RESEARCH NOTE



Exploration of CAG triplet repeat in nontranslated region of SCA12 gene

WASEEM GUL LONE^{1,2}, IMRAN ALI KHAN³, SUBHADRA POORNIMA⁴, NOOR AHMAD SHAIK⁵, ANGMUTHU KANIKANNAN MEENA⁶, KAIPA PRABHAKAR RAO² and QURRATULAIN HASAN^{3*}

¹Department of Genetics, Mahaveer Hospitals, AC Guards, Hyderabad 500 004, India ²Department of Genetics, Osmania University, Hyderabad 500 007, India

³Department of Clinical Laboratory Sciences, College of Applied Medical Sciences, King Saud University, Riyadh 11433, Kingdom of Saudi Arabia

⁴Department of Genetics and Molecular Medicine, Kamineni Hospitals, LB Nagar, Hyderabad 500 068, India ⁵Department of Genetic Medicine, Faculty of Medicine, Princess Al-Jawhara Centre of Excellence in Research

of Hereditary Disorders, King Abdul-Aziz University, Jeddah 21589, Kingdom of Saudi Arabia

⁶Department of Neurology, Nizams Institute of Medical Sciences, Panjagutta, Hyderabad 500 082, India

[Gul Lone W., Khan I. A., Poornima S., Shaik N. A., Meena A. K., Rao K. P. and Hasan Q. 2016 Exploration of CAG triplet repeat in nontranslated region of *SCA12* gene. *J. Genet.* **95**, 427–432]

Introduction

Spinocerebellar ataxias (SCAs) are a heterogeneous group of neurodegenerative disorders characterized by progressive cerebellar ataxia of gait and limbs, dysarthria, dysphagia and other neurological signs. The genetic classification of the autosomal dominant types of SCAs is associated with more than 30 loci. Several of these SCAs (SCA1, 2, 3, 6, 7, 17) are due to cytosine adenine guanine (CAG) repeat expansions in the coding regions of the corresponding genes translated into abnormally long polyglutamine stretches (Stevanin et al. 2000). Most of the CAG repeat disorders are characterized by autosomal dominant heredity and anticipation (i.e. earlier onset age and increasing severity in successive generations). However, our earlier study did not find correlation between age of onset and the expanded CAG repeat number at various SCA loci (Gul et al. 2014). Apart from these, a novel mutation in SCA12 type was described which consisted of trinucleotide repeat expansion in the noncoding region of SCA12 gene (Holmes et al. 1999).

SCA12 gene is located on chromosome at position 5q31q33. The CAG tract lies at 5'UTR and encodes the brainspecific subunit of the serine-threonine protein phosphatase, PP2A. The normal, nonpathogenic number of CAG repeats has been established in different ethnical populations, ranges 9–18 (Fujigasaki *et al.* 2001) and 7–28 (Holmes *et al.* 1999). The pathogenic CAG repeat expansions was observed to be 66–78 in patients of German origin (Holmes *et al.* 1999) and of 55–69 in Indian SCA12 families (Fujigasaki *et al.* 2001; Srivastava *et al.* 2001). Clinically, the distinguishing feature is early prominent action tremor in the limbs, but can occur in the trunk, neck, lips and tongue (Cho *et al.* 2008; O'Hearn *et al.* 2012). Postural tremor (tremor at rest) and intention tremor (tremor with purposeful movements) are also observed. Signs of cerebellar dysfunction (e.g. ataxia and dysmetria) tend to be less prominent and less disabling in individuals with SCA12 than in other types of SCA.

In different populations, the disease prevalence of these ataxias has shown to be associated with the presence of large normal alleles (LNA) at the respective loci (Squitieri *et al.* 1994; Takano *et al.* 1998; Goldberg *et al.* 1993). Screening of populations for establishing normal and expanded ranges for that particular geographical region helps in proper molecular diagnosis. The repeat numbers which are considered as LNA fall in 5–10% of the upper repeat number, when they are arranged/plotted in ascending order as described earlier (Alluri *et al.* 2007). In the present study, all the repeats were from the healthy individuals (n = 200) and from nondisease alleles of patients (n = 376).

An attempt was also made to establish the normal repeat ranges and LNA at SCA12 locus to enable proper molecular diagnosis. The role of triplet repeat expansion beyond threshold size in disease pathogenesis has been defined; however, the exact role of lower repeats in disease pathogenesis is still unknown in these types of disorders. Notably, the lower repeats may also result in abnormal phenotype, as it has been noted in diseases such as spinal and bulbar muscular atrophy (SBMA) (Kooy *et al.* 1999). The lower repeats were not found in controls which indicate that lower repeats at these loci may

Keywords. CAG repeats; large normal alleles; polyglutamine; neurodegenerative diseases; spinocerebellar ataxia type 12.

^{*}For correspondence. E-mail: qhasan2000@yahoo.com.

Waseem Gul Lone and Imran Ali Khan contributed equally for this work.

play a role in diseases pathogenesis. Various studies (Deutsch *et al.* 2006; Chen *et al.* 2009) revealed that lower repeats at *SCA12* genelocus are susceptible to Alzheimer's disease and essential tremor. Till date, there are no studies carried out on CAG repeat analysis in *SCA12* gene from southern region of India. Therefore, we aimed to determine the CAG repeat expansions in *SCA12* gene from southern population of India.

Materials and methods

Ethics committee

Ethical approval was obtained from the ethical committee of Kamineni Hospitals and Nizams Institute of Medical Sciences, Hyderabad, India. Clinical, personal and family history was obtained from patients by a personal interview with patients/attenders, in a well-designed proforma. A three generation pedigree was recorded with relevant medical history of the family members. The study was carried on 288 cases which included 186 males and 102 females aged between 5 and 76 years with the mean age (\pm SD) of 35.6 \pm 15.9 years.

Patient recruitment

Patients diagnosed for various neurodegenerative diseases with triplet repeat expansions were recruited. These cases were diagnosed based on the clinical symptoms and MRI, ENMG, EEG and CT scan findings. All these cases were referred to the Department of Genetics and Molecular Medicine, Kamineni Hospitals, Hyderabad, for molecular diagnosis. All these diseases showed overlapping symptoms making it difficult for the clinicians to identify the disease precisely.

Samples

For this study, we selected 188 patients referred with various triplet repeat diseases from different regions of Telangana and Andhra Pradesh. For molecular analysis, 2 mL of the venous blood was collected from patients. Healthy controls (n = 100) were selected from general population without any family history of neurological diseases, who had visited the Kamineni hospitals.

DNA and molecular analysis

For molecular diagnosis, DNA was extracted from peripheral blood leucocytes by using salting out technique, which was routinely used in our laboratory (Khan et al. 2015). CAG repeat analysis was performed for 288 samples including 100 normal individuals and 188 patients. Analyses for CAG repeats were performed using thermal PCR using the following pairs (forward and reverse) of oligonucleotides as described by Bahl et al. (2005). Amplification of the fragment was performed with forward primer 5'-GCCAGCG CACTCACCCTC-3' and reverse primer 5'-CTGCTGGGA AAGAGTCGTGC-3'. Oligonucleotides were synthesized by Bioserve Biotechnology (Hyderabad, India). Total volume of 25 μ L of PCR reaction (0.2 mL Eppendorf tube) was carried out for 35 cycles with an initial denaturation at 95°C for 5 min followed by cycling at 95°C for 30 s, annealing for 40 s at 60°C. The PCR products obtained were analysed by electrophoresis, on a 2% agarose gel stained with ethidium bromide. Molecular weight analysis and sizing of alleles were performed by using UVI Tech software (UK).

Statistical analysis

Possible differences between patients and controls in the normal repeat frequency distributions were assessed using two nonparametric tests: the two sample Kolmogorov–Smirnov test and the Mann–Whitney U test. Allele frequencies at each locus were estimated by the gene counting method. Statistical analyses of differences in controls and patient groups are in the frequency of LNA, performed with the Fisher exact test (*t*-test). To perform statistical analyses of the differences in the frequencies of large normal allele between patients and controls, we defined large normal allele as in Takano *et al.* (1998). In this study, alleles longer than 15 CAG repeats were considered as large alleles and less than 15 CAG repeats were identified as normal alleles.

Results

In this case–control study, only one case (0.5%) (figures 1&2; table 1) was found to have 56 repeats in *SCA12* gene which is an expansion in the pathological range according to defined



Figure 1. Ethidium bromide-stained 2% agarose gel showing amplified PCR products of *SCA12* gene. Lane 1, patient sample showed 56 CAG repeats (expansion); lane 2, patient sample showed 26 CAG repeats; lane 3, patient sample showed 13 CAG repeats; lane 4, patient sample showed 13 CAG repeats; lane 5, patient sample showed 27/19 CAG repeats; lane 6, control sample showed 12 CAG repeats; lane 7, DNA size marker 100 bp.

pathogenic repeat range (45-78) (Holmes et al. 1999). The age of onset in patient III-2 (figure 2) was 50 years with the clinical characteristics of tremor, oscillations in fingers and hands, upper limbs were more affected than the lower limbs (table 1). The patient showed an autosomal dominant pattern of inheritance (figure 2). Her mother had the same symptoms and died at the age of 59 (II-1, figure 2), hence the sample is not available for molecular analysis. Although, the repeat expansions in the pathological range based on published criteria was found in only one case, however >26 repeats were found in four patients and none in controls, which includes two cases having an expansion at SCA1 locus (table 1). All these patients exhibit clinical feature of SCA12 type, i.e. tremor. Out of these five cases, three cases had expansion at different SCA loci, i.e. SCA12 and SCA1. The other two cases had age at onset in the second decade of their life and could not be categorized as SCA12 type since their repeats at SCA12 locus do not fall in the expanded range based on the published expansion range.

A total of 575 pooled chromosomes from 288 individuals were analysed for the distribution of the repeat number alleles (excluding expanded allele). At SCA12 locus alleles ranged between 6–27 CAG repeats with 92% of them having repeats between 7–21 CAG repeats (figure 3) and this did not differ markedly from the ranges reported previously (Bahl *et al.* 2005) from the south Indian population (8–23 CAG repeats). The most frequent allele in the patient group (20.2%) and control group (16%) was 10 and 11 CAG repeats, respectively.

At SCA12 locus, LNA is considered as >15 repeats (Takano *et al.* 1998; Silveira *et al.* 2002; Brussino *et al.* 2010). The percentage of LNA in patient group and control group was 32% and 35.5%, respectively based on published criteria. Since, 17 CAG repeats were found more in controls, LNA in the present study was taken as >18 (figure 3). The percentage of LNA in patient group was 19.33% and 15.5% in the controls based on the present study criteria (figure 4).



Figure 2. Three generation pedigree of the patient with SCA12 expansion.

			10	7						
Expansion based										Clinical details
on reference range	Age	Sex	Age of onset	Family history	Consanguinity	Repeats observed	Gait	Speech	Eye	Any other
Yes	60	Female	50	Yes	No	56/12	+	+	I	Tremor and oscillations in fingers and hands,
No	24	Male	21	No	No	27	+	I	+	upper limbs are more affected than lower limbs Rest tremor, cerebellar
No	36	Male	29	No	No	26	+	I	+	signs and hyperrefiexia Intention tremor with
No	27	Male	24	Yes	No	26/17	+	+	Ι	deteriorated writing skills Tremor in hands and giddy while
No	45	Male	40	Yes	Yes	26/15	+	+	I	watking, Expansion at OCA 1 rocus (+0.20 repeats) Deteriorated writing, tremor while holding something, cerebellar atrophy. Expansion at SCA 1 locus (49/28 repeats)
Characteristic featur	es of	SCA12 tv	me in natients	with 56 (reneat f	alling in the refer	ence nathological r	anoe) (6 and 23	rene:	ts (reneats not falling in the reference nathological range).

Waseem Gul Lone et al.



Figure 3. Distribution of alleles in patients and controls at SCA12 locus.

The prevalence of SCAs in various populations has been correlated with the percentage of LNA. To study the frequency of triplet repeat expansion in our two groups relative to the frequency of normal alleles of larger size at the SCA12 locus, we used the criteria of Takano *et al.* (1998). In the present study, LNA allele distribution were assessed from the healthy individuals (n = 100) and from nondisease alleles (n = 188). At the SCA12 locus, the frequency of LNA >15 repeats in the present study was 0.33. The percentage of LNA at SCA12 locus was 32% and 35.5% in patients and controls, respectively (figure 4). There is no published literature on the frequency of LNA at SCA12 locus from India. The frequency of LNAs in the present study is significantly higher than the frequencies reported in Portuguese and Brazilian populations (table 2).



Figure 4. LNA at SCA12 locus based on published and the present study criteria.

In the present study, lower repeats at SCA12, i.e. six CAG repeats were identified in two patients and none in controls. These two patients had age at onset at 30 years with gait and speech disturbances. However, tremor was not identified in these patients. One patient is walking with support and the other has lumbar spondylosis and sensory neuropathy (table 3).

Discussion

Earlier studies reported SCA12 as rare among SCA subtypes in Indian population (Fujigasaki *et al.* 2001). The recent literature showed a notable incidence of SCA12 in north Indian population, accounts for 7% of all SCAs (Srivastava *et al.* 2001). In another study, it accounts for 16% of all SCAs (Bahl *et al.* 2005) in north India. The database study from north India also showed frequency of 12% (Faruq *et al.* 2009). The present study on SCA12 is the first study from southern India which reports the first case of SCA12.

Table 2. Frequency of LN alleles in the present study compared with other populations.

Gene	Present	Portuguese group	Brazilian group
	study	(Silveira <i>et al.</i> 2002)	(Silveira <i>et al.</i> 2002)
SCA12 (>15)	0.33	0.03 (<0.0001)	0.05 (0.0001)

P values indicated in parentheses are obtained from chi-square analysis when comparison was between the present study and other studies using MedCalc 7.6.0.0.

The frequencies of LNAs in this study at SCA12 locus were quite high when compared with Portuguese and Brazilian populations.

CAG repeats in SCA12 gene

Table 3.	Clinical details of the	patients with lower re	peats than the normal	range at SCA12 locus.
				. /

				Family		Repeats				Clinical details
Gene	Age	Sex	Onset	history	Consanguinity	observed	Gait	Speech	Eye	Any other
SCA12 SCA12	36 32	Male Male	30 30	No No	No No	12/6 11/6	+ +	+ +	_	Walking with support Lumbar spondylosis, sensory neuropathy

Lower repeats at SCA 12, i.e. six CAG repeats were identified in the two patients and none in the controls. These two patients had age at onset at third decade with gait and speech disturbances.

The CAG repeat length in the normal chromosomes at the SCA12 loci was found to be highly polymorphic in Indian population and ranged from 7–32 CAG repeats (Bahl *et al.* 2005, 2009), which indicates wide heterogeneity of *SCA12* gene in Indian population. In this study, alleles had repeat range between 6–27 with 92% of chromosomes having repeats between 7–21 (figure 3) and this did not differ markedly from the ranges reported previously (8–23 CAG repeats) from north Indian population (Bahl *et al.* 2005).

The causative mutation remains unidentified in 187 patients in the present study at SCA12 locus due to the fact that molecular diagnosis was based on the published pathological ranges, which are mostly arbitrary based on association studies (van de Warrenburg et al. 2002; Martindale et al. 2012). Although in the current study, expansions in the pathological range based on published criteria was found in only one case, however >26 repeats were found in four patients and none of them in controls. These patients were reassessed for clinical features (table 1) and were found to have characteristic features of SCA12 in two cases and the other two had an expansion at SCA1 locus. This raises the question regarding the pathological threshold of the expanded repeats. Hence, we earlier proposed to reevaluate the threshold of abnormal repeat expansion at various SCA loci which includes SCA12 gene also (Gul et al. 2014).

LNA are relatively unstable and undergo expansion to reach an intermediate range from which further expansion to the disease range takes place. Therefore, a study on the percentage of LNA in any population would be an indirect reflection of the prevalence of the disease in that population. This was found to be true in African and Israeli populations (Goldman *et al.* 1996; Mor-Cohen *et al.* 1997). A similar study involving White and Japanese ADCA patients and normal individuals observed that the prevalence of SCAs is highly correlated with the frequency of LNAs in the normal population (Takano *et al.* 1998). The frequencies of LNAs in this study at SCA12 locus were quite high when compared with other populations (table 2). This study supports more prevalence of SCA12 in south Indian population.

Expression of expanded polyglutamine proteins leads to numerous cellular abnormalities, mediated at the protein level by a toxic gain in function conferred, at least in part, by misfolding of the repeat domain (Gatchel and Zoghbi 2005). Yet, protein misfolding due to expansion of polyglutamine tracts may not be the only mechanism for neurodegenerative diseases, repeats lower than normal range may also result in abnormal phenotype, as has been noted in diseases such as SBMA (Kooy *et al.* 1999). In the present study, lower repeats were identified in two patients and none of them in controls which indicate that lower repeats at these loci may play a role in disease pathogenesis. Lower repeats at *SCA12* gene locus are risk factors for Alzheimer's disease and essential tremor disorders (Chen *et al.* 2009). This study is the first study from India where six CAG repeats at SCA12 locus have been found in a clinically diagnosed ataxia patient.

Conclusions

This study indicates that an unexpectedly high proportion of the ataxic patients have repeat lengths in the *SCA12* gene which is longer than the control group, thereby raising the question if the currently used cut-off length for normal/ elongated CAG repeat should be reconsidered or changed. The frequencies of LNAs in this study at SCA12 locus were quite high when compared with other populations. The study supports more prevalence of SCA12 in South Indian population. It also suggests that repeats more than the normal range may also play an important role in disease pathogenesis.

Acknowledgements

Authors show gratitude towards the families and physicians who have acquiesced the samples for these studies. Authors are also grateful to Kamineni Hospitals, LB Nagar, Hyderabad for providing the facilities.

Limitation: The genetic analysis of *SCA12* gene in this study was from south Indian population and CAG repeats vary in different ethnic groups and the assumptions and hypothesis made in this manuscript may not be applicable to other population.

References

- Alluri R. V., Komandur S., Wagheray A., Chaudhuri J. R., Sitajayalakshmi, Meena A. K. *et al.* 2007 Molecular analysis of CAG repeats at five different loci: alternative hypothesis for disease pathogenesis. *Mol. Cells* 24, 338–342.
- Bahl S., Virdi K., Mittal U., Sachdeva M. P., Kalla A. K., Holmes S. E. *et al.* 2005 Evidence of a common founder for SCA12 in the Indian population. *Ann. Hum. Genet.* **69**, 528–534.
- Bahl S., Ahmed I., Indian Genome Variation Consortium and Mukerji M. 2009 Indian Genome Variation Consortium Utilizing

linkage disequilibrium information from Indian Genome Variation Database for mapping mutations: SCA12 case study. *J. Genet.* **88**, 55–60.

- Brussino A., Graziano C., Giobbe D., Ferrone M., Dragone E., Arduino C. *et al.* 2010 Spinocerebellar ataxia type 12 identified in two Italian families may mimic sporadic ataxia. *Mov. Disord.* 25, 1269–1273.
- Chen C. M., Hou Y. T., Liu J. Y., Wu Y. R., Lin C. H., Fung H. C. et al. 2009 PPP2R2B CAG repeat length in the Han Chinese in Taiwan: association analyses in neurological and psychiatric disorders and potential functional implications. Am. J. Med. Genet. B. Neuropsychiatr: Genet. 150, 124–129.
- Cho J. W., Kim S. Y., Park S. S. and Jeon B. S. 2008 Spinocerebellar ataxia type 12 was not found in Korean Parkinsonian patients. *Can. J. Neurol. Sci.* 35, 488–490.
- Deutsch S. I., Rosse R. B. and Lakshman R. M. 2006 Dysregulation of tau phosphorylation is a hypothesized point of convergence in the pathogenesis of Alzheimer's disease, frontotemporal dementia and schizophrenia with therapeutic implications. *Prog. Neuropsychopharmacol. Biol. Psychiatr.* **30**, 1369–1380.
- Faruq M., Scaria V., Singh I., Tyagi S., Srivastava A. K. and Mukerji M. 2009 SCA-LSVD: a repeat-oriented locus-specific variation database for genotype to phenotype correlations in spinocerebellar ataxias. *Hum. Mutat.* **30**, 1037–1042.
- Fujigasaki H., Verma I. C., Camuzat A., Margolis R. L., Zander C., Lebre A. S. *et al.* 2001 SCA12 is a rare locus for autosomal dominant cerebellar ataxia: a study of an Indian family. *Ann. Neurol.* **49**, 117–121.
- Gatchel J. R. and Zoghbi H. Y. 2005 Diseases of unstable repeat expansion: mechanisms and common principles. *Nat. Rev. Genet.* **6**, 743–755.
- Goldman A., Ramsay M. and Jenkins T. 1996 Ethnicity and myotonic dystrophy: a possible explanation for its absence in sub-Saharan Africa. *Ann. Hum. Genet.* **60**, 57–65.
- Goldberg Y. P., Kremer B., Andrew S. E., Theilmann J., Graham R. K., Squitieri F. *et al.* 1993 Molecular analysis of new mutations for Huntington's disease: intermediate alleles and sex of origin effects. *Nat. Genet.* 5, 174–179.
- Gul Lone W., Poornima S., Meena A. K., Rao K. P. and Hasan Q. 2014 Role of dynamic and mitochondrial mutations in neurodegenerative diseases with ataxia: lower repeats and LNAs at multiple loci as alternative pathogenesis. *J. Mol. Neurosci.* 837–847.
- Holmes S. E., O'Hearn E. E., McInnis M. G., Gorelick-Feldman D. A., Kleiderlein J. J., Callahan C. *et al.* 1999 Expansion of a novel CAG trinucleotide repeat in the 5' region of PPP2R2B is associated with SCA12. *Nat. Genet.* 23, 391–392.

- Khan I. A., Shaik N. A., Kamineni V., Jahan P., Hasan Q. and Rao P. 2015 Evaluation of gestational diabetes mellitus risk in south Indian women based on MTHFR (C677T) and FVL (G1691A) mutations. *Front. Pediatr.* **3**, 34.
- Kooy R. F., Reyniers E., Storm K., Vits L., van Velzen D, de Ruiter P. E. *et al.* 1999 CAG repeat contraction in the androgen receptor gene in three brothers with mental retardation. *Am. J. Med. Genet.* 85, 209–213.
- Martindale J. E., Seneca S., Wieczorek S. and Sequeiros J. 2012 Spinocerebellar ataxias: an example of the challenges associated with genetic databases for dynamic mutations. *Hum. Mutat.* **33**, 1359–1365.
- Mor-Cohen R., Magal N., Gadoth N., Shohat T. and Shohat M. 1997 Correlation between the incidence of myotonic dystrophy in different groups in Israel and the number of CTG trinucleotide repeats in the myotonin gene. *Am. J. Med. Genet.* **71**, 156–159.
- O'Hearn E., Holmes S. E. and Margolis R. L. 2012 Spinocerebellar ataxia type 12. *Handb. Clin. Neurol.* **103**, 535–547.
- Silveira I., Miranda C., Guimarães L., Moreira M. C., Alonso I., Mendonça P. *et al.* 2002 Trinucleotide repeats in 202 families with ataxia: a small expanded (CAG)n allele at the SCA17 locus. *Arch. Neurol.* 59, 623–629.
- Squitieri F., Andrew S. E., Goldberg Y. P., Kremer B., Spence N., Zeisler J. *et al.* 1994 DNA haplotype analysis of Huntington disease reveals clues to the origins and mechanisms of CAG expansion and reasons for geographic variations of prevalence. *Hum. Mol. Genet.* **3**, 2103–2114.
- Srivastava A. K., Choudhry S., Gopinath M. S., Roy S., Tripathi M., Brahmachari S. K. *et al.* 2001 Molecular and clinical correlation in five Indian families with spinocerebellar ataxia 12. *Ann. Neurol.* 50, 796–800.
- Stevanin G., Dürr A. and Brice A. 2000 Clinical and molecular advances in autosomal dominant cerebellar ataxias: from genotype to phenotype and physiopathology. *Eur. J. Hum. Genet.* 8, 4–18.
- Takano H., Cancel G., Ikeuchi T., Lorenzetti D., Mawad R., Stevanin, G. *et al.* 1998 Close associations between prevalences of dominantly inherited spinocerebellar ataxias with CAG-repeat expansions and frequencies of large normal CAG alleles in Japanese and Caucasian populations. *Am. J. Hum. Genet.* 63, 1060–1066.
- van de Warrenburg B. P., Sinke R. J., Verschuuren-Bemelmans C. C., Scheffer H., Brunt E. R., Ippel P. F. *et al.* 2002 Spinocerebellar ataxias in the Netherlands: prevalence and age at onset variance analysis. *Neurology* 58, 702–708.

Received 20 August 2014, in final revised form 3 July 2015; accepted 6 July 2015 Unedited version published online: 9 July 2015 Final version published online: 7 April 2016