RESEARCH NOTE

Analysis of single nucleotide polymorphisms of *PRNP* gene in twenty-four ethnic groups of India

MAINAK SENGUPTA[1], AMRITA CHAKRABORTY[1], INDIAN GENOME VARIATION CONSORTIUM[2] and KUNAL RAY[1]*

[1]*Molecular and Human Genetics Division, Indian Institute of Chemical Biology (a unit of CSIR), Kolkata 700 032, India*
[2]*Nodal Laboratory, Institute of Genomics and Integrative Biology (a unit of CSIR), New Delhi 110 007, India*

Introduction

Human prion protein (PRNP), a copper binding sialoglycoprotein, is the causative agent for transmissible spongiform encephalopathies (TSEs), a group of neurodegenerative diseases that are generally associated with aggregation of amyloid plaques within the central nervous system disrupting the normal tissue structure (Liberski *et al*. 2001; Aguzzi *et al*. 2008). While more than 20 PRNP pathogenic mutations have been reported in patients, polymorphisms in the gene have been suggested to influence the incidence of the diseases with a Met129Val polymorphism (rs1799990) being the most important among them (Soldevila *et al*. 2006). This particular variant has been found to be associated with different prion related diseases as well as cognitive behaviour and long-term memory. Here, as a part of the effort from Indian Genome Variation Consortium (IGVC), we attempt to establish the baseline allele and genotype frequency of four *PRNP* SNPs including rs1799990 among 24 ethnic groups representing the Indian population. The allele and genotype frequency of rs1799990 is found to be different among different ethnic groups. This pilot study would serve as a platform for future epidemiological works with respect to PRNP SNPs in India.

Although the exact function of PRNP is not elucidated till date, experimental data suggests its involvement in the removal of reactive oxygen species in neurons, leading to neuroprotection. It has also been shown that the prion protein might influence neuronal Cu uptake (Brown 1999) or incorporation into enzymes such as Cu/Zn-superoxide dismutase (Brown and Besinger 1998). PRNP itself is hypothesized to act as a superoxide dismutase (Wong *et al*. 2000a). About 85% cases of human prion diseases have been found to occur sporadically by an unknown mechanism whereas approximately 10%–15% are inherited (Yu *et al*. 2004). *PRNP* SNPs are thought to play a key role in the disease incidence. At codon 129, homozygosity either in methionine or valine has been found to increase the risk for sporadic, iatrogenic and variant Creutzfeldt–Jakob disease (CJD) among Caucasians while the heterozygosity imparts protection (Croes *et al*. 2004; Yu *et al*. 2004). In fact, all patients of variant CJD were found to be 129 MM homozygotes (Brown 2001). Again, prevalence of heterozygous genotype among the survivors of Kuru; another human prion disease, implicates the protective aspect of the particular genotype (Mead *et al*. 2003). The putative functionality of the SNP is yet to be unearthed. Although the copper-binding ability and the level of superoxide dismutase activity do not seem to be altered with the variations in the genotype, studies by Wong *et al*. (2000b) revealed that copper binding results in different allele-dependent conformations and thus structures. The *PRNP* gene is also a candidate for antioxidant activity-mediated association with cognitive ageing, impairment, dementia, brain morphology, Alzheimer's disease (Kachiwala *et al*. 2005), neural plasticity (Buchmann *et al*. 2008) and long term memory (Papassotiropoulos *et al*. 2005). It has been found that Met/Met and Met/Val individuals can recall 17% more words than Val/Val individuals at 24 h following learning. Link of the PRNP codon 129 polymorphism with individual differences in cognitive decline has been reported (Del Bo *et al*. 2003). Rujescu *et al*. (2002) suggested an association of methionine homozygosity at codon 129 with white matter tissue reduction and enlargement of CSF compartments in right-handed male schizophrenic patients and controls. However, no direct association between schizophre-

*For correspondence. E-mail: kunalray@gmail.com; kray@iicb.res.in.
[Sengupta M., Chakraborty A., Indian Genome Variation Consortium and Ray K. 2010 Analysis of single nucleotide polymorphisms of *PRNP* gene in twenty-four ethnic groups of India. *J. Genet.* **89**, 247–251]

Keywords. prion; *PRNP* gene; single nucleotide polymorphism (SNP).

nia and the Met129Val polymorphism has been found either among Caucasians (Rujescu *et al*. 2002) or Han Chinese (Tsai *et al*. 2001). Again, Met/Val heterozygotes were significantly over-represented in 39 individuals with primary progressive aphasia (PPA), a rare condition of unknown cause, compared to more than 400 controls (Li *et al*. 2005). Thus the prospective role of variants of PRNP in overall neuroprotection as well as neurodevelopment underscores the delineation of the genotype and allele frequency data of the common polymorphisms of *PRNP* gene in different world populations especially that of Met129Val polymorphism. However, till date, there is no information on the overall genotypic distribution of Met129Val from India except for one report (Mead *et al*. 2003) on 'Non-UP' ($n = 88$) and 'UP' ($n = 64$) Indians. In this report Mead *et al*. (2003) have presented frequency of M129V polymorphism in different population groups of the world showing reducing cline towards East Asia. Here, we aim to establish the baseline genotype and allele frequency of four common SNPs of *PRNP* gene including the Met129Val change among 552 individuals representing 24 ethnic populations belonging to four linguistic groups of India as part of the effort taken up by the IGVC.

Materials and methods

Study populations

This study was conducted on DNA samples collected from 552 unrelated individuals from 24 ethnic populations inhabiting six different geographical regions (north, northeast, east, south, west and central) of mainland India belonging to four linguistic families of the Indian population, namely Indo–European, Dravidian, Tibeto–Burman and Austro–Asiatic. They were selected from a larger pool of 1871 individuals from 55 ethnicities based on system structure analysis of the populations indicating levels of admixture and thus identifying a reduced number of reference populations for future disease-association studies (Indian Genome Variation Consortium 2008). Majority of the F_{ST} values among these 55 populations were significantly greater than zero ($P < 0.05$) indicating population differentiation but the extent of differentiation overall was low (Indian Genome Variation Consortium 2008). The study was approved by internal review committee on research using humans as per the regulations of the Indian Council of Medical Research. While selecting the sample donor, a uniform bar-coded, detailed questionnaire was developed, containing information pertaining to family history of diseases of the donor and no diseased individual was enrolled in the study (Indian Genome Variation Consortium 2005).

Collection of blood samples and genomic DNA preparation

For the study of analysis of SNPs in the Indian population as a part of IGVC, 5–10 mL of blood from unrelated individuals was drawn with appropriate informed consent. DNA was isolated from each individual using salting-out method (Johns and Paulus-Thomas 1989) using sodium perchlorate followed by isopropanol precipitation. Identification of populations as well as collection of samples were carried out as described by Indian Genome Variation Consortium (2005).

Genotyping of SNPs

The SNPs were selected on the basis of heterozygosity, potential functionality and coverage of the gene. The relevant information about the selected SNPs is furnished in table 1. Genotyping of the selected SNPs in the study cohort was done using the homogenous MassExtend (hME) assay run on the Sequenom MassARRAY system (San Diego, USA) (http://www.tcgaresearch.org/service_massarray.htm) based on allele-specific primer extension followed by MALDI-TOF mass spectrometry (Indian Genome Variation Consortium 2008), thus ensuring the specificity and sensitivity. Prior to validation in the entire sample pool, the polymorphic status of the selected SNPs were checked in an initial discovery panel as well as individual populations, and went through a series of strict quality controls. This part of the work was conducted at Institute for Genomics and Integrative Biology (IGIB), New Delhi, the nodal laboratory of IGVC, and The Centre for Genomic Application (TCGA), New Delhi (Indian Genome Variation Consortium 2008).

Statistical and in silico analyses

Analysis of the SNP genotypes for the study was done by the TYPERanalyzer software of the MassARRAY (http://www.sequenom.com/Genetic-Analysis/Systems/Hardware-and-Options) system for the SNPs typed in SEQUENOME platform. Allele frequencies and heterozygosities at each variant site were computed by the genotype-counting method. Coefficient of pair-wise LD (linkage disequilibrium; r^2) and haplotype analysis for the SNPs were done through HAPLOVIEW (http://www.broad.mit.edu/mpg/haploview/) software package. Hardy–Weinberg equilibrium (HWE) for each SNP in the individual populations was calculated by exact tests for HWE using the UtilExec HWE2 software (http://www.genemapping.cn/util.htm).

Results and discussion

A total of 552 individuals were screened for four SNPs in the *PRNP* gene viz. rs2756271, rs6116471, rs6116474 and rs1799990 to determine the allele and genotype frequencies of each SNP within different Indian subpopulations. Our results revealed that all four SNPs were polymorphic among Indians thereby suggesting their informative nature as molecular markers for any future epidemiological studies in the entire Indian spectrum (table 2; see tables 1–3 in electronic supplementary material at http://www.ias.ac.in/jgenet/). HWE did hold well in majority of the populations for each SNP. However, two of the Tibeto–Burman ethnic groups viz. TB-NE-LP1 and TB-N-IP1 did not follow HWE for rs1799990 and rs6116471, respectively (table 2 and see tables 1–3 in

Table 1. Relevant information of the four *PRNP* SNPs selected in this study.

SNP	Nucleotide change (amino acid change)	Position	Spacing (in bp)	Ancestral allele
rs2756271	4605262 C/T	5′- to gene	1	C
rs6116471	4609534 A/G	Intron 1	4272	A
rs6116474	4617207 T/C	Intron 1	7673	T
rs1799990	4620251 A/G (Met129Val)	Exon 2	3044	Not available

Nucleotide change coordinates are represented by the contig positions of the SNPs. In the 'spacing' column, the distance of a particular SNP (in bp) is given with respect to the immediates next SNP.

Table 2. Allele and genotype frequenices of the non-synonymous SNP rs1799990 (Met129Val).

Population group (n)	Allele frequency (A)	Genotype frequency			HW *P* values	Standard errors of the allele frequency estimates
		AA	AG	GG		
IE-E-IP1 (22)	0.82	0.68	0.27	0.05	0.538	0.058
IE-E-LP2 (22)	0.89	0.77	0.23	0.00	1	0.047
IE-E-LP4 (9)	0.83	0.67	0.33	0.00	1	0.089
IE-NE-IP1 (22)	0.82	0.64	0.36	0.00	1	0.058
IE-NE-LP1 (23)	0.78	0.57	0.43	0.00	0.538	0.061
IE-N-IP2 (23)	0.87	0.74	0.26	0.00	1	0.049
IE-N-LP1 (23)	0.7	0.52	0.35	0.13	0.365	0.068
IE-N-LP5 (23)	0.76	0.57	0.39	0.04	1	0.063
IE-N-LP9 (23)	0.78	0.61	0.35	0.04	1	0.061
IE-N-SP4 (23)	0.72	0.52	0.39	0.09	1	0.066
IE-W-LP1 (23)	0.89	0.78	0.22	0.00	1	0.046
IE-W-LP2 (23)	0.89	0.78	0.22	0.00	1	0.046
IE-W-LP3 (23)	0.66	0.41	0.5	0.09	1	0.069
IE-W-LP4 (23)	0.61	0.26	0.70	0.04	1	0.071
DR-C-IP2 (11)	0.86	0.73	0.27	0.00	1	0.073
DR-S-IP4 (23)	0.83	0.74	0.17	0.09	0.098	0.055
DR-S-IP2 (23)	0.87	0.74	0.26	0.00	1	0.049
DR-S-IP3 (23)	0.76	0.57	0.39	0.04	1	0.062
AA-C-IP5 (7)	0.79	0.71	0.14	0.14	0.23	0.108
AA-E-IP3 (21)	0.55	0.24	0.62	0.14	0.394	0.77
TB-NE-LP1 (22)	0.89	0.86	0.05	0.09	0.008*	0.047
TB-N-IP1 (21)	0.76	0.57	0.38	0.05	1	0.066
TB-N-SP1 (23)	0.78	0.61	0.35	0.04	1	0.061
OG-W-IP (23)	0.87	0.79	0.17	0.04	0.31	0.049
ASW (53)	0.566	0.34	0.453	0.208	0.582	0.048
CEU (113)	0.655	0.442	0.425	0.133	0.535	0.031
CHB (84)	0.952	0.905	0.095	0.00	1	0.016
CHD (85)	0.988	0.976	0.024	0.00	1	0.009
GIH (88)	0.767	0.58	0.375	0.045	0.772	0.031
JPT (86)	0.983	0.965	0.035	0.00	1	0.009
LWK (90)	0.706	0.456	0.500	0.044	0.076	0.034
MEX (50)	0.610	0.380	0.46	0.16	0.774	0.049
MKK (143)	0.650	0.490	0.322	0.189	0.0007*	0.028

Table 2 *(contd)*

Population group (n)	Allele frequency (A)	Genotype frequency			HW *P* values	Standard errors of the allele frequency estimates
		AA	AG	GG		
TSI (87)	0.615	0.391	0.448	0.161	0.652	0.037
YRI (113)	0.668	0.442	0.451	0.106	1	0.031

n, Number of individuals where genotyping was successful for this SNP. Originally each Indian ethnic group consisted of 23 individuals. The frequency of the allele majorly represented (i.e. major allele) in maximum Indian populations is furnished. The Indian ethnic group TB-NE-LP1 and the HapMap population MKK are not following HWE for this SNP (marked by astricks). For the nomenclature of the Indian ethnic groups, refer to Indian Genome Consortium (2008). Briefly, Indo–European, IE; Dravidian, DR; Austro-Asiatic, AA; Tibeto-Burman, TB; represents the four major linguistic groups of India. OG, Outgroup; E, east; W, west; N north; NE, northeast; S, south; C, central, demarcate the geographical locations of the ethnic groups in India. IP, isolated population; LP, large population; SP, special population, represent population types. The HapMap populations are as follows: ASW, African ancestry in Southwest USA; CEU, Utah residents with northern and western European ancestry from the CEPH collection; CHB, Han Chinese in Beijing, China; CHD, Chinese in Metropolitan Denver, Colorado; GIH, Gujarati Indians in Houston, Texas; JPT, Japanese in Tokyo, Japan; LWK, Luhya in Webuye, Kenya; MEX, Mexican ancestry in Los Angeles, California; MKK, Maasai in Kinyawa, Kenya; TSI, Toscans in Italy; YRI, Yoruba in Ibadan, Nigeria.

electronic supplementary material). Considering the association of rs1799990 with different prion related disorders, we were primarily interested in assessing the genotype and allele frequencies of Met129Val in the different Indian ethnic groups (table 2). Our results showed that homozygous methionine constituted the major proportion of the genotypes among 21 out of 24 Indian ethnic groups where as in case of three ethnic groups (IE-W-LP3, IE-W-LP4 and AA-E-IP3), heterozygous genotype prevailed. The heterozygous genotype was least represented among TB-NE-LP1, a Tibeto–Burman large population group, residing in the north eastern part of the country. The Val129 was the minor allele in all the populations studied but the distribution differed considerably among the different populations ranging from 0.11 (in IE-E-LP2, IE-W-LP1, IE-W-LP2 and TB-NE-LP1) to 0.45 (in AA-E-IP3). Comparison with the HapMap populations (table 2) revealed that in general the Indian data were similar to that of the Gujaratis settled in USA (GIH) but was considerably different from the rest of the Asian population viz. Han Chinese from Bejing (CHB), Chinese settled in USA (CHD) and Japanese from Tokyo (JPT). It is worthwhile to mention here that the ethnic group IE-N-SP4 from Uttar Pradesh (UP) presented the same rare allele frequency (0.28) to that of the UP Indians mentioned by Mead *et al*. (2003). The detailed allele and genotype analysis for rs2756271, rs6116471 and rs6116474 are given in tables 1–3 in electronic supplementary material. TGTA was found to be the major haplotype in majority of the ethnic groups as well as when they were grouped in 'clusters' as described previously (figure 3 in Indian Genome Variation Consortium 2008) (see table 4 in electronic supplementary material). Pair-wise linkage disequilibrium assessment revealed no LD between the selected SNPs in any of the 'clusters' (see figure 1 in electronic supplementary material). This small study thus draws the baseline frequency for four SNPs in *PRNP* gene among the 24 ethnic groups scattered throughout the country, belonging to different linguistic groups, castes and creeds. This could serve as a platform data for future epidemiological works regarding neural development and plasticity with respect to *PRNP* gene.

Acknowledgements

This study has been supported by the Council of Scientific and Industrial Research (CSIR), Govt. of India (Grant no. CMM 0016). MS is supported by predoctoral fellowship from University Grant Commission (UGC).

References

Aguzzi A., Sigurdson C. and Heikenwailder M. 2008 Molecular mechanisms of priori pathogenesis *Annu. Rev. Pathol.* **3**, 11–40.

Brown D. R. 1999 Prion protein expression aids cellular uptake and veratridine-induced release of copper. *J. Neurosci. Res*. **58**, 717–725.

Brown D. R. 2001 BSE did not cause variant CJD: an alternative cause related to post-industrial environmental contamination. *Med. Hypotheses* **57**, 555–560.

Brown D. R. and Besinger A. 1998 Prion protein expression and superoxide dismutase activity. *Biochem. J*. **334**, 423–429.

Buchmann A., Mondadori C. R., Hanggi J., Aerni A., Vrticka P., Luechinger R. *et al*. 2008 Prion protein M129V polymorphism affects retrieval-related brain activity. *Neuropsychologia* **46**, 2389–2402.

Croes E. A., Alizadeh B. Z., Bertoli-Avella A. M., Rademaker T., Vergeer-Drop J., Dermaut B. *et al*. 2004 Polymorphisms in the prion protein gene and in the doppel gene increase susceptibility for Creutzfeldt-Jakob disease. *Eur. J. Hum. Genet*. **12**, 389–394.

Del Bo R., Comi G. P., Giorda R., Crimi M., Locatelli F., Martinelli-Boneschi F. *et al*. 2003 The 129 codon polymorphism of the prion protein gene influences earlier cognitive performance in Down syndrome subjects. *J. Neurol*. **250**, 688–692.

Indian Genome Variation Consortium 2005 The Indian Genome Variation database (IGVdb): a project overview. *Hum. Genet*. **118**, 1–11

Indian Genome Variation Consortium 2008 Genetic landscape of the people of India: a canvas for disease gene exploration. *J. Genet*. **87**, 3–20.

Johns M. and Paulus-Thomas J. 1989 Purification of human genomic DNA from whole blood using sodium perchlorate in place of phenol. *Anal. Biochem*. **180**, 276–278.

Kachiwala S. J., Harris S. E., Wright A. F., Hayward C., Starr J. M., Whalley L. J. *et al*. 2005 Genetic influences on oxidative stress and their association with normal cognitive ageing. *Neurosci. Lett*. **386**, 116–120.

Liberski P. P., Bratosiewicz J. Waliś A., Kordek R., Jeffrey M. and Brown P. 2001 A special report. Prion protein (Prp)–amyloid plaques in the transmissible spongiform encephalopathies, or prion diseases revisited. *Folia Neuropathol*. **39**, 217–235.

Li X., Rowland L. P., Mitsumoto H., Przedborski S., Bird T. D., Schellenberg G. D. *et al*. 2005 Prion protein codon 129 genotype prevalence is altered in primary progressive aphasia. *Ann. Neurol*. **58**, 858–864.

Mead S., Stumpf M. P., Whitfield J., Beck J. A., Poulter M., Campbell T. *et al*. 2003 Balancing selection at the prion protein gene consistent with prehistoric kurulike epidemics. *Science* **300**, 640–643.

Papassotiropoulos A., Wollmer M. A., Aguzzi A., Hock C., Nitsch R. M. and de Quervain D. J. 2005 The prion gene is associated with human long-term memory. *Hum. Mol. Genet*. **14**, 2241–2246.

Rujescu D., Meisenzahl E. M., Giegling I., Kirner A., Leinsinger G., Hegerl U. *et al*. 2002 Methionine homozygosity at codon 129 in the prion protein is associated with white matter reduction and enlargement of CSF compartments in healthy volunteers and schizophrenic patients. *Neuroimage* **15**, 200–206.

Soldevila M., Andrés A. M., Ramírez-Soriano A., Marquès-Bonet T., Calafell F., Navarro A. and Bertranpetit J. 2006 The prion protein gene in humans revisited: lessons from a worldwide resequencing study. *Genome Res*. **16**, 231–239

Tsai M. T., Su Y. C., Chen Y. H. and Chen C. H. 2001 Lack of evidence to support the association of the human prion gene with schizophrenia. *Mol. Psychiatry* **6**, 74–78.

Wong B. S., Pan T., Liu T., Li R., Gambetti P. and Sy M.S. 2000a Differential contribution of superoxide dismutase activity by prion protein *in vivo. Biochem. Biophys. Res. Commun*. **273**, 136–139.

Wong N. K., Renouf D. V., Lehmann S. and Hounsell E. F. 2000b Glycosylation of prions and its effects on protein conformation relevant to amino acid mutations. *J. Mol. Graph*. **18**, 126–134.

Yu S. L., Jin L., Sy M. S., Mei F. H., Kang S. L., Sun G. H. *et al*. 2004 Polymorphisms of the PRNP gene in Chinese populations and the identification of a novel insertion mutation. *Eur. J. Hum. Genet*. **12**, 867–870.

Received 10 November 2009, in revised form 26 January 2010; accepted 28 February 2010

Published on the Web: 5 August 2010