I. INTRODUCTION
INTRODUCTION

The unraveling of biological phenomenon is a continual process wherein every new finding increases the degree of complexity and opens new areas that need to be explored. Decades of work add to the existing knowledge of the problem under study with its basic nature still remaining a puzzle. We find a number of such phenomena that were considered unsolved years before and their solution still not in sight. One such problem is that of sex determination and differentiation. The determination of the two sexes and the cascade of events leading to differentiation of sexual phenotypes are processes that have intrigued researchers for long. In mammals, the sex determination pathway is decided in the embryo with the inheritance of the X or Y chromosome from the father though the process is initiated during organogenesis. At present numerous genes have been implicated in the process of sex determination and differentiation. The expression of SRY located on the Y chromosome in the gonadal ridge triggers the entire cascade of genes that define the male sex-determining pathway (Sinclair et al., 1990) within the initial indifferent gonad. The bipotential gonad is known to express a number of genes like SF1, LIM1, LHX9, EMX2, and WT1 (Val et al., 2003, Shawlot et al., 1999, 1995, Birk et al., 2000, Yoshida et al., 1997, Call et al., 1990) that probably help in the selection of the specific cascade that determine the sex of the individual. SOX9, FGF9, WT1, SF1, and GATA-4 are expressed at later stages after the initiation of the testis-determining cascade (Cameron and Sinclair, 1997, Colvin et al., 2001, Val et al., 2003, Call et al., 1990, Tevosian et al., 2002). The interactions between these genes lead to the activation of AMH (Durlinger et al., 2002), which in turn initiates the testis differentiation process. The female sex determination and differentiation pathway wherein genes like WNT4 and DAX1 ensure down regulation of the SRY expression (Vainio et al., 1999, Weiss and Jameson, 2003) culminates in the generation of ovarian determining factor and the subsequent retention of the mullerian duct. While
these genes are primarily responsible for the execution of both the pathways others such as DMRT1 (Raymond et al., 1999), and Vanin 1 (Bowles et al., 2000) that affect the differentiation of the gonads and cell migration respectively have also been identified. The cascade is thus emerging as a network wherein some genes have multiple roles to play at different stages and others are expressed during a short window during which they provide the critical trigger. Identification of new candidate genes that form part of the network will definitely help in a better understanding of the processes involved.

Apart from an in-depth study of the genes themselves, the other aspect that needs to be investigated is the role of the non-coding region of the genome. This region that constitutes 98% of the human genome mainly consists of introns, non-coding RNAs and repetitive DNA. It is quite inconceivable that such a vast portion of the genome will be functionless. In recent years, these components of the genome have been implicated in the normal genome structure, organization and in the regulation of the expression of genes that are located in their vicinity. Polymorphic tandem repeats, located downstream to the human telomerase gene, are now identified as responsible for the regulation of its expression (Wang et al., 2003). L1 and Alu repeats are found to exert a positive transcriptional enhancer effect on a human zinc finger gene (Landry et al., 2001). Other repeat elements like Short tandem repeats (STRs) and Simple Sequence Repeats (SSRs) have also been implicated in the regulation of gene expression.

Bkm is a simple sequence repeat identified from the Banded Krait and primarily constitutes of a tetranucleotide repeat (GATA) (Singh and Jones, 1982). Extensively conserved in the eukaryotes, these sequences are concentrated on the heteromorphic sex chromosome. In snakes they were identified from the W chromosome (Singh et
while in Drosophila, *Bkm* sequences are confined mainly to the region 19F-20AB of the X chromosome that is implicated in male sterility (Singh et al., 1981). In mice, they are predominantly present on the Y chromosome and showed a sex-specific hybridization pattern. The *Bkm* satellite sequence was further implicated in the sex determination process in mice on the elucidation of the mechanism of the sex reversed mice (XXsxr mice). In depth analysis of the sex reversed mice showed that the transfer of a *Bkm*-associated region of the Y onto the X chromosome was responsible for the sex reversal (Singh and Jones, 1982). In combination with the fact that other repeats constituting the genome did not show such extensive association with the sex chromosomes, *Bkm* sequences were believed to perform significant functions in the condensation and decondensation processes of the sex chromosomes and in the process may be instrumental in the tissue and stage specific expression of the associated genes (Singh and Jones, 1982). Therefore, *Bkm* sequences were used as a probe to identify repetitive DNA and unique sequences, from the mouse and human genomic libraries, which may be functionally relevant in the sex determination and differentiation process. Consequently, screening of the mouse and human genomic libraries led to the identification of repetitive DNA clones; M34 from mouse and C102 from human, both of which showed sex specificity and extensive distribution along the length of the sex determining Y chromosome (Singh et al., 1988, 1994, Singh and Majumdar 1993). Further, human testis cDNA library was also screened using *Bkm* positive genomic clone in order to identify testis-specific transcripts (Rajyashri and Singh, 1995). This work led to the identification of a transcript that localized to the X chromosome and showed extensive conservation across species with the mouse homolog being predominantly expressed in the testis. The major gene(s) involved in gonadogenesis must be present in both sexes but expressed in a sex-specific manner as is evident by the fact that incubation of crocodilian eggs at higher temperature (32.5°C) leads to the development of only males and their incubation at
lower temperatures (30°C) to only females. Since organization of gonads is apparently similar in the animal kingdom, it also suggests that the gene (s) involved in gonadogenesis may be highly conserved. Therefore, the further characterization of this gene was undertaken.

This gene, identified as *WDR13*, was a novel member of the WD family of proteins (Singh *et al.*, 2003). The WD family of proteins constitutes members containing the WD repeat motif. This motif, first identified in Gβ protein, consists of a core region of 40 amino acids enclosed in by GH (glycine-histidine) and WD (tryptophan-aspartatic acid) (Voorn and Ploegh 1992, Duronio *et al.*, 1992, Neer *et al.*, 1994). The members of this family are found to be present in organisms across the species, from *C. elegans*, *S. cerevisiae* to *H. sapiens* (Neer *et al.*, 1994, Smith *et al.*, 1999). They hence seem to be conserved during evolution and are not confined to any specific function. The known members of the family are generally regulatory proteins associated with signal transduction, transcription, and RNA processing, development and cytoskeletal assembly. A number of WD proteins have been identified from human that are implicated in functions as wide ranging as *WDR8* in bone ossification (Koshizuka *et al.*, 2001) to *WDC156* in spermatogenesis (Ito *et al.*, 2001) and are also implicated in diseases like cancers (Claudio *et al*; 1999) and Down’s Syndrome (Michaud *et al*; 2000). These earlier studies thus gave a clear indication that the presence of the WD motif as such does not specify a function except for the fact the motif acts as a platform for the interaction of different proteins and hence denotes a regulatory role.

In most biological processes as complex as sex determination, it is selective and sequential gene expression that defines the process. This expression profile is maintained by specific regulatory mechanisms. The study of mechanisms of gene regulation thus gives insight into the function of the gene itself. The key word here is
regulation. The specialized functions that define any differentiated cell in a multicellular organism are the result of cascades that begin from the transcription of specific genes, followed by their regulated translation. How are the genes and the transcripts specified and selected? What decides the differential translation of these transcripts in the different cells and tissues? An effort to answer these questions may provide an insight into the way each cell or tissue type gets differentiated and may also explain the exact mechanism by which a specific gene executes its function. As mentioned before, WDR13, a novel member of the WD family of proteins was identified and preliminary studies with the gene indicated a differential expression profile in the mice, with a predominant expression in the testis. Therefore, suggesting its possible involvement in the sex differentiation process. Hence an understanding of the mechanisms that control the expression of the gene will definitely help to elucidate its function in a rational manner. Studies to further characterize the gene and elucidate the regulatory mechanism involved in maintaining its expression was thus undertaken.

The objectives of the present study can be listed as:

- Expression profiling of gene WDR13 in the human and mouse systems
- Functional Characterization and Analysis of the regulatory mechanisms that govern the expression of the gene
- Evolutionary and Phylogenetic analysis of WDR13