Psychiatric Genetics 1999; **9**:7–11

Possible association between schizophrenia and a CAG repeat polymorphism in the spinocerebellar ataxia type 1 (SCA1) gene on human chromosome 6p23

E-J. Joo[a], J.H. Lee[a], T.D. Cannon[a,b] and R. Arlen Price[a]

[a]*Department of Psychiatry,* [b]*Department of Psychology, University of Pennsylvania, Philadelphia, PA 19104, USA*

Correspondence to Dr R. Arlen Price, Center for Neurobiology and Behavior, University of Pennsylvania, Clinical Research Building 135b, 415 Curie Boulevard, Philadelphia, PA 19104, USA. E-mail: arlen@bgl.psycha.upenn.edu

Received 22 October 1997; accepted 8 January 1998

The gene for spinocerebellar ataxia type 1 (SCA1) is a potential candidate gene for schizophrenia because of previous positive linkage findings in this region (6p22-24), and because the reported correlation between SCA1 onset and the number of CAG repeats suggests anticipation. To test the involvement of this gene in the development of schizophrenia, we examined genotypes of the SCA1 CAG repeat polymorphism for 49 Caucasian patients with schizophrenia, and 88 Caucasian controls. We found a significant association between the frequencies of alleles of this gene and schizophrenia (χ^2 = 18.40, df = 8, P = 0.018). Among 13 alleles, one allele (31 trinucleotide repeat) was significantly more frequent in patients with schizophrenia than in controls (χ^2 = 9.57, df = 1, P = 0.002). This association was sustained after applying a Bonferroni correction for multiple testing (P = 0.05/13 = 0.004), and the chi-square results were shown to be robust through Monte Carlo simulation. We observed no allelic association with three flanking microsatellite markers (D6S288, D6S1605, and D6S337), suggesting that our result was not due to population stratification. Further studies of this locus are needed to confirm this finding, and to determine a potential role for this gene in the development of schizophrenia. © **1999 Lippincott Williams & Wilkins.**

Keywords: chromosome 6, trinucleotide repeat polymorphism, psychosis, schizophrenia, disequilibrium, association, SCA1

INTRODUCTION

A genetic contribution to the development of schizophrenia is well documented by previous genetic epidemiological studies (Cannon *et al.*, 1998). Schizophrenia is a complex trait, having locus heterogeneity, an unclear mode of inheritance, variable penetrance, and possible gene-environment interactions. This complexity probably accounts for the equivocal results of many gene mapping studies, and the common failure of subsequent studies to replicate initial reports.

Several studies of chromosome 6p22-24 have reported positive LOD scores, suggesting linkage with schizophrenia. Straub *et al.* (1995) reported the strongest results with D6S296, D6S274, and D6S285 on chromosome 6p23. Other positive linkage results within this region have been reported by different groups (Moises *et al.*, 1995; Schwab *et al.*, 1995; Wang *et al.*, 1995), supporting chromosome 6p22-24 as a candidate region for susceptibility to schizophrenia. Not surprisingly, other research groups have failed to replicate this positive finding (Antonarakis *et al.*, 1995; Gurling *et al.*, 1995; Mowry *et al.*, 1995; Daniels *et al.*, 1997). However, two integrated studies provided additional support for a susceptibility locus for schizophrenia on chromosome 6p22-24. A collaborative study, which included 713 schizophrenia pedigrees, found a maximum LOD score of 2.68 for this region (Schizophrenia linkage collaborative group for chromosomes 3, 6 and 8, 1996). Another study by Turecki *et al.* (1997) reported significant pooled P-values, using meta-analysis of the results of linkage studies for chromosome 6p markers.

Among several genes located on chromosome 6p23, the gene for SCA1 (spinocerebellar ataxia type 1) is an appealing candidate, because it is located between markers D6S260 and D6S274 (Jodice *et al.*, 1993; Orr *et al.*, 1993), and contains an expansion of a CAG trinucleotide repeat within the coding region (Orr *et al.*, 1993; Banfi *et al.*, 1994). The number of trinucleotide repeats in this gene was reported to be correlated with age at onset of SCA1. The more severe phenotype usually found in later generations was associated with an earlier onset of this disease (Orr *et al.*, 1993; Ranum *et al.*, 1994). This pattern

suggested that the expansion of this CAG nucleotide repeat could be a mechanism for the anticipation observed in SCA1.

To date, two studies have examined the association between SCA1 and schizophrenia. Wang *et al.* (1996) reported a significant association between schizophrenia and a specific allele of SCA1 (29 CAG repeats), which was transmitted more frequently to patients with schizophrenia from heterozygous parents than other alleles. However, Morris-Rosendahl *et al.* (1997) could not find any significant difference in allelic frequencies in a French sample between unrelated patients with schizophrenia and controls.

To evaluate a possible genetic role of the SCA1 CAG trinucleotide repeat polymorphism in the development of schizophrenia, we performed an association study with 49 unrelated Caucasian patients with schizophrenia, and 88 unrelated Caucasian controls for the SCA1 CAG repeat polymorphic region. After a positive result was found, three adjacent polymorphic microsatellite markers flanking the SCA1 CAG repeat region were examined, to exclude population structure as a possible source of the observed association.

MATERIALS AND METHODS

Subjects

All patients with schizophrenia, and the matching controls, were recruited and evaluated at the Mental Health Clinical Research Center (MHCRC) in the Department of Psychiatry at the University of Pennsylvania, Philadelphia, USA. Psychiatric interviews using the SCID (Structured Clinical Interview for DSM-IIIR, Spitzer *et al.*, 1987) were performed for patients with schizophrenia (patient edition) and controls (non-patient edition). A physical examination and routine laboratory tests, including a drug toxicity screen, were completed for all subjects. DSM-IIIR criteria were used for the diagnosis of schizophrenia. Patients with schizophrenia, and control subjects who had other Axis I diagnoses, substance-abuse history, neurological history, head trauma, and medical diagnoses that may affect mental functions, were excluded from the study (for additional details see Shtasel *et al.*, 1991; Gur *et al.*, 1994). Control subjects were also excluded from the study if they had any Axis I diagnosis in their first degree relatives. The mean age of the patients with schizophrenia was 33 (SD = 11.2, $n = 49$) years, and the corresponding age for controls was 30 (SD = 12.9, $n = 88$) years. The mean age at onset of schizophrenia was 22 (SD = 6.8) years.

Genotyping

DNA was extracted from blood samples. Primers for SCA1 CAG repeat region were synthesized as described previously (Rep-1: 5′-AACTGGAAATGTGGACGTAC, Rep-2: 5′-CAACATGGGCAGTCTGAG; Orr *et al.*, 1993). Three flanking markers, D6S288, D6S1605, and D6S337, were chosen and purchased from Research Genetics (USA). The primers were end-labeled with ^{33}P-γ-ATP (Amersham) using T_4 polynucleotide kinase. A 10 μl polymerase chain reaction volume was used, consisting of 25 ng genomic DNA, 2.5 nmol each of dNTP, 20 mM Tris-HCl (pH 8.4), 50 mM KCl, 1.5 mM $MgCl_2$, 3 pmol forward primer, 3 pmol reverse primer, and 0.62 U Taq polymerase. The reaction was done under the following conditions: initial denaturation at 94°C for 5 min, two cycles of denaturation at 94°C for 1 min, annealing at 55°C for 2 min, extension at 72°C for 2 min, and 34 cycles at 94°C for 30 s, at 55°C for 45 s, at 72°C for 30 s, and a final extension time of 7 min at 72°C. Genotypes were visualized by electrophoresis of the polymerase chain reaction products on a 6% sequencing acrylamide gel, and exposure of the vacuum-dried gel to X-ray film for at least 48 h. Genotypes were scored independently by three observers who were blind to the diagnostic status of the subject. Four subjects were sequenced, to confirm product specificity and number of trinucleotide repeats.

Statistical analysis

A contingency χ^2 test was performed to compare all allele frequencies between patients with schizophrenia and controls. We included the allele if its calculated expected value was greater than 1. To determine whether any specific allele was associated with schizophrenia, we tested each allele separately by applying a 2×2 contingency χ^2 test. $P < 0.05$ was used as the significance level of an individual test, after adjusting for the number of alleles tested. Following a significant result for SCA1, we tested flanking markers for association, to determine whether the observed association might be due to unknown population substructure.

To ensure against potential problems associated with the conventional chi-square statistic applied to a contingency table with sparse data, we employed a Monte Carlo method to compute chi-square statistics, as implemented by Clump (Sham and Curtis, 1995), and simultated 2000 replicates to compute empirical *P*-values. Once the association was confirmed, the etiologic fraction attributable to the significant allele was computed (Schlesselman, 1982).

We then calculated Pearson correlation to determine whether repeat number predicted age at onset for patients with schizophrenia. We used two different models for the CAG repeats: for a dominant model, the size of the larger allele of each individual was used, and for an additive model, the sum of both allele sizes was used.

RESULTS

We found a significant difference in allele frequencies between patients with schizophrenia and controls ($\chi^2 = 18.40$, df = 8, $P = 0.018$). The association remained significant, whether we included nine alleles with the expectation of 1 or greater, or all 13 alleles. Table 1 shows the distribution of the alleles of SCA1 CAG repeats for patients with schizophrenia and controls. The size of the trinucleotide repeat ranged from 15 to 35 for patients with schizophrenia, and from 26 to 39 for controls. All subjects with schizophrenia and controls were in the normal range with respect to the number of CAG repeats on SCA1 gene (Orr *et al.*, 1993; Jodice *et al.*, 1994). The 29 and 31 repeat alleles were the two most common alleles in our subjects, as well as in Wang's multiplex families with schizophrenia (1996) and French subjects (Morris-Rosendahl *et al.*, 1997). Sequencing of four individuals confirmed the size of the repeat and detected 3-4 CAT elements near the center of this region, consistent with published sequence (GeneBank accession # S64648).

When each allele was tested by 2×2 contingency χ^2 test, one allele (31 trinucleotide repeat) was significantly more frequent in patients with schizophrenia than in controls ($\chi^2 = 9.57$, df = 1, $P = 0.002$). This association remained significant after applying a Bonferroni correction for 13 tests (adjusted critical value [$P < 0.05$] for 13 tests is 0.004). The Monte Carlo chi-square statistics (T1 in Sham and Curtis) confirm the results obtained from the conventional chi-square statistic. The comparison of cases *vs* controls for nine alleles with expected values greater than 1, yielded a *P*-value of 0.0125, while the comparison of allele 31 *vs* all other alleles yielded a *P*-value of 0.003. The etiologic fraction that is attributable to allele 31 was 0.097 (95% CI = 0.026–0.168), suggesting an allele of weak or modest effect.

TABLE 1. Distribution of alleles of SCA1 gene CAG repeat polymorphism

Number of repeat[a]	Schizophrenics $n = 98$[b]	Controls $n = 176$
15	1 (0.010)[c]	0 (0.000)
26	0 (0.000)	2 (0.011)
27	0 (0.000)	3 (0.017)
28	1 (0.010)	10 (0.057)
29	33 (0.337)	65 (0.369)
30	41 (0.418)	60 (0.341)
31	12 (0.122)	5 (0.028)
32	8 (0.082)	19 (0.108)
33	1 (0.010)	4 (0.023)
34	0 (0.000)	1 (0.006)
35	1 (0.010)	2 (0.011)
36	0 (0.000)	4 (0.023)
39	0 (0.000)	1 (0.006)

[a]Number of trinucleotide repeat, including both CAG and CAT, within the trinucleotide expansion region in SCA1 gene.
[b]n = number of chromosome.
[c]The number in parenthesis is the frequency of each allele.

The correlation coefficients between the size of the trinucleotide repeat and age at onset of schizophrenia were not significant ($r = -0.08$, $P = 0.597$, for the dominant model; $r = -0.19$, $P = 0.199$, for the additive model).

To determine whether the linkage disequilibrium extended over a larger interval, we examined three polymorphic flanking markers: D6S288, D6S1605, and D6S337. Several YAC clones contain D6S288, SCA1 CAG repeat region, and D6S337 together in this order from centromere to telomere (Olavesen *et al.*, 1995), and D6S1605 was reported very close to D6S288 (0.1 cM, Colette *et al.*, 1996). Thus, these three markers and the SCA1 CAG repeat region appear to be located between marker D6S274 and marker D6S260, within approximately 1 cM on the *P* arm of chromosome 6. No significant associations with schizophrenia were found for these three flanking markers (D6S288: $\chi^2 = 2.84$, df = 3, $P = 0.42$; D6S1605: $\chi^2 = 6.02$, df = 9, $P = 0.74$; D6S337: $\chi^2 = 5.74$, df = 3, $P = 0.12$).

DISCUSSION

We found a significant association of schizophrenia with a CAG trinucleotide repeat polymorphism in the SCA1 gene on chromosome 6p23. The patients with schizophrenia were five times more likely to have the 31 repeat allele than controls (OR = 4.77, 95% CI = 1.50–16.11). However, this allele was not the most common for either group; $q = 0.122$ for patients with schizophrenia; $q = 0.028$ for controls, and is likely to be an allele of a weak or modest effect.

Two possible interpretations of our positive association between SCA1 CAG repeat region and schizophrenia should be considered. One possibility is that there is a true linkage disequilibrium between this region and schizophrenia. Another possibility is

that there is a population stratification in our subjects, which leads to a false positive association.

A possible contribution of population stratification to our positive result appears to be unlikely, since allele frequencies for the three flanking markers did not differ significantly between patients with schizophrenia and controls, as it would have been expected if the association was due to population stratification. In addition, our control subjects were sampled from the same population as the schizophrenia patients, and were matched carefully with patients with schizophrenia with respect to race and age.

The associated allele (31 repeat) itself may be a susceptibility allele for schizophrenia, but not a major gene. However, this explanation seems unlikely since an earlier study reported a significant association with a different allele (29 repeat), even though it is still possible that multiple alleles at this locus are associated with schizophrenia. The real gene would be in the non-CAG repeat region of SCA1 gene, or in closely flanking sequence. Our results of no association with flanking markers support this hypothesis.

Anticipation has been suggested in some multiplex families with schizophrenia. The expansion of an unstable trinucleotide repeat is a known mechanism of anticipation in other diseases (e.g., Huntington's disease and SCA1). Several studies of possible CAG repeat expansions throughout the genome in patients with schizophrenia have reported mixed results (Morris *et al.*, 1995; O'Donovan *et al.*, 1995; Petronis *et al.*, 1996). Morris *et al.* (1995) reported that the distribution of repeat sizes in female patients with schizophrenia *vs* female controls was significantly different. Further, the larger CAG expansions were associated with a younger age at onset of schizophrenia.

We evaluated a potential contribution of the trinucleotide repeat expansion in the SCA1 gene to an earlier age at onset of schizophrenia, and found no significant correlation. Although our sample size may have limited power, our result suggests that the influence of the trinucleotide repeat in the SCA1 gene on age at onset of schizophrenia is neither great or general. However, we cannot exclude a possible correlation between CAG repeat size and age at onset of schizophrenia in a subset of patients with multiple familial cases of schizophrenia.

SCA1 is a candidate gene for schizophrenia, because of the location of the SCA1 gene on chromosome 6p23, and the significant allelic association with schizophrenia in our study, as well as the study by Wang *et al.* (1996). Our results suggest that it is more likely that as yet undetected variability in the SCA1 gene or closely flanking region, rather than a direct influence of the repeat expansion itself, is responsible for the linkage disequilibrium between this gene and schizophrenia. Further studies of this genetic region will be necessary to confirm our findings, and to clarify a more specific pathophysiological mechanism of this genetic region in the development of schizophrenia.

Acknowledgments

This research was supported by Grants MH-43880 and MH-52857 from the National Institute of Mental Health, and a General Clinical Research Center grant (M01RR0040) from the National Institute of Health. We gratefully acknowledge the many contribution to this work by Dr Danielle Reed, Dr Weizhen Xu, Guang Ming Yuan, Andrew Krakowski, Dr Ruben Gur, and Dr Raquel Gur.

REFERENCES

American Psychiatric Association (1987). *DSM-IIIR: Diagnostic and Statistical Manual of Mental Disorders*. 3rd ed. Washington, DC: American Psychiatric Association.

Antonarakis SE, Blouin J-L, Pulver AE, Wolyniec P, Lasseter VK, Nestadt G, *et al.* (1995). Schizophrenia susceptibility and chromosome 6p24-22. *Nat Genet* **11**:235–236.

Banfi S, Servadio A, Chung MY, Kwiatkowski TJ, McCall AE, Duvick LA, *et al.* (1994). Identification and characterization of the gene causing type 1 spinocerebellar ataxia. *Nat Genet* **7**:513–520.

Cannon TD, Kaprio J, Lonnqvist J, Huttunen MO, Koskenvuo M (1998). The genetic epidemiology of schizophrenia in a Finnish twin cohort: A population-based modeling study. *Arch Gen Psychiatry* **55**:67–74.

Colette D, Sabine F, Cecile F, Delphine S, Nathalie D, Alain V, *et al.* (1996). A comprehensive genetic map of the human genome based on 5,264 microsatellites. *Nature* **380**:152–154.

Daniels JK, Spurlock G, Williams NM, Cardno AG, Jones LA, Murphy KC, Asherson P, *et al.* (1997). Linkage study of chromosome 6p in sib-pairs with schizophrenia. *Am J Med Genet* **74**:319–323.

Gur RE, Mozley PD, Shtasel DL, Cannon TD, Gallacher F, Turetsky B, *et al.* (1994). Clinical subtypes of schizophrenia differ in brain and cerebrospinal fluid volume. *Am J Psychiatry* **151**:343–350.

Gurling H, Kalsi Gursharan, Chen AH, Green M, Butler R, Read T, *et al.* (1995). Schizophrenia susceptibility and chromosome 6p24-22. *Nat Genet* **11**:234–235.

Jodice C, Frontali M, Persichetti F, Novelletto A, Pandolfo M, Spadaro M, *et al.* (1993). The gene for spinocerebellar ataxia 1 (SCA1) is flanked by two closely linked highly polymorphic microsatellite loci. *Hum Mol Genet* **2**:1383–1387.

Jodice C, Malaspina P, Persichetti F, Novelletto A, Spadaro M, Giunti P, *et al.* (1994). Effect of trinucleotide repeat length and parental sex on phenotypic variation in spinocerebellar ataxia 1. *Am J Hum Genet* **54**:959–965.

Moises HW, Yang L, Kristbjarnarson H, Wiese C, Byerley W, Macciardi F, *et al.* (1995). An international two-stage genome-wide search for schizophrenia susceptibility genes. *Nat Genet* **11**:321–324.

Morris AG, Gaitonde E, McKenna PJ, Mollon JD, Hunt DM (1995). CAG repeat expansions and schizophrenia: association with disease in females and with early age at onset. *Hum Mol Genet* **4**: 1957–1961.

Morris-Rosendahl DJ, Burgert E, Uyanik G, Mayerova A, Duval F, Macher JP, Crocq MA (1997). Analysis of the CAG repeats in the SCA1 and B37 genes in schizophrenic and bipolar I disorder patients: Tentative association between B37 and schizophrenia. *Am J Med Genet* **74**:324–330.

Mowry BJ, Nancarrow DJ, Lennon DP, Sandkuijl LA, Crowe RR, Silverman JM, *et al.* (1995). Schizophrenia susceptibility and chromosome 6p24-22 [correspondence]. *Nat Genet* **11**:233–234.

O'Donovan MC, Guy C, Craddock N, Murphy KC, Cardno AG, Jones LA, *et al.* (1995). Expanded CAG repeats in schizophrenia and bipolar disorder [correspondence]. *Nat Genet* **10**:380–381.

Olavesen MG, Davies AF, Broxholme SJ, Wixon JL, Senger G, Nizetic D, *et al.* (1995). An integrated map of human chromosome 6p23. *Genome Res* **5**:342–358.

Orr HT, Chung MY, Banfi S, Kwiatkowski TJ, Servadio A, Beaudet AL, *et al.* (1993). Expansion of an unstable trinucleotide CAG repeat in spinocerebellar ataxia type 1. *Nat Genet* **4**:221–226.

Petronis A, Bassett AS, Honer WG, Vincent JB, Tatuch Y, Sasaki T, *et al.* (1996). Search for unstable DNA in schizophrenia families with evidence for genetic anticipation. *Am J Hum Genet* **59**:905–911.

Ranum LP, Chung MY, Banfi S, Bryer A, Schut LJ, Ramesar R, *et al.* (1994). Molecular and clinical correlations in spinocerebellar ataxia type I: evidence for familial effects on the age at onset. *Am J Hum Genet* **55**:244–252.

Schizophrenia linkage collaborative group for chromosomes 3, 6 and 8 (1996). Additional support for schizophrenia linkage on chromosomes 6 and 8: A multicenter study. *Am J Med Genet* **67**:580–594.

Schlesselman JJ (1982). *Case-control Studies: Design, Conduct, Analysis*. New York: Oxford University Press.

Schwab SG, Albus M, Hallmayer J, Hönig S, Borrmann M, Lichtermann D, *et al.* (1995). Evaluation of a susceptibility gene for schizophrenia on chromosome 6p by multipoint affected sib-pair linkage analysis. *Nat Genet* **11**:325–327.

Sham PC, Curtis D (1995). Monte Carlo tests for associations between disease and alleles at highly polymorphic loci. *Ann Hum Genet* **59**:97–105.

Shtasel DL, Gur RE, Mozley PD, Richards J, Taleff MM, Heimberg C, *et al.* (1991). Volunteers for biomedical research: recruitment and screening of normal controls. *Arch Gen Psychiatry* **48**:1022–1025.

Spitzer RL, Willams JB, Gibbon M (1987). *Instruction Manual for the Structured Clinical Interview for DSM-IIIR (SCID).* New York: New York State Psychiatric Institute.

Straub RE, MacLean CJ, O'Neill FA, Burke J, Bernadette M, Duke F, *et al.* (1995). A potential vulnerability locus for schizophrenia on chromosome 6p24-22: Evidence for genetic heterogeneity. *Nat Genet* **11**:287–293.

Turecki G, Rouleau GA, Joober Ridha, Mari J, Morgan K (1997). Schizophrenia and chromosome 6p. *Am J Med Genet* **74**:195–198.

Wang S, Sun CE, Walczak CA, Ziegle JS, Kipps BR, Goldin LR, Diehl SR (1995). Evidence for a susceptibility locus for schizophrenia on chromosome 6pter-p22. *Nat Genet* **10**:41–46.

Wang S, Detera-Wadleigh SD, Coon H, Sun CE, Goldin LR, Duffy DL, *et al.* (1996). Evidence of linkage disequilibrium between schizophrenia and the SCA1 CAG repeat on chromosome 6p23. *Am J Hum Genet* **59**:731–736.