SUMMARY

Wilson disease (WD, OMIM# 277900) is a rare inborn error of copper accumulation inherited in an autosomal recessive manner. It is caused by mutations in the copper transporting P-type ATPase gene (*ATP7B*) (Thomas GR et al., 1995a). The disease is manifested at a median age of 12 to 23 years as hepatolenticular degeneration resulting from accumulation of copper primarily in the liver and brain. Patients mostly present with hepatic, neuropsychiatric or mixed symptoms (Das SK and Ray K, 2006). Varied clinical and biochemical manifestations often complicate diagnostic workup and diagnosis can be challenging especially for patients who are asymptomatic or present with liver disease. Failure to diagnose WD before its substantial progression is the leading cause of morbidity and mortality. But if patients can be diagnosed early enough, lifelong de-coppering with appropriate chelating agents can lead to gradual improvement and make them largely symptom free for the rest of their lives (Walshe JM, 2007). There is no single clinical or biochemical test to exclude WD. Genetic analysis still remains to be the confirmatory approach to detect the disease especially in sporadic cases with no family history and ambiguous clinical indications.

When the work for this dissertation was initiated, several studies have revealed the *ATP7B* mutation spectrum from North Indian (Kumar S et al., 2005), South Indian (Santhosh S et al., 2006) and East Indian (Gupta A et al., 2005, 2007a) (our group) WD population. However, the WD mutation pattern from West India was yet to be elucidated. Hence, the genetic, clinical and functional studies leading to this dissertation have sequentially focussed on the following areas (i) further characterization of molecular defects present in *ATP7B* gene in East Indian WD patients and exploration of the *ATP7B* mutation spectrum from West India (ii) screening of the *ATP7B* regulatory region in patients with single detectable coding mutations or lack of any, and followed it up with functional characterization of variants in the promoter and 5'UTR, (iii) Investigating the genotype-phenotype relationships in WD by generating a matrix for exploration of the clinico-genetic interface, followed by evaluation of candidate modifier loci, and (iv) functional evaluation of variants detected in the *ATP7B* coding region by localization and trafficking studies.

A total of 199 WD patients (174 East Indian and 25 West Indian) mostly with neurological problems from 178 unrelated families comprising of 687 individuals have been included in this study prospectively. This group is an extension of a previous study from our lab where 25 East Indian patients had been characterized with both *ATP7B* mutations and single mutations found in a few (Gupta et al., 2007a). *ATP7B* mutation screening has been undertaken in 174 patients (East Indian: 149, West Indian: 25). Finally, a comprehensive mutation profile of all 199 patients is described. In this study, we detected 29 reported and 22 novel mutations after screening most of the *ATP7B* coding region (exons 2 to 21) and flanking intron-exon boundaries. Of the reported ones, 19 defects had been reported from other populations but detected for the first time in our patient cohort and 10 variants were previously described by our group (Gupta A et al., 2007a) in addition to finding those during the study related to my dissertation. Adding up all mutations, our patient cohort harbors a total of 10

common and 48 rare lesions in the coding region of *ATP7B*. Three mutations (p.G1061E, p.N1270S and p.A1049A-fs) are shared between East and West. The common mutations have been found to account for 74% of characterized mutant alleles with p.C271X (63/260) and p.G1101R (7/31) being the most prevalent in East and West Indian patient cohorts, respectively. Comparison of mutation pools across India suggested that the exons 2, 8, 9, 13, 14, 15, 18 and Intron 4 harbor the prevalent mutations and should be screened with priority in the Indian WD population. In addition to the mutations, 28 innocuous variants could be detected during this study. Mutations in both the alleles could be characterized in 114 patients, while in 63 patients only one mutation could be detected in the coding region. No coding mutations could be detected in 22 patients. Further studies on these patients with missing mutation(s) are described in relevant sections below.

WD is one of the few genetic diseases, the progression of which can be controlled if diagnosed early enough. Knowledge of common mutations in a population would pave the way for rapid diagnosis, even in some suspect case without family information, before the manifestation of overt clinical symptoms. But mutation screening is a time consuming process. So, a parallel strategy was employed to identify the carriers and presymptomatic individuals in 41 East Indian and 19 West Indian WD affected families with multiple sibs to provide relevant information faster without waiting for determination of the underlying causal mutation. By studying the inheritance of linked polymorphic microsatellite markers D13S314 and D13S316 in the affected pedigrees, we could detect 5 presymptomatic, 40 carrier and 36 normal individuals. Mutation information, when available was utilized to resolve the ambiguous cases.

Most genetic studies reported earlier focussed on studying the *ATP7B* coding region but the gene promoter, though characterized, remained largely unexplored. As already described, 85 patients (63 + 22) in our pool could not be characterized for either one or both *ATP7B* mutations even after screening much of the coding region. In order to identify the missing mutations, we screened the *ATP7B* promoter and both the 5'- and 3'-UTRs in these patients. We detected 5 novel and 6 reported variants in the promoter and 5'UTR. Promoter cloning, site-directed mutagenesis and dual luciferase assay was carried out for functional evaluation of these variants. Five variants and one haplotype in the 5' regulatory region could be implicated in WD pathogenesis based on significant down regulation of the promoter activity. We carried out segregation studies in pedigrees to dissect the phase of alleles with respect to promoter variants and downstream coding mutations. We further hypothesized that *ATP7B* promoter variations altering its activity might also be present in patients where both coding mutations have been detected and influence the expression of the mutated protein. The differences in promoter strengths could have the potential to add on to the observed clinical heterogeneity in addition to the influence of other confounding genetic and environmental factors. Thus, the role of promoter variants in the regulation of *ATP7B* merits further investigation.

Till date, no nucleotide substitutions in the *ATP7B* 3'UTR have been associated with WD phenotype. This study reports 3 novel variations in this region in WD patients, however, further experiments are required to ascertain their status as mutations or rare innocuous variants. We reached a mutation detection frequency of 77% in our study.

We devised a Neurological Involvement Score (NIS) to explore the spectrum of neurological association across cognition, behaviour and motor domains in WD patients with preferential neuropsychiatric manifestation. The scale was found to have high internal consistency (Cronbach's α 0.802) and inter-rater reliability (ICC 0.888). Thus NIS was highly reproducible. Subsequently we undertook a novel unified approach of genotype-phenotype (G2P) correlation by utilizing NIS as a simple tool to explore the neuropsychiatric arena of WD patients in light of the mutations identified across different ATP7B protein domains. Using age at onset and neurological involvement score, we generated a G2P matrix. The scheme of G2P matrix provided opportunities to examine and compare the patient phenotypes with respect to same genotype and also across different protein domains. We observed that the truncation mutations could be equally deleterious irrespective of the site of termination being early or late in the polypeptide chain. Compound heterozygotes bearing combination of missense and truncation mutations indicated missense mutations to retain more functionality in the mutated protein as reported previously (Nicastro E et al., 2009). Specific mutations in the nucleotide binding domain were found to attribute hepatic predominance.

Patients with same/similar mutation background would be expected to have similar presentation. The most interesting observation was made when we encountered a few patients who depicted stark difference in their clinical manifestation as compared to other patients in the same mutation cluster. We counted the patients with divergent phenotypes as "outliers" in the matrix. Possible roles of *COMMD1* and *ATOX1*, two genes in copper homeostasis pathway, were also investigated for their potential to act as modifier loci in the outliers, though no variant could be implicated for the cause. Since the 'basal level heterogeneity' of presenting symptoms in WD is not exclusively explained by the *ATP7B* mutations, and cases of misdiagnosis and phenotypic overlap have been reported implicating additional genes such as *NPC1*, *SLC33A1*, and *SLC30A10* (Connemann BJ et al., 2012; Huppke P et al., 2012a; Quadri M et al., 2012), we propose that these loci should be studied for their putative role in imparting WD features by some novel pathophysiological mechanisms.

The diagnostic delay emerged to be a confounding factor for G2P analysis. The mean diagnostic delay in our study was 19.5 months (0.2 to 180 months) and there is opportunity for further improvement.

Missense mutations might have a spectrum of consequences on the mutant protein. In this context, we made an attempt to explore the effect of missense mutations on cellular localization and anterograde trafficking of ATP7B. GFP fusion constructs for 9 different ATP7B mutants selected from different domains were investigated for their localization and

copper responsive redistribution. While 6 of 9 mutants showed normal behavior, 2 mutants (p.G1061E and p.G1101R) were found to be grossly mislocalized to endoplasmic reticulum while 1 mutant (p.S1362A) was found to be unresponsive to copper induction. Mutations might affect several properties of ATP7B which could alter its precise conformation, catalytic and kinase mediated phosphorylation, ATPase activity, copper binding and transport rate across membranes etc. Each of these factors might have deleterious consequences on normal copper distribution and may lead to WD. Thus, a detailed functional exploration of the effect of missense variants on other properties of the protein would provide deeper insight of their effect on the clinical phenotype of the patients.

The studies leading to this dissertation have resulted in the following publication as noted below and additional manuscript describing unpublished work is under preparation:

Mukherjee S, Dutta S, Majumdar S, Biswas T, Jaiswal P, Sengupta M, Bhattacharya A, Gangopadhyay PK, Bavdekar A, Das SK, Ray K. *Genetic defects in Indian Wilson disease patients and genotype-phenotype correlation*. Parkinsonism and Related Disorders 2013; (*In press*); DOI, 10.1016/j.parkreldis.2013.09.021