American Journal of Medical Genetics (Neuropsychiatric Genetics) 96:873–876 (2000)

Long Repeat Tracts at *SCA8* in Major Psychosis

John B. Vincent,[1,2]* Qiu-Ping Yuan,[3] Martin Schalling,[3] R Adolfsson,[4] M. Helena Azevedo,[5] Antonio Macedo,[5] Amy Bauer,[6] Camille DallaTorre,[6] Helena M. Medeiros,[6] Michele T. Pato,[6] Carlos N. Pato,[6] Timothy Bowen,[7] Carol A. Guy,[7] Michael J. Owen,[7] Michael C. O'Donovan,[7] Andrew D. Paterson,[1,2] Arturas Petronis,[1] and James L. Kennedy[1]

[1]*Neurogenetics Section, Clarke Division, Centre for Addiction and Mental Health, Toronto, Ontario, Canada*
[2]*Department of Genetics, The Hospital for Sick Children, Toronto, Ontario, Canada*
[3]*Neurogenetics Unit, Department of Molecular Medicine, Karolinska Hospital, Stockholm, Sweden*
[4]*Department of Psychiatry, University of Umeå, Umeå, Sweden*
[5]*Department of Psychiatry, and Center for Neuroscience, University of Coimbra, Coimbra, Portugal*
[6]*Department of Psychiatry, State University of New York at Buffalo, Buffalo, New York*
[7]*Division of Psychological Medicine, University of Wales College of Medicine, Cardiff, United Kingdom*

Expansion at a recently identified unstable trinucleotide repeat on chromosome 13q21 has been reported as the molecular cause for spinocerebellar ataxia type 8 (SCA8). The trinucleotide repeat, which consists of a [CTA]n repeat and adjacent [CTG]n repeat, was reported to have a pathogenic range of 107–127 CTG repeats (or 110–130 combined CTA and CTG repeats) in a large ataxia kindred. This repeat region was also cloned by our group from a bipolar affective disorder (BPAD) patient, who has approximately 600 combined repeats, and large alleles (>100 repeats) were reported to be present in 0.7% of controls and 1.5% of major psychosis patients (n = 710 and n = 1,120, respectively). We have followed up these findings by screening three new samples of BPAD and schizophrenia (SCZ) patients and controls, including 272 individuals from 14 BPAD families from Sweden, 130 individuals from 32 SCZ and BPAD families/trios from the Azores Islands, and 206 SCZ individuals from the United Kingdom and Ireland, and 219 matched controls. We found large repeat alleles above the SCA8 pathogenic range in individuals from 3 of 32 Azorean pedigrees and in 1 of 206 SCZ individuals from the United Kingdom, and repeat alleles within the SCA8 pathogenic range in 1 of 14 Swedish families. Although the rarity of major psychosis patients carrying the *SCA8* expansion mutation would require a much larger sample size to reach statistical significance, these results support the previously reported observation of increased occurrence of large repeats at *SCA8* in major psychosis. Am. J. Med. Genet. (Neuropsychiatr. Genet.) 96:873–876, 2000.
© 2000 Wiley-Liss, Inc.

KEY WORDS: **SCA8; schizophrenia; bipolar affective disorder; trinucleotide repeat expansion**

INTRODUCTION

A large CAG/CTG trinucleotide repeat, identified by using the repeat expansion detection technique (RED) [Schalling et al., 1993] was cloned from an affected individual of a family segregating a form of autosomal dominant spinocerebellar ataxia termed SCA8 [Koob et al., 1999]. This repeat, consisting of $[CTA]_n[CTG]_n$, was mapped to chromosome 13q21 and is believed to be at the 3′ end of a gene, *SCA8*, which is expressed at low levels in cerebellum and for which no open reading frame has been identified, and for which 99% of the population has repeats totalling 16–37 in length [Koob et al., 1999]. There are several reasons to believe that this gene is the cause of SCA8: i) expanded repeats occur in a large ataxia kindred [Koob et al., 1999], with 110–130 combined repeats (107–127 CTG repeats) cosegregating with disease; ii) linkage analysis of this repeat locus produces a maximum Lod score of 6.8; iii) polymerase chain reaction (PCR) analysis detected large repeat alleles at *SCA8* in 7 of 102 families with undefined hereditary ataxias [Koob et al., 1999]. However, we have observed that large trinucleotide repeat

Contract grant sponsor: Medical Research Council of Canada; Contract grant number: MT15007; Contract grant sponsor: National Alliance for Research on Schizophrenia and Depression; Contract grant sponsor: National Institute for Mental Health; Contract grant numbers: MH33990, MH52618, MH58693-01A1, and POIMH56193.

*Correspondence to: John B. Vincent, Rm 9102, Department of Genetics, The Hospital for Sick Children, 555 University Ave., Toronto, Ontario M5G 1X8, Canada.
E-mail: jvincent@genet.sickkids.on.ca

Received 14 June 2000; Accepted 4 August 2000

© 2000 Wiley-Liss, Inc.

alleles at *SCA8* (38–1,300 combined repeats, including some within the narrow pathogenic range of 110–130 combined repeats), occur relatively frequently in individuals who are not only unaffected with ataxia but also have no family history of the disease.

This repeat region was also independently cloned from a bipolar affective disorder (BPAD) individual with a very large CAG/CTG RED product [Vincent et al., 1998, 2000a]. Southern hybridization revealed the presence of approximately 600 combined repeats for this individual. The flanking sequence (GenBank AF087653) showed sequence and localization identity with the Koob et al. [1999] *SCA8* sequence (GenBank AF126748). Screening of 1,120 unrelated patients affected with either schizophrenia (SCZ), BPAD, schizoaffective disorder, childhood onset psychosis, or childhood onset depression (screened for absence of major medical or neurological illness) and 710 unrelated unaffected controls (screened for absence of major psychiatric illness) by using PCR and genomic Southern hybridization (for very large repeat alleles that fail to amplify by PCR) showed that large alleles in and above the suggested pathogenic range at *SCA8* are, in fact, common in the general population [Vincent et al., 2000a]. Allele sizes ranging from 45 to 1,300 combined repeats were present in 1.1% of controls and 2.1% in the psychiatric disorders group. Four individuals (1 control and 3 from the psychiatric group) had repeats within the pathogenic range reported for SCA8 of 110–130 combined repeats. Similar findings have now been reported by several other groups [Stevanin et al., 2000; Worth et al., 2000]. Stevanin et al. [2000] reported 3 of 188 healthy controls with combined repeat sizes of 107, 111, and 123 repeats and an ataxia kindred in which expansion and disease did not cosegregate. Worth et al. [2000] detected alleles with 100, 101, 103, 133, and 174 combined repeats in 653 non-ataxia control individuals. These studies give the rate of occurrence of large alleles in control individuals as 1.6% and 0.77%, respectively. Although these figures may underestimate the frequency of large alleles because apparent homozygotes were not checked by using more reliable methods for detecting very large alleles, they compare well with the rates observed by our group [Vincent et al., 2000a].

This work has now been followed up in three independent studies: i) 206 UK and Irish Caucasian unrelated SCZ individuals, including 157 individuals from multiplex SCZ families and 219 group-matched control individuals from the United Kingdom; ii) Fourteen Swedish Caucasian families ascertained for BPAD, unipolar affective disorder (UPAD), or other major psychiatric disorders; iii) 11 Caucasian multiplex families or affected relative pairs and 21 proband/parent trios originating from the Azores Islands and segregating for either SCZ or BPAD were ascertained, and repeat alleles at *SCA8* were examined. Local ethical committee approval was obtained, and blood was drawn after obtaining informed written consent from each individual.

In the UK/Irish study *SCA8* alleles with 79, 88, and 150 repeats were present among the SCZ individuals, and a single large allele of 80 repeats alleles was present in the controls. In the Swedish study, DNA samples from 272 family members were available. Genotyping at *SCA8* was performed for 87 affected individuals (22 BPAD, 43 UPAD, and 22 other psychiatric disorders), including 4 spouses (1 BPAD, 2 UPAD, and 1 other psychiatric disorder) and 185 healthy members involving 39 spouses and 146 first- and second-degree relatives. Altogether, 23 individuals, including 11 affected (1 BPAD, 7 UPAD, and 3 others) and 12 unaffected, had medium size alleles ranging from 70 to 100 combined repeats. In one family from the Swedish study, two affected children, one with BPAD and the other with UPAD, both carrying 125 combined repeats (age at interview (AAI): 42 and 44 years, respectively), inherited the expanded alleles from their currently healthy father (AAI: 67 years), also with a 125 repeat allele. These three individuals were carefully reexamined by experienced psychiatrists. Clinical symptoms similar to those of spinocerebellar ataxia or any other movement disorder were not found in any individuals. Three of the 11 families and 21 trios (with 130 family members) from the Azores Islands have members with large alleles with >100 repeats. In two of the families, individuals are either unaffected (n = 8), diagnosis unknown (n = 4), affected by schizophrenia (n = 3), schizotypal personality disorder (STPD; n = 1) or schizoaffective disorder (SAD; n = 1). Four of the five affected individuals have large repeat alleles. Two of the three schizophrenia individuals had alleles with ~170 and ~450 combined repeats, the SAD individual had ~170 repeats, and the STPD individual had ~400 repeats. None of the unaffected individuals had large repeat alleles, although the father of the ~170 repeat schizophrenia individual had ~640 repeats, but he was not interviewed and diagnosis remains unknown. The third family has two BPAD individuals, two unipolar depression (UPD) individuals, and one unaffected individual. Both BPAD individuals have large repeat alleles (~104 and ~200 repeats), but neither the UPD subjects nor the unaffected individuals have large repeat alleles. No other neurological symptoms were observed in any of the families.

If only one large allele (>109 combined repeats) individual per family is included, this puts the frequency of major psychosis individuals with "expansion" alleles among the combined Swedish and Azorean samples at 8.7%, and among the entire group (n = 1,372, including the sample from Vincent et al., 2000) at 1.1%, and among controls at 0.43% (n = 929). For large alleles between 38 and 109 combined repeats, the frequency for affected individuals is 1.5% (n = 1,372) and 0.97% for control individuals (n = 929). Results are summarized in Table I. Comparison of frequencies of affected and unaffected individuals with large repeats, either >45 combined repeats or >109 combined repeats, shows no significant differences in the Vincent et al. [2000a] sample ($\chi^2 = 2.0$, $P = 0.16$; $\chi^2 = 0.094$, $P = 0.76$, respectively), or in the Cardiff sample ($\chi^2 = 0.31$, $P = 0.58$; $\chi^2 = 0$, $P = 1$, respectively), or in the combined Vincent et al. and Cardiff samples ($\chi^2 = 3.2$, $P = 0.074$; $\chi^2 = 0.47$, $P = 0.49$, respectively). However, as we have stated previously, because of the relatively low frequency of the large alleles, a number in excess of

TABLE I. Numbers of Large Repeat Alleles (Above the 99th Centile in the Normal Population) for Major Psychosis Patients and Controls*

	No. of individuals screened	No. of individuals with large alleles below pathogenic range (45–109 repeats)	No. of individuals with large alleles within pathogenic range (110–130 repeats)	No. of individuals with large alleles above pathogenic range (131–1,300 repeats)
Vincent et al. [2000a], unrelated	Affected: 1,120	13	3	7
	Unaffected: 710	4	1	3
Cardiff (UK) unrelated individuals	Affected: 206	2	0	1
	Unaffected: 219	1	0	0
Swedish families (n = 14)	Affected: 87	11	2[a]	0
	Unaffected: 185	12	1[a]	0
Azores Islands families, affected relative pairs (n = 11) and trios (n = 21)	Affected: 46	1	0	5[b]
	Unaffected/not interviewed: 84	0	0	1[b]
Total	2,657	44	7	17

*Values are within, above, and below the pathogenic range of SCA8 [as defined by Koob et al., 1999] in the previous study by Vincent et al. [2000a] and in three new studies. Ascertainment and diagnostic criteria are described elsewhere [Bowen et al., 1998; Lindblad et al., 1998; Pato et al., 2000]. The number of large repeat alleles in the UK and Swedish studies is likely to be an underestimate, because PCR is unable to amplify large repeats efficiently, and Southern hybridization was not used to confirm the apparent homozygotes detected by PCR. Combined CTA and CTG repeat lengths are given. One of 14 Swedish families and 3 of 32 Azorean families/trios segregate large repeat alleles within or above the pathogenic range for SCA8.
[a]Members of the same family.
[b]Members of two families and one trio.

10,000 would be necessary to detect an effect (odds ratio ~1.8) at an 80% power level [Vincent et al., 2000a]. No analysis was performed for pooled data from the Swedish and Azorean family samples, because the much higher frequencies for large repeat alleles among these samples indicate that they are from genetically distinct populations.

There are two main issues raised by these studies: the involvement of the large *SCA8* repeats in spinocerebellar ataxia and the contribution of large *SCA8* repeats in major psychosis. Given the relatively high frequency of large repeats (45–1,300) at *SCA8*, it seems likely that long repeat tracts alone are insufficient to precipitate spinocerebellar ataxia, even within the narrow pathogenic range of 110–130 combined repeats [Koob et al., 1999], and another mutation or allelic variant at this locus is required for SCA8 pathogenicity. It is noteworthy that the pathogenic range for SCA8 of 107–127 CTG repeats or 110–130 combined repeats implies that only 3 CTA repeats were present in these alleles. All 6 non-expanded, non-ataxia alleles sequenced in our previous study [Vincent et al., 2000a] possessed either 8 or 9 CTA repeats, as do BAC clone sequences (GenBank AC013772; AC013803) spanning *SCA8*. We identified a third BAC clone, H_NH0121J06 (AL160391), containing the *SCA8* locus [Vincent et al., 2000a], which has 10 CTA repeats. Sequence analysis of the very large repeats is extremely difficult because of the secondary structure; however, in four large non-ataxia *SCA8* alleles, the CTA lengths were 7, 8, 9, and 17 repeats, with respective CTG repeats stretches of 105, 120, 103, and 85 (the latter being the cloned repeat, λ7a, truncated from ~600 combined repeats). The repeat interrupts observed between the CTA and CTG repeats by Koob et al. [1999] (AF126748; AF126749) were also not present in any of our sequenced alleles, nor in 1,120 affected individuals and 710 controls (as determined by allele-specific oligonucleotide hybridization at the CTA-CTG repeat junction) [Vincent et al., 2000a], nor in the three BAC clone sequences in the GenBank database (see above). It is apparent that the association between expansion mutation at *SCA8* and spinocerebellar ataxia is not straightforward. We propose that CTA length and/or repeat interrupts may be crucial in distinguishing pathogenic and nonpathogenic alleles. Another possibility is that the expanded alleles with interrupts maybe in linkage disequilibrium with another mutation, either at *SCA8*, or at the overlapping sense-strand gene termed Kelch-like (*KLHL1*) [Nemes et al., 2000], or at another nearby gene. Alternatively, epigenetic factors may play an important role in discriminating pathogenic and nonpathogenic alleles. Until this second etiologic factor at *SCA8* is delineated, diagnostic analysis at this locus would be inappropriate.

The relatively high co-occurrence of large alleles and psychosis observed in this and our previous study suggests that this may be a susceptibility locus. Further analysis is warranted, however, because the large alleles are relatively scarce; extensive screening of major psychosis patients will be required to establish sufficient sample numbers for segregation and relative risk analysis. Psychosis is frequently a feature of a number of neurodegenerative disorders, most notably Huntington's disease, but is almost exclusively present alongside rather than separate from the neurodegenerative symptoms [see Vincent et al., 2000b, for review]. It is conceivable that allelic variants of the expansion mutation at the *SCA8* locus could predispose ataxia and major psychosis exclusive of each other, possibly operating through different molecular mechanisms and/or affecting different regions of the brain.

The possible involvement of a noncoding gene as a susceptibility factor for major psychosis that may function at the RNA level through regulation of a coding gene on the opposite strand adds a further level of complexity to an already deeply complex etiological puzzle. A similar story may be emerging at a locus on chromo-

some 1q42.1, where a translocation that cosegregates with SCZ in a pedigree disrupts a two-gene system, with a coding gene (*DISC1*) on the sense strand and a noncoding gene (*DISC2*) on the antisense strand [Millar et al., 2000]. Although the possible involvement of either *SCA8*/*KLHL1* or *DISC1*/*DISC2* loci in SCZ and BPAD may represent only a small minority of cases, these findings may have implications for the search for other susceptibility loci for major psychosis and for other complex genetic disorders.

REFERENCES

Bowen T, Guy CA, Craddock N, Williams NM, Spurlock G, Murphy K, Jones LA, Gray M, Sanders RD, McCarthy G, Chandy KG, Fantino E, Kalman K, Gutman GA, Gargus JJ, Williams J, McGuffin P, Owen MJ, O'Donovan MC. 1998. Further support for an association between a polymorphic CAG repeat in the hKCa3 gene and schizophrenia. Mol Psychiatry 3:266–269.

Koob MD, Moseley ML, Schut LJ, Benzow KA, Bird TD, Day JW, Ranum LPW. 1999. An untranslated CTG expansion causes a novel form of spinocerebellar ataxia. Nat Genet 21:379–384.

Lindblad K, Nylander P-O, Zander C, Yuan Q-P, Stahle L, Engstrom C, Balciuniene J, Pettersson U, Breschel T, McInnis M, Ross CA, Adolfsson R, Schalling M. 1998. Two commonly expanded CAG/CTG repeat loci: involvement in affective disorders? Mol Psychiatry 3:405–410.

Millar JK, Wilson-Annan JC, Anderson S, Christie S, Taylor MS, Semple CAM, Devon RS, St. Clair DM, Muir WJ, Blackwood DHR, Porteous DJ. 2000. Disruption of two novel genes by a translocation cosegregating with schizophrenia. Hum Mol Genet 9:1415–1423.

Nemes JP, Benzow KA, Koob MD. 2000. The SCA8 transcript is an antisense RNA to a brain-specific transcript encoding a novel actin-binding protein (KLHL1) Hum Mol Genet 9:1543–1551.

Pato CN, Macedo A, Dourado A, Vincent JB, Bauer A, Xu G, Coelho I, Dourado A, Valente J, Pato MT, Azevedo MH, Kennedy JL. 2000. Detection of expansion regions in bipolar families. Am J Med Genet (Neuropsychiatr Genet) (in press).

Schalling M, Hudson TJ, Buetow KH, Housman DE. 1993. Direct detection of novel expanded trinucleotide repeats in the human genome. Nat Genet 4:135–139.

Stevanin G, Herman A, Durr A, Jodice C, Frontali M, Agid Y, Brice A. 2000. Are (CTG)n expansions at the SCA8 locus rare polymorphisms? Nat Genet 24:213.

Vincent JB, Parikh SV, Petronis A Meltzer HY, Kennedy JL. 1998. Unstable DNA in major psychosis: cloning of a new unstable trinucleotide repeat region on chromosome 13. Am J Hum Genet 63:A344.

Vincent JB, Neves-Pereira ML, Paterson AD, Yamamoto E, Parikh SV, Macciardi F, Gurling HMD, Potkin S, Pato CN, Azevedo MH, Kovacs M, Davies M, Lieberman JA, Meltzer HY, Petronis A, Kennedy JL. 2000a. An unstable trinucleotide repeat region on chromosome 13 implicated in spinocerebellar ataxia: a common expansion locus. Am J Hum Genet 66:819–829.

Vincent JB, Paterson AD, Strong E, Petronis A, Kennedy JL. 2000b. The unstable trinucleotide repeat story of major psychosis. Am J Med Genet (Semin Med Genet) 97:77–97.

Worth PF, Houlden H, Giunti P, Davis MB, Wood NW. 2000. Large, expanded repeats in SCA8 are not confined to patients with cerebellar ataxia. Nat Genet 24:214–215.