Genetic variation in *PCDH11X* is associated with susceptibility to late onset Alzheimer's disease

Authors: Minerva M. Carrasquillo, Fanggeng Zou, V. Shane Pankratz, Samantha L. Wilcox, Li. Ma, Louise P. Walker, Samuel G. Younkin, Curtis S. Younkin, Linda H. Younkin, Gina D. Bisceglio, Nilufer Ertekin-Taner, Julia E. Crook, Dennis W. Dickson, Ronald C. Petersen, Neill R. Graff-Radford, Steven G. Younkin

SUPPLEMENTARY METHODS

Study populations and ascertainment. All subjects in the Mayo clinical case-control series were diagnosed by a neurologist at the Mayo Clinic in Jacksonville, Florida (JS series) or Rochester, Minnesota, (RS series). The neurologist confirmed a Clinical Dementia Rating score of 0 for all JS and RS subjects enrolled as controls; cases had diagnoses of possible or probable AD made according to NINCDS-ADRDA criteria[1]. In the autopsy-confirmed series (AUT), all brains were evaluated by Dr. Dennis Dickson and came from the brain bank he maintains at the Mayo Clinic in Jacksonville, FL. In the AUT series the diagnosis of definite AD was also made according to NINCDS-ADRDA criteria[1]. All AD brains analyzed in the study had a Braak score of 4.0 or greater. Brains employed as controls had a Braak score of 2.5 or lower but often had brain pathology unrelated to AD and pathological diagnoses that included vascular dementia, frontotemporal dementia, dementia with Lewy bodies, multi-system atrophy, amyotrophic lateral sclerosis, and progressive supranuclear palsy. Age and gender data for the cases and controls in each series included in the stage I and stage II analyses are shown in **Supplementary Table 1**.

Evaluation of 95% vs. 90% call rate cut-offs. In **Supplementary Table 2,** we evaluate the effect of eliminating samples and SNPs using call rate cut-offs of 95% as compared to 90%. Changing the cut-off from 90% to 95% removed 49 samples (2.3%) and 3,454 SNPs (1.1%). Among the SNPs removed were two SNPs in the top 25 (ranks 9 and 15) in which there was genotyping error, as noted in **Supplementary Table 3**. Further comparison of the 90% and 95% cut-offs showed that SNPs ranked 1-8 were identical; SNPs ranked 10-14, and 16-18 at 90% were retained in the top 25 at 95%; six new SNPs entered the top 25 at 95% with rank changes (90%/95%) of 27/19, 41/20, 45/21, 39/23, 36/24, and 30/25; four SNPs left the top 25 with rank changes of 20/29, 23/32, 24/27, and 25/28. It is important to emphasize that changing the call rate cut-off had relatively little effect on the point estimates of the odds ratios. The largest change in rank, which was 45/21 for rs2398513, occurred because of a change in OR from 0.76 (0.66-0.86) at 90% to 0.74 (0.65-0.85) at 95%. Most importantly, increasing the call rate cut-off from 90% to 95% or 98% had essentially no effect on the OR for rs5984894 which was 1.37 (1.19-1.58) and 1.36 (1.18-1.57) at call rates of 90% and 95% respectively. We elected to use a 90% cut-off because it retains samples thereby improving power. In addition it avoids the elimination of SNPs with highly significant association that can be genotyped accurately but have call rates between 90% and the more stringent call rate of 95%. This approach requires that we check the 25 most significant SNPs for genotyping accuracy as we have done, and it is not surprising that 3 of these SNPs had problems with genotyping as noted above.

Evaluation of population substructure. To evaluate the effect of population substructure on the unadjusted *P* values for the 25 SNPs shown in **Supplementary Table 4**, stage I GWAS data were analyzed using the principal components methodology implemented in EIGENSTRAT[2]. Adjustment for population substructure was performed in PLINK by including the top ten axes of variation generated by EIGENSTRAT as covariates in logistic regression analyses using an allelic dosage model. Logistic regression was employed for this analysis

Nature Genetics: doi:10.1038/ng.305

because PLINK does not support the use of covariates when testing for allelic association with a χ^2 test, and because the allelic dosage model in logistic regression is equivalent to the allelic association approach when HWE is met[3]. The default settings in EIGENSTRAT generate axes of variation after performing five iterations to remove outliers. Outliers are defined as individuals whose ancestry is at least 6 standard deviations from the mean on any of the top ten inferred axes. Because we wanted to assess the likelihood that the significance of our unadjusted *P* values could be inflated in our follow-up and combined series, where we had insufficient data to identify outliers by principal components analysis, we evaluated the effect of adjusting our GWAS data both with and without removal of outliers. The results of this analysis are shown in **Supplementary Table 4**. As expected, the unadjusted *P* values for logistic regression using an allelic dosage model were similar to those obtained when testing for allelic association using a χ^2 test (**Supplementary Table 3**). Comparison of the unadjusted *P* values obtained before outlier removal in **Supplementary Table 4** with the adjusted *P* values obtained after outlier removal shows that adjustment for population substructure generally had a modest effect. All of the *APOE* SNPs that showed unadjusted genome-wide significance continued to show genome-wide significance after removal of outliers and adjustment for population substructure. Outlier removal and adjustment for population substructure diminished the significance of the unadjusted *P* values for seventeen SNPs, improved significance for six SNPs, and the significance of the remaining two SNPs stayed the same. The OR for rs5984894, the *PCDH11X* SNP that showed highly significant association in the combined series, increased trivially from 1.37 (95% CI 1.19-1.58) to 1.39 (95% CI 1.19-1.61) but the reduced power that resulted from removing 101 outliers caused the significance of association to remain at $P=1.8\text{x}10^{-5}$.

The effect of population substructure on rs5984894 was also evaluated by including the top ten axes of variation generated by EIGENSTRAT as additional covariates in the multivariable logistic regression analysis of rs5984894 with sex as a covariate described below. This adjustment had essentially no effect on the results obtained for the combined stage I GWAS data or the component JS, RS, or AUT series. (**Supplementary Table 5** online, model 5). Thus, population substructure did not inflate the significance of stage I GWAS results, and it is unlikely that it inflated the highly significant associations observed in stage II and in the stage I + II combined data.

Evaluation of rs5984894 by multivariable logistic regression. In our initial analyses, we used logistic regression with male sex as a covariate to obtain ORs for male hemizygotes, female heterozygotes, and female homozygotes, and test them for significance. In this model, the ORs for female homozygotes and female heterozygotes compare these groups to female non-carriers and the coefficient for male hemizygotes compares this group to male non-carriers (Supplementary Table 5, model 1). Sex was not a significant covariate in the combined stage I, II, or I + II datasets. Overall in stage I, the male hemizygotes ($P=0.04$), female heterozygotes ($P=0.02$), and female homozygotes ($P=0.0002$) had ORs of 1.33 (95% CI 1.02-1.74), 1.43 (95% CI 1.06-1.92) and 1.92 (95% CI 1.36-2.70) respectively with a global *P* value of $5.7\text{x}10^{-5}$. Overall in stage II, the male hemizygotes ($P=0.74$), female heterozygotes ($P=0.15$) and female homozygotes ($P=0.0002$) had ORs of 1.04 (95% CI 0.82-1.33), 1.19 (95% CI 0.94-1.50) and 1.70 (95% CI 1.29-2.24) respectively with a global *P* value of $4.8\text{x}10^{-6}$. When all data in stage I and II were combined, the male hemizygotes ($P=0.07$), female heterozygotes ($P=0.01$), and female homozygotes ($P=2.0\text{x}10^{-7}$) had ORs of 1.18 (95% CI 0.99-1.41), 1.26 (95% CI 1.05-1.51), and 1.75 (95% CI 1.42-2.16) respectively with a global *P* value of $3.9\text{x}10^{-12}$.

To compare all subjects in the 60-80 year age group with those in the over 80 year group, NCRAD subjects in the 60-80 year range were removed from the stage II dataset and included in the stage I dataset. In the 60-80 year group, the male hemizygotes ($P=0.03$), female heterozygotes ($P=0.07$) and female homozygotes ($P=0.0002$) had ORs of 1.28 (95% CI 1.02-1.61), 1.25 (95% CI 0.98-1.60) and 1.74 (95% CI 1.31-2.32)

respectively with a global *P* value of 1.8×10^{-6}. In the over 80 year group, the male hemizygotes (*P*=0.66), female heterozygotes (*P*=0.08), and female homozygotes (*P*=0.0004) had ORs of 1.07 (95% CI 0.80-1.42), 1.26 (95% CI 0.97-1.65), and 1.76 (95% CI 1.29-2.40) respectively with a global *P* value of 4.3×10^{-6}. Thus the association of female heterozygotes and homozygotes was nearly identical in the two age groups, whereas male hemizygotes showed significant (*P*=0.03) association in the 60-80 year group but little evidence for association in the over 80 year group. To pursue this suggestive age-dependence, we evaluated the age over 60 x male hemizygote interaction in the combined series but it did not achieve significance (*P*=0.66).

To verify that male and female non-carriers had no significant difference in risk, we compared male non-carriers, male hemizygotes, female heterozygotes, and female homozygotes to female non-carriers in the combined stage I, II, and I + II datasets (**Supplementary Table 5**, model 2). Compared to female non-carriers, male non-carriers showed no significant increase in risk in the combined data from stage I (*P*=0.90), II (*P*=0.37), or I + II (*P*=0.14) where the ORs were 0.98 (95% CI 0.72-1.33) and 0.89 (95% CI 0.69-1.15), and 0.86 (95% CI 0.71-1.05) respectively.

To evaluate the risk in female homozygotes as compared to all other groups, we made female homozygotes the referent group and compared them to male non-carriers, female non-carriers, male hemizygotes, and female heterozygotes in the combined stage I, II, and I + II datasets (**Supplementary Table 5**, model 3). Compared to female homozygotes all other groups were at significantly less risk both in stage I and stage II. When, for example, female heterozygotes were compared to female homozygotes, they were at significantly less risk in stages I, II and I + II, with *P* values of 0.05, 0.004 and 0.0005, respectively.

We note that models 1, 2, and 3, are similar in that each has four degrees of freedom and the same global χ^2 and *P* values. When male sex is included as a covariate (model 1), this covariate represents the risk of male non-carriers as compared to female non-carriers. This is evidenced by the identical values for the sex covariate in model 1 and for male non-carriers in model 2 (where the male sex covariate is replaced by the male non-carrier variable). We note that model 2 gives the risk for male hemizygotes as compared to female non-carriers and that the risk of male hemizygotes compared to female non-carriers in model 1 must be calculated by summing the coefficients for male sex and male hemizygosity.

In addition to sex, we evaluated age over 60 (number of years above the age of 60), and the presence of *APOE* ε4 as covariates (**Supplementary Table 5**, model 4) in our analysis of the combined stage I + II dataset. In the combined data, age ($P=4.9 \times 10^{-7}$) and *APOE* ε4 ($P<2.2 \times 10^{-16}$) were significant covariates with ORs of 1.02 (95% CI 1.01-1.03) and 6.21 (95% CI 5.45-7.08) respectively. We note that the OR of 6.20 for the presence of *APOE4* in the combined stage I + II dataset is greater than in typical LOAD case control series because it includes cases from the NCRAD series of LOAD families where the frequency of *APOE* ε4 is especially high. When *APOE* ε4, and age over 60 were added as covariates (model 4), the associations found when sex was the only covariate (model 1) persisted. In the combined data, the ORs for female heterozygotes, female homozygotes, and male hemizygotes changed relatively little from 1.26 (95% CI 1.05-1.51), and 1.75 (95% CI 1.42-2.16), and 1.18 (95% CI 0.99-1.41) respectively to 1.23 (95% CI 1.01-1.51), 1.68 (95% CI 1.33-2.12), and 1.14 (95% CI 0.94-1.39).

We also investigated series and series x carrier genotype interactions as covariates in our logistic regression analyses. Consistent with the results of the Breslow Day tests for series to series heterogeneity, which gave *P* values of 0.95, 0.22, and 0.43 in the allelic association analyses of stages I, II, and I + II, these analyses

provided insufficient evidence to conclude that there was a significant effect of series on the associations observed (data not shown).

To be complete, we evaluated rs5984894 using no covariates. As shown in **Supplementary Table 5** (model 6), the results for male hemizygotes, female heterozygotes, and female homozygotes were essentially the same as in the model using sex as a covariate.

Power considerations. Power was calculated separately for the stage I GWAS, and for the stage II follow-up analyses. For each of these calculations, we utilized the power formula developed by Slager and Schaid for the Armitage trend test[4]. For the stage I GWAS, we computed the smallest odds ratio for an ordinal test of association that would be detectable with 80% power. We computed these detectable odds ratios for a variety of minor allele frequencies, and while accounting for multiple testing in two ways: by applying a Bonferroni correction for 320,000 SNPs ($P<1.56\text{x}10^{-7}$), and by empirically selecting the 25 most significant SNPs ($P\approx 25/320{,}000=7.81\text{x}10^{-5}$). Our targeted stage I sample size was 970 cases and 1,495 controls. We computed the minimum detectable ORs for this sample size, and also for the 844 cases and 1,255 subjects that met quality control criteria. These estimates are contained in **Supplementary Table 6a**. For the stage II follow-up analyses, we computed the smallest detectable odds ratio for an ordinal test of association that would be detectable with 80% power after correcting for 25 comparisons in our follow-up series of 1,547 cases and 1,209 controls. These estimates are contained in **Supplementary Table 6b**. After identifying an X-linked SNP as being our primary interest, we computed the power of the available stage II sample to detect effects of the magnitude observed in the stage I analyses, with Bonferroni correction for 25 tests. The stage II study had 22% power to detect the male hemizygote OR of 1.33 observed in stage I, and 80% power to detect an OR of 1.64. The stage II study had 96% power to detect the per-allele OR of 1.38 that was observed in stage I. Although our study was adequately powered to detect the association overall, and among females alone, it had insufficient power to reliably replicate the association observed among males. Therefore, it remains unclear whether the observed age dependence in male hemizygotes or the observed differences between male hemizygotes and female heterozygotes are genuine.

REFERENCES

1. McKhann, G. et al. Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology* **34**, 939-44 (1984).
2. Price, A.L. et al. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* **38**, 904-9 (2006).
3. Schaid, D.J. & Jacobsen, S.J. Biased tests of association: comparisons of allele frequencies when departing from Hardy-Weinberg proportions. *Am J Epidemiol* **149**, 706-11 (1999).
4. Slager, S.L. & Schaid, D.J. Case-control studies of genetic markers: power and sample size approximations for Armitage's test for trend. *Hum Hered* **52**, 149-53 (2001).

Supplementary Table 1. Series composition. The number of cases and controls, mean age, and percentage that are female are shown for each series. Mean age is given as age-at-diagnosis of AD for cases and age-at-examination for controls. The standard deviation (SD) from the mean is given in parenthesis.

	n		Mean Age (SD)		% Females	
Series	Cases	Controls	Cases	Controls	Cases	Controls
Stage I						
JS 60-80	353	331	74.5 (4.4)	72.5 (4.8)	61.2	59.8
RS 60-80	245	701	73.8 (4.9)	74.0 (3.6)	58.4	51.4
AUT 60-80	246	223	73.5 (5.3)	71.8 (5.5)	50.0	37.2
Stage I combined	844	1255	74.0 (4.8)	73.2 (4.4)	57.1	51.1
Stage II						
JS 80+	237	260	83.8 (3.2)	85.1 (4.0)	62.0	60.8
RS 80+	276	624	86.0 (4.4)	84.0 (3.1)	66.7	57.1
AUT 80+	332	116	87.4 (4.8)	84.6 (6.0)	68.7	54.3
NCRAD 60+	702	209	75.2 (6.8)	78.3 (8.9)	64.8	61.7
Stage II combined	1547	1209	81.1 (7.8)	83.3 (5.6)	65.5	58.4
Stage I + II combined	2391	2464	78.6 (7.7)	78.2 (7.1)	62.6	54.7

Nature Genetics: doi:10.1038/ng.305

Supplementary Table 2. Effect of 90% and 95% cut-offs for SNP and sample call rate on *P* value rank.

Cut-off SNPs: N / Ntotal / (%) Subjects: N / Ntotal / (%)			90% SNP and Sample Call Rate 313,504 / 318,237 / (98.5%) 2,117 / 2,465 / (85.9%)			95% SNP and Sample Call Rate 310,050 / 318,237 / (97.4%) 2,068 / 2,465 / (83.9%)		
Chr	SNP	Position (bp)	Rank	OR (95% CI)	*P* value	Rank	OR (95% CI)	*P* value
19	rs2075650	50,087,459	**1**	2.90 (2.50-3.37)	1.3×10^{-46}	**1**	2.90 (2.49-3.37)	1.2×10^{-45}
19	rs157580	50,087,106	**2**	0.51 (0.44-0.58)	4.1×10^{-23}	**2**	0.50 (0.44-0.58)	5.6×10^{-23}
19	rs439401	50,106,291	**3**	0.55 (0.48-0.63)	6.6×10^{-18}	**3**	0.56 (0.49-0.64)	5.7×10^{-17}
19	rs6859	50,073,874	**4**	1.68 (1.48-1.90)	4.7×10^{-16}	**4**	1.65 (1.46-1.88)	5.2×10^{-15}
19	rs8106922	50,093,506	**5**	0.67 (0.58-0.76)	1.8×10^{-9}	**5**	0.67 (0.58-0.76)	3.6×10^{-9}
19	rs405509	50,100,676	**6**	0.70 (0.62-0.79)	1.1×10^{-8}	**6**	0.70 (0.61-0.79)	1.5×10^{-8}
19	rs10402271	50,021,054	**7**	1.39 (1.22-1.58)	5.0×10^{-7}	**7**	1.40 (1.23-1.60)	3.0×10^{-7}
12	rs11044668	19,593,783	**8**	0.70 (0.61-0.80)	5.4×10^{-7}	**8**	0.70 (0.61-0.80)	6.7×10^{-7}
9	rs3858095	93,944,438	**9**	1.44 (1.24-1.66)	8.5×10^{-7}	Removed by cutoff criteria		
11	rs2746600	33,671,217	**10**	1.38 (1.21-1.58)	1.2×10^{-6}	**9**	1.39 (1.22-1.59)	8.1×10^{-7}
13	rs7318037	81,367,146	**11**	1.37 (1.21-1.56)	1.3×10^{-6}	**13**	1.36 (1.19-1.55)	3.5×10^{-6}
19	rs377702	50,054,507	**12**	1.35 (1.19-1.53)	2.4×10^{-6}	**11**	1.35 (1.19-1.54)	2.6×10^{-6}
12	rs10841260	19,597,931	**13**	1.35 (1.19-1.54)	2.8×10^{-6}	**12**	1.36 (1.19-1.54)	3.2×10^{-6}
8	rs2318144	58,277,297	**14**	1.98 (1.48-2.65)	2.9×10^{-6}	**10**	2.04 (1.52-2.74)	1.4×10^{-6}
1	rs3007421	6,452,776	**15**	1.81 (1.41-2.32)	2.9×10^{-6}	Removed by cutoff criteria		
14	rs856675	84,405,968	**16**	1.56 (1.29-1.89)	4.5×10^{-6}	**16**	1.54 (1.27-1.87)	1.0×10^{-5}
10	rs701864	95,154,196	**17**	1.36 (1.19-1.56)	4.7×10^{-6}	**22**	1.34 (1.17-1.54)	1.8×10^{-5}
19	rs1114832	50,328,041	**18**	1.59 (1.30-1.95)	4.8×10^{-6}	**17**	1.57 (1.28-1.92)	1.2×10^{-5}
1	rs639222	64,062,886	**19**	1.76 (1.37-2.25)	6.0×10^{-6}	**14**	1.77 (1.38-2.26)	5.9×10^{-6}
X	rs1279795	123,152,101	**20**	1.37 (1.19-1.58)	9.0×10^{-6}	29	1.36 (1.18-1.57)	2.3×10^{-5}
1	rs649608	64,086,284	**21**	1.76 (1.37-2.27)	1.0×10^{-5}	**15**	1.78 (1.38-2.30)	7.6×10^{-6}
15	rs8039031	34,954,382	**22**	0.70 (0.60-0.82)	1.1×10^{-5}	**18**	0.70 (0.60-0.82)	1.3×10^{-5}
19	rs1048699	50,342,226	**23**	1.57 (1.28-1.93)	1.2×10^{-5}	32	1.55 (1.26-1.91)	3.1×10^{-5}
X	rs5984894	91,280,393	**24**	1.37 (1.19-1.58)	1.3×10^{-5}	27	1.36 (1.18-1.57)	2.2×10^{-5}
12	rs6486961	19,518,082	**25**	1.32 (1.17-1.50)	1.4×10^{-5}	28	1.32 (1.16-1.49)	2.3×10^{-5}
12	rs2366452	93,480,866	27	0.76 (0.67-0.86)	2.2×10^{-5}	**19**	0.75 (0.66-0.86)	1.3×10^{-5}
21	rs2835370	36,807,495	41	1.62 (1.29-2.04)	3.7×10^{-5}	**20**	1.66 (1.32-2.10)	1.6×10^{-5}
19	rs2965101	49,929,652	45	0.76 (0.66-0.86)	4.5×10^{-5}	**21**	0.74 (0.65-0.85)	1.7×10^{-5}
16	rs12923427	17,482,566	39	1.37 (1.18-1.59)	3.4×10^{-5}	**23**	1.39 (1.19-1.62)	1.9×10^{-5}
18	rs7245160	70,417,826	36	0.69 (0.58-0.82)	3.3×10^{-5}	**24**	0.68 (0.57-0.81)	1.9×10^{-5}
5	rs6886050	125,823,414	30	1.74 (1.34-2.25)	2.5×10^{-5}	**25**	1.76 (1.35-2.28)	1.9×10^{-5}

Nature Genetics: doi:10.1038/ng.305

Supplementary Table 3. Allelic association of top 25 SNPs from stage I. *P* values and odds ratios (OR) with their 95% confidence interval (95% CI) for allelic association as implemented in PLINK. Stage I combined series: JS, RS, and AUT subjects with an age at diagnosis/entry of 60-80 years (844 cases and 1,255 controls after quality control). Stage II combined series: JS, RS, and AUT subjects with an age at diagnosis/entry of over 80 years, and NCRAD with an age at diagnosis/entry of over 60 years (1,547 AD cases and 1,209 controls). Stage I + II combined: 2,391 cases, 2,464 controls. Chr: chromosome. MAF: minor allele frequency. *Visual inspection showed that these SNPs had unsatisfactory genotype clusters in stage I that caused inaccurate genotyping.

Chr	rs Number	Position (bp)	Stage I Combined				Stage II Combined				Stage I + II Combined			
			MAF		*P* value	OR (95% CI)	MAF		*P* value	OR (95% CI)	MAF		*P* value	OR (95% CI)
			Cases	Controls			Cases	Controls			Cases	Controls		
1	rs3007421*	6,452,776	0.09	0.05	$1.8x10^{-6}$	1.83 (1.42-2.36)	0.12	0.12	$7.8x10^{-1}$	1.02 (0.87-1.20)	0.11	0.09	$7.7x10^{-5}$	1.32 (1.15-1.51)
1	rs639222	64,062,886	0.09	0.05	$5.2x10^{-6}$	1.77 (1.38-2.26)	0.07	0.07	$9.8x10^{-1}$	1.00 (0.81-1.23)	0.08	0.06	$2.2x10^{-3}$	1.28 (1.09-1.50)
1	rs649608	64,086,284	0.08	0.05	$9.0x10^{-6}$	1.77 (1.37-2.28)	0.07	0.07	$5.4x10^{-1}$	0.93 (0.75-1.16)	0.07	0.06	$1.1x10^{-2}$	1.24 (1.05-1.46)
8	rs2318144*	58,277,297	0.06	0.03	$4.2x10^{-6}$	1.96 (1.47-2.63)	0.13	0.14	$5.3x10^{-1}$	0.95 (0.81-1.11)	0.11	0.08	$1.4x10^{-4}$	1.31 (1.14-1.50)
9	rs3858095*	93,944,438	0.30	0.23	$3.3x10^{-6}$	1.41 (1.22-1.63)	0.29	0.30	$4.6x10^{-1}$	0.96 (0.85-1.08)	0.29	0.27	$4.0x10^{-3}$	1.14 (1.04-1.25)
10	rs701864	95,154,196	0.34	0.27	$5.0x10^{-6}$	1.37 (1.19-1.56)	0.32	0.33	$2.8x10^{-1}$	0.94 (0.84-1.05)	0.32	0.3	$1.3x10^{-2}$	1.12 (1.02-1.22)
11	rs2746600	33,671,217	0.38	0.30	$1.2x10^{-6}$	1.39 (1.22-1.58)	0.34	0.34	$8.5x10^{-1}$	0.99 (0.88-1.11)	0.35	0.32	$2.4x10^{-3}$	1.14 (1.05-1.24)
11	rs3740938	102,092,272	0.05	0.08	$1.4x10^{-5}$	0.56 (0.43-0.73)	0.06	0.06	$9.4x10^{-1}$	1.01 (0.80-1.26)	0.06	0.07	$1.6x10^{-3}$	0.77 (0.65-0.91)
12	rs11044668	19,593,783	0.24	0.31	$1.5x10^{-6}$	0.71 (0.62-0.82)	0.29	0.28	$4.8x10^{-1}$	1.04 (0.93-1.18)	0.27	0.29	$1.3x10^{-2}$	0.89 (0.82-0.98)
12	rs10841260	19,597,931	0.41	0.34	$3.2x10^{-6}$	1.35 (1.19-1.54)	0.38	0.38	$5.0x10^{-1}$	0.96 (0.86-1.08)	0.39	0.36	$8.6x10^{-3}$	1.12 (1.03-1.21)
13	rs7318037	81,367,146	0.40	0.33	$2.9x10^{-6}$	1.36 (1.20-1.55)	0.34	0.35	$4.6x10^{-1}$	0.96 (0.86-1.07)	0.36	0.34	$2.8x10^{-2}$	1.10 (1.01-1.19)
14	rs856675	84,405,968	0.14	0.10	$5.1x10^{-6}$	1.56 (1.29-1.89)	0.12	0.12	$8.7x10^{-1}$	0.99 (0.84-1.16)	0.13	0.11	$3.8x10^{-3}$	1.20 (1.06-1.36)
15	rs8039031	34,954,382	0.17	0.23	$1.5x10^{-5}$	0.71 (0.60-0.83)	0.21	0.21	$5.6x10^{-1}$	0.96 (0.84-1.10)	0.2	0.22	$1.5x10^{-3}$	0.85 (0.77-0.94)
19	rs10402271	50,021,054	0.39	0.32	$8.4x10^{-7}$	1.38 (1.22-1.57)	0.39	0.30	$3.5x10^{-12}$	1.49 (1.33-1.67)	0.39	0.31	$2.9x10^{-17}$	1.44 (1.32-1.56)
19	rs377702	50,054,507	0.45	0.38	$5.5x10^{-6}$	1.34 (1.18-1.52)	0.42	0.39	$3.1x10^{-2}$	1.13 (1.01-1.26)	0.43	0.39	$4.3x10^{-6}$	1.21 (1.12-1.32)
19	rs6859	50,073,874	0.54	0.41	$1.7x10^{-15}$	1.66 (1.47-1.89)	0.52	0.40	$1.1x10^{-20}$	1.67 (1.50-1.86)	0.53	0.41	$5.5x10^{-34}$	1.65 (1.52-1.79)
19	rs157580	50,087,106	0.25	0.40	$3.0x10^{-22}$	0.51 (0.45-0.59)	0.27	0.39	$1.4x10^{-20}$	0.58 (0.52-0.65)	0.26	0.39	$7.4x10^{-42}$	0.55 (0.51-0.60)
19	rs2075650	50,087,459	0.34	0.15	$4.8x10^{-46}$	2.89 (2.49-3.36)	0.32	0.11	$9.5x10^{-79}$	3.93 (3.39-4.57)	0.33	0.13	$3.7x10^{-120}$	3.29 (2.97-3.65)
19	rs8106922	50,093,506	0.29	0.38	$1.2x10^{-9}$	0.66 (0.58-0.76)	0.30	0.42	$2.2x10^{-20}$	0.59 (0.53-0.66)	0.3	0.4	$7.6x10^{-27}$	0.63 (0.58-0.69)
19	rs405509	50,100,676	0.41	0.50	$6.2x10^{-9}$	0.69 (0.61-0.78)	0.42	0.55	$1.1x10^{-21}$	0.59 (0.53-0.66)	0.42	0.53	$6.2x10^{-27}$	0.64 (0.59-0.70)
19	rs439401	50,106,291	0.26	0.39	$2.4x10^{-17}$	0.56 (0.48-0.64)	0.27	0.37	$2.8x10^{-15}$	0.63 (0.56-0.71)	0.27	0.38	$1.6x10^{-31}$	0.60 (0.55-0.65)
19	rs1114832	50,328,041	0.13	0.08	$5.9x10^{-6}$	1.59 (1.30-1.94)	0.12	0.10	$5.0x10^{-2}$	1.19 (1.00-1.42)	0.12	0.09	$5.9x10^{-6}$	1.35 (1.19-1.54)
19	rs1048699	50,342,226	0.12	0.08	$1.5x10^{-5}$	1.57 (1.28-1.93)	0.11	0.09	$2.2x10^{-2}$	1.23 (1.03-1.47)	0.12	0.09	$3.6x10^{-6}$	1.37 (1.20-1.56)
X	rs5984894	91,280,393	0.53	0.45	$1.2x10^{-5}$	1.38 (1.19-1.59)	0.51	0.46	$5.7x10^{-4}$	1.23 (1.10-1.39)	0.52	0.46	$3.8x10^{-8}$	1.29 (1.18-1.41)
X	rs1279795	123,152,101	0.53	0.45	$7.2x10^{-6}$	1.38 (1.20-1.59)	0.46	0.47	$4.3x10^{-1}$	0.95 (0.84-1.08)	0.48	0.46	$3.0x10^{-2}$	1.11 (1.01-1.21)

Nature Genetics: doi:10.1038/ng.305

Supplementary Table 4. Evaluation of population substructure on top 25 SNPs from stage I. *P* values and odds ratios (OR) with their 95% confidence interval (95% CI) are shown for the 25 SNPs with most significant association in the stage I GWAS (844 AD patients vs. 1,255 controls after quality control). Results were obtained by logistic regression using an allelic dosage model as implemented in PLINK. Adjustment for population substructure was performed in PLINK by including the top ten axes of variation generated by EIGENSTRAT as covariates in logistic regression analyses using an allelic dosage model. Correction was performed with and without removal of outliers using the default settings (5 iterations to remove outliers) in EIGENSTRAT. EIGENSTRAT defines outliers as individuals whose ancestry is at least 6 standard deviations from the mean of any one of the top ten inferred axes of variation.

			No Outliers Removed				Outliers Removed			
			Unadjusted		Adjusted		Unadjusted		Adjusted	
Chr	rs Number	Position (bp)	*P* value	OR (95% CI)	*P* value	OR (95% CI)	*P* value	OR (95% CI)	*P* value	OR (95% CI)
1	rs3007421	6,452,776	3.6×10^{-6}	1.81 (1.41-2.32)	3.8×10^{-5}	1.70 (1.32-2.19)	1.1×10^{-5}	1.77 (1.37-2.29)	6.3×10^{-5}	1.70 (1.31-2.20)
1	rs639222	64,062,886	7.7×10^{-6}	1.76 (1.38-2.26)	3.1×10^{-5}	1.71 (1.33-2.20)	8.0×10^{-5}	1.68 (1.30-2.17)	6.8×10^{-5}	1.70 (1.31-2.20)
1	rs649608	64,086,284	1.4×10^{-5}	1.75 (1.36-2.26)	1.0×10^{-4}	1.68 (1.29-2.18)	3.4×10^{-4}	1.62 (1.25-2.11)	3.1×10^{-4}	1.64 (1.25-2.14)
8	rs2318144	58,277,297	1.5×10^{-5}	1.88 (1.41-2.49)	2.1×10^{-5}	1.88 (1.41-2.51)	1.3×10^{-5}	1.91 (1.43-2.55)	3.4×10^{-5}	1.85 (1.38-2.48)
9	rs3858095	93,944,438	7.8×10^{-6}	1.38 (1.20-1.59)	1.6×10^{-5}	1.37 (1.19-1.59)	4.6×10^{-5}	1.35 (1.17-1.56)	2.7×10^{-5}	1.37 (1.18-1.59)
10	rs701864	95,154,196	6.3×10^{-6}	1.36 (1.19-1.56)	1.9×10^{-5}	1.35 (1.18-1.55)	2.2×10^{-5}	1.35 (1.17-1.55)	1.2×10^{-5}	1.36 (1.19-1.57)
11	rs2746600	33,671,217	2.1×10^{-6}	1.37 (1.20-1.56)	7.9×10^{-7}	1.40 (1.22-1.59)	1.3×10^{-6}	1.39 (1.22-1.59)	1.4×10^{-6}	1.40 (1.22-1.60)
11	rs3740938	102,092,272	2.6×10^{-5}	0.57 (0.44-0.74)	2.1×10^{-5}	0.56 (0.43-0.73)	8.7×10^{-5}	0.59 (0.45-0.77)	5.5×10^{-5}	0.58 (0.44-0.76)
12	rs11044668	19,593,783	2.1×10^{-6}	0.71 (0.62-0.82)	1.9×10^{-5}	0.73 (0.64-0.85)	1.8×10^{-5}	0.73 (0.63-0.84)	3.5×10^{-5}	0.74 (0.64-0.85)
12	rs10841260	19,597,931	4.0×10^{-6}	1.35 (1.19-1.53)	9.6×10^{-6}	1.34 (1.18-1.52)	7.7×10^{-5}	1.30 (1.14-1.48)	6.1×10^{-5}	1.31 (1.15-1.49)
13	rs7318037	81,367,146	3.6×10^{-6}	1.36 (1.19-1.54)	5.7×10^{-6}	1.36 (1.19-1.54)	1.4×10^{-5}	1.34 (1.17-1.53)	4.7×10^{-6}	1.37 (1.20-1.56)
14	rs856675	84,405,968	4.6×10^{-6}	1.58 (1.30-1.92)	2.4×10^{-5}	0.66 (0.55-0.80)	4.1×10^{-5}	1.53 (1.25-1.87)	1.1×10^{-4}	1.50 (1.22-1.84)
15	rs8039031	34,954,382	1.4×10^{-5}	0.70 (0.60-0.82)	3.9×10^{-6}	0.68 (0.58-0.80)	1.4×10^{-5}	0.70 (0.59-0.82)	1.4×10^{-5}	0.69 (0.59-0.82)
19	rs10402271	50,021,054	8.9×10^{-7}	1.39 (1.22-1.58)	5.1×10^{-7}	1.40 (1.23-1.60)	2.0×10^{-6}	1.38 (1.21-1.58)	1.0×10^{-6}	1.40 (1.22-1.60)
19	rs377702	50,054,507	5.2×10^{-6}	1.35 (1.18-1.53)	5.4×10^{-7}	1.40 (1.23-1.59)	1.2×10^{-5}	1.34 (1.17-1.52)	5.1×10^{-6}	1.36 (1.19-1.55)
19	rs6859	50,073,874	4.4×10^{-15}	1.67 (1.47-1.90)	3.9×10^{-15}	1.69 (1.48-0.92)	1.6×10^{-13}	1.63 (1.43-1.86)	6.1×10^{-14}	1.66 (1.45-1.89)
19	rs157580	50,087,106	3.3×10^{-21}	0.51 (0.45-0.59)	4.1×10^{-21}	0.51 (0.44-0.58)	7.0×10^{-19}	0.53 (0.46-0.61)	3.7×10^{-20}	0.51 (0.44-0.59)
19	rs2075650	50,087,459	9.3×10^{-42}	3.01 (2.57-3.54)	7.8×10^{-43}	3.13 (2.66-3.68)	1.5×10^{-40}	3.03 (2.57-3.57)	3.1×10^{-42}	3.19 (2.70-3.77)
19	rs8106922	50,093,506	3.8×10^{-9}	0.67 (0.59-0.77)	3.4×10^{-9}	0.67 (0.58-0.76)	1.7×10^{-9}	0.66 (0.58-0.76)	1.5×10^{-9}	0.65 (0.57-0.75)
19	rs405509	50,100,676	1.0×10^{-8}	0.69 (0.61-0.79)	6.2×10^{-9}	0.69 (0.60-0.78)	1.1×10^{-8}	0.69 (0.61-0.78)	2.9×10^{-8}	0.69 (0.61-0.79)
19	rs439401	50,106,291	1.2×10^{-16}	0.56 (0.48-0.64)	2.4×10^{-16}	0.56 (0.48-0.64)	2.3×10^{-15}	0.57 (0.49-0.65)	5.4×10^{-16}	0.55 (0.48-0.64)
19	rs1114832	50,328,041	7.0×10^{-6}	1.59 (1.30-1.95)	9.0×10^{-6}	1.59 (1.30-1.96)	2.5×10^{-6}	1.65 (1.34-2.03)	3.9×10^{-6}	1.64 (1.33-2.02)
19	rs1048699	50,342,226	1.5×10^{-5}	1.58 (1.29-1.95)	2.2×10^{-5}	1.58 (1.28-1.95)	6.9×10^{-6}	1.63 (1.32-2.02)	1.1×10^{-5}	1.62 (1.31-2.01)
X	rs5984894	91,280,393	1.8×10^{-5}	1.37 (1.19-1.58)	1.3×10^{-5}	1.38 (1.20-1.60)	2.0×10^{-5}	1.38 (1.19-1.60)	1.8×10^{-5}	1.39 (1.19-1.61)
X	rs1279795	123,152,101	7.9×10^{-6}	1.39 (1.20-1.60)	5.9×10^{-5}	1.35 (1.17-1.56)	7.3×10^{-6}	1.40 (1.21-1.62)	2.2×10^{-5}	1.38 (1.19-1.60)

Nature Genetics: doi:10.1038/ng.305

Supplementary Table 5. Analysis of rs5984894 by multivariable logistic regression. Model 1 uses male sex as a covariate to obtain ORs for male hemizygotes, female heterozygotes, and female homozygotes, and test them for significance. In this model, the ORs for female homozygotes and female heterozygotes compare these groups to female non-carriers and the coefficient for male hemizygotes compares this group to male non-carriers. In model 2, we compared male non-carriers, male hemizygotes, female heterozygotes, and female homozygotes to female non-carriers, to verify that male and female non-carriers had no significant difference in risk. In model 3, we made female homozygotes the referent group and compared them to male non-carriers, female non-carriers, male hemizygotes, and female heterozygotes to determine rigorously the significance of the risk in female homozygotes as compared each of the other groups. Model 4 evaluates the effect of age over 60 and the presence of *APOE* ε4 as covariates in addition to male sex. Model 5 evaluates the effect of population substructure on rs5984894 by including as covariates the top ten axes of variation generated by EIGENSTRAT (not shown); it also uses sex as a covariate as in model 1. Model 6 evaluates rs5984894 using no covariates.

Model	Series	Age	Sex		AgeOver60		APOE 4+		Male Non-Carriers				Female Non-Carriers				Male Hemizygotes				Female Heterozygotes				Female Homozygotes				Global Statistics		
			OR (95% CI)	*P*	OR (95% CI)	*P*	OR (95% CI)	*P*	AD (n)	Con (n)	OR (95% CI)	*P*	AD (n)	Con (n)	OR (95% CI)	*P*	AD (n)	Con (n)	OR (95% CI)	*P*	AD (n)	Con (n)	OR (95% CI)	*P*	AD (n)	Con (n)	OR (95% CI)	*P*	χ^2	df	*P* value
	Stage I																														
1	JS	60-80	1.28 (0.78-2.11)	0.33	Not Included	N/A	Not Included	N/A	66	70	N/A	N/A	50	68	Referent	N/A	69	57	1.28 (0.79-2.09)	0.31	106	87	1.66 (1.04-2.63)	0.03	59	41	1.96 (1.14-3.36)	0.01	8.1	4	0.09
1	RS	60-80	1.00 (0.58-1.72)	0.99	Not Included	N/A	Not Included	N/A	51	185	N/A	N/A	25	91	Referent	N/A	45	136	1.20 (0.76-1.90)	0.43	74	185	1.46 (0.87-2.44)	0.16	40	72	2.02 (1.12-3.64)	0.02	9.8	4	0.04
1	AUT	60-80	0.79 (0.41-1.53)	0.48	Not Included	N/A	Not Included	N/A	56	71	N/A	N/A	24	24	Referent	N/A	62	56	1.40 (0.85-2.32)	0.19	59	38	1.55 (0.77-3.12)	0.22	38	19	2.00 (0.91-4.40)	0.09	11.0	4	0.03
1	ALL	60-80	0.98 (0.72-1.33)	0.90	Not Included	N/A	Not Included	N/A	173	326	N/A	N/A	99	183	Referent	N/A	176	249	1.33 (1.02-1.74)	0.04	239	310	1.43 (1.06-1.92)	0.02	137	132	1.92 (1.36-2.70)	0.0002	24.7	4	5.7×10^{-5}
2	ALL	60-80	Not Included	N/A	Not Included	N/A	Not Included	N/A	173	326	0.98 (0.72-1.33)	0.90	99	183	Referent	N/A	176	249	1.31 (0.96-1.78)	0.09	239	310	1.43 (1.06-1.92)	0.02	137	132	1.92 (1.36-2.70)	0.0002	24.7	4	5.7×10^{-5}
3	ALL	60-80	Not Included	N/A	Not Included	N/A	Not Included	N/A	173	326	0.51 (0.38-0.69)	1.3×10^{-5}	99	183	0.52 (0.37-0.73)	0.0002	176	249	0.68 (0.50-0.93)	0.01	239	310	0.74 (0.55-1.00)	0.05	137	132	Referent	N/A	24.7	4	5.7×10^{-5}
5	JS	60-80	1.28 (0.76-2.15)	0.36	Not Included	N/A	Not Included	N/A	65	68	N/A	N/A	45	60	Referent	N/A	67	54	1.31 (0.79-2.17)	0.29	106	81	1.81 (1.11-2.97)	0.02	59	40	2.06 (1.17-3.63)	0.01	9.2	4	0.06
5	RS	60-80	0.88 (0.50-1.54)	0.65	Not Included	N/A	Not Included	N/A	50	180	N/A	N/A	24	84	Referent	N/A	44	133	1.21 (0.76-1.92)	0.43	69	177	1.26 (0.73-2.16)	0.40	39	70	1.78 (0.97-3.28)	0.06	8.4	4	0.08
5	AUT	60-80	0.75 (0.37-1.52)	0.43	Not Included	N/A	Not Included	N/A	51	70	N/A	N/A	21	24	Referent	N/A	58	53	1.68 (0.98-2.87)	0.06	57	36	1.66 (0.79-3.47)	0.18	36	19	2.23 (0.98-5.06)	0.05	13.6	4	0.009
5	ALL	60-80	0.96 (0.70-1.33)	0.81	Not Included	N/A	Not Included	N/A	166	318	N/A	N/A	90	168	Referent	N/A	169	240	1.35 (1.03-1.78)	0.03	232	294	1.49 (1.09-2.04)	0.01	134	129	1.97 (1.38-2.81)	0.0002	27.2	4	1.8×10^{-5}
	Stage II																														
1	JS	80+	1.33 (0.73-2.44)	0.35	Not Included	N/A	Not Included	N/A	44	48	N/A	N/A	33	48	Referent	N/A	44	50	0.96 (0.54-1.71)	0.89	78	72	1.58 (0.91-2.72)	0.10	33	36	1.33 (0.70-2.55)	0.38	2.7	4	0.60
1	RS	80+	0.75 (0.46-1.22)	0.25	Not Included	N/A	Not Included	N/A	45	143	N/A	N/A	44	105	Referent	N/A	46	123	1.19 (0.74-1.91)	0.48	78	179	1.04 (0.67-1.62)	0.86	62	65	2.28 (1.39-3.73)	0.001	23.4	4	1.0×10^{-4}
1	AUT	80+	0.44 (0.19-1.01)	0.05	Not Included	N/A	Not Included	N/A	46	22	N/A	N/A	52	11	Referent	N/A	57	28	0.97 (0.49-1.92)	0.94	118	33	0.76 (0.36-1.61)	0.47	55	12	0.97 (0.39-2.39)	0.95	9.3	4	0.05
1	NCRAD	80+	1.62 (0.67-3.89)	0.29	Not Included	N/A	Not Included	N/A	26	13	N/A	N/A	26	21	Referent	N/A	20	19	0.53 (0.21-1.31)	0.17	49	21	1.88 (0.87-4.07)	0.11	34	12	2.29 (0.95-5.48)	0.06	7.4	4	0.12
1	NCRAD	60-80	0.78 (0.41-1.51)	0.46	Not Included	N/A	Not Included	N/A	106	29	N/A	N/A	84	18	Referent	N/A	94	19	1.35 (0.71-2.57)	0.36	162	42	0.83 (0.45-1.52)	0.54	96	15	1.37 (0.65-2.89)	0.41	3.6	4	0.46
1	NCRAD	60+	1.11 (0.67-1.84)	0.67	Not Included	N/A	Not Included	N/A	132	42	N/A	N/A	110	39	Referent	N/A	114	38	0.95 (0.58-1.58)	0.86	211	63	1.19 (0.75-1.88)	0.47	130	27	1.71 (0.98-2.97)	0.06	4.5	4	0.35
1	ALL	60+	0.89 (0.69-1.15)	0.37	Not Included	N/A	Not Included	N/A	267	255	N/A	N/A	239	203	Referent	N/A	261	239	1.04 (0.82-1.33)	0.74	485	347	1.19 (0.94-1.50)	0.15	280	140	1.70 (1.29-2.24)	0.0002	30.0	4	4.8×10^{-6}
2	ALL	60+	Not Included	N/A	Not Included	N/A	Not Included	N/A	267	255	0.89 (0.69-1.15)	0.37	239	203	Referent	N/A	261	239	0.93 (0.72-1.20)	0.57	485	347	1.19 (0.94-1.50)	0.15	280	140	1.70 (1.29-2.24)	0.0002	30.0	4	4.8×10^{-6}
3	ALL	60+	Not Included	N/A	Not Included	N/A	Not Included	N/A	267	255	0.52 (0.40-0.68)	1.8×10^{-6}	239	203	0.59 (0.45-0.78)	0.0002	261	239	0.55 (0.42-0.71)	9.8×10^{-6}	485	347	0.70 (0.55-0.89)	0.004	280	140	Referent	N/A	30.0	4	4.8×10^{-6}
	Stage I + II																														
1	JS	60+	1.30 (0.89-1.91)	0.18	Not Included	N/A	Not Included	N/A	110	118	N/A	N/A	83	116	Referent	N/A	113	107	1.13 (0.78-1.64)	0.51	184	159	1.62 (1.14-2.30)	0.01	92	77	1.67 (1.10-2.53)	0.02	9.1	4	0.06
1	RS	60+	0.83 (0.58-1.19)	0.31	Not Included	N/A	Not Included	N/A	96	328	N/A	N/A	69	196	Referent	N/A	91	259	1.20 (0.86-1.67)	0.28	152	364	1.19 (0.85-1.66)	0.32	102	137	2.11 (1.45-3.08)	9.2×10^{-5}	31.4	4	2.5×10^{-6}
1	AUT	60+	0.51 (0.31-0.82)	0.01	Not Included	N/A	Not Included	N/A	102	93	N/A	N/A	76	35	Referent	N/A	119	84	1.29 (0.87-1.92)	0.21	177	71	1.15 (0.71-1.87)	0.58	93	31	1.38 (0.78-2.44)	0.27	27.5	4	1.5×10^{-5}
1	NCRAD	60+	1.11 (0.67-1.84)	0.67	Not Included	N/A	Not Included	N/A	132	42	N/A	N/A	110	39	Referent	N/A	114	38	0.95 (0.58-1.58)	0.86	211	63	1.19 (0.75-1.88)	0.47	130	27	1.71 (0.98-2.97)	0.06	4.5	4	0.35
1	ALL	60+	0.86 (0.71-1.05)	0.14	Not Included	N/A	Not Included	N/A	440	581	N/A	N/A	338	386	Referent	N/A	437	488	1.18 (0.99-1.41)	0.07	724	657	1.26 (1.05-1.51)	0.01	417	272	1.75 (1.42-2.16)	2.0×10^{-7}	59.4	4	3.9×10^{-12}
2	ALL	60+	Not Included	N/A	Not Included	N/A	Not Included	N/A	440	581	0.86 (0.71-1.05)	0.14	338	386	Referent	N/A	437	488	1.02 (0.84-1.24)	0.82	724	657	1.26 (1.05-1.51)	0.01	417	272	1.75 (1.42-2.16)	2.0×10^{-7}	59.4	4	3.9×10^{-12}
3	ALL	60+	Not Included	N/A	Not Included	N/A	Not Included	N/A	440	581	0.49 (0.41-0.60)	2.1×10^{-12}	338	386	0.57 (0.46-0.71)	2.0×10^{-7}	437	488	0.58 (0.48-0.71)	1.4×10^{-7}	724	657	0.72 (0.60-0.87)	0.0005	417	272	Referent	N/A	59.4	4	3.9×10^{-12}
1	ALL	60-80	0.86 (0.67-1.11)	0.26	Not Included	N/A	Not Included	N/A	279	355	N/A	N/A	183	201	Referent	N/A	270	268	1.28 (1.02-1.61)	0.03	401	352	1.25 (0.98-1.60)	0.07	233	147	1.74 (1.31-2.32)	0.0002	32.1	4	1.8×10^{-6}
1	ALL	80+	0.85 (0.63-1.14)	0.28	Not Included	N/A	Not Included	N/A	161	226	N/A	N/A	155	185	Referent	N/A	167	220	1.07 (0.80-1.42)	0.66	323	305	1.26 (0.97-1.65)	0.08	184	125	1.76 (1.29-2.40)	0.0004	30.3	4	4.3×10^{-6}
4	ALL	60+	0.87 (0.70-1.08)	0.21	1.02 (1.01-1.03)	4.9×10^{-7}	6.21 (5.45-7.08)	$<2.2\times10^{-16}$	440	581	N/A	N/A	338	386	Referent	N/A	437	488	1.14 (0.94-1.39)	0.19	724	657	1.23 (1.01-1.51)	0.04	417	272	1.68 (1.33-2.12)	1.4×10^{-5}	897.2	6	1.5×10^{-190}
6	ALL	60+	Not Included	N/A	Not Included	N/A	Not Included	N/A	440	581	Referent	N/A	338	386	Referent	N/A	437	488	1.11 (0.95-1.31)	0.19	724	657	1.37 (1.19-1.58)	1.3×10^{-5}	417	272	1.91 (1.59-2.28)	2.0×10^{-12}	57.2	3	2.4×10^{-12}

Nature Genetics: doi:10.1038/ng.305

Supplementary Table 6. Power calculations**. (a)** Smallest odds ratios (ORs) greater than 1.0 that are detectable at 80% power in the stage I genome-wide association study. **(b)** Smallest odds ratios (ORs) greater than 1.0 that are detectable with 80% power in stage II (follow-up), correcting for 25 statistical tests, with 1,547 cases and 1,209 controls. MAF: minor allele frequency.

a

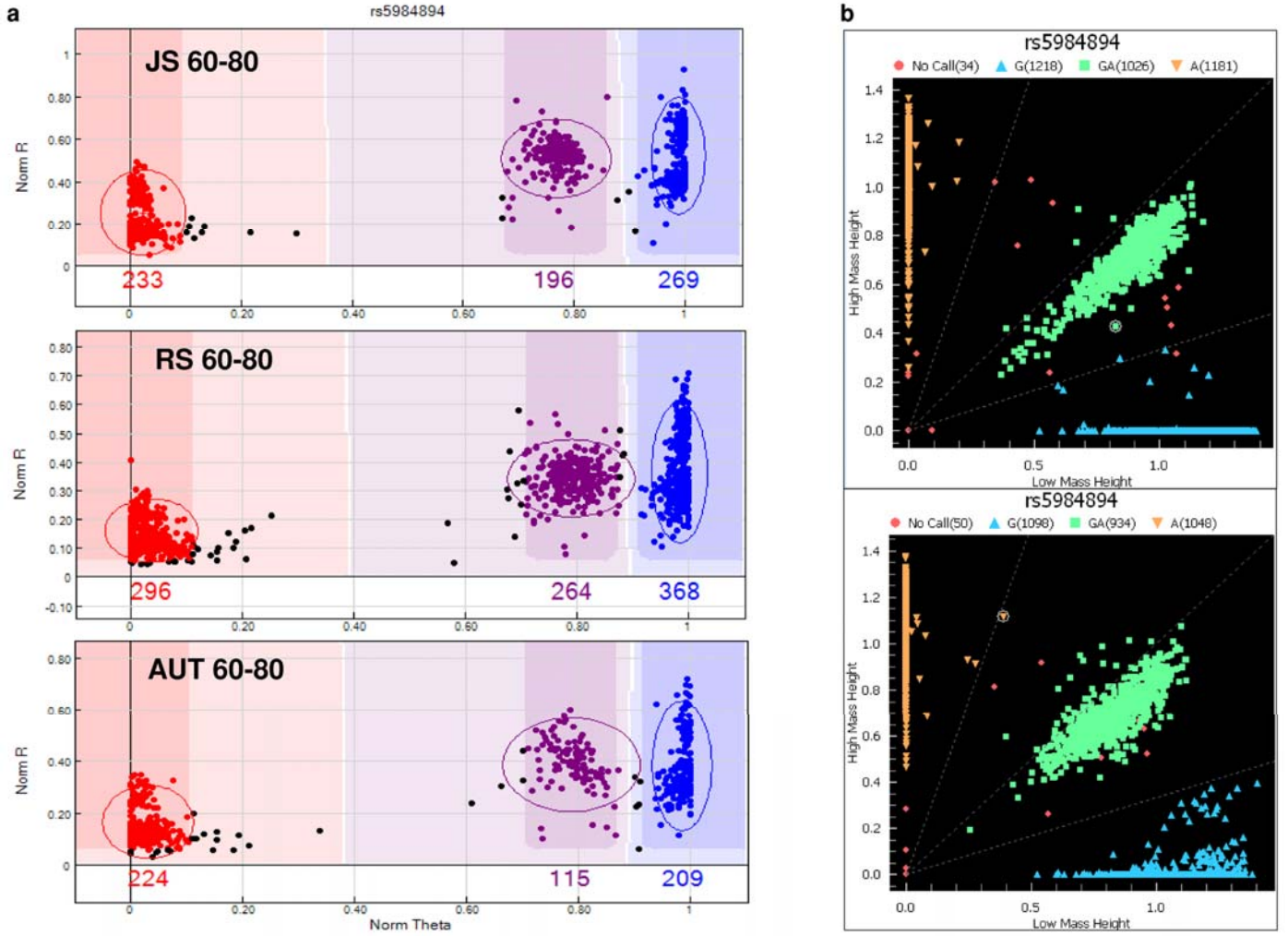
	OR			
	Bonferroni corrected for 320,000 tests: $P < 1.56 \times 10^{-7}$		Empirically selecting the 25 most significant SNPs: $P \cong 7.81 \times 10^{-5}$	
MAF	970 cases 1,495 controls	844 cases 1,255 controls	970 cases 1,495 controls	844 cases 1,255 controls
0.05	1.88	1.96	1.68	1.71
0.10	1.63	1.68	1.49	1.51
0.20	1.47	1.51	1.36	1.38
0.30	1.41	1.45	1.32	1.34
0.40	1.39	1.43	1.31	1.32
0.50	1.39	1.43	1.30	1.32

b

MAF	OR
0.05	1.53
0.10	1.38
0.20	1.28
0.30	1.24
0.40	1.23
0.50	1.23

Nature Genetics: doi:10.1038/ng.305

Supplementary Figure 1. rs5984894 genotype cluster plots**. (a)** BeadStudio 2.0 cluster plots for rs5984894 in each of three stage I subseries (JS 60-80, RS 60-80 and AUT 60-80). Black symbols identify samples where no call was made. Note that we generated cluster plots for each of the stage I subseries individually. This was done to minimize the potential of obtaining scattered clusters as a consequence of combining samples that were prepared differently (JS DNA template = non-WGA DNA from blood; RS DNA template = WGA DNA from blood; AUT DNA template = WGA DNA from brain). Another reason for applying the clustering algorithm individually to each subseries is that the BeadStudio program is more efficient when given smaller sample sets. **(b)** Typer 4.0 cluster plots for rs5984894. Since only up to ten 384-well plates can be uploaded at once into Typer 4.0, the first 3,425 samples from the NCRAD stage II subseries are shown on the top plot and the remaining stage II samples (JS 80+, RS 80+, AUT 80+ and additional NCRAD samples) are shown on the bottom plot. Note that in addition to the NCRAD samples used in the case-control association analyses that we report, many more NCRAD samples were genotyped for future family based association analyses.



Nature Genetics: doi:10.1038/ng.305