Discussion

SCAs are a heterogeneous group of inherited neurodegenerative disorder, which the cerebellum is primarily affected. Most of them are associated with unstably expanded tandem repeats in pathogenic genes. In this study, we analyzed the frequency distributions of 9 SCAs in Taiwanese, including recently identified SCA8, SCA10, SCA12, and SCA17, which were not surveyed before 2004. Our data showed that the genes responsible for 33% of dominant SCA in Taiwanese remain to be determined, similar to the previous studies in which pathogenic genes were unknown for approximately 20-40% of SCAs (Moseley et al. 1998; Takano et al. 1998; Soong et al. 2001). The prevalences of SCA1 (4%), SCA3 (36%), SCA8 (0%), and DRPLA (0%) in our study were similar to those reported: SCA1 5%, SCA3 47-48%, SCA8 0%, and DRPLA 0-1% in Taiwanese and Chinese (Soong et al. 2001; Tang et al. 2000). Although the SCA2 alleles were reported as 6% in Chinese (Tang et al. 2000) and 11% in Taiwanese (Soong et al. 2001), no SCA2 allele was detected in our study, probably due to differences in the number of sampled chromosomes, sampled individuals and/or sampled populations. In addition, similar prevalence and frequency distributions of large normal alleles of SCA1, SCA3, SCA6, and DRPLA (**Table 4**) indicate that mutations of these genes might origin from the common ancestor, while the larger normal alleles have more chance to mutate (Leeflang et al. 1995).

Abnormal expansions ranging from 75 to 92 repeats of *SCA8* alleles were detected in four patients with PD (**Table 5**). The expanded *SCA8* alleles ranging from 68 (Stevanin et al. 2000) to more than 1000 repeats (Ikeda et al.

2004) were characterized in familial and sporadic ataxia patients. In addition, the larger expanded allele has been reported in patients with psychiatric disorders and other neurological disorders, such as PD, Friedreich's ataxia and AD (Izumi et al. 2003; Juvonen et al. 2000; Sobrido et al. 2001; Stevanin et al. 2000; Tazon et al. 2002; Vincent et al. 2000; Worth et al. 2000). *SCA8* transcripts distribute throughout the brain in spite of details not being verified (Benzow and Koob 2002). Additionally, no survey was conducted to investigate the interaction between *SCA8* and other gene transcripts. It cannot be excluded that *SCA8* transcripts could interact with products of disease genes on the pathogenic pathways of PD, and *vice versa*. Thus, PD phenotypes may be the manifestation of expanded *SCA8* alleles in these four PD patients. Although no large expanded allele was found in our normal controls, *SCA8* mutation has been found in rare instances in healthy control individuals, suggesting that association between expansion mutation at *SCA8* and SCA is not straightforward. However, in most instances the expanded allele closely correlates with SCA8 disease, and rarely identical haplotypes existed between affected and normal chromosomes (Ikeda et al. 2004), implying CTG expansion should play a predominant role in pathogenesis.

Abnormal CAG expansions in the SCA17 *TBP* gene were detected in a PD patient (46 repeats) (**Table 5**). The *TBP* alleles range from 30 to 43 repeats in our normal controls. In the general population, the reported normal *TBP* alleles range in size from 25 to 44 repeats (Gostout et al. 1993; Zuhlke et al. 2001) and the expanded alleles in patients with ataxia from 43 to 63 repeats (Koide et al. 1999; Fujigasaki et al. 2001a; Nakamura et al. 2001; Silveira et al. 2002; Zuhlke et al. 2003), but with reduced penetrance (Zuhlke et al.

2003). Recently Hungtington's disease-like phenotype due to CAG repeat expansion in the *TBP* gene was reported (Stevanin et al. 2003). In addition, non-cerebellar symptoms may develop in some specific SCA types (Gwinn-Hardy et al. 2000; Gwinn-Hardy et al. 2001; O'Hearn et al. 2001), as well as the PD manifestation of *SCA8* mutation described in this study. As *TBP* is a critical factor regulating the initiation of transcription and ubiquitously expressed in all cells, the 46-repeat allele in PD patient may be linked to the neurological manifestations observed.

The overlapping organization and evolutionary conservation of *SCA8* and *KLHL1* genes (Nemes et al. 2000; Benzow and Koob 2002) hint the close relation of them. KLHL1 protein is the homologue of *Drosophila* kelch protein, functioning as maintaining cytoskeleton order, and it is recently reported that KLHL1 enhances neurite outgrowth mediated via the interaction with glycogen synthase kinase 3β (Seng et al. 2006) and modulates P/Q-type calcium channel function (Aromolaran et al. 2007). Nevertheless, no studies focused on the *trans* antisense effect of *SCA8* to *KLHL1*. We co-expressed *KLHL1* and *SCA8* carrying various numbers of CTG repeats to investigate whether *SCA8* regulates the expression of *KLHL1* and whether this regulation is mediated in a CTG repeats-dependent manner. The expression of *KLHL1* was reduced as expected when *SCA8* was expressed in the same cells (29 ~ 55% of the levels of *KLHL1*-EGFP gene, **Figure 4B**), suggesting that *SCA8* is a negative regulator of *KLHL1*. The synthesis of overlapping transcripts potentially interferes with the RNA processing at different levels. Epigenetic modification, transcriptional interference, impaired splicing, or RNA export as well as mechanisms triggered by double-stranded RNA such as RNA editing and RNA

interference/micro-RNA synthesis may represent consequences of antisense transcription (Werner and Berdal 2005). Among the identified natural antisense transcripts, the 5' end overlapping *SCA8*/*KLHL1* organization is similar to the relation between homeobox-containing transcription factor *Msx1* gene and its antisense (*Msx1*-AS) compartment, while non-coding *Msx1*-AS RNA is complementary to region extending from 3' end of exon 2 to the middle of intron 1 (more than 2 kb) (Blin-Wakkach et al. 2001). Although the precise mechanism for *Msx1* sense–AS RNA interactions remains unclear, the nature that complementary region extends across intron 1 and exon 2 implying the duplex RNA might be an obstacle to splicing. However, the overlapping region of *SCA8*/*KLHL1* does not comprise the exon-intron boundary and the levels of both *KLHL1* and *SCA8* transcripts were not significantly reduced by RT-PCR and gel semi-quantitation (洪, 2005), the repression of *KLHL1* by *SCA8* was likely mediated through post-transcriptional interferences such as nuclear retention and translational blockage.

SCA8 was first defined as a non-coding CUG-expanded RNA disease (Koob et al. 1999). Although the precise transcription start site remains to be determined, *SCA8* has been proven to be an intact transcript which have up to six exons and alternative polyadenylation sites. But sequence analysis performed formerly revealed no significant translatable ORFs existing in any of splice isoforms (Nemes et al. 2000). The standard translation initiation recruits eIF4F to the 5' m^7GpppN cap of mRNA, followed by the ribosome binding and scanning for the proper region to start translation. On the other hand, some eukaryotic mRNAs and viral RNAs have the ability to recruit the translational machinery to the internal portion of RNAs mediated through

the IRES elements. Judged form the structure of *SCA8* transcripts and the long 5' UTR containing a G/C rich region (Nemes et al. 2000), which could serve as a recognition site of *trans* interacting factors, the translational ability of SCA8 remains to be uncovered. In our previous study, sequence analysis revealed the existence of three small ORFs in the *SCA8* transcripts and *SCA8* RNA was demonstrated to possess the bipartite cap independent IRESs using dicistronic constructs (Lin et al. submitted). Owing to short intervention between ORF1 and ORF2, ORF2 is considered non-translatable through the IRES activity. Thus, GFP tagged ORF1 and ORF3 were transiently transfected into HEK293 cells to verify if indeed the *SCA8* transcripts are translatable. Both GFP-tagged ORF1 and ORF3 are expressed, while ORF3 expressing at much lower level (3% of the cap-dependent expression) and could not be detected on the immunoblotting (**Figure 5**). The strength of IRES is mediated by the arrangement of ORFs (Hennecke et al. 2001) and physiological conditions leading to greatly reduced cap-dependent protein synthesis (Fernandez et al. 2001; Kim and Jang 2002; Sherrill et al. 2004). Therefore, the relatively lower expression level of *SCA8* ORF3 compared to ORF1 might be due to the genomic arrangement, or the robust expression of ORF3 would occur under cellular stresses.

Although the expression level of SCA8 ORF3 was very low in our *in vitro* system, the uneven cytoplasmic distribution of ORF3-23R-EGFP could be observed (**Figure 6**). Since overexpression of EGFP alone resulted in diffusely distribution of green fluorescence, the uneven distribution of ORF3-23R-EGFP should be derived from the content of ORF3. Additionally, with the increase of CTG repeats, more and larger protein aggregates formed (**Figure 6**) and stalled in the stacking layer of polyacrylamide gel (**Figure 7**),

further indicating that CTG-encoding polyleucine is prone to aggregate. Cytoplasmic and intranuclear accumulation of mutant disease proteins are often found in neurodegenerative diseases, including Huntington's disease, SCA1, SCA2 SCA3, SCA7, SCA17, and DRPLA (Scherzinger et al. 1997; Skinner et al. 1997; Koyano et al. 2000; Paulson et al. 1997; Holmberg et al. 1998; Nakamura et al. 2001; Miyashita et al. 1998), but the role of mutant protein aggregates remains controversial. Although so far no diseases are linked to mutant polyleucine proteins, recently it was found that the polyleucine stretch conferred more toxicity than the polyglutamine stretch and showed a high propensity for aggregation distributing closely around the nucleus (Dorsman et al. 2002; Oma et al. 2004). But the threshold of polyleucine to induce cellular abnormality was not determined. Polyglutamine diseases arise when the hydrophobic glutamine tract in the pathogenic protein elongates over a threshold length, which is usually around 30-40 glutamine residues. Besides, expression of 30-residue polyleucine in COS-7 cells resulted in perinuclear as well as cytoplasmic aggregates (Oma et al. 2004), suggesting polyleucine stretch is more hydrophobic than polyglutamine stretch. Consequently, with the same amount of hydrophobic amino acids when beyond the threshold, polyleucine stretch would lead to more aggregates. On the other hand, considering *SCA8* genes with 23 CTG repeats non-pathogenic, the relatively low expression level in our cellular system as well as in neurons of both healthy and affected individuals (Benzow and Koob 2002) may prevent the SCA8 phenotypes in spite of the uneven distributions. With the increase of CTG repeats, long polyleucine stretch renders ORF3 more hydrophobicity, which might in turn serves as a nucleus for aggregate growth, one of widely accepted processes involved in aggregates formation (Harper and Lansbury

1997; Ferrone 1999). However, whether mutant ORF3 aggregates cause cell death was not surveyed in this study, and the role of soluble ORF1 aggregates should also be determined in the future.

In an effort to understand the effect of long CTG repeat tracts on cell survival under specific stresses, we had developed a number of otherwise isogenic HEK293 derived cell lines expressing transcripts with various numbers of CTG repeats. In this study, the finding that isogenic HEK293 cell lines with mutant *SCA8* were more sensitive to high dosage of staurosporine is consistent with that staurosporine enhanced cytotoxicity caused by expanded polyglutamine protein (Cooper et al. 1998). Although staurosporine is an apoptosis-inducing agent, it also affects the cell cycle and causes cell cycle arrest in the G1 phase (Orr et al. 1998), the increased susceptibility to cell death induced by it may be mediated in part through effects on the cell cycle. In addition, isogenic HEK293 cells with mutant *SCA8* were also increased cell susceptibility to the proteasome inhibitor, MG-132. There is evidence that ataxins 1, 3 and 7, are susceptible to be ubiquitinated and targeted by the proteasome for degradation and clearance (Cummings et al. 1999; Matilla et al. 2001; Chai et al. 2004). However, mutant ataxins form misfolded structure, which in turn impare the recognition and degradation process by the proteasome. Accordingly, intra-cellular aggregates produced by the expanded *SCA8* gene might derange proteasome function, although no appropriate methods could detect the aggregate formation now.

We have developed a SCA8 transgenic mouse model in which human *SCA8* gene is expressed under the control of cerebellar Purkinje cell-specific

promoter, pcp2/L7. Although human *SCA8* and mouse analog are evolutionary conserved, the mouse *SCA8* gene does not consist of exons C, B, and A (Benzow and Koob 2002), in which the last one contains the CTG trinucleotide repeats relevant to the pathogenesis of SCA8. Thus, we introduced human *SCA8* with either normal-ranged 23 or expanded 157 CTG/CTA combined repeats into mouse genome to see if the expanded CTG repeats would lead the mice to manifest the SCA8 phenotypes. In our SCA8 transgenic mouse model, however, we cannot detect histological abnormalities in the Purkinje cells (**Figure 13**). Biomarkers of other cell types could be further used to examine the cerebellum, such as glial fibrillary acidic protein (GFAP), a marker for the microglia, used to detect the activated astrocytes. In the previous studies of anti-addictive drug, Ibogaine, GFAP was served as a useful biomarker for neurotoxicity (O'Callaghan et al. 1996; Xu et al. 2000). In addition, abundant reactive microglia have been shown to surround the β-amyloid plaques in the AD brain (McGeer and McGeer 1995), suggesting that microglia also plays a role in the neurodegenerative disorders.

Previously in our study of HEK293 cell model, RNA FISH experiments revealed ribonuclear foci formation in cells carrying expanded 88 and 157 combined repeats (Lin et al. 2007). In our mouse model, we haven't examined the formation of ribonuclear foci, which now are considered as a hallmark of affected muscle in DM1 and DM2 patients. A CTG expansion in the 3'-UTR of *DMPK* gene, sequesters CUGBPs from their normal cellular functions, leading to abnormal RNA splicing of several genes. It is suggested that decreased levels of muscleblind (MBNL) and increased levels of CUGBP are relevant to the formation of CUG-containing ribonuclear foci

and might mediate the RNA-induced toxicity (de Haro et al. 2006). Yet the precise role of the ribonuclear foci remains to be determined since even though ribonuclear foci and the CUG binding protein, MBNL, are colocalized, no pathological phenotypes can be detected in *Drosophila* MBNL null models (Houseley et al. 2005). Since SCA8 and DM1 share molecular similarity and no ribonuclear foci formation has been reported in SCA8 patients or other models, we are going to determine whether the CUG-containing ribonuclear foci could be detected in our SCA8 transgenic mouse model using RNA-FISH strategy.

Although Pcp2/L7 promoter restricts the expression of its downstream gene to Purkinje and retinal bipolar cells, the expression of the *SCA8* transgene in our model can be found in other regions of the brain. Endogenous Pcp2/L7 protein is expressed exclusively in Purkinje cells and retinal bipolar neurons (Nordquist et al. 1988; Oberdick et al. 1988; Berrebi et al. 1991). It has been shown that 3.1-kb stretch of sequence upstream of *Pcp2/L7* restricted the gene expression to Purkinje cells (Vandaele et al. 1991), while the 5' upstream region of *Pcp2/L7* DNA direct the gene activity both in Purkinje and retinal bipolar cells (Oberdick et al. 1990). Further investigations indicates that the nonspecific expression was resulted from the control of a 2.88-kb *Pcp2/L7* DNA fragment upstream from exon 4 (Barski et al. 2000), suggesting that the regulatory element for specific expression might not be included in this region. Since *SCA8* transgene in our model is controlled only by a 0.85-kb fragment upstream from exon 2, it may not be surprising that the expression is not restricted to the cerebellum. Nevertheless, the expression pattern in other tissues should be further determined.

The clasping phenotype has been used as an evaluation criterion to study the disease progression in HD mice (Reddy et al. 1999; van Dellen et al. 2001) in spite of the causing mechanisms remaining unclear. This dystonic posturing of the hind limbs in our SCA8 transgenic mouse model is worth a target for uncovering the disease pathway and drugs screening for relieving the distress.

In this study, gene expression profiles in SCA8 transgenic mice were examined by using microarray and proteomic analyses to identify potential mechanisms responsible for the pathogenesis of SCA8. Although no significantly behavioral abnormality was observed in SCA8 transgenic mice, from the results of microarray assay and real-time quantitative RT-PCR, we found that expressions of Calml4, Igf2, Igfbp2, Ttr, Pthlh, and Rbp1 were deranged. Igf2 (insulin-like growth factor 2), Igfbp2 (insulin-like growth factor binding protein 2), and Ttr (transthyretin) had been found to be implicated in Alzheimer's disease (Stein and Johnson 2002), the most common neurodegenerative disease, and increased expressions of Igf2 and Ttr were linked to the protection of neurons from β-amyloid induced toxicity (Dore et al. 1997; Serot et al. 1997). Thus, decreased levels of these proteins in SCA8 transgenic mice might lead to more harmful effects caused by mutant SCA8. Proteomic analysis showed that expressions of calbindin and calretinin were decreased in SCA8-157R trangenic mice. These proteins belong to the large family of EF-hand calcium-binding proteins (Rogers and Resibois 1992) and calcium-binding proteins-expressing neurons are thought to be resistant to injury-promoting perturbations in calcium homeostasis (Mattson et al. 1991). In addition, reduced immunoreactivity to calcium-binding proteins was shown in other SCAs (Kumada et al. 2000; Vig et al.

1996; Vig et al. 1998; Vig et al. 2000; Yang et al. 2000), suggesting that down-regulated of calbindin and calretinin might increase cerebellar susceptibility to neurotoxicity. Although 2-D electrophoresis offers an efficient and fast tool to investigate the protein profile, MALDI-TOF analysis only provide predicted candidate proteins and could not reflect the actual biological status of them. Therefore, immuoblotting and IHC should be performed for further validation of these candidate proteins.