The SERPIN E2 gene and chronic obstructive pulmonary disease.

Single-Nucleotide Polymorphism *rs498055* on Chromosome 10q24 Is Not Associated with Alzheimer Disease in Two Independent Family Samples

To the Editor: In the January issue of the *American Journal of Human Genetics,* Grupe and colleagues[1] published evidence suggesting genetic association between SNP *rs498055* on chromosome 10q24, located in a putative homologue of ribosomal protein S3a (*RPS3A* [MIM 180478]), and risk for Alzheimer disease (AD [MIM 104300]) in four of six independent case-control samples. The authors reached this conclusion after testing nearly 1,400 SNPs, using an exploratory case-control sample, followed by assessments of a number of independent data sets of different size, origin, and ascertainment. Although three of the replication samples showed significant risk effects for the G allele of *rs498055,* this effect was not confirmed in two smaller series of neuropathologically confirmed AD cases and controls. None of the other 68 "hits" uncovered in the first pass received the same degree of consistent replication as did *rs498055.* Overall, the effect of the putative risk allele was modest (yielding odds ratios [ORs] between ~1.3 and 1.4) and—according to the authors' conclusion—likely reflects linkage disequilibrium (LD) with another genetic variant nearby. Here, we have set out to independently assess the association between *rs498055* and AD risk in two large and carefully characterized samples of AD-affected families comprising nearly 1,900 subjects from 654 pedigrees. However, in contrast to the findings of Grupe and colleagues, we observed no evidence of association between *rs498055* and AD in any of our analyses.

Using high-efficiency fluorescence polarization (HEFP) technology, we genotyped this SNP in two family-based AD samples: (1) 1,439 subjects from 437 multiplex AD-affected families recruited as part of the National Institute of Mental Health (NIMH) Genetics Initiative AD Study Sample (average age at onset [± SD] of affected individuals was 72.4 ± 7.7 years) and (2) 489 subjects from 217 independent families, mostly consisting of discordant sibships, recruited as part of the Consortium on Alzheimer's Genetics (CAG) (average age at onset was 71.2 ± 9.1 years). These samples, as well as the genotyping procedures, are described in detail in the work of Bertram et al.[2]; PCR and HEFP primer sequences for *rs498055* are available on request. Average genotyping efficiency across both samples was 98.4%, with a genotyping error rate <1% (on the basis of ~10% duplicated samples). Power analyses (fig. 1) in the combined sample showed that, at a disease-allele frequency of 0.47 (i.e., the average frequency of the G allele in U.S. controls reported by Grupe et al.) and $\alpha = .05$, power was 64% for an OR of 1.3 and was 83% for an OR of 1.4 (see fig. 1 for more details). Naturally, power was lower for the two samples considered separately, but it was still 40%–60% for the NIMH sample alone, comparable to the power of the replication samples in the study by Grupe et al.

In contrast to the findings of Grupe et al., we did not observe any significant evidence of association between *rs498055* and AD risk, neither in the two samples individually nor after combining both data sets (table 1), overall or when stratified by age at onset (with age 65 years as cutoff) or apolipoprotein E (*APOE* [MIM 107741]) ε4-carrier status. Interestingly, and in contrast to the overtransmission of the G allele noted by Grupe et al., in our two samples, this allele was generally undertransmitted to affected individuals, which approached statistical significance in two of our stratified analyses in the combined sample ($P = .09$ in "late-onset" families, and $P = .06$ in "*APOE* ε4-positive" families [table 1]). Finally, we also tested for association between *rs498055* and age at onset of AD (age at last examination of unaffected individuals), using FBAT-LOGRANK, FBAT-Wilcoxon, and FBAT-Flemington-Harrington[3,4] in the unstratified samples. However, none of these tests showed even marginally significant *P* values (data not shown). Since quantitative trait analyses are expected to be more powerful than analyses of binary traits if the underlying association is true,[3] these results strengthen our overall negative conclusion.

Our study is the first to independently assess the potential association between *rs498055* and AD that had emerged from a semisystematic screen of 1,397 SNPs on chromosome 10.[1] The fact that we failed to replicate the previous findings is noteworthy for several reasons. First, *rs498055* is located within the chromosome 10q24 linkage peak reported elsewhere for this collection of NIMH families.[5] Thus, our sample should be particularly well suited to detect disease associations underlying this linkage signal. Second, our study is the first to analyze this SNP with use of family-based methodologies in which affected subjects are compared with related unaffected subjects from the same family. Results from such analyses are more robust to bias due to population admixture or other sources of skewed genotype distributions in cases or controls; this is of particular note, given the differences in allele frequencies reported for the two control populations from the St. Louis area in the work of Grupe et al. (see below). Despite the strengths of our approach, it is possible that we have missed a putative risk effect at *rs498055* because of insufficient power, especially when aiming to detect minor effects with ORs of ⩽1.3 (fig. 1). However, the differences between our findings and those of Grupe et al. are unlikely to result from lack of power alone, since we see underrather than overtransmission of the G allele in both samples. It is unclear whether these discrepancies are caused by chance or by differential patterns of LD across the various samples.

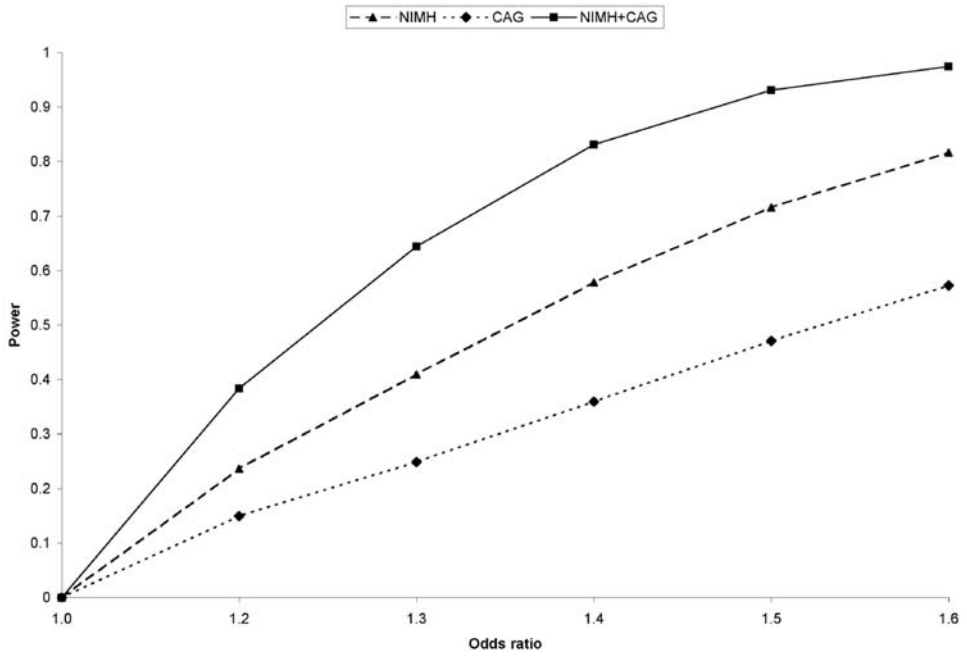

Figure 1. Power to detect a range of effect sizes in the family samples analyzed. Power estimates were done with PBAT (v3.1).[12] Estimates are based on approximation and are calculated for an additive disease model based on parameters published by Grupe et al.[1] (i.e., disease-allele frequency of 0.47 and OR 1.2–1.6), with the exception of disease prevalence, which was set to 10%. Although the precise prevalence of AD is unknown and difficult to estimate, power does not change appreciably when prevalence is varied from 5% to 15% (data not shown). ORs are for heterozygous carriers of the disease allele versus homozygous noncarriers. Families were modeled after the observed pedigree structure for each sample, with both parents set as "missing." Note that PBAT can currently handle a maximum offspring number of only four; however, 68 (16%) of NIMH pedigrees actually have more than four genotyped and phenotyped offspring, so that the power for "NIMH" and "NIMH+CAG" is likely to be underestimated (see the PBAT Web site for more details).

The long arm of chromosome 10 has been a focus of work for many AD genetics laboratories since the discovery of significant linkage with AD phenotypes by three independent groups, including ours (AD6 [MIM 605526]).[5–7] These publications were followed by two additional studies suggesting the presence of an AD risk and/or age-at-onset–modifying gene on this chromosome.[8,9] Although nearly 30 positional candidate genes have since been assessed as potential AD risk factors underlying these linkage signals and several positive association findings have been published, no gene has received consistent support from independent follow-up studies,[10] and none shows evidence of conclusive and significant summary effects in systematic meta-analyses of all published and available genotype data (AlzGene).

Unfortunately, the present failure to replicate the promising results of Grupe and colleagues is consistent with this overall pattern. There are several possible reasons for the differences between our findings and theirs. First, the difference might be due to chance, because the initial finding is a false-positive result. The 69 hits among 1,397 SNPs in the exploratory data set is close to the expected value by chance alone, as is the confirmation of 5 of these 69 signals in at least one of the two direct follow-up samples. However, we agree with the authors that a significant overrepresentation of the same allele in three of five confirmation samples, as observed for *rs498055,* may be unlikely to occur by chance alone. Second, the difference might have arisen by chance because our finding is a false-negative result. Although this is possible, it should be noted that our sample is as large or larger than many replication samples in the field. In addition, the difference is unlikely to result from insufficient power alone, because the putative risk allele, if anything, is undertransmitted in our samples. Third, the differences may relate to our use of family-based methods, which are more robust to bias due to population admixture. Although the degree to which admixture may lead to spurious association findings in case-control samples is controversial,[11] the issue is a concern here, given the marked difference in allele frequencies across the two independent Washington University control samples (47% for the case-control sample—similar to the other U.S. control sample—and 44% for the controls used in comparison with the linkage sample), a difference substantial enough that the allelic association between *rs498055* and AD in the linkage sample (49.8% risk-allele frequency) would not have been significant had the other Washington University control set (or the University of California–San Diego controls) been used. Finally, the differences across studies may be due to differences in pat-

Table 1. Association Analyses of *rs498055* in Two Independent Family Samples

Sample[a]	G Allele Frequency	FBAT Statistic Result: No. of Informative Families	FBAT Statistic Result: *Z* Score[b]	FBAT Statistic Result: *P*
All Families:				
NIMH	.519	123	−1.096	.27
CAG	.504	84	−.194	.85
Combined	.514	207	−1.001	.32
Families with late-onset disease:				
NIMH	.537	81	−1.541	.12
CAG	.531	66	−.808	.42
Combined	.535	147	−1.703	.09
APOE ε4-positive families:				
NIMH	.494	107	−1.488	.14
CAG	.515	31	−1.380	.17
Combined	.499	138	−1.929	.06

NOTE.—Association tests were performed using FBAT (v1.5.5) with an additive transmission model, the empirical variance function, and an equal-weight offset correction for affected and unaffected individuals (see the FBAT Web site for more details).

[a] Families were classified as "late onset" when all sampled affected individuals had age at onset of >65 years and were classified as "*APOE* ε4 positive" when at least one affected individual per family carried the ε4 allele. The smaller strata of remaining families (i.e., those displaying an earlier age at onset or those in which none of the affected individuals carried an ε4 allele) also failed to show evidence of significant association (data not shown).

[b] For the G allele of *rs498055,* which was reported as the putative risk allele by Grupe et al.[1] (positive values indicate overtransmission to affected individuals). Note that the direction of transmission is consistent for both family samples analyzed here and is opposite to that seen in the previous publication.[1]

terns of LD across the various samples, which are impossible to assess as long as the precise nature of the putative risk allele at this locus remains unknown.

Clearly, additional analyses of sufficiently powered and independent samples are needed to assert whether *rs498055,* or a polymorphism in LD with it, makes a relevant contribution to AD risk. At least in the two family samples investigated here—one of which shows linkage to the same chromosomal interval as *rs498055*—this SNP is not a major determinant of AD risk.

LARS BERTRAM, MONICA HSIAO, CHRISTOPH LANGE, DEBORAH BLACKER, AND RUDOLPH E. TANZI

Acknowledgments

The authors thank all families for participating in this study. This work was sponsored by grants from the NIMH, the National Institute on Aging (NIA) (at the Alzheimer Disease Research Center), and the Alzheimer Association. L.B. is supported by NIA grant 1R01 AG023667-01 and by a Young Investigator Award from the National Alliance for Research on Schizophrenia and Depression.

Web Resources

The URLs for data presented herein are as follows:

AlzGene, http://www.alzgene.org/ (Bertram L, McQueen MB, Mullin K, Blacker D, Tanzi R. The AlzGene Database–Published AD Candidate Genes; Alzheimer Research Forum; last accessed 3/1/06)

FBAT, http://www.biostat.harvard.edu/~fbat/fbat.htm

Online Mendelian Inheritance in Man (OMIM), http://www.ncbi.nlm.nih.gov/Omim/ (for *RPS3A,* AD, *APOE,* and AD6)

PBAT, http://www.biostat.harvard.edu/~clange/default.htm

References

1. Grupe A, Li Y, Rowland C, Nowotny P, Hinrichs AL, Smemo S, Kauwe JSK, et al (2006) A scan of chromosome 10 identifies a novel locus showing strong association with late-onset Alzheimer disease. Am J Hum Genet 78:78–88
2. Bertram L, Hiltunen M, Parkinson M, Ingelsson M, Lange C, Ramasamy K, Mullin K, Menon R, Sampson AJ, Hsiao MY, Elliott KJ, Velicelebi G, Moscarillo T, Hyman BT, Wagner SL, Becker KD, Blacker D, Tanzi RE (2005) Family-based association between Alzheimer's disease and variants in UBQLN1. N Engl J Med 352:884–894
3. Lange C, DeMeo DL, Laird NM (2002) Power and design considerations for a general class of family-based association tests: quantitative traits. Am J Hum Genet 71:1330–1341
4. Jiang H, Harrington D, Raby BA, Bertram L, Blacker D, Weiss ST, Lange C (2006) Family-based association test for time-to-onset data with time-dependent differences between the hazard functions. Genet Epidemiol 30:124–132
5. Bertram L, Blacker D, Mullin K, Keeney D, Jones J, Basu S, Yhu S, McInnis MG, Go RC, Vekrellis K, Selkoe DJ, Saunders AJ, Tanzi RE (2000) Evidence for genetic linkage of Alzheimer's disease to chromosome 10q. Science 290:2302–2303
6. Ertekin-Taner N, Graff-Radford N, Younkin LH, Eckman C, Baker M, Adamson J, Ronald J, Blangero J, Hutton M, Younkin SG (2000) Linkage of plasma Aβ42 to a quantitative locus on chromosome 10 in late-onset Alzheimer's disease pedigrees. Science 290:2303–2304

7. Myers A, Holmans P, Marshall H, Kwon J, Meyer D, Ramic D, Shears S, Booth J, DeVrieze FW, Crook R, Hamshere M, Abraham R, Tunstall N, Rice F, Carty S, Lillystone S, Kehoe P, Rudrasingham V, Jones L, Lovestone S, Perez-Tur J, Williams J, Owen MJ, Hardy J, Goate AM (2000) Susceptibility locus for Alzheimer's disease on chromosome 10. Science 290:2304–2305
8. Li Y-J, Scott WK, Hedges DJ, Zhang F, Gaskell PC, Nance MA, Watts RL, et al (2002) Age at onset in two common neurodegenerative diseases is genetically controlled. Am J Hum Genet 70:985–993
9. Lee JH, Mayeux R, Mayo D, Mo J, Santana V, Williamson J, Flaquer A, Ciappa A, Rondon H, Estevez P, Lantigua R, Kawarai T, Toulina A, Medrano M, Torres M, Stern Y, Tycko B, Rogaeva E, St George-Hyslop P, Knowles JA (2004) Fine mapping of 10q and 18q for familial Alzheimer's disease in Caribbean Hispanics. Mol Psychiatry 9:1042–1051
10. Bertram L, Tanzi RE (2004) Alzheimer's disease: one disorder, too many genes? Hum Mol Genet Spec No 1 13:R135–R141
11. Hoggart CJ, Parra EJ, Shriver MD, Bonilla C, Kittles RA, Clayton DG, McKeigue PM (2003) Control of confounding of genetic associations in stratified populations. Am J Hum Genet 72:1492–1504
12. Lange C, DeMeo D, Silverman EK, Weiss ST, Laird NM (2004) PBAT: tools for family-based association studies. Am J Hum Genet 74:367–369

From the Genetics and Aging Research Unit, MassGeneral Institute for Neurodegenerative Diseases, Department of Neurology (L.B.; M.H.; R.E.T.), and Gerontology Research Unit, Department of Psychiatry (D.B.), Massachusetts General Hospital, Harvard Medical School, Charlestown, MA; and Departments of Biostatistics (C.L.) and Epidemiology (D.B.), Harvard School of Public Health, Boston

Address for correspondence and reprints: Dr. Lars Bertram, Genetics and Aging Research Unit, MGH-East, MassGeneral Institute for Neurodegenerative Diseases, 114 16th Street, Charlestown, MA 02129. E-mail: bertram@helix.mgh.harvard.edu

Am. J. Hum. Genet. 2006;79:180–183.

0002-9297/2006/7901-0022$15.00

Reply to Bertram et al.

To the Editor: The study by Bertram and colleagues (in this issue)[1] failed to replicate, in two family-based sample sets, the association of *rs498055* with Alzheimer disease (AD [MIM 104300]) that we observed in four large, well-characterized case-control sample sets.[2] Although the result is disappointing, there are several differences between the studies that may have contributed to these discrepant findings. First, there are significant differences in the study designs. Bertram et al. used two family-based sample sets that included subjects with both early- and late-onset AD (e.g., 320 families with late-onset AD and 117 families with early/mixed-onset in the National Institute of Mental Health [NIMH] sample set) of different ethnicities (94% white; 6% others),[3] which resulted in 147 informative families with late-onset AD for both sample sets combined. The characterization of their unaffected controls was based on self-assessment or a telephone interview, a procedure sufficient when "unaffecteds" are used solely to determine phase in linkage studies, but likely to significantly impact power in association studies, especially when familial loading is high, as it is in the sample of Bertram et al. Indeed, the authors acknowledge this in one of their previous publications by pointing out that the characterization of controls "may miss some mild cases of dementia" and lead "to a decrease in power."[3] In contrast, our study included only clinically evaluated, late-onset cases and nondemented controls of white origin. Second, the use of a family-based sample that was ascertained on the basis of multiple affected relatives is likely to particularly adversely impact power to detect a risk allele of relatively high frequency and small effect size, such as *rs498055*. Under these circumstances, the allele frequency in unaffected relatives also increases,[4] with consequent loss of power in comparison with case-control studies such as our own. To investigate this more fully, we compared the allele frequencies for a known genetic risk factor for AD, apolipoprotein E (*APOE* [MIM 107741]), and for the putative risk factor under debate, *rs498055,* in our combined case-control series and in the NIMH linkage families used by us in the study described by Myers et al.[5] In this context, it is worth noting that 355 of 372 individuals from the linkage sample–derived cases in our recent publication overlap with affected individuals in the NIMH family sample set described by Bertram et al. For the comparison, we identified the subgroup of NIMH families with genotypes for at least one unaffected and one affected individual and then selected at random one unaffected and one affected individual from each of these families. Table 1 illustrates clearly that the frequency of the *APOE4* allele is substantially higher in unaffected individuals from the linkage families than in unaffected individuals from the case-control series (30.4% vs. 12.5%) and that, although the *APOE4* allele frequency is highest in the linkage cases, the difference between the unrelated cases and controls is much greater than that between familial cases and related controls (35.6% vs. 12.5% compared with 42.8% vs. 30.4%). As a result, the odds ratio (OR) for the *APOE4* allele in the case-control series is 3.8, compared with only 1.7

Table 1. Frequency of AD Risk Alleles Is Higher in Unaffected Individuals from Multiply Affected Families Than in Unrelated Controls

	No. (%) of Subjects			
	Cases		Controls	
Locus and Allele	Linkage	CC	Linkage	CC
APOE:				
2	7 (2.3)	80 (4.2)	16 (5.2)	181 (8.5)
3	169 (54.9)	1,144 (60.2)	197 (64.4)	1,691 (79)
4 (Risk allele)	132 (42.8)	676 (35.6)	93 (30.4)	268 (12.5)
rs498055:				
A	133 (46.2)	915 (47.9)	140 (46.4)	1,174 (54.4)
G (Risk allele)	155 (53.8)	995 (52.1)	162 (53.6)	984 (45.6)

NOTE.—CC = case-control study.

for the family-based samples. Similarly for *rs498055,* the difference in frequencies between the cases and controls is greater for the unrelated samples than for the linkage families (table 1). Thus, the failure of Bertram et al. to replicate our results does not necessarily indicate that the original association was a false-positive result. We concur with Bertram et al. that the significant association of *rs498055* in four of six samples "may be unlikely to occur by chance"[1(p181)] (in this issue). However, it is possible that our initial study provided an overestimate of the allelic OR for *rs498055*. If this were true and the OR were <1.3, then the study by Bertram et al. would clearly be underpowered. Further replication in well-characterized sample sets is required to assess whether the association is genuine. Ideally, this should be done with large case-control sample sets, to achieve maximum power. For this particular marker, we estimate that 360 cases and 360 controls are needed to achieve 80% power in a replication study (one-sided $\alpha = .05$), assuming an allelic OR of 1.3 and a risk-allele frequency of 45.6%. A meta-analysis of all studies should then be performed to determine whether *rs498055* is associated with late-onset AD. In addition, it might be interesting to test the other reported significant markers from this region in additional sample sets.

Andrew Grupe, Yonghong Li, Charles Rowland, Tony Hinrichs, Peter Holmans, John Hardy, Michael O'Donovan, Michael J. Owen, Julie Williams, and Alison Goate

Acknowledgments

Funding for this work was partly provided by National Institutes of Health grant RO1 AG16208 (to A. Goate), the Medical Research Council, UK (to J.W., M.J.O., and M.O.), and the Alzheimer's Research Trust (to J.W. and M.J.O). J.H. is supported by the National Institutes of Health intramural program and by the VERUM Foundation (DIADEM project).

Web Resource

The URL for data presented herein is as follows:

Online Mendelian Inheritance in Man (OMIM), http://www.ncbi.nlm.nih.gov/Omim/ (for AD and *APOE*)

References

1. Bertram L, Hsiao M, Lange C, Blacker D, Tanzi RE (2006) Single-nucleotide polymorphism *rs498055* on chromosome 10q24 is not associated with Alzheimer disease in two independent family samples. Am J Hum Genet 79:180–183 (in this issue)
2. Grupe A, Li Y, Rowland C, Nowotny P, Hinrichs AL, Smemo S, Kauwe JSK, et al (2006) A scan of chromosome 10 identifies a novel locus showing strong association with late-onset Alzheimer disease. Am J Hum Genet 78:78–88
3. Bertram L, Hiltunen M, Parkinson M, Ingelsson M, Lange C, Ramasamy K, Mullin K, Menon R, Sampson AJ, Hsiao MY, Elliott KJ, Velicelebi G, Moscarillo T, Hyman BT, Wagner SL, Becker KD, Blacker D, Tanzi RE (2005) Family-based association between Alzheimer's disease and variants in UBQLN1. N Engl J Med 352:884–894
4. Risch N, Teng J (1998) The relative power of family-based and case-control designs for linkage disequilibrium studies of complex human diseases I: DNA pooling. Genome Res 8:1273–1288
5. Myers A, Wavrant De Frieze F, Holmans P, Hamshere M, Crook R, Compton D, Marshall H, et al (2002) A full genome screen for Alzheimer's disease: stage two analysis. Neuropsychiatric Genetics 114:235–244

From Celera Diagnostics, Alameda, CA (A. Grupe; Y.L.; C.R.); Departments of Psychiatry, Neurology, and Genetics, Washington University, St. Louis (T.H.; A. Goate); Biostatistics and Bioinformatics Unit and Department of Psychological Medicine, Wales College of Medicine, Cardiff University, Cardiff (P.H.; M.O.; M.J.O.; J.W.); and National Institute on Aging, Bethesda (J.H.)

Address for correspondence and reprints: Dr. Alison Goate, Department of Psychiatry, B8134, Washington University School of Medicine, 660 S. Euclid Avenue, St. Louis, MO 63110. E-mail: goate@icarus.wustl.edu

Am. J. Hum. Genet. 2006;79:183–184.

0002-9297/2006/7901-0023$15.00

The *SERPINE2* Gene and Chronic Obstructive Pulmonary Disease

To the Editor: In the February 2006 issue of the *Journal,* DeMeo et al.[1] identified *SERPINE2* as a positional candidate gene for susceptibility to chronic obstructive pulmonary disease (COPD [MIM 606963]) and reported on the association of polymorphic variants of this gene with early-onset disease in a family-based study and with severe disease in a case-control study. With early prior information provided by the authors, we have independently tested for an association of the *SERPINE2* gene with COPD in the largest case-control study reported to date. Our study consists of 1,018 COPD cases and 911 controls prospectively recruited from six European centers. We have provided details about the patients elsewhere.[2] The study population was screened for genotypes at the Medical Research Council (United Kingdom) Gene Services Unit for five SNPs (table 1) in the *SERPINE2* gene. All the SNPs evaluated were reported in the study by DeMeo et al. as associated with disease, with three of the five associated with disease in both the family and case-control study cohorts they assessed.

Table 1. LD between *SERPINE2* SNPs Expressed as r^2

	r^2 for SNP			
SNP	*rs1438831*	*rs920251*	*rs6747096*[a]	*rs3795879*
rs920251	.952 (1.0)	…	…	…
rs6747096[a]	.140	.148	…	…
rs3795879[a]	.140 (.145)	.145 (.145)	.964	…
ss49785625[a]	.020	.023	.054	.055

NOTE.—The r^2 values in parentheses are values obtained from HapMap and compared with our own data in controls. *ss49785625* and *rs6747096* are not in HapMap.

[a] SNP reported by DeMeo et al.[1] to be associated with disease in both family and case-control cohorts.

We examined linkage disequilibrium (LD) between the SNPs (table 1) and evaluated SNP and haplotype associations as described elsewhere.[2] DeMeo et al. did not report specific LD values between SNPs or noncontiguous SNPs contributing to haplotypes. SNPs and genotype frequencies in the study population are shown in table 2. We found no significant deviation from Hardy-Weinberg equilibrium in frequencies for any of the SNPs.

We found no association between any of the *SERPINE2* SNPs and disease, in examining both the allelic and genotype distributions, although our study was well powered to detect associations of the magnitude observed by DeMeo et al., and we would have expected to see these frequency differences with the SNPs that we studied. We also failed to find a relationship between any haplotypes of these SNPs and disease (data not shown). It was of interest that the allele and genotype frequencies observed in our control and patient groups were virtually identical to those observed in control subjects by DeMeo et al., indicating a common distribution of *SERPINE2* variants in the European and North American populations studied. Our previous study has also shown that there is no evidence of population stratification in our sample.

Patients evaluated in both the family-based and case-control studies reported by DeMeo et al. represent a severe subset of the disease spectrum. To determine whether the association with *SERPINE2* noted by DeMeo et al. was related to disease severity, we also analyzed SNP allele and genotype frequencies in the subgroup of our patients with forced expiratory volume at 1 s $\leqslant$45% (n = 388), a group that represents severe disease, but we failed to observe any association.

Our inability to replicate the observations of DeMeo et al. in a more highly powered case-control study may be related to differences in the disease phenotype of the patients studied, because our patients included those with and without emphysema. The possibility, however, that the associations reported by DeMeo et al. represent false-positive results must also be considered. In this respect, it is of note that, in the study by DeMeo et al., different associations were reported for SNPs that are in linkage disequilibrium with one another. For example, *rs3795879* and *rs3795877* have an r^2 value of 1 in HapMap, yet different associations with quantitative spirometric phenotypes were reported for the family study. Similarly, *rs1438831* and *rs920251* are in complete LD, with an r^2 value of 1 in HapMap and 0.95 in our study; however, in DeMeo et al.'s case-control study, the allele and genotype frequencies of *rs920251* were found to be significantly associated with disease (P values of 0.015 and 0.011, respectively), whereas no similar association was observed for *rs1438831*. In both instances, the almost complete linkage between these pairs of SNPs would be expected to result in similar associations.

These results underline the importance of replication in other large independent studies before *SERPINE2* can be unequivocally assigned as a candidate gene for COPD. It is becoming apparent that, to detect modest genetic effects for complex diseases, several independent studies may be required and the data may need to be subjected to meta-analysis. For example, this approach has been used to study Alzheimer disease (see Alzheimer's Association Web site). Similar approaches need to be adopted for COPD. It would also be helpful to have similar criteria adapted for phenotypic selection and to plan prospective studies on this basis.

Table 2. *SERPINE2* Genotype and Allele Frequencies in Controls and COPD Cases

SNP and Sample	Frequency of Allele		Frequency of Genotype		
	C	T	CC	CT	TT
rs1438831:					
COPD case	.66	.34	.43	.45	.12
Control	.66	.34	.43	.46	.11
	A	G	AA	AG	GG
rs920251:					
COPD case	.35	.65	.13	.45	.42
Control	.35	.65	.12	.46	.42
	A	G	AA	AG	GG
rs6747096:					
COPD case	.79	.21	.61	.35	.04
Control	.79	.21	.63	.33	.04
	C	T	CC	CT	TT
rs3795879:					
COPD case	.78	.22	.60	.36	.05
Control	.79	.21	.62	.33	.05
	A	G	AA	AG	GG
ss49785625:					
COPD case	.54	.46	.30	.48	.22
Control	.53	.47	.29	.48	.23

Sally Chappell, Leslie Daly, Kevin Morgan, Tamar Guetta Baranes, Josep Roca, Roberto Rabinovich, Ann Millar, Seamas C. Donnelly, Vera Keatings, William MacNee, Jan Stolk, Pieter S. Hiemstra, Massimo Miniati, Simonetta Monti, Clare M. O'Connor,* and Noor Kalsheker*

Acknowledgments

We thank Ed Silverman and Dawn DeMeo for providing us with detailed information about the *SERPINE2* SNPs. This work was supported by European Union 5th Framework Programme grant QLG1-CT-2001-01012.

Web Resources

The URLs for data presented herein are as follows:

Alzheimer's Association, http://www.alz.org/
Online Mendelian Inheritance in Man (OMIM), http://www.ncbi.nlm.nih.gov/Omim/ (for COPD)

References

1. DeMeo DL, Mariani TJ, Lange C, Srisuma S, Litonjua AA, Celedon JC, Lake SL, Reilly JJ, Chapman HA, Mecham BH, Haley KJ, Sylvia JS, Sparrow D, Spira AE, Beane J, Pino-Plata V, Speizer FE, Shapiro SD, Weiss ST, Silverman EK (2006) The *SERPINE2* gene is associated with chronic obstructive pulmonary disease. Am J Hum Genet 78:253–264
2. Chappell S, Daly L, Morgan K, Guetta Baranes T, Roca J, Rabinovich R, Millar A, Donnelly S, Keatings V, MacNee W, Stolk J, Hiemstra P, Miniati M, Monti S, O'Connor CM, Kalsheker N (2006) Cryptic haplotypes of *SERPINA1* confer susceptibility to chronic obstructive pulmonary disease. Hum Mutat 27:103–109

From the University of Nottingham and Division of Clinical Chemistry, Molecular Medical Sciences, Institute of Genetics, University Hospital, Queens Medical Centre, Nottingham, United Kingdom (S.C.; K.M.; T.G.B.; N.K.); Departments of Public Health Medicine and Epidemiology (L.D.) and Medicine and Therapeutics (S.C.D; C.M.O.), The Conway Institute, University College Dublin, Dublin; Service de Pneumologia, Hospital Clinico y Provincial de Barcelona, Barcelona (J.R.; R.R.); University of Bristol and Lung Research Group, Department of Clinical Science at North Bristol, Southmead Hospital, Bristol, United Kingdom (A.M.); Letterkenny General Hospital, Letterkenny, Ireland (V.K.); University of Edinburgh, Respiratory Medicine, ELEGI Colt Laboratories, Edinburgh (W.M.); Department of Pulmonology, Leiden University Medical Center, Leiden, The Netherlands (J.S.; P.S.H.); and CNR Institute of Clinical Physiology, Pisa, Italy (M.M.; S.M.)

Address for correspondence and reprints: Dr. Noor Kalsheker, The University of Nottingham, Division of Clinical Chemistry, Institute of Genetics, Queens Medical Centre, Nottingham, NG7 2UH, United Kingdom. E-mail: noor.kalsheker@nottingham.ac.uk

Am. J. Hum. Genet. 2006;79:184–186.

0002-9297/2006/7901-0024X$15.00

* These two authors contributed equally to this work.

Reply to Chappell et al.

To The Editor: We appreciate the efforts of Chappell and colleagues[1] to replicate our *SERPINE2* findings. We identified *SERPINE2* as a candidate gene for chronic obstructive pulmonary disease (COPD [MIM 606963]) on the basis of our gene-expression results (in both murine and human lung) and our genetic association analysis results in two study populations. Chappell et al. found no evidence for association of five *SERPINE2* SNPs with COPD in their case-control study. As in many complex-disease genetic association studies in general, and in previous COPD genetic association studies in particular,[2] the results are inconsistent.

There are many potential explanations for these inconsistent results, including population stratification, genetic heterogeneity, false-positive and/or false-negative results, differences in the number of SNPs genotyped, and phenotypic heterogeneity.[3] In comparing the results of our two research groups for association analysis of *SERPINE2* SNPs with COPD, phenotypic heterogeneity is of particular importance. COPD is a syndrome composed of both emphysema and airway disease, with variable contributions of these processes in different individuals with COPD. Review of chest CT scans of probands from the Boston Early-Onset COPD Study—the population in which we performed family-based association analysis of COPD-related phenotypes—revealed that the vast majority of these probands had emphysema.[4] Moreover, the COPD cases in our case-control replication population were clearly selected for emphysema as part of the National Emphysema Treatment Trial (NETT). In addition, the Boston Early-Onset COPD Study probands and the NETT cases had very severe COPD. Thus, our test and replication populations were severely affected with COPD, typically with a substantial degree of emphysema. As noted by Chappell et al., our cases represent "a severe subset of the disease spectrum,"[1(p185)] and their cases represent a broader spectrum of severity, including individuals with and without emphysema. The differences in disease severity and emphysema may be important contributors to their nonreplication of our association findings. Also of note, although Chappell et al. genotyped five SNPs in *SERPINE2,* they did not genotype several other SNPs for which we observed replicated associations and LOD score reduction in conditional linkage models.

Chappell et al. also comment about apparently inconsistent association results in our family-based and case-control association analyses among SNPs in tight linkage disequilibrium (LD). Modest differences in the statistical significance of the association analysis results were noted for several SNPs that are in strong but not complete LD in our study populations. There are reasonable explanations for these modest differences. (1) The SNP pairs mentioned are not in complete LD; in our combined case-control cohort, the r^2 values were 0.93 for *rs3795879* and *rs3795877* and 0.91 for *rs1438831* and *rs920251*. (2) Despite excellent genotype completion rates, there were slight differences in missing data between these SNP pairs. Of note, these were not the only *SERPINE2* SNPs significantly associated with COPD-related phenotypes in our study; we observed 18 significantly associated *SERPINE2* SNPs in the family-based association analysis and 7 significantly associated SNPs in the case-control analysis.

We fully agree with Chappell et al. that replication of significant associations is essential—which is why we included in our article the replication of our family-based association analysis results in a separate case-control study. This is also the reason why we provided early access to significantly associated SNPs to the Chappell and Kalsheker group.

Is *SERPINE2* a *confirmed* COPD susceptibility gene? Certainly not. Before the impact of *SERPINE2* on COPD susceptibility is fully known, more genetic association studies as well as functional studies will be needed. However, we contend that *SERPINE2* remains a valid COPD candidate gene. Finally, we agree with Chappell et al. that agreement on phenotypic definitions and collaboration between re-

search groups are crucial for the future of genetic studies of COPD and other complex diseases.

Web Resources

The URL for data presented herein is as follows:

Online Mendelian Inheritance in Man (OMIM), http://www.ncbi.nlm.nih.gov/Omim/ (for COPD)

References

1. Chappell S, Daly L, Morgan K, Guetta Baranes T, Roca J, Rabinovich R, Millar A, Donnelly SC, Keatings V, MacNee W, Stolk J, Hiemstra PS, Miniati M, Simonetta M, O'Connor CM, Kalsheker N (2006) The *SERPINE2* gene and chronic obstructive pulmonary disease. Am J Hum Genet 79:184–185 (in this issue)
2. Hersh CP, DeMeo DL, Lange C, Litonjua AA, Reilly JJ, Kwiatkowski D, Laird N, Sylvia JS, Sparrow D, Speizer FE, Weiss ST, Silverman EK (2005) Attempted replication of reported chronic obstructive pulmonary disease candidate gene associations. Am J Respir Cell Mol Biol 33:71–78
3. Silverman EK, Palmer LJ (2000) Case-control association studies for the genetics of complex respiratory diseases. Am J Respir Cell Mol Biol 22:645–648
4. Silverman EK, Chapman HA, Drazen JM, Weiss ST, Rosner B, Campbell EJ, O'Donnell WJ, Reilly JJ, Ginns L, Mentzer S, Wain J, Speizer FE (1998) Genetic epidemiology of severe, early-onset chronic obstructive pulmonary disease: risk to relatives for airflow obstruction and chronic bronchitis. Am J Respir Crit Care Med 157:1770–1778

From the Channing Laboratory and Division of Pulmonary and Critical Care Medicine, Department of Medicine, Brigham and Women's Hospital, and Harvard Medical School, Boston (D.L.D.; E.K.S)

Address for correspondence and reprints: Dr. Dawn L. DeMeo, Channing Laboratory, 181 Longwood Avenue, Boston, MA 02115. E-mail: redld@channing.harvard.edu

Am. J. Hum. Genet. 2006;79:186–187.

0002-9297/2006/7901-0025$15.00