



ARTICLE

# Analysis of CAG and CCG repeats in *Huntingtin* gene among HD patients and normal populations of India

Srimanta Pramanik<sup>1</sup>, Priyadarshi Basu<sup>1</sup>, Prasanta K Gangopadhaya<sup>2</sup>, Krishna K Sinha<sup>3</sup>, DK Jha<sup>3</sup>, S Sinha<sup>3</sup>, Shyamal K Das<sup>2</sup>, Bibek K Maity<sup>2</sup>, Subhas C Mukherjee<sup>4</sup>, Sushanta Roychoudhuri<sup>5</sup>, Partha P Majumder<sup>6</sup> and Nitai P Bhattacharyya<sup>1</sup>

<sup>1</sup>Crystallography and Molecular Biology Division, Saha Institute of Nuclear Physics; <sup>2</sup>Bangur Institute of Neurology, Calcutta; <sup>3</sup>Advanced Diagnostic Centre, Ranchi; <sup>4</sup>Calcutta Medical College; <sup>5</sup>Indian Institute of Chemical Biology; <sup>6</sup>Indian Statistical Institute, Calcutta, India

We have analysed the distribution of CAG and adjacent polymorphic CCG repeats in the *Huntingtin* gene in 28 clinically diagnosed unrelated Huntington's disease (HD) patients and in normal individuals belonging to different ethnic groups of India. The range of expanded CAG repeats in HD patients varied from 41 to 56 repeats, whereas in normal individuals this number varied between 11 and 31 repeats. We identified six CCG alleles from a total of 380 normal chromosomes that were pooled across different ethnic populations of India. There were two predominant alleles: (CCG)<sub>7</sub> (72.6%) and (CCG)<sub>10</sub> (20%). We report here for the first time one four-repeat CCG allele which has not been found in any population so far. We found 30 haplotypes (two loci CAG–CCG) for 380 normal chromosomes. In the present study, no statistically significant preponderance of expanded HD alleles was found on either (CCG)<sub>7</sub> or (CCG)<sub>10</sub> backgrounds. Our studies suggest that the overall prevalence of HD in Indian populations may not be as high as in Western populations. Further studies are necessary to identify the origin of HD mutation in these populations. *European Journal of Human Genetics* (2000) 8, 678–682.

**Keywords:** Huntington's disease; Indian populations; haplotype

## Introduction

Huntington's disease (HD: OMIM 143100) is a progressive neurodegenerative autosomal dominant disorder associated with motor disturbance, cognitive decline and psychiatric manifestations which usually has late onset.<sup>1</sup> HD is caused by an expansion of (CAG) repeats, located within the coding region of the *huntingtin* gene on the short arm of chromosome 4. In normal individuals, the number of CAG repeats varies between 10 and 35. HD results when CAG repeats exceed 35 in number.<sup>2</sup>

There is a wide variation in the prevalence of HD. In most countries of western Europe, United States and Canada, the prevalence is 5–10 affected persons per 100 000 individuals.<sup>3,4</sup> On the other hand, a much reduced prevalence of HD has been observed among the Chinese,<sup>5</sup> Japanese,<sup>6,7</sup> African Blacks<sup>8</sup> and Finnish populations.<sup>9</sup> Prevalence of HD among

immigrants from the Indian sub-continent (Pakistan, Punjab and Gujarat) was estimated to be similar to what has been observed for European populations.<sup>10</sup> HD prevalence among ethnically diverse populations of India is unknown.

Molecular basis of the differences in the prevalence of HD among various populations is unclear. For myotonic dystrophy, where CTG repeat in the 3' untranslated region of the myotonin protein kinase gene (*DM-PK*) undergoes expansion in patients, frequency of larger CTG repeats in the normal range is positively correlated with disease prevalence in populations from Africa and Israel.<sup>11,12</sup> Similarly, prevalence of various dominant spinocerebellar ataxias (SCAs) is positively correlated with the frequency of larger CAG repeats in the normal range in Caucasian and Japanese populations.<sup>13</sup> A similar positive correlation between frequency of large alleles (> 20 repeats) and HD prevalence has also been noted.<sup>14,15</sup> These normal larger repeats are prone to expand to the disease range and are possibly a reservoir for expanded alleles. There is an interesting relationship, which is variable across major human morphological groups, between the CAG repeat numbers and another closely linked CCG repeat locus in the *huntingtin* gene. In Caucasian populations, large

Correspondence: Dr Nitai P Bhattacharyya, Crystallography and Molecular Biology Division, Saha Institute of Nuclear Physics, 1/ AF Bidhan Nagar, Calcutta 700 064, India. Tel: +91 33 337 0679, extn 311; Fax: +91 33 337 4637; E-mail: nitai@hp2.saha.ernet.in  
Received 24 December 1999; revised 13 April 2000; accepted 27 April 2000

CAG repeat alleles, including those in the disease range, are strongly associated with (CCG)<sub>7</sub>,<sup>15,16-18</sup> while in Mongoloid populations this association is with (CCG)<sub>10</sub>.<sup>19</sup>

India represents one of the most ethnically and genetically diverse regions of the world.<sup>20</sup> Although there are reports of HD in India, no molecular genetic study on HD patients from Indian populations have so far been conducted.<sup>21</sup> In the present study, we report results of analyses of the CAG and CCG repeat distributions in 28 clinically diagnosed unrelated HD patients from eastern India and also in normal individuals belonging to different ethnic groups of India to determine the origin of HD mutation and to estimate the possible prevalence of the disease in Indian populations.

## Materials and methods

### Patient sampling

Patients visiting Bangur Institute of Neurology and Medical College Hospital, Calcutta and the Advanced Diagnostic Centre, Ranchi were identified and clinically examined by some of us. Blood samples were collected from the patients with appropriate consent. Patients and patients' relatives were interviewed to obtain family histories. Twelve patients had at least one affected first-degree relative.

### DNA isolation and PCR amplification

DNA was isolated from frozen blood samples of patients using a standard protocol. C1 and C2 primers were used to determine the CAG repeat length in the HD locus.<sup>16</sup> The amplified products were resolved on 6% denaturing polyacrylamide gel against allelic markers and positive controls of known sizes and hybridised with (CTG)<sub>10</sub>. Sizes of the allelic markers were determined by analysing on an ABI prism Genescan system using fluorescent-labelled C1 (TET-labelled) primer and also by silver staining of 8% polyacrylamide gel with appropriate DNA ladders. B1 and B2 primers were used to determine the CCG length. PCR conditions were the same as described by Andrew *et al.*<sup>16</sup> C1 and B2 primers were used to determine the length encompassing both the CAG and CCG repeats.<sup>16</sup>

### Population sampling

Blood samples were collected from various ethnic groups of India with prior informed consent and DNA was isolated using a standard protocol.<sup>22</sup> The population samples were collected from the states of Orissa, West Bengal, Uttar Pradesh and Tripura. Of these, the Lodhas and Santals are Austro-Asiatic speaking tribals and the Tripuris are a Tibeto-Burman speaking tribal group. The remaining are Indo-European speaking caste populations at different levels of social hierarchy. Further details about the populations studied have been described by Majumder *et al.*<sup>23</sup>

## Results

### Analysis of CAG repeats in HD patients and normal individuals

Twenty-eight clinically identified unrelated HD patients (20 males and 8 females) were analysed for (CAG)<sub>n</sub> expansion

at the HD locus. Age at presentation among affected individuals varied between 17 and 70 years. The range of the CAG repeat number on expanded HD chromosomes was 41–56, while the range of the normal allele on the homologous chromosome was 13–29 repeats. We noted an inverse correlation between age at presentation and CAG repeat number. In seven of the 12 HD patients with positive family history, the HD chromosome was derived paternally, while in the remaining patients it was maternally derived.

CAG repeat size distributions are presented in Table 1, separately for 12 ethnic populations as also for the pooled ( $n = 506$ ) sample. Size of the CAG repeats ranged between 11 and 31. Altogether 16 alleles were found in the normal individuals, of which the 16 repeat allele was the most frequent in every population. The overall observed heterozygosity was 0.7; the variation across populations was 0.54 in the Tripuri (TR1) to 0.78 in the Bengali Brahmin (BR2). CAG repeat distribution in the pooled sample was positively skewed (+ 1.27). Maximum number of alleles (12) was found in the Bengali Brahmin (BR2) population, whereas minimum number of alleles (6) was found in two populations, Tanti (TA1) and Santal (SA1). The distribution of the CAG repeats in the pooled 506 chromosomes revealed that 7.5% (38 alleles) were in the > 20 repeat range. Alleles having CAG size of 32–36 repeats were not found in any population. In the Rajput (RA1) population, one 31 repeat allele was found which was the highest CAG repeat size in our study populations. In the Uttar Pradesh Brahmin (BR1) and Mahishya (MA1) populations, one 11 repeat allele was found which was the lowest CAG repeat size.

The mean CAG repeat size was found to be  $(16.8 \pm 2.08)$  in the pooled set of 506 normal chromosomes. There was variation in the mean CAG size among the ethnic populations.

### Distribution of CCG repeats in HD patients and normal individuals

We identified three CCG repeat alleles, repeat numbers 7 (83.9%), 9 (5.3%) and 10 (10.7%) in 28 HD patients.

In normal individuals, we identified six CCG alleles in 380 normal chromosomes. We found one 4 repeat CCG allele among the Lodha (LO1) which has not been found in any population so far. A representative photograph showing all the six CCG alleles is depicted in Figure 1. The 7 repeat allele was uniformly the most frequent (pooled frequency = 72.6%) across populations. The 10 repeat allele was the next most frequent one (pooled frequency = 20%) in all the populations, except Santal (SA1), among whom the frequency of the 9 repeat allele was greater. In the Tripuri (TR1) population, the frequency of the 10 repeat allele was almost equal to the 7 repeat allele. Three 6 repeat alleles were found in the Tripuri (TR1) population which were absent in all other populations. The 8 repeat allele was absent in all the populations except in the Mahishya (MA1), Tanti (TA1) and

**Table 1** Distribution of CAG repeats in the *huntingtin* gene in various ethnic populations of India

Populations (2n)	Distribution of CAG repeats in the HD gene					Allele range	Total no. of alleles	Heterozygosity	CAG repeat size (mean $\pm$ s.d.)
	11–12	13–15	16–18	19–21	$\geq 22$				
BR2 (44)	0.000	0.068	0.591	0.205	0.136	13–26	12	0.78	17.9 $\pm$ 2.95
LO1 (64)	0.000	0.077	0.718	0.186	0.015	14–23	8	0.62	16.8 $\pm$ 1.54
SA1 (40)	0.000	0.250	0.750	0.000	0.000	13–18	6	0.69	15.9 $\pm$ 1.21
TR1 (60)	0.000	0.017	0.867	0.100	0.017	15–22	7	0.54	16.7 $\pm$ 1.20
TA1 (26)	0.000	0.077	0.654	0.269	0.000	14–21	6	0.70	17.0 $\pm$ 1.89
GA1 (10)	0.000	0.300	0.500	0.200	0.000	13–21	6	0.76	16.2 $\pm$ 2.31
MA1 (46)	0.022	0.152	0.717	0.087	0.022	11–25	9	0.70	16.4 $\pm$ 1.96
BA1 (52)	0.000	0.077	0.788	0.095	0.038	14–24	10	0.64	16.7 $\pm$ 1.88
AG1 (20)	0.000	0.100	0.650	0.150	0.100	15–24	7	0.75	17.5 $\pm$ 2.49
BR1 (50)	0.02	0.08	0.72	0.16	0.02	11–25	9	0.74	17.0 $\pm$ 2.28
CH1 (18)	0.000	0.167	0.667	0.056	0.111	13–23	8	0.71	16.8 $\pm$ 2.33
RA1 (76)	0.000	0.053	0.802	0.105	0.039	15–31	9	0.69	17.1 $\pm$ 2.23
Pooled (506)	0.003	0.093	0.733	0.127	0.032	11–31	16	0.70	16.8 $\pm$ 2.08

Abbreviations of populations:

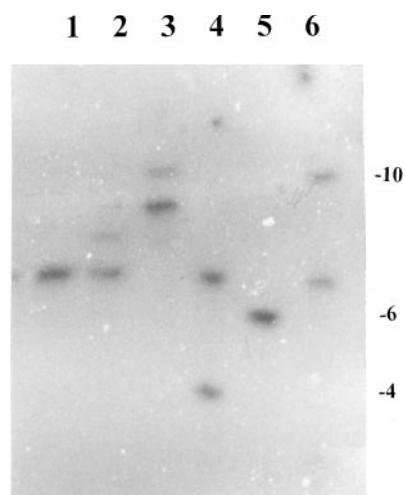
BR2: West Bengal Brahmin; LO1: Lodha; SA1: Santal; TR1: Tripuri; TA1: Tanti; GA1: Gaud; MA1: Mahishya; BA1: Bagdi; AG1: Agharia; BR1: Uttar Pradesh Brahmin; CH1: Chamar; RA1: Rajput

Bengali Brahmin (BR2) populations, each of which had one 8 repeat allele.

A significant finding is that 11 and 12 repeat alleles which have been found in Western populations, were absent in the study populations. Observed heterozygosity in the pooled set of chromosomes was 0.42.

#### Haplotype analysis

We could unambiguously determine 2-loci CAG–CCG haplotypes for 22 of the 28 patients who were either homozygous at the CCG locus or data were available from appropriate relatives. For the remaining 6 patients, we determined the



**Figure 1** Representative photograph showing CCG alleles of different repeat numbers in the *huntingtin* gene. CCG repeat number is indicated on the right of the photograph.

haplotypes by using primers that span each individual tract of CAG and CCG repeats alone and comparing this result with PCR amplimers which were generated by using primers (C1 and B2) that span both adjacent repeats. The result shows that in three cases expanded CAG repeat was associated with (CCG)<sub>10</sub> and in 25 cases HD chromosome was associated with (CCG)<sub>7</sub>.

In normal individuals, we unambiguously determined the haplotypes for those who were homozygous either at the CAG or CCG locus. This method helped us to correctly determine CAG–CCG haplotypes in 282 normal chromosomes. We determined the phase in another 98 chromosomes by methods described above. Altogether, 30 haplotypes were found in this manner for 380 normal chromosomes (Table 2). Haplotype (CAG)<sub>16</sub>–(CCG)<sub>7</sub> has the maximum frequency (0.337), whereas (CAG)<sub>16</sub>–(CCG)<sub>10</sub> has the second highest frequency (0.113).

We calculated the percentage of HD chromosomes associated with CCG allele sizes of 7 and 10 repeats and compared it with that of other published data<sup>15,19</sup> (Table 3). In Western populations, 94.4% of HD chromosomes were associated with 7 repeat CCG allele, whereas 5.6% were associated with 10 repeat CCG allele. In populations with low prevalence of HD (eg Chinese, Japanese, Blacks, Finnish), 38.9% of HD chromosomes segregated with 7 repeat CCG allele, while 61.1% segregated with 10 repeat CCG allele. In another report on Japanese populations, 84.5% of expanded alleles were associated with 10 repeat CCG allele.<sup>19</sup> In Indian populations, most of the HD chromosomes (89.3%) segregated with 7 repeat CCG allele, while 10.7% segregated with 10 repeat CCG allele.

We also calculated the mean CAG repeat size on normal chromosomes with CCG repeat numbers of 7 and 10

**Table 2** Frequency of various haplotypes

Haplotype no.	Haplotype	No. of haplotypes	Frequency
1	(CAG) <sub>11</sub> -(CCG) <sub>7</sub>	2	0.0053
2	(CAG) <sub>13</sub> -(CCG) <sub>9</sub>	3	0.0079
3	(CAG) <sub>13</sub> -(CCG) <sub>10</sub>	3	0.0079
4	(CAG) <sub>14</sub> -(CCG) <sub>7</sub>	10	0.0263
5	(CAG) <sub>14</sub> -(CCG) <sub>10</sub>	1	0.0026
6	(CAG) <sub>15</sub> -(CCG) <sub>7</sub>	18	0.0474
7	(CAG) <sub>15</sub> -(CCG) <sub>10</sub>	1	0.0026
8	(CAG) <sub>16</sub> -(CCG) <sub>4</sub>	1	0.0026
9	(CAG) <sub>16</sub> -(CCG) <sub>6</sub>	3	0.0079
10	(CAG) <sub>16</sub> -(CCG) <sub>7</sub>	128	0.3368
11	(CAG) <sub>16</sub> -(CCG) <sub>8</sub>	2	0.0052
12	(CAG) <sub>16</sub> -(CCG) <sub>9</sub>	10	0.0263
13	(CAG) <sub>16</sub> -(CCG) <sub>10</sub>	43	0.1131
14	(CAG) <sub>17</sub> -(CCG) <sub>7</sub>	39	0.1026
15	(CAG) <sub>17</sub> -(CCG) <sub>9</sub>	5	0.0132
16	(CAG) <sub>17</sub> -(CCG) <sub>10</sub>	22	0.0579
17	(CAG) <sub>18</sub> -(CCG) <sub>7</sub>	17	0.0447
18	(CAG) <sub>18</sub> -(CCG) <sub>9</sub>	2	0.0052
19	(CAG) <sub>18</sub> -(CCG) <sub>10</sub>	4	0.0105
20	(CAG) <sub>19</sub> -(CCG) <sub>7</sub>	25	0.0658
21	(CAG) <sub>19</sub> -(CCG) <sub>8</sub>	1	0.0026
22	(CAG) <sub>20</sub> -(CCG) <sub>7</sub>	12	0.0316
23	(CAG) <sub>21</sub> -(CCG) <sub>7</sub>	12	0.0316
24	(CAG) <sub>21</sub> -(CCG) <sub>10</sub>	2	0.0052
25	(CAG) <sub>22</sub> -(CCG) <sub>7</sub>	4	0.0105
26	(CAG) <sub>23</sub> -(CCG) <sub>7</sub>	4	0.0105
27	(CAG) <sub>23</sub> -(CCG) <sub>9</sub>	1	0.0026
28	(CAG) <sub>25</sub> -(CCG) <sub>7</sub>	2	0.0052
29	(CAG) <sub>26</sub> -(CCG) <sub>7</sub>	2	0.0052
30	(CAG) <sub>31</sub> -(CCG) <sub>7</sub>	1	0.0026

respectively and compared our data with that of Western populations and populations with low prevalence of HD, eg Chinese, Japanese, Blacks, Finnish.<sup>15</sup> The mean CAG repeat sizes on chromosomes with CCG repeat of 7 were  $19.3 \pm 3.8$  and  $16.8 \pm 1.5$  in Western populations and populations with low prevalence of HD, respectively. In Indian populations the mean CAG size was found to be  $17.1 \pm 2.3$  in this case. The mean CAG repeat sizes on chromosomes with CCG repeat of

10 were  $16.7 \pm 2.5$  and  $16.2 \pm 1.9$  in Western populations and populations with low prevalence of HD, respectively. In Indian populations the mean CAG repeat size was found to be  $16.4 \pm 1.2$  in this case.

### Discussion

Analysis of the CAG repeat locus among 28 unrelated clinically diagnosed HD patients from the eastern part of India showed that the range of expansion of CAG repeats was between 41 and 56 in the *huntingtin* gene. The smallest size of the CAG repeat which causes HD is not defined precisely. In Western populations and in the Japanese, the lowest number has been reported to be around 37. In our study, the smallest size of the CAG repeat is found to be 41. The ranges of the expanded repeat in HD was similar to those obtained in other populations.<sup>2,4,19</sup>

We have found an inverse correlation between the age at presentation and CAG repeat expansion in HD patients which is consistent with earlier findings from Caucasian and Japanese populations.

Distribution of CCG repeats in the pooled data is similar to that noted in previous studies.<sup>15,16,19</sup> We have observed for the first time one (CCG)<sub>4</sub> repeat allele. We did not observe any 11 and 12 repeat CCG alleles which, however, have been found in Caucasian populations.

In Caucasian HD patients, the expanded alleles were strongly associated (94.4%) with (CCG)<sub>7</sub> allele, while in normal populations the frequency of (CCG)<sub>7</sub> was 67.8%.<sup>15</sup> In the Japanese populations, 84.5% of the expanded alleles were associated with (CCG)<sub>10</sub> and in normal populations, this frequency was 37.3%.<sup>19</sup> This difference was explained by the difference in the origin of ancestral lineages of the two groups. Masuda *et al* proposed that HD mutation observed in the Japanese could be a new mutation that occurred after migration from western Europe.<sup>19</sup>

**Table 3** Frequency of CCG alleles in normal individuals and HD patients in populations of different geographical origins

CCG repeat size	Normal individuals				HD patients (%)			
	Western European <sup>a</sup> (n=479)	Total no. of CCG alleles (%)			Western European <sup>a</sup> (n=178)	Total no. of CCG alleles (%)		
		Populations with low prevalence of HD <sup>a</sup> (n=417)	Japanese <sup>b</sup> (n=185)	Indian (2n=380)		Populations with low prevalence of HD <sup>a</sup> (n=72)	Japanese <sup>b</sup> (n=58)	Indian (n=28)
4	0	0	0	1 (0.3)	0	0	0	0
6	0	5 (1.2)	1 (0.5)	3 (0.8)	0	0	0	0
7	325 (67.8)	207 (49.6)	115 (62.2)	276 (72.6)	168 (94.4)	28 (38.9)	9 (15.5)	25 (89.3)
8	2 (0.4)	9 (2.2)	0	3 (0.8)	0	0	0	0
9	21 (4.4)	9 (2.2)	0	21 (5.5)	0	0	0	0
10	130 (27.1)	182 (43.6)	69 (37.3)	76 (20.0)	10 (5.6)	44 (61.1)	49 (84.5)	3 (10.7)
11	0	5 (1.2)	0	0	0	0	0	0
12	1 (0.2)	0	0	0	0	0	0	0

<sup>a</sup>Data taken from Squitieri *et al*.<sup>15</sup>

<sup>b</sup>Data taken from Masuda *et al*.<sup>19</sup>

However, in the present study, no statistically significant preponderance of expanded HD alleles was found in either (CCG)<sub>7</sub> or (CCG)<sub>10</sub> backgrounds ( $\chi^2 = 4.12$ ,  $df = 2$ ;  $P > 0.05$ ). Thus, it is unlikely that there is preferential expansion of CAG repeats on (CCG)<sub>7</sub> chromosomes as was proposed by earlier workers.<sup>16-18</sup>

Squitieri *et al* have shown association between CAG repeat size on normal chromosomes with the frequency of HD.<sup>15</sup> According to their study, CAG size on normal chromosomes of African Blacks (mean =  $16.2 \pm 2.5$ ), Chinese (mean =  $16.4 \pm 1.5$ ), Japanese (mean =  $16.6 \pm 1.3$ ) and Finnish (mean =  $17.1 \pm 1.8$ ) populations (countries with low prevalence of HD) have been found to be smaller compared to Western populations (mean =  $18.4 \pm 3.7$ ) where the prevalence rate is higher. Mean CAG repeat size varies in our study populations from  $15.9 \pm 1.21$  to  $17.9 \pm 2.95$ . In the pooled data, this value is  $16.8 \pm 2.08$ , which is close to the frequency observed in African and Oriental populations. At the same time, Squitieri *et al* have also suggested that there is a variation in the distribution of CAG size on normal chromosomes of different ancestries.<sup>15</sup> They have found fewer numbers of chromosomes with > 20 CAG repeats in populations with lower prevalence of HD compared with Western populations. In Indian populations, 7.5% (38 alleles) were in the > 20 CAG repeat range which was lower compared with Western populations as found by Squitieri *et al*. It thus appears that overall prevalence of HD in Indian populations is expected to be low, even though there may be variation across ethnic groups. It is interesting to note that one of the two Austro-Asiatic speaking tribal populations (Santal and Lodha) had the lowest mean CAG repeat size and does not have large normal alleles greater than 18 CAG repeats. It has recently been proposed that these Austro-Asiatic speaking groups moved to India from Africa and settled 60 000 years ago.<sup>23</sup> Our studies on CTG repeat variation in these populations in the *DM-PK* gene (Basu *et al*) (submitted) and CAG repeat variation in the *huntingtin* gene (present manuscript) support this notion.

#### Acknowledgements

We are grateful to Dr Bidyut Roy, Mrs Monami Roy, Mr Badal Dey and Mr Madan Chakraborty of ISI, Calcutta, for collection of blood samples and isolation of genomic DNA. We thank Mr Utpal Basu and Mr Ashim Chakraborty of SINP, Calcutta, for technical support and photography. This work was partially supported by a grant from the Department of Biotechnology, Government of India, to NPB.

#### References

- 1 Martin JB, Gusella JF: Huntington's disease: Pathogenesis and management. *N Engl J Med* 1986; **315**: 1267-1276.
- 2 The Huntington's Disease Collaborative Research Group: A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. *Cell* 1993; **72**: 971-983.

- 3 Hayden MR: Huntington's chorea. Springer-Verlag: New York, 1981.
- 4 Harper PS: Huntington's disease. WB Saunders: London, 1991.
- 5 Leung CM, Chan YW, Chang CM, Yu YL, Chen CN: Huntington's disease in Chinese: a hypothesis of its origin. *J Neurol Neurosurg Psychiatry* 1992; **55**(8): 681-684.
- 6 Narabayashi H: Huntington's chorea in Japan: Review of literature. In: *Advances in Neurology*. Vol 1. Raven Press: New York, 1973.
- 7 Kanazawa I: Prevalence rate of Huntington's chorea in Ibaraki prefecture. Ministry of Health & Welfare of Japan: 1983; pp 151-156.
- 8 Hayden MR, Macgregor J, Beighton PH: The prevalence of Huntington's chorea in South Africa. *S Afr Med J* 1980; **58**(5): 193-196.
- 9 Palo J, Somer H, Ikonen E, Karila L, Peltonen L: Low prevalence of Huntington's disease in Finland. *Lancet* 1987; **2**: 805-806.
- 10 Shiwach RS, Lindenbaum RH: Prevalence of Huntington's disease among UK immigrants from the Indian subcontinent. *Br J Psychiatry* 1990; **157**: 598-599.
- 11 Goldmann A, Ramsay M, Jenkins T: Ethnicity and myotonic dystrophy: a possible explanation for its absence in sub-Saharan Africa. *Am J Hum Genet* 1996; **60**: 57-65.
- 12 Mor-Cohen R, Magal N, Gadoth N, Shohat T, Shohat M: 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* 1997; **71**(2): 156-159.
- 13 Takano H, Cancel G, Ikeuchi T *et al*: 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* 1998; **63**: 1060-1066.
- 14 Rubinsztein DC, Amos W, Leggo J *et al*: Mutational bias provides a model for the evolution of Huntington's disease and predicts a general increase in disease prevalence. *Nat Genet* 1994; **7**: 525-530.
- 15 Squitieri F, Andrew SE, Goldberg YP *et al*: 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* 1994; **3**(12): 2103-2114.
- 16 Andrew SE, Goldberg YP, Theilmann J, Zeisler J, Hayden MR: A CCG repeat polymorphism adjacent to the CAG repeat in the Huntington disease gene: implications for diagnostic accuracy and predictive testing. *Hum Mol Genet* 1994; **3**(1): 65-67.
- 17 Barron LH, Rae A, Holloway S *et al*: A single allele from the polymorphic CCG rich sequence immediately 3' to the unstable CAG trinucleotide in the IT15 cDNA shows almost complete disequilibrium with Huntington's disease chromosomes in the Scottish population. *Hum Mol Genet* 1994; **3**: 173-175.
- 18 Novelletto A, Persichetti F, Sabbadini G *et al*: Analysis of the trinucleotide repeat extension in Italian families affected with Huntington's disease. *Hum Mol Genet* 1994; **3**: 93-98.
- 19 Masuda N, Goto J, Murayama N, Watanabe M, Kondo I, Kanazawa I: Analysis of triplet repeats in the huntingtin gene in Japanese families affected with Huntington's disease. *J Med Genet* 1995; **32**: 701-705.
- 20 Majumder PP: People of India: Biological diversity and affinities. *Evol Anthropol* 1998; **6**: 100-110.
- 21 Chuttani PN: Huntington's chorea in India. *J A Physician India* 1955; **3**: 40.
- 22 Miller SA, Dykes DD, Polesky HF: A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res* 1988; **16**: 1215.
- 23 Majumder PP, Roy B, Banerjee S *et al*: Human-specific insertion/deletion polymorphisms in Indian populations and their possible evolutionary implications. *Eur J Hum Genet* 1999; **7**: 435-446.