ORIGINAL ARTICLE

Frequencies of spinocerebellar ataxia subtypes in Thailand: window to the population history?

Thanyachai Sura¹, Jakris Eu-ahsunthornwattana¹, Supak Youngcharoen¹, Manisa Busabaratana¹, Donniphat Dejsuphong¹, Objoon Trachoo¹, Supischa Theerasasawat¹, Atchara Tunteeratum¹, Chinnakrit Noparutchanodom² and Supoch Tunlayadechanont³

Spinocerebellar ataxias (SCAs) are a heterogeneous group of disorders with almost 30 subtypes. The prevalence and relative frequency of each subtype vary among different populations. In this article, we report the relative frequency of six SCA subtypes in the Thai population and attempt to explain the observed pattern when compared with other populations in this region. We searched for SCA type 1, SCA2, SCA3, SCA6, SCA7 and dentatorubral–pallidoluysian atrophy mutations using GeneScan analysis in 340 patients from 182 families, in which at least one person had a clinical diagnosis of SCA. We analyzed the relative frequencies of SCA subtypes on a family basis, and compared these with the data in the Chinese and Indian populations. SCA3 was found in 19.2% of the patients (Agresti–Coull 95% confidence interval: 14.1–25.6%), SCA1 in 11.5% (7.6–17.1%) and SCA2 in 10.4% (6.7–15.8%). SCA6 was found in three families, with a relative frequency of 1.6% (0.3–5.0%). Compared with the related populations, the Thai SCA3 frequency was less than that of the Chinese, whereas it was higher than that in most of the Indian studies. The reverse is true for the SCA1/SCA2 frequency. A similar study in Singapore, where there was a clear history of population admixture, also showed the frequencies between those of the Chinese and the Indian populations. Although SCA3 was the most common identifiable SCA subtype in Thailand, SCA1 and SCA2 were also relatively common. Our results also supported some degree of admixture with the Indians in the Thai population and justify further study in the area.

Journal of Human Genetics (2009) 54, 284–288; doi:10.1038/jhg.2009.27; published online 27 March 2009

Keywords: Asians (MeSH: Asian continental ancestry group); Indochina; Machado–Joseph disease; population admixture; prevalence; spinocerebellar ataxias

INTRODUCTION

Spinocerebellar ataxias (SCAs), also known clinically as autosomal dominant cerebellar ataxias, are a group of genetically and phenotypically heterogeneous ataxic disorders resulting from the progressive degeneration of the cerebellum and its connections, with an estimated prevalence of up to five to seven cases per 100 000 people.^{1–3} The most widely accepted classification of SCAs at present corresponds to the order of gene description¹ and consists of at least 28 subtypes, corresponding to defects in at least 28 distinct loci.² Among these, SCA3 (also known as Machado–Joseph disease: MJD) is the most common subtype worldwide, whereas SCA1, SCA2, SCA6, SCA7 and SCA8 are the other common subtypes, each with over 2% prevalence; the rest are believed to be rare.³

The overall prevalence and frequency of genetic subtypes of SCA vary among different populations, similar to the case in Cuba (predominantly SCA2), Portugal, Brazil and Azores (SCA3/MJD), Mexico (SCA10) and Japan (dentatorubral–pallidoluysian atrophy:

DRPLA).^{1–3} Estimations of the subtype frequencies are therefore important in the planning of diagnostic strategy in a specific population. Furthermore, as the frequency of each SCA subtype also differs between the Chinese-related and Indian-related populations (see Discussion), and the variation is believed to be the result of founder populations,^{3,4} it could potentially be useful for assessing the mix between the Chinese and the Indians in the Indochina peninsula.

MATERIALS AND METHODS

Study design

This is a retrospective review of the results collected as part of genetic consultation from August 1994 to July 2006. The study has been approved by the ethical committee of Ramathibodi Hospital.

Subjects

Our medical genetic division has offered molecular diagnostic service for SCAs since 1994. By July 2006, we had obtained DNA samples of 340 individuals

¹Division of Medical Genetics and Molecular Medicine and Academic Center for Medical Genetics, Department of Medicine, Faculty of Medicine, Ramathibodi Hospital, Mahidol University, Bangkok, Thailand; ²Department of Medicine, Ratchaburi Hospital, Ratchaburi, Thailand and ³Division of Neurology, Department of Medicine, Faculty of Medicine, Ramathibodi Hospital, Mahidol University, Bangkok, Thailand

Correspondence: Dr J Eu-ahsunthornwattana, Division of Medical Genetics and Molecular Medicine, Department of Medicine, Ramathibodi Hospital, Rama VI Rd., Phaya Thai, Bangkok 10400, Thailand.

E-mail: jakris_e@iname.com

Received 8 September 2008; revised 23 February 2009; accepted 8 March 2009; published online 27 March 2009

from 182 families. Each family has at least one person who had symptoms of progressive ataxia unexplained by other metabolic, nutritional or infective causes, and without space-occupying lesion on the cranial magnetic resonance imaging. These patients were clinically diagnosed as having SCAs by their attending neurologists and were referred to the genetic service for molecular confirmation and for counseling about the disease and the risks to themselves and their relatives. As part of the counseling, patients were advised to share the information with their relatives, who can then opt for a counseling session with a view to undergo predictive testing if they so wish. This is the reason why some of the non-index cases were asymptomatic at the time of testing. If more than one symptomatic person in a family attended the service at the same time, we assigned only one person as the index case and the remaining as non-index cases (see also Discussion).

Table 1 Relative frequency of SCA subtypes in Thailand

Region	Population ^a	Subtype					Total
		SCA1	SCA2	SCA3/MJD	SCA6	Other	
Bangkok	5 716 248	16 (10%)	16 (10%)	31 (20%)	3 (2%)	92 (58%)	158
Central	10 966 538	3 (18%)	1 (6%)	4 (24%)	0 (0%)	9 (53%)	17
Northeast	21 385 647	1 (20%)	2 (40%)	0 (0%)	0 (0%)	2 (40%)	5
North	11 871 934	1 (50%)	0 (0%)	0 (0%)	0 (0%)	1 (50%)	2
All index cases		21 (12%)	19 (10%)	35 (19%)	3 (2%)	104 (57%)	182
95% CI ^b		8–17%	7–16%	14–26%	0–5%	50–64%	
All subjects		31 (9%)	29 (9%)	74 (22%)	7 (2%)	199 (59%) ^c	340
95% CI ^b		6–13%	6–12%	18–26%	1–4%	53–64%	

Abbreviations: CI, confidence interval; SCA, spinocerebellar ataxia.

Figures given are numbers of cases in each category, along with the row-percentages in parentheses. Subjects in the 'Other' column are those who were clinically diagnosed with SCA, but were tested negative for SCA1, SCA2, SCA3/MJD, SCA6, SCA7 and DRPLA, apart from the last row (see row 'All subjects').

^aThe population in each region is based on an official Citizen Registry,¹⁴ and may be underestimated in metropolitan areas, especially, Bangkok, whereas overestimated in the more rural areas, particularly, the Northeast.

^bFor the row-percentage of each SCA subtype, calculated using the Agresti–Coull method.¹³

^cThis group also included asymptomatic, at-risk individuals who were relatives of the symptomatic cases and were thus offered the tests. They were, therefore, less likely to be affected.

DNA analysis

Genomic DNA of each subject was extracted from the peripheral blood sample by the standard phenol–chloroform method. PCR was performed with the primers Rep1 and Rep2 for SCA1⁵ (annealing temperature, 62 °C), SCA2-A and SCA2-B for SCA2⁶ (57 °C), MJD52 and MJD25 for SCA3/MJD⁷ (57 °C), SCA6F and SCA6R for SCA6⁸ (57 °C), 4U1024 and 4U716 for SCA7⁹ (52 °C) and CTG-B37-F and CTG-B37-R for DRPLA^{10,11} (47 °C). The 5'-end of each forward primer was fluorescent-tagged, allowing us to determine the repeat size of each amplicon by GeneScan analysis (ABI 310 Genetic analyzer; Applied Biosystems, Foster City, CA, USA). Subjects with more than 40 CAG repeats in SCA1, 34 repeats in SCA2, 55 repeats in SCA3/MJD, 21 repeats in SCA6, 37 repeats in SCA7 or 48 CAG repeats in DRPLA locus were classified as having positive results for that subtype.

Between 1994 and 2004, only diagnostic testing for SCA1 was offered. When we introduced tests for SCA3/MJD (2004) and SCA6 (late 2005), the stored DNA samples of the patients who had negative test result for SCA1 were reanalyzed for SCA3/MJD and SCA6. More recently, the remaining negative samples were again retested when the tests for SCA2, SCA7 and DRPLA became available.

Data entry and analysis

The test results, along with the individual and family identification data, were entered into a spreadsheet application. Validation and cleaning was carried out before importation into R-2.8.0¹² for statistical analysis. After initial data integrity check, we calculated the proportion of each SCA subtype with its Agresti–Coull 95% confidence interval¹³ in all subjects as well as in only the index cases, and compared these with those of earlier studies in related populations. We also calculated the ratio between the risk of having SCA3 and that of having either SCA1 or SCA2 (that is, the odds of having SCA3 rather than SCA1/2) in each population, along with its non-parametric bootstrap bias-corrected and accelerated (BCa) 95% confidence interval, and compared this across the populations.

RESULTS

Among the 340 individuals tested, 182 were index cases in their families and 158 were the relatives of the index cases. They came from 182 families, with the numbers of individuals tested in each family ranging from 1 to 20. All of the subjects were Thai; the majority of them were from the Bangkok area (see Table 1). As we have relatively few subjects outside the Bangkok area, stratification analysis by region was not very informative. The relative frequency among the Bangkok cases was similar to the overall frequency.

Table 2 Frequency distributions of SCA1, -2, -3, -6, -7 and DRPLA in related populations

Country	Year	Subtype						Other	Total
		SCA1	SCA2	SCA3/MJD	SCA6	SCA7	DRPLA		
China—mainland ¹⁵	Pub. 2000	4 (3.3)	5 (4.1)	41 (33.6)	0 (0.0)	0 (0.0)	0 (0.0)	72 (59.0)	122
Hong Kong ^{16 a}	2001–2003	3 (21.4) ^a	0 (0.0) ^a	10 (71.4) ^a	0 (0.0) ^a	0 (0.0) ^a	1 (7.1) ^a	NA	14 ^a
Taiwan ¹⁷	Pub. 2001	4 (3.3)	8 (6.5)	35 (28.5)	10 (8.1)	2 (1.6)	1 (0.8)	63 (51.2)	123
Taiwan—Central/South ¹⁸	Pub. 2004	1 (1.1)	9 (10.0)	26 (28.9)	0 (0.0)	1 (1.1)	NA	53 (58.9)	90
India—mostly North ¹⁹	Pub. 2000	3 (7.1)	10 (23.8)	2 (4.8)	0 (0.0)	0 (0.0)	0 (0.0)	27 (64.3)	42
India—East ²⁰	1997–1999	6 (10.5)	10 (17.5)	4 (7.0)	1 (1.8)	0 (0.0)	0 (0.0)	36 (63.2)	57
India—East (mainly Bihar) ²¹	1996–1999	4 (14.3)	16 (57.1)	0 (0.0)	0 (0.0)	0 (0.0)	NA	8 (28.6)	28
India—South ²²	2003–2006	34 (31.8)	24 (22.4)	15 (14.0)	NA	NA	NA	34 (31.8)	105
India—not specified ²³	Pub. 2001	12 (15.6)	19 (24.7)	2 (2.6)	0 (0.0)	2 (2.6)	0 (0.0)	42 (54.5)	77
India—East (West Bengal) ²⁴	1996–2000	2 (14.3)	4 (28.6)	5 (35.7)	0 (0.0)	0 (0.0)	0 (0.0)	3 (21.4)	14
Singapore ^{25 b}	Pub. 2002	4 (3.3) ^b	12 (10.0) ^b	15 (12.5) ^b	2 (1.7) ^b	0 (0.0) ^b	2 (1.7) ^b	85 (82.5) ^b	120 ^b
Thailand (this study)	1994–2006	21 (11.5)	19 (10.4)	35 (19.2)	3 (1.6)	0 (0.0)	0 (0.0)	104 (57.1)	182

Abbreviations: NA, not available/not applicable; SCA, spinocerebellar ataxia.

Where possible, the years during which the cases were recruited are given; otherwise the publication years are given (denoted by 'pub.').

Numbers from earlier studies have been adjusted to include all tested cases (including the cases that appeared to be sporadic) on a one index-case per family basis. The numbers in parentheses are percentage relative to total numbers of cases. All families tested are quoted in the total column except for the following:

^aThe Hong Kong study, in which only the total number of cases/families who were tested positive for either SCA1, -2, -3, -6, -7, -8, -12 or DRPLA was reported, the total number of families with positive results was used.

^bThe Singapore study, in which 204 cases were tested.²⁵ These were estimated to have come from at least 120 families (EK Tan, personal communication, 22 August 2006).

The frequency of each SCA subtype is shown in Table 1. The most prevalent subtype detected was SCA3/MJD, which affected 19% of the families, followed by SCA1 (11.5%) and SCA2 (10.4%). SCA6 was relatively rare, with just less than 2% in relative frequency. About 57% of the index cases could not be assigned a molecular diagnosis based on the tests currently offered, implying other disease subtypes.

Comparison of the subtype frequencies with other populations is shown in Table 2 and graphically in Figure 1.

DISCUSSION

In line with earlier studies in most populations, we found that SCA3/MJD is the most prevalent subtype of SCAs in the Thai population. However, this accounted for just about 19% of the families compared

Figure 1 Geographical variation in SCA subtypes in south, southeast and east Asia (**a**), and the odds of having the 'Chinese SCA' (SCA3) rather than 'Indian SCAs' (SCA1 and -2) cases in these populations, with their non-parametric bootstrap BCa 95% confidence intervals (**b**). The Hong Kong study was excluded from the geographical plot as the number of negative ('other') cases had not been reported. The upper confidence limits were 10 and 12 for China and Hong Kong, respectively. The confidence interval for the East India/Bihar study could not be calculated due to the total absence of SCA3 case within the sample. See Table 2 for data sources.

with about 30% in the Chinese-related populations other than those in Singapore. On the contrary, apart from the studies carried out in West Bengal, which reported the relative frequency of SCA3/MJD of approximately 36%,²⁴ the reported relative frequency of SCA3/MJD ranged from 0 to about 14% in most studies in Indian populations.^{19,20,21,23} The West Bengal study is in fact an interesting exception, as earlier studies have shown a genetic relationship between the population in this area and the Tibeto-Burman speakers from South-east and East Asia.^{26–29} Also, interestingly, in Singaporeans who are a mixture of ethnic Chinese, Malays and Indians, the relative frequency of SCA3/MJD was even lower than that in ours, at about 13%.

At 11.5%, the relative frequency of SCA1 in our population was near the lower end of the relative frequencies in the Indians (7–32%),^{19,20,21–23} but was much higher than those in the Chinese/Japanese populations (approximately 3%).^{15,17,18,25,30} SCA2, which was almost as common as SCA1 in our population, also showed a similar pattern: its relative frequency of 10.4% was much lower than that in the Indians (18–57%)^{19,20,21–23} while being still marginally higher than the upper range of the frequencies in the Chinese populations (4–10%).^{15,17,18,25,30}

These three common subtypes accounted for 41% of the cases. Of the remaining three tested subtypes, only SCA6 was found. SCA6 is generally rare in this region, apart from in a specific ethnic group in Taiwan. Our relative frequency of SCA6 was comparable with that of the Singaporeans.

The frequency of index patients who were not detected as SCA1, -2, -3, -6, -7 or DRPLA in this study was comparable with that from other studies in the region. This probably reflected the prevalence of one or more untested subtypes such as SCA12.^{23,25}

The number of all individuals who had negative results could be partly inflated as asymptomatic (thus possibly unaffected) family members were included. As the frequency of SCA subtypes in all cases would also be influenced by the size of the families, we believe the analysis based on the index cases as representatives for their families as carried out here would yield a more accurate result.

An interesting pattern emerges if one compares, among populations in this region, the odds of having SCA3, which is the predominant type in Chinese populations, rather than having either SCA1 or SCA2, which are more common among the Indian populations. In our series, the odds were 0.88, which was similar to those in the other two populations with evident admixture (Singapore²⁵ and West Bengal;²⁴ see Figure 1b). With a range of 0.83 to 0.94, the odds in these three populations were much higher than those in the Indian populations (highest value=0.28 in South India²²), whereas being much lower than those in the Chinese populations (lowest value=2.60 in Taiwan¹⁸). We believe that this pattern could support the hypothesis that there is an admixture between the Chinese- and Indian-related populations in the Thai population. However, some caution is needed against making conclusions based on this alone, as many studies had small sample sizes, making the estimations rather unstable. Furthermore, although this pattern will arise in the presence of a population admixture, the reverse is not necessarily true: it could arise, for example, if the Thai SCAs have a background independent of the Indian or Chinese populations, or because of a random founding population within the main population, or even a random error from sampling.

Over the last decade, genetic studies have linked the populations in the Indochina peninsula, including the Thai people, to the Han Chinese, particularly those in the southern part of China.^{31–33} This is quite expected as the region is close to China and there has earlier been historical, archeological and linguistic evidence to support the

relationship.^{34,35} At the same time, given the similar proximity of the region to India, and the historical, archeological and cultural evidence of relationship with the Indian subcontinent,^{34,36} it is very likely that there is some degree of mixing between the Chinese and Indian populations in the Indochina peninsula, although there has earlier been little genetic evidence to support this. In this context, we believe that the observation from this study at least indirectly adds genetic support for this historical and cultural linking.

The majority of our subjects came from Bangkok and the central part of Thailand. Although this could mean that people in this area are relatively overrepresented in our sample, the degree of overrepresentation may not be as high, as the official population data are known to underestimate the number of people in the urban area, particularly Bangkok and its surroundings, and overestimate the number of people in a more rural area. This is because of the 'unofficial' migration that has not been accounted for in the official registry. Furthermore, because of the relatively small size of the country and the homogeneity of the population, we do not think that the true relative frequencies in the Thai population would differ much from our estimation.

On the basis of the subtype frequencies from this study, it would be prudent to look for SCA3 first, followed by SCA1 and -2 in a Thai patient with presumed diagnosis of SCA unless the clinical features suggest otherwise, or when a "panel" test is used. In the latter case, these three common types should be included in the panel as a minimum. As for the population inference, it is not possible to infer the direction of migration or the timing from this study. Nevertheless, we believe this provides enough justification for further, more definitive molecular studies to look into the link between the populations on the Indochina peninsula and the Indians.

ACKNOWLEDGEMENTS

This study was supported by the annual research grant of the Faculty of Medicine Ramathibodi Hospital, Mahidol University, Thailand. The map in Figure 1a was created using the R packages MAPS 2.0–40 and MAPDATA 2.0–23.

- 1 Manto, M.-U. The wide spectrum of spinocerebellar ataxias (SCAs). *Cerebellum* **4**, 2–6 (2005).
- 2 Dueñas, A. M., Goold, R. & Giunti, P. Molecular pathogenesis of spinocerebellar ataxias. *Brain* **129**, 1357–1370 (2006).
- 3 Schöls, L., Bauer, P., Schmidt, T., Schulte, T. & Riess, O. Autosomal dominant cerebellar ataxias: clinical features, genetics, and pathogenesis. *Lancet Neurol.* **3**, 291–304 (2004).
- 4 Martins, S., Calafell, F., Gaspar, C., Wong, V. C., Silveira, I., Nicholson, G. A. *et al.* Asian origin for the worldwide-spread mutational event in Machado-Joseph disease. *Arch. Neurol.* **64**, 1502–1508 (2007).
- 5 Orr, H. T., Chung, M. Y., Banfi, S., Kwiatkowski, T. J. Jr, Servadio, A., Beaudet, A. L. *et al.* Expansion of an unstable trinucleotide CAG repeat in spinocerebellar ataxia type 1. *Nat. Genet.* **4**, 221–226 (1993).
- 6 Pulst, S. M., Nechiporuk, A., Nechiporuk, T., Gispert, S., Chen, X. N., Lopes-Cendes, I. *et al.* Moderate expansion of a normally biallelic trinucleotide repeat in spinocerebellar ataxia type 2. *Nat. Genet.* **14**, 269–276 (1996).
- 7 Kawaguchi, Y., Okamoto, T., Taniwaki, M., Aizawa, M., Inoue, M., Katayama, S. *et al.* CAG expansions in a novel gene for Machado-Joseph disease at chromosome 14q32.1. *Nat. Genet.* **8**, 221–227 (1994).
- 8 Michlewski, G. & Krzyzosiak, W. J. Molecular architecture of CAG repeats in human disease related transcripts. *J. Mol. Biol.* **340**, 665–679 (2004).
- 9 David, G., Abbas, N., Stevanin, G., Dürr, A., Yvert, G., Cancel, G. *et al.* 1997. Cloning of the SCA7 gene reveals a highly unstable CAG repeat expansion. *Nat. Genet.* **17**, 65–70 (1997).
- 10 Li, S.-H., McInnis, M. G., Margolis, R. L., Antonarakis, S. E. & Ross, C. A. Novel triplet repeat containing genes in human brain: cloning, expression, and length polymorphisms. *Genomics* **16**, 572–579 (1993).
- 11 Koide, R., Ikeuchi, T., Onodera, O., Tanaka, H., Igarashi, S., Endo, K. *et al.* Unstable expansion of CAG repeat in hereditary dentatorubral-pallidoluysian atrophy (DRPLA). *Nat. Genet.* **6**, 9–13 (1994).
- 12 R Development Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing (2008) (Available online at URL: <http://www.R-project.org>).

- 13 Agresti, A. & Coull, B. A. Approximate is better than 'exact' for interval estimation of binomial proportions. *Am. Stat.* **52**, 119–126 (1998).
- 14 Bureau of Registry Administration, Department of Provincial Administration. (Notification of the Bureau of Central Registry, Department of Provincial Administration, on the number of citizens throughout the Kingdom divided into the Bangkok Metropolis and each other province, according to the Citizen Registry documentation as of 31st December, B. E. 2550 (2007). *Royal Gazette* **125** (suppl 25 Ngor), 7–10 (2008) (in Thai).
- 15 Tang, B., Liu, C., Shen, L., Dai, H., Pan, Q., Jing, L. *et al.* Frequency of SCA1, SCA2, SCA3/MJD., SCA6, SCA7, and DRPLA CAG trinucleotide repeat expansion in patients with hereditary spinocerebellar ataxia from Chinese kindreds. *Arch. Neurol.* **57**, 540–544 (2000).
- 16 Lau, K. K., Lam, K., Shiu, K. L., Au, K. M., Tsoi, T. H., Chan, A. Y. W. *et al.* Clinical features of hereditary spinocerebellar ataxia diagnosed by molecular genetic analysis. *Hong Kong Med. J.* **10**, 255–259 (2004).
- 17 Soong, B. W., Lu, Y. C., Choo, K. B. & Lee, H. Y. Frequency analysis of autosomal dominant cerebellar ataxias in Taiwanese patients and molecular characterization of spinocerebellar ataxia type 6. *Arch. Neurol.* **58**, 1105–1109 (2001).
- 18 Tsai, H.-F., Liu, C.-S., Leu, T.-M., Wen, F.-C., Lin, S.-J., Liu, C.-C. *et al.* Analysis of trinucleotide repeats in different SCA loci in spinocerebellar ataxia patients and in normal population of Taiwan. *Acta Neurol. Scand.* **109**, 355–360 (2004).
- 19 Saleem, Q., Choudhry, S., Mukerji, M., Bashyam, L., Padma, M. V., Chakravarthy, A. *et al.* Molecular analysis of autosomal dominant hereditary ataxias in the Indian population: high frequency of SCA2 and evidence for a common founder mutation. *Hum. Genet.* **106**, 179–187 (2000).
- 20 Basu, P., Chattopadhyay, B., Gangopadhyay, P. K., Mukherjee, S. C., Sinha, K. K., Das, S. K. *et al.* Analysis of CAG repeats in SCA1, SCA2, SCA3, SCA6, SCA7 and DRPLA loci in spinocerebellar ataxia patients and distribution of CAG repeats at the SCA1, SCA2 and SCA6 loci in nine ethnic populations of eastern India. *Hum. Genet.* **106**, 597–604 (2000).
- 21 Sinha, K. K., Worth, P. F., Jha, D. K., Sinha, S., Stinton, V. J., Davis, M. B. *et al.* Autosomal dominant cerebellar ataxia: SCA2 is the most frequent mutation in eastern India. *J. Neurol. Neurosurg. Psychiatr.* **75**, 448–452 (2004).
- 22 Krishna, N., Mohan, S., Yashavantha, B. S., Rammurthy, A., Kiran Kumar, H. B., Mittal, U. *et al.* SCA 1, SCA 2 & SCA 3/MJD mutations in ataxia syndromes in southern India. *Indian J. Med. Res.* **126**, 465–470 (2007).
- 23 Srivastava, A. K., Choudhry, S., Gopinath, M. S., Roy, S., Tripathi, M., Brahmachari, S. K. *et al.* Molecular and clinical correlation in five Indian families with spinocerebellar ataxia 12. *Ann. Neurol.* **50**, 796–800 (2001).
- 24 Chakravarty, A. & Mukherjee, S. C. Autosomal dominant cerebellar ataxias in ethnic Bengalees in West Bengal—an Eastern Indian state. *Acta Neurol. Scand.* **105**, 202–208 (2002).
- 25 Zhao, Y., Tan, E. K., Law, H. Y., Yoon, C. S., Wong, M. C. & Ng, I. Prevalence and ethnic differences of autosomal-dominant cerebellar ataxia in Singapore. *Clin. Genet.* **62**, 478–481 (2002).
- 26 Su, B., Xiao, C., Deka, R., Seielstad, M. T., Kangwanpong, D., Xiao, J. *et al.* Y chromosome haplotypes reveal prehistorical migrations to the Himalayas. *Hum. Genet.* **107**, 582–590 (2000).
- 27 Basu, A., Mukherjee, N., Roy, S., Sengupta, S., Banerjee, S., Chakraborty, M. *et al.* Ethnic India: a genomic view, with special reference to peopling and structure. *Genome Res.* **13**, 2277–2290 (2003).
- 28 Cordaux, R., Weiss, G., Saha, N. & Stoneking, M. The northeast Indian passageway: a barrier or corridor for human migrations? *Mol. Biol. Evol.* **21**, 1525–1533 (2004).
- 29 Kashyap, V. K., Guha, S., Sitalaximi, T., Bindu, G. H., Hasnain, S. E. & Trivedi, R. Genetic structure of Indian populations based on fifteen autosomal microsatellite loci. *BMC Genet.* **7**, 28 (2006).
- 30 Takano, H., Cancel, G., Ikeuchi, T., Lorenzetti, D., Mawad, R., Stevanin, G. *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.* **63**, 1060–1066 (1998).
- 31 Chu, J. Y., Huang, W., Kuang, S. Q., Wang, J. M., Xu, J. J., Chu, Z. T. *et al.* Genetic relationship of populations in China. *Proc. Natl Acad. Sci. USA* **95**, 11763–11768 (1998).
- 32 Tajima, A., Pang, I.-H., Fucharoen, G., Fucharoen, S., Matsuo, M., Tokunaga, K. *et al.* Three major lineages of Asian Y chromosomes: implications for the peopling of east and southeast Asia. *Hum. Genet.* **110**, 80–88 (2002).
- 33 Hill, C., Soares, P., Mormina, M., Macaulay, V., Meehan, W., Blackburn, J. *et al.* Phylogeography and ethnogenesis of aboriginal Southeast Asians. *Mol. Biol. Evol.* **23**, 2480–2491 (2006).
- 34 Wyatt, D. K. *Thailand: A Short History* (Trasvin Publications, Chiang Mai, 1984).
- 35 Bellwood, P. The origins and dispersals of agricultural communities in Southeast Asia. In *Southeast Asia: From Prehistory to History* (eds Glover, I. & Bellwood, P.) 21–40 (RoutledgeCurzon, New York, 2004).
- 36 Bellina, B. & Glover, I. The archaeology of early contact with India and the Mediterranean world, from the fourth century BC to the fourth century AD. In *Southeast Asia: From Prehistory to History* (eds Glover, I. & Bellwood, P.) 68–88 (RoutledgeCurzon, New York, 2004).