = 1 were used for the association analysis.
SE = standard error.

To assess whether these findings were applicable to ethnic groups other than Caribbean Hispanics, we also resequenced *SORL1* in 211 patients of Northern European ancestry (Table 4). We detected 13 rare missense variations and a 3bp deletion eliminating a highly conserved residue *p*.N174 (Table 5). Seven of these variations are predicted to be damaging, including 3 novel variations. Of the 14 rare variations identified, 7 overlapped with the mutations detected in the Caribbean Hispanic patients, including 2 of the coding mutations, rs2298813 and rs117260922. Their frequencies were higher than or comparable to the Caucasian population in the 1000 Genomes database, but much lower than observed in Caribbean Hispanics.

We also compared the minor allele frequencies of the 17 coding-*SORL1* SNVs discovered in the Hispanics with those observed in the whole genome sequencing and the exome chip data from the Alzheimer's Disease Neuroimaging Initiative (ADNI) data set (https://ida.loni.usc.edu/login.jsp?project = ADNI&page = HOME).[25] We used baseline phenotypes from ADNI samples to compute frequencies of LOAD and mild cognitive impairment compared to controls. The frequency of the common SNP rs2298813 (A528T; see Table 1) was concordant with the observations in the Hispanic cohort, but the allele frequencies were much lower than in the Caribbean Hispanics. The rare SNP rs143571823 (T947M) was heterozygous in 1 ADNI control and was not found in any case. SNP rs117260922 (E270K) was not observed in the entire ADNI data set. Differences in allele frequencies between Caribbean Hispanics and the Caucasians in the ADNI study could be conferred by differences in sequencing technologies, capture platforms, sequencing depth, and variant calling algorithms in the 2 experiments. We evaluated the effects of the 45 rare *SORL1* missense mutations observed in the ADNI data set at a sample minor allele frequency < 0.01 using the SKAT test. The SKAT test of rare missense mutations in demented versus healthy controls in the ADNI samples was significant ($p = 0.037$).

Caribbean Hispanics are known to be an admixed population; therefore, we also investigated the association of the rs2298813 in a meta-analyses LOAD study of African Americans.[26] The SNP was significant in African Americans at $p = 0.01$ and is observed with a higher frequency in cases compared with controls.

Functional Analyses

We tested the impact of 3 most significant *SORL1* mutations on Aβ production. Clonal HEK293sw cell lines stably overexpressing similar quantities of wild-type and mutant *SORL1* were generated. Aβ40 and Aβ42 levels were then measured in conditioned media from the cells. Wild-type and mutant *SORL1* were expressed at the same levels, yet the E270K and T947M mutants both resulted in a significant increase in Aβ40 secretion (E270K, $171 \pm 5.6\%$ of control value, $p < 0.001$; T947M, $202 \pm 11.6\%$ of control value, $p < 0.01$; n = 3 independent replications) and Aβ42 secretion (E270K, $214 \pm 5.7\%$ of control value, $p < 0.001$; T947M, $221 \pm 8.4\%$ of control value, $p < 0.001$; n = 3 independent replications; Fig 2). The A528T mutant increased Aβ42 secretion moderately ($158 \pm 11.1\%$ of control value, $p < 0.01$; n = 3 replications), but did not change the Aβ40 secretion ($103 \pm 3.3\%$ of control value, $p > 0.05$; n = 3 independent replications).

TABLE 4. Demographic and Clinical Characteristics of Sequenced Individuals

Characteristics	Caribbean Hispanic Affected Subjects, n = 154	Caribbean Hispanic Unaffected Subjects, n = 80	Northern European Caucasian Affected Subjects, n = 211
Mean age at onset or last examination, yr ± SD	77.0 ± 8.9	83.9 ± 3.8	73.0 ± 7.8
Mean years of education ± SD	4.3 ± 4.6	7.0 ± 4.0	Not available
Women, No. (%)	107 (69.5)	57 (71.3)	107 (50.7)
APOE ε4, %	22.8	11.9	38.0

APOE = apolipoprotein E; SD = standard deviation.

TABLE 5. Coding *SORL1* Variations Detected in 211 Cases of North European Ancestry

SNP	BP, HG19	A1[a]	A2	Cases, No.	MAF	1,000g-CEU Frequency	ESP Frequency	Function[b]	AA Change	Polyphen[c]	Conservation[d]
11-121348942	del 121348942-121348944	del	ACA	2	0.005	NA	NA	Deletion	173_174del	NA	C
rs117260922[e]	121367627	A	G	14	0.033	0.0059	0.007459	NS	E270K	D	C
rs150609294	121384931	C	A	1	0.002	0	0.001	NS	N371T	D	C
rs2298813[e]	121393684	A	G	18	0.045	0.0471	0.060366	NS	A528T	P	C
rs146903951[e]	121440937	C	T	3	0.007	0.0059	0.007921	NS	F1099L	B	C
rs62617129[e]	121444958	G	A	3	0.007	0.0059	0.006075	NS	I1116V	B	N
11-121458817	121458817	T	G	1	0.002	NA	NA	NS	Q1301H	B	C
11-121460792	121460792	G	T	1	0.002	NA	NA	NS	F1374L	D	C
rs199717181	121474988	A	G	1	0.002	0	0.000231	NS	G1536S	D	C
rs138580875	121475859	C	G	1	0.002	NA	0.000615	NS	W1563C	D	C
rs62622819[e]	121485599	A	T	5	0.012	0.0235	0.005999	NS	H1813Q	P	N
rs1792120[e]	121491782	G	A	1	0.002	0	0.01669	NS	V1967I	B	N
rs140327834	121495816	T	A	2	0.005	0	0.00423	NS	D2065V	D	C
rs74811057[e]	121495870	G	A	1	0.002	0	0.013229	NS	K2083R	B	C

[a]Minor allele.
[b]SNV function: NS = nonsynonymous SNV.
[c]Polyphen prediction: B = benign; D = damaging; P = possibly damaging.
[d]Phylop conservation prediction: C = conserved; N = not conserved.
[e]Found in Caribbean Hispanics.
1000g-CEU- Central European Ancestry subset of the 1000 Genomes Project.
AA = Amino Acid; ESP = The NHLBI GO Exome Sequencing Project; HG19 = UCSC version of the Human Genome-hg19; MAF = minor allele frequency; NA = not available; SNP = single nucleotide polymorphism; SNV = single nucleotide variant.

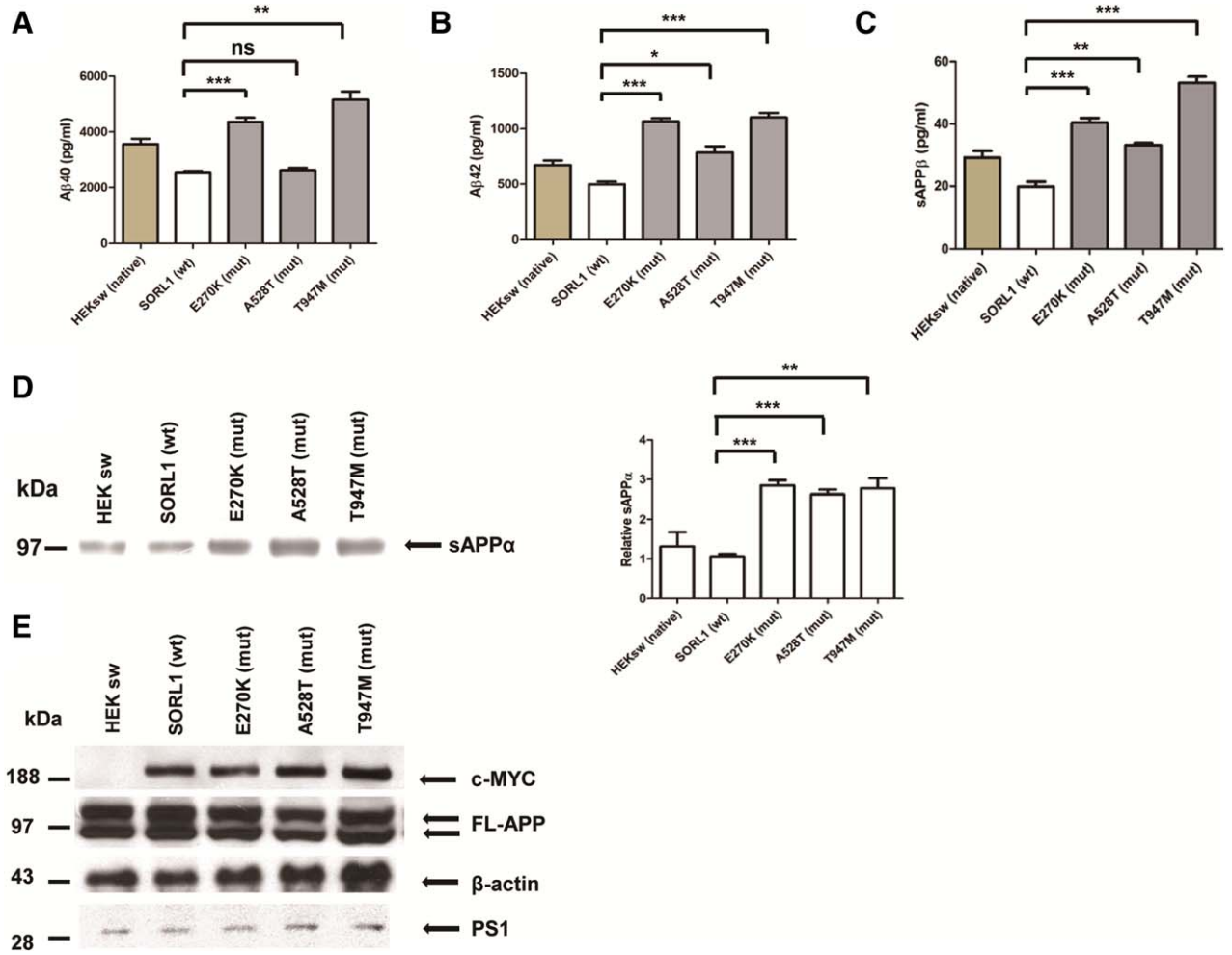

FIGURE 2: Overexpression of SORL1 mutants leads to elevated Aβ secretion. (A–C) Measurement of secreted Aβ40, Aβ42 and sAPPβ from culture medium in stable HEK293 cells expressing the APP Swedish mutant (HEKsw) together with either wild-type (wt) SORL1 or mutant (mut) SORL1. Aβ levels were normalized to the protein levels of the cell lysates. Error bars = standard error of the mean. *$p<0.05$, **$p<0.01$, ***$p<0.001$, ns = not significant after Bonferroni correction; n = 3 independent replications. (D) Cultured media from cells were collected and subjected to Western blot and probed with 6E10 antibody to detect sAPPα. Bar graphs were normalized to control. **$p<0.01$, ***$p<0.001$, after Bonferroni correction; n = 3 independent replications. (E) Cell lysates were harvested to perform Western blot of full-length amyloid precursor protein (FL-APP) and PS1. β-Actin was used as loading control; n = 3 independent replications.

All 3 mutations caused significant increases in sAPPα and sAPPβ secretion compared to wild-type *SORL1*. Thus, values were for sAPPα: E270K, $266 \pm 13.0\%$ of control value, $p < 0.001$; A528T, $246 \pm 12.0\%$ of control value, $p < 0.001$; T947M, $259 \pm 25.2\%$ of control value, $p < 0.01$; n = 3 independent replications; and for sAPPβ: E270K, $204 \pm 7.2\%$ of control value, $p < 0.001$; A528T, $167 \pm 3.5\%$ of control value, $p < 0.01$; T947M, $268 \pm 10.3\%$ of control value, $p < 0.001$; n = 3 independent replications (see Fig 2). The *SORL1* mutants did not alter the levels of either total cellular APP holoprotein or PS1. All 3 mutants did increase the amounts of biotinylatable cell-surface APP (E270K, $286 \pm 36.2\%$ of control value, $p < 0.05$; A528T, $365 \pm 7.8\%$ of control value, $p < 0.01$, T947M, $294 \pm 20.1\%$ of control value, $p < 0.05$; n = 3 independent replications; Fig 3A).

To understand how these mutants altered APP processing, we assessed the physical interaction of the mutants with APP. Coimmunoprecipitation experiments showed that all 3 mutations bound APP less well (E270K, $\sim 41 \pm 5.1\%$ of control value, $p < 0.05$; A528T, $\sim 43 \pm 5.9\%$ of control value, $p < 0.05$; T947M, $\sim 34 \pm 3.5\%$ of control value, $p < 0.01$; n = 3 independent replications, Fig 4). However, the mechanism by which this reduced APP:*SORL1* interaction differed significantly. The E270K and A528T mutants displayed normal levels of *SORL1* at the cell surface (E270K, $101 \pm 7.0\%$ of control value, $p > 0.05$; A528T, $105 \pm 10.1\%$ of control value; n = 3 replications; see Fig 3B), but failed to physically interact with APP on the cell surface, presumably due to the effect of the mutant on *SORL1* conformation. In sharp contrast, the

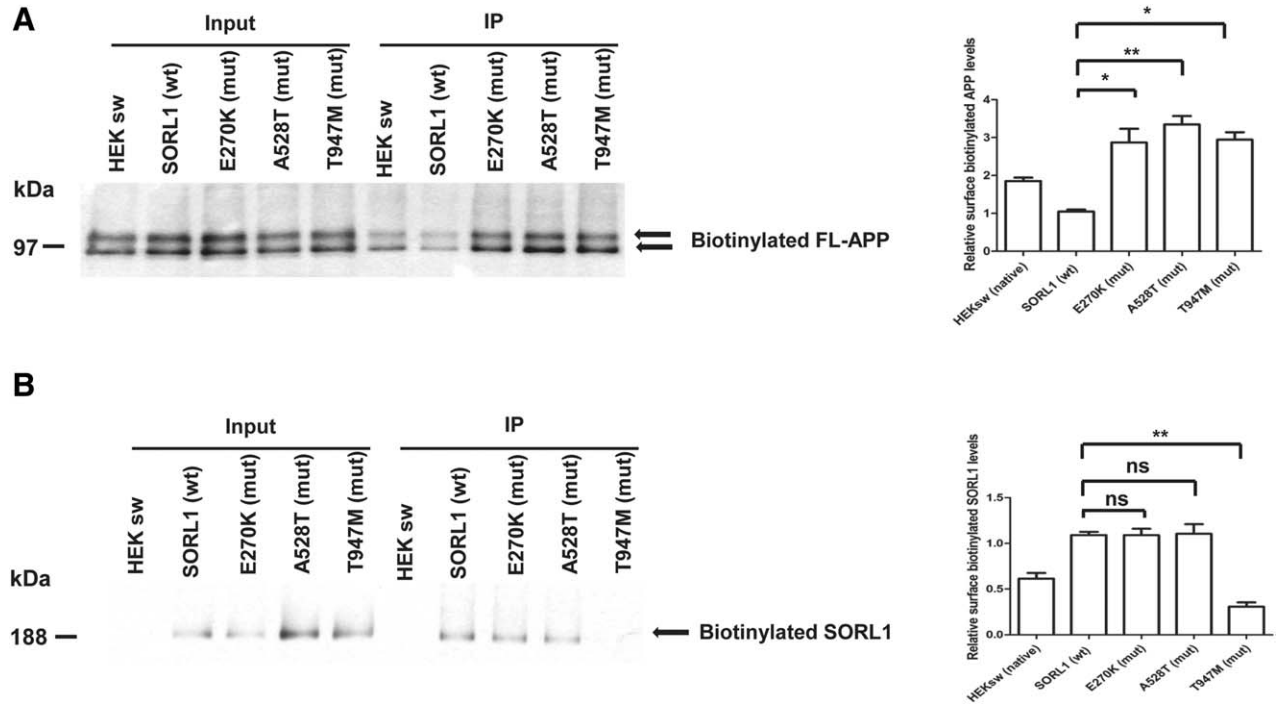

FIGURE 3: The expression of SORL1 mutants (mut) leads to changes of cell surface amyloid precursor protein (APP) and SORL1 levels. Cell surface proteins were biotinylated and precipitated. Surface levels of APP and SORL1 were analyzed by Western blot. (A) APP levels at the cell surface are elevated in all 3 mutants. *$p < 0.05$, **$p < 0.01$, after Bonferroni correction, n = 3 replications. (B) SORL1 surface levels are decreased in the T947M mutant. **$p < 0.01$, ns = not significant, after Bonferroni correction, n = 3 replications. FL = full length; IP = immunoprecipitated; wt = wild type.

T947M mutant showed decreased amounts of *SORL1* at the cell surface (~27 ± 4.5% of control value, $p < 0.05$; n = 3 independent replications; see Fig 3B). The reduced abundance of this mutant at the cell surface clearly accounts for its failure to interact with APP at the cell surface.

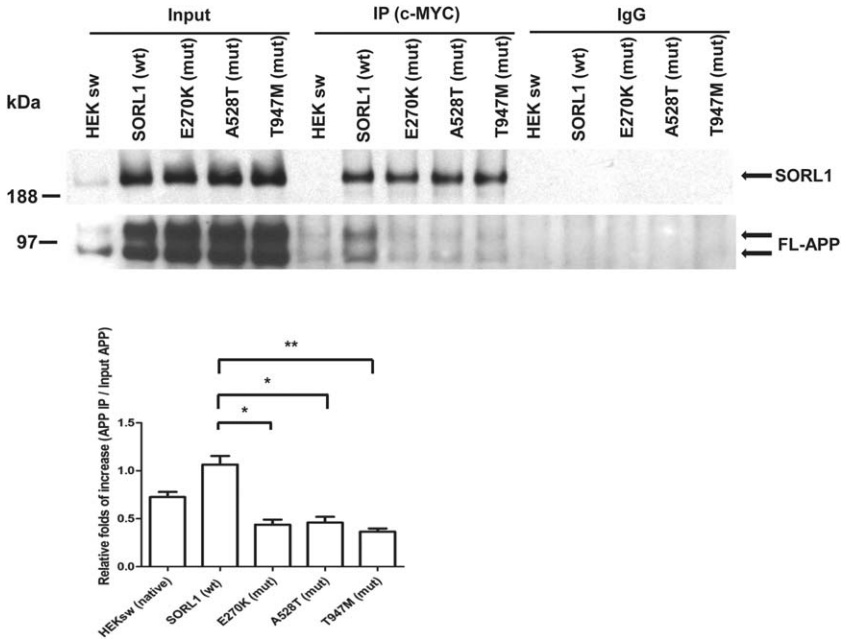

FIGURE 4: All 3 SORL1 mutants (mut) have a reduced binding affinity to amyloid precursor protein (APP). SORL1 was pulled down from cell lysates with a c-MYC antibody and the amount of coprecipitated full-length APP (FL-APP) was measured. *$p < 0.05$, **$p < 0.01$, after Bonferroni correction, n = 3 replications. IgG = immunoglobulin G; IP = immunoprecipitated; wt = wild type.

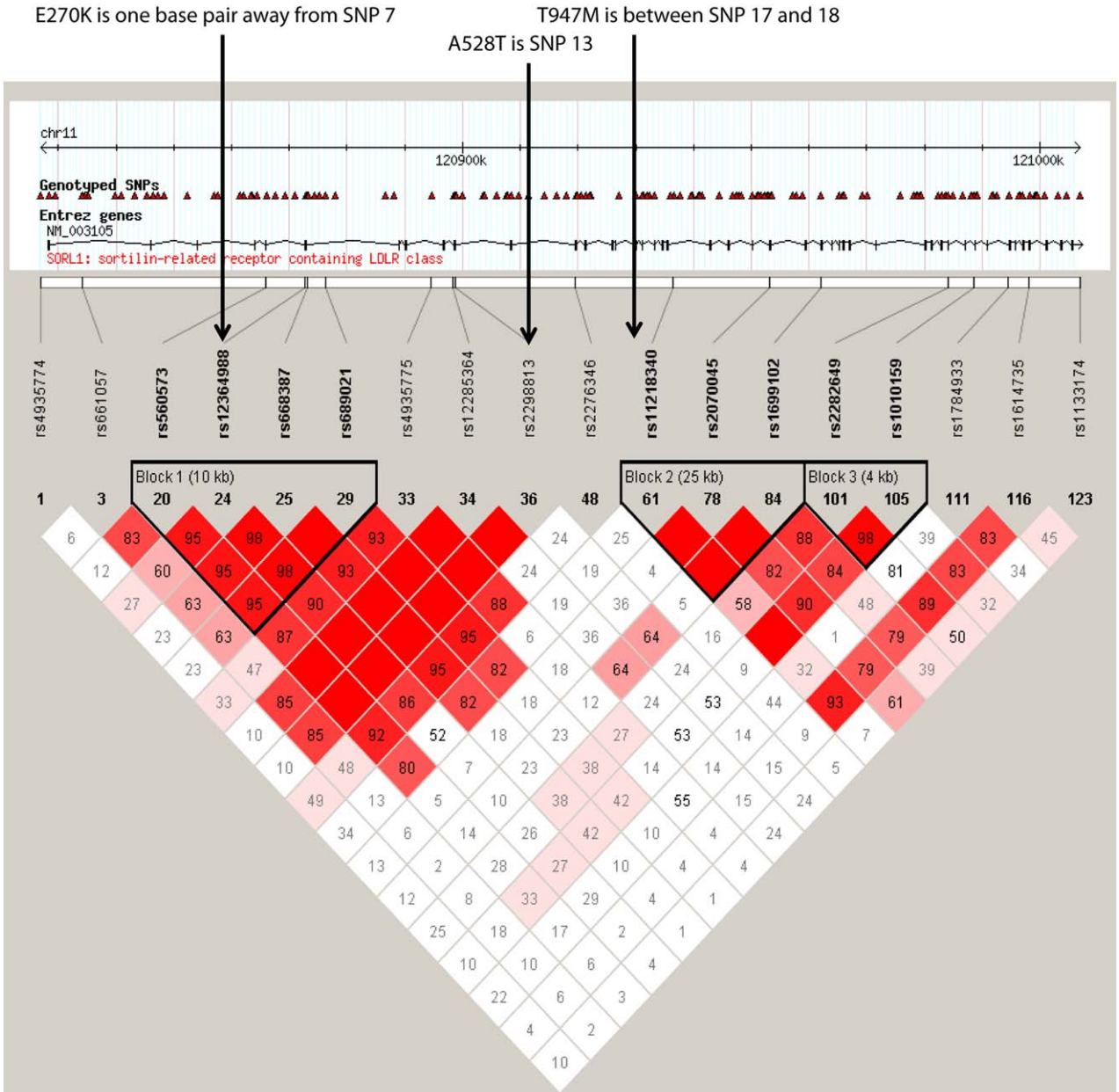

FIGURE 5: Position of the coding mutations relative to the single nucleotide polymorphisms (SNPs) significantly associated with Alzheimer disease (Rogaeva et al[3]).

Discussion

Our results indicate that there may be both common and rare variants in *SORL1* in some population groups that increase the risk of LOAD. The association with *SORL1* has been confirmed in genetic studies of autopsy-confirmed LOAD[6] and in 2 meta-analyses involving several thousand patients and controls.[4,9] Although 3 rare putative variants were identified in European patients with an early onset, autosomal dominant form of Alzheimer disease, no confirmatory functional assessment was performed,[27] and those variants were not detected in the present study. This suggests that the association between *SORL1* and LOAD may be related to the presence of multiple rare coding mutations, some of which may be population specific.

We based our conclusions about the pathogenic nature of the mutations identified here on 2 levels of evidence as suggested here.[28] At the gene level, we demonstrated statistical evidence of an excess of multiple rare, damaging mutations that segregated significantly among cases compared to controls. Previously, we found that reduced expression of SORL1 increased the processing of APP into Aβ-generating compartments.[3] At the variant level, the evidence for pathogenesis of these variants was

based on statistical association and segregation within affected families among Caribbean Hispanics, bioinformatics information indicating evolutionary conservation consistent with the deleterious mutations, and functional studies in HEK293 cell lines indicating the effects of these mutations on APP processing.

The functional mutations investigated in the current study were either absent or much less frequent in patients of Northern European ancestry and in the ADNI data set. Although the frequency of rs2298813, the most common variant, was still increased in cases compared with controls, the difference was not at the level observed in the Caribbean Hispanics and did not reach statistical significance. This may have resulted from the low frequency of this SNP or the small sample size. In contrast, among African Americans the allele frequency was similar to that among Caribbean Hispanics and the variant rs2298813 was found to be significantly associated with LOAD.

It is possible that within the Caribbean Hispanic population this mutation, rs2298812, and the other rare mutations increase risk of disease because they are more penetrant and because there is a strong pattern of inbreeding[29] compared to the other populations investigated. Similar observations have been made in persons with *BRCA1* and *LRRK2* mutations. *BRCA1* mutations are more penetrant among large families of Ashkenazi ancestry than in the general population,[30] and the penetrance of the *LRRK2* G2019S mutation can vary by ethnic group among patients with Parkinson disease.[31]

The 3 variants in *SORL1* identified in the present study show increased secretion of Aβ when transfected into HEK293 cell lines. Interestingly, the rs2298813 (A528T) variant was the most common among the Hispanics and present in 9% of unaffected healthy controls but 15.6% in familial cases. Intriguingly, all 3 of these variants map onto or close to SNPs that were associated with LOAD in the original report by Rogaeva et al.[3] Thus, rs117260922 (E270K) is 1 nucleotide from SNP7 (rs12364988), rs2298813 (A528T) is SNP13, and rs143571823 (T947M) is located within a 3KB region between SNP17 (rs55634) and SNP18 (rs11218340) and is in tight linkage disequilibrium with both SNPs (Fig 5).

The molecular mechanisms underlying this apparently consistent effect of mutants on disease risk appears different between the 3 mutations. The E270K and the A528T mutants have similar levels of *SORL1* at the cell surface as wild-type SORL1-expressing cells. This result suggests that these 2 mutations do not affect the maturation and trafficking of *SORL1* to the cell surface. In contrast, the T947M mutant appears to reach the cell surface less well than wild-type *SORL1* or the other *SORL1* mutants. This suggests that the T947M mutant may act by causing misfolding of SORL1 in the endoplasmic reticulum and its destruction by quality control mechanisms before the SORL1 protein can reach the cell surface.

Taken together, these data indicate that inherited mutants impair interaction of *SORL1* with FL-APP, and thereby fail to direct FL-APP into the retromer-recycling endosome pathway. As a result, in cells expressing mutant SORL1, more of the FL-APP is able to drift into the early and then late endosomes, where it is sequentially cleaved by β-secretase and then by γ-secretase to generate increased amounts of Aβ as demonstrated here. Coding *SORL1* mutations associated with LOAD in this study likely account in part for the GWAS signals. We demonstrated that a common effect of such mutations is to alter Aβ production via changes in APP processing. However, it is conceivable that other rare mutations may alter different aspects of APP/Aβ metabolism. A recently described[27] rare mutation (G511R) seemingly alters Aβ binding to SORL1 and may affect the ability of SORL1 to direct lysosomal targeting of nascent Aβ peptides.[32] When available, the first line of mechanism-based, disease-modifying therapies for carriers of *SORL1* mutations should likely be focused on modulating APP processing and Aβ production.

Acknowledgment

This work was supported by federal grants from the National Institute of Aging (NIA), National Institute of Health (NIH) (P50AG08702, R37AG15473, RO1AG037212) and Marilyn and Henry Taub Foundation to R.M., and the Canadian Institutes of Health Research, National Institute of Health, Wellcome Trust, Medical Research Council, National Institute of Health Research, Ontario Research Fund, and Alzheimer Society of Ontario to P.S.G.-H.

Authorship

B.N.V., Y.Z., and J.H.L. contributed equally to the article.

Potential Conflicts of Interest

Nothing to report.

References

1. Bohm C, Seibel NM, Henkel B, et al. SorLA signaling by regulated intramembrane proteolysis. J Biol Chem 2006;281:14547–14553.

2. Lintzel J, Franke I, Riedel IB, et al. Characterization of the VPS10 domain of SorLA/LR11 as binding site for the neuropeptide HA. Biol Chem 2002;383:1727–1733.

3. Rogaeva E, Meng Y, Lee JH, et al. The neuronal sortilin-related receptor SORL1 is genetically associated with Alzheimer disease. Nat Genet 2007;39:168–177.
4. Lambert JC, Ibrahim-Verbaas CA, Harold D, et al. Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. Nat Genet 2013;45:1452–1458.
5. Lee JH, Barral S, Reitz C. The neuronal sortilin-related receptor gene SORL1 and late-onset Alzheimer's disease. Curr Neurol Neurosci Rep 2008;8:384–391.
6. Lee JH, Cheng R, Honig LS, et al. Association between genetic variants in SORL1 and autopsy-confirmed Alzheimer disease. Neurology 2008;70:887–889.
7. Lee JH, Cheng R, Schupf N, et al. The association between genetic variants in SORL1 and Alzheimer disease in an urban, multiethnic, community-based cohort. Arch Neurol 2007;64:501–506.
8. Meng Y, Lee JH, Cheng R, et al. Association between SORL1 and Alzheimer's disease in a genome-wide study. Neuroreport 2007; 18:1761–1764.
9. Miyashita A, Koike A, Jun G, et al. SORL1 is genetically associated with late-onset Alzheimer's disease in Japanese, Koreans and Caucasians. PloS One 2013;8:e58618.
10. Reitz C, Cheng R, Rogaeva E, et al. Meta-analysis of the association between variants in SORL1 and Alzheimer disease. Arch Neurol 2011;68:99–106.
11. Tan EK, Lee J, Chen CP, et al. SORL1 haplotypes modulate risk of Alzheimer's disease in Chinese. Neurobiol Aging 2009;30:1048–1051.
12. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 2009;25:1754–1760.
13. McKenna A, Hanna M, Banks E, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res 2010;20:1297–1303.
14. Li H, Handsaker B, Wysoker A, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics 2009;25:2078–2079.
15. Bansal V. A statistical method for the detection of variants from next-generation resequencing of DNA pools. Bioinformatics 2010; 26:i318–i324.
16. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res 2010;38:e164.
17. Adzhubei I, Jordan DM, Sunyaev SR. Predicting functional effect of human missense mutations using PolyPhen-2. Curr Protoc Hum Genet 2013;Chapter 7:Unit7.20.
18. Pollard KS, Hubisz MJ, Rosenbloom KR, Siepel A. Detection of nonneutral substitution rates on mammalian phylogenies. Genome Res 2010;20:110–121.
19. Wu MC, Lee S, Cai T, et al. Rare-variant association testing for sequencing data with the sequence kernel association test. Am J Hum Genet 2011;89:82–93.
20. Hiekkalinna T, Schaffer AA, Lambert B, et al. PSEUDOMARKER: a powerful program for joint linkage and/or linkage disequilibrium analysis on mixtures of singletons and related individuals. Hum Hered 2011;71:256–266.
21. Schaid DJ, Rowland CM, Tines DE, et al. Score tests for association between traits and haplotypes when linkage phase is ambiguous. Am J Hum Genet 2002;70:425–434.
22. Yu G, Nishimura M, Arawaka S, et al. Nicastrin modulates presenilin-mediated notch/glp-1 signal transduction and betaAPP processing. Nature 2000;407:48–54.
23. Hasegawa H, Sanjo N, Chen F, et al. Both the sequence and length of the C terminus of PEN-2 are critical for intermolecular interactions and function of presenilin complexes. J Biol Chem 2004;279:46455–46463.
24. Chen F, Hasegawa H, Schmitt-Ulms G, et al. TMP21 is a presenilin complex component that modulates gamma-secretase but not epsilon-secretase activity. Nature 2006;440:1208–1212.
25. Saykin AJ, Shen L, Foroud TM, et al. Alzheimer's Disease Neuroimaging Initiative biomarkers as quantitative phenotypes: genetics core aims, progress, and plans. Alzheimers Dement 2010;6:265–273.
26. Reitz C, Jun G, Naj A, et al. Variants in the ATP-binding cassette transporter (ABCA7), apolipoprotein E 4,and the risk of late-onset Alzheimer disease in African Americans. JAMA 2013;309:1483–1492.
27. Pottier C, Hannequin D, Coutant S, et al. High frequency of potentially pathogenic SORL1 mutations in autosomal dominant early-onset Alzheimer disease. Mol Psychiatry 2012;17:875–879.
28. MacArthur DG, Manolio TA, Dimmock DP, et al. Guidelines for investigating causality of sequence variants in human disease. Nature 2014;508:469–476.
29. Vardarajan BN, Schaid DJ, Reitz C, Lantigua R, Medrano M, Jiménez-Velézquez IZ, Lee JH, Ghani M, Rogaeva E, St George-Hyslop P, Mayeux RP. Inbreeding among Caribbean Hispanics from the Dominican Republic and its effects on risk of Alzheimer disease. Genet Med. 2014 Nov 13 PubMed PMID: 25394174.
30. Petrucelli N, Daly MB, Feldman GL. Hereditary breast and ovarian cancer due to mutations in BRCA1 and BRCA2. Genet Med 2010; 12:245–259.
31. Sierra M, Gonzalez-Aramburu I, Sanchez-Juan P, et al. High frequency and reduced penetrance of LRRK2 G2019S mutation among Parkinson's disease patients in Cantabria (Spain). Mov Disorders 2011;26:2343–2346.
32. Caglayan S, Takagi-Niidome S, Liao F, et al. Lysosomal sorting of amyloid-β by the SORLA receptor is impaired by a familial Alzheimer's disease mutation. Sci Transl Med 2014;6:223ra20.