SYNOPSIS

Molecular basis of spinocerebellar ataxia 1 (SCA1) in a selected south Indian population

Spinocerebellar ataxias (SCAs) comprise a highly heterogeneous group of autosomal dominantly inherited ataxias. To date, more than 30 SCA genes or loci have been identified. Recognized mechanisms leading to spinocerebellar neurodegeneration include polyglutamine expansions (SCAs 1, 2, 3, 6, 7, 17 and DRPLA), noncoding expansions (SCA10 and SCA12) as well as conventional mutations in genes encoding cytoskeletal proteins (βIII spectrin, SCA5), ion channels (voltage gated potassium channel Kv3.3, SCA13), protein kinases (tau tubulin kinase 2, SCA11; protein kinase C gamma, SCA14), intracellular calcium channels (inositol 1,4,5-triphosphate receptor 1, SCA15), fibroblast growth factors (FGF14, SCA27) and ATPases (AFG3L2, SCA28). Expansion of translated CAG repeat sequences in a subset of SCAs lead to abnormally long polyglutamine (polyQ) tract in the encoded proteins named ataxins 1, 2 and 3, alpha 1A voltage-dependent calcium channel, ataxin 7, TATA-box binding protein and atrophin 1, respectively, hence included in the group regarded as polyglutaminopathies. They all show, as common features, the progressive neurodegeneration of neuronal subsets in distinct brain areas and the formation of polyQ-containing protein aggregate forming characteristic nuclear or cytoplasmic inclusions. The age of onset and severity of disease symptoms inversely correlate with the length of the glutamine repeat.

Amongst the SCAs, SCA1 was the first to be genetically characterized with an expansion of CAG trinucleotide repeats located within exon 8 of the ATXN1 gene (Banfi et al. 1994). In normal individuals, the repeat number varies between 4 and 39 repeats, which are occasionally interrupted by 1-3 CAT triplets. In contrast, in SCA1 patients, the repeats are expanded beyond 39 and are uninterrupted (Quan, Janas, and Popovich 1995). An inverse correlation is known to be observed between the repeat size and age of onset (H. Y. Zoghbi and Orr 2000). Genetic anticipation or the tendency of affected individuals in each successive generation of kindred to show earlier onset with greater disease severity is also known in some studies (T. Matilla et al. 1993). A mechanism of progressive expansion particularly from the paternal alleles is a feature of SCA1 (H. T. Orr et al. 1993).

India, SCA1 accounts for 22% of ADCA (Autosomal Dominant Cerebellar Ataxia) (Mittal et al. 2005). Previous studies on prevalence of SCA1 in the Indian population were carried out in patients visited in various hospitals in the country (Basu et al. 2000; Saleem et al. 2000; Sinha et al. 2004; Mittal et al. 2005; Krishna et al. 2007). These studies showed high prevalence of SCA1 in South India (Mittal et al. 2005; Krishna et al. 2007). One such study explored the genetic basis, such as founder mutations, frequency of large normal alleles (>30 repeats) and CAT interruptions in the patients (Mittal et al. 2005). Small pockets of villages with a high prevalence of SCA1 have been reported in Tamil Nadu in South India (Rengaraj et al. 2005). However, genotypic characterizations have not been conducted in these isolated populations. Furthermore, there are no reports on longitudinal studies of disease progression with its genetic correlates amongst SCA1 patients within the Indian population. Studies investigating disease onset in individuals living in homogenous conditions helps in understanding the phenotypic variability in individuals with shared genetic background and environment influences.

*ATXN1* encodes a polyglutamine (polyQ) containing protein (ATXN1) which localizes to both cytoplasm and nucleus (Orr 2012). Although its exact function is not known, it is hypothesized to play a role in regulation of transcription and RNA processing/metabolism (Matilla-Dueñas, Goold, and Giunti 2008). Studies in cerebella of SCA1 mice and patients have shown transcriptional dysregulation of several intracellular calcium signaling genes prior to the onset of disease phenotypes indicating that cellular changes antedate neuronal death (Lin et al. 2000; Serra et al. 2004; Cvetanovic et al. 2011).

Recent studies in cerebella of SCA1 mice and SCA1 patients (Rodriguez-Lebron et al. 2013; Persengiev et al. 2011) have shown miRNAs deregulated in SCA1 pathogenesis. MicroRNA or miRNA’s are 18-27 nucleotide long non-coding RNA’s, that negatively regulate the expression of target genes by binding to the 3’ UTR of mRNA’s and inhibit their translation or promotes mRNA degradation (Valencia-Sanchez et al., 2006). Dicer is an essential enzyme in miRNA biogenesis. Dicer knockouts in animals provide us a very compelling role of miRNA in neurodegeneration. Dicer ablation in Purkinje cell of mice leads to cell death and ataxia (Schaefer et al., 2007). MiRNA’s are increasingly being implicated in neurodegenerative diseases such Alzheimer’s (Hebert et al., 2008), Huntington’s (Marti et al.) and Parkinson’s disease (Martins et al.). However, the impact of
miRNAs in modulating the onset, progression or severity of the disease in SCA1 patients remains unknown. Growing interest to study the progression of neurodegeneration, and to identify specific molecular pathways has prompted researchers to resort to the use of easily accessible tissue such as peripheral blood for conducting studies. Peripheral blood has been considered a rich source for biomarkers (protein, mRNA and miRNA markers) for studying neurodegenerative diseases rather than brain tissue samples. Brain samples come with various drawbacks as most patient studies are done with post autopsied brain which are prone to data fluctuations. It is also difficult to obtain brain samples when compared to peripheral blood. Studies in peripheral blood of patients with Alzheimer’s disease (Schipper et al., 2007), and Huntington disease (Gaughwin et al.) have shed new light on the use of peripheral blood in deeper understanding of the disease pathogenesis.

**Objectives**

1. Genetic characterization of a small community following the practice of consanguineous marriage for generations, with a history of hereditary ataxia and clinical characterization of the cohort to systematically document the phenotypes

**Rationale**

SCA1 is associated with a considerable variability of clinical symptoms and anticipation and studies based on cohorts with SCA1 outlining phenotypic characteristics is lacking in India. Such studies are important in understanding the disease characteristics in individuals with shared environmental influences, in this study we thus chose an inbreeding community living in close proximity. Since SCAs comprise of disorders with overlapping phenotypes, identification of a particular subtype by genetic tests is very essential. The first step to understand the subtype of hereditary ataxia affecting the cohort is thus to genetically characterize the cohort. It was also pertinent to test for the purity of CAG repeats, as CAT interruptions are known to occur in normal individuals and not in SCA1 patients. Subsequently, neurological phenotypes of the individuals was also characterized to gain insights in to the spectrum of the disease and to compare the manifestation of disease characteristics among individuals living in environmentally homogenous conditions

2. Correlation of genotypes with neurological phenotypes
**Rationale:** It is known that CAG repeats are negatively correlated to the SCA1 disease onset. Genetic anticipation is also known in SCA1. Thus it is important to observe the correlation of CAG repeats and genetic anticipation in this cohort, to understand if the trend of the disease is as reported earlier.

3. Small RNA sequencing in PBMCs of select presymptomatic and symptomatic patients. **Rationale:** The present evidence supporting the role of miRNA in SCAs is mainly correlative. It is based on the following findings: i) Changes in transcript levels of several calcium signaling genes have been found in the cerebella of patients and mouse models; ii) changes in miRNAs are found in symptomatic SCA1 brains; iii) Disruption of miRNA in purkinje cell in mice leads to ataxia iv) Brain is inaccessible to conduct such studies and peripheral blood has shown promise in monitoring disease related changes. In this study we test if miRNA changes indeed occur in PBMCs of SCA1 patients

4. RNA sequencing of select presymptomatic and symptomatic patients. **Rationale:** Expression of intracellular calcium genes is affected in SCA1 mice prior to the onset of symptoms. It is thus important to understand expression differences in SCA1 patients specifically the presymptomatic individuals when compared to normal individuals. RNA sequencing was carried out to capture global differences in gene expression in PBMCs of presymptomatic and symptomatic patients when compared to normal individuals from the cohort.

**Results**

**Identification of a SCA1 cohort in Indian population and the phenotypes of the patients:**

Patients and their families were chosen from Adukkamparai, a village located in an area of South India, where a high prevalence of hereditary ataxia was recorded previously (Rengaraj et al. 2005). A total of 100 individuals from 15 families were recruited in this study and we identified 16 individuals from 6 families with the clinical symptoms of ataxia in the first visit. Among the symptomatic individuals 4 had mild disease, 7 had moderate illness and 5 had severe symptoms during our baseline visit. In our second visit after four-years, 7 who were symptomatic in the baseline visit showed disease progression. Clinical signs and symptoms varied among the individuals of the cohort.
**ATXN1 genotyping in SCA1 patients:**

Genotyping for repeat expansions identified 21 individuals with repeat expansions suggestive of presymptomatic stage. Among these, 4 individuals were more than 50 years at the time of evaluation. The repeat sizes in these four individuals were 32/43, 28/42; 25/43 and 24/43. These patients including the oldest among them, a 60 year old female (O2), did not have any any features that were seen in other pre-symptomatics. On the second visit after four years, 12 of the 21 pre-symptomatics did not show any progress of the disease and remained pre-symptomatic, but 5 of them had progressed to develop mild disease and four patients were lost to follow-up.

**CAT interruptions and founder effect of ATXN1 mutations in SCA1 patients:**

To ascertain genetic characteristics of SCA1 in this cohort, analysis for CAT interruptions and founder mutation was carried out in both the presymptomatic and symptomatic individuals. It was thus established that all the presymptomatic including the four elderly presymptomatic and all symptomatic patients harbored pure CAG alleles and had a similar founder. As reported previously age of onset was inversely correlated to number of CAG repeats (correlation coefficient -0.45). In addition, it was observed that paternal inheritance led to an earlier age of onset. To conclude, CAG repeat size and inheritance were insufficient to explain the phenotypic heterogeneity and progression of the disease in these individuals suggesting the role of additional genetic or epigenetic modifications.

**Gene expression analysis in SCA1:**

To study the molecular aspect of SCA1 specifically transcriptional dysregulation as reported in previous studies (Lin et al. 2000; Serra et al. 2004; Cvetanovic et al. 2011), blood samples were collected from specific set of patients (pre-symptomatic and symptomatic) and normal individuals. Differential expression studies (RNA-seq, miRNA sequencing and QPCR) were performed on mononuclear cells isolated from the samples.

RNA sequencing resulted in 89 genes (75 down and 14 up) differentially expressed in the presymptomatic when compared to normal age matched individuals and 21 genes (20 up and 1 down) differentially expressed in symptomatic when compared to normal age matched individuals. To cross-validate these results, QPCR was done with increased sample number for the twelve best hits (on the basis of fold change and raw read count). Of the twelve genes, QPCR study found that *Clusterin* was down regulated in severe cases when compared to normal individuals and up-
regulated in pre-symptomatic individuals when compared to age matched normal individuals. This correlation between its up-regulation to disease severity is suggestive of it acting in a pathway, which counters disease progression. To conclude, we indeed found dysregulation of genes in the presymptomatic stage of the disease. More number of RNA samples needs to be sequenced to validate these results and to understand deregulated networks involved in the pathogenesis of SCA1.

Since studies in mice suggest calcium signaling genes as downstream targets of ATXN1 (Lin et al. 2000; Serra et al. 2004), six genes involved in intracellular calcium signaling were selected (ITPR1, ORAI1, STIM1, SERCA1, RYR2 and VEGFA) and their expression levels quantified for all disease categories (Pre-symptomatic, Mild/moderate, Severe) by QPCR. We found ITPR1 and ORAI1, which were significantly down regulated in Mild/moderate category versus Severe.

**MiRNA analysis in SCA1 patients:**

Additionally, small RNA sequencing was conducted to understand the role of miRNAs in SCA1. Previous studies have shown the role of specific miRNAs in SCA1 human brains and in mice models. Significant changes in miRNA expression in the symptomatic SCA1 cases when compared to normal individuals (7 up and 3 down) were observed. Co-expressed miRNA clusters were also observed in SCA1 patients. Over-expression of three miRNAs observed in the symptomatic SCA1 patients specifically targets the expression of specific mRNA’s, which are down regulated in the RNA sequencing experiment.

Together, differentially expressed miRNAs and mRNAs in SCA1 were captured in PBMCs of SCA1 patients indicating the reliability of using blood for the study of neurodegenerative progression in SCA1. To summarize, changes in both miRNA and mRNA expression is observed during disease progression in pre-symptomatic and symptomatic SCA1 patients when compared to normal individuals, thus adding another layer of complexity to our understanding of the pathogenesis of SCA1.