SEARCH FOR PUTATIVE GENETIC FACTORS INVOLVED IN THE TEMPERATURE DEPENDENT SEX DETERMINATION IN COMMON INDIAN MUGGER (*CROCODYLUS PALUSTRIS*)

AMIT ANAND

CCMB

CENTRE FOR CELLULAR AND MOLECULAR BIOLOGY

HYDERABAD. 500007

JANUARY 2009

Search for putative genetic factors involved in the temperature dependent sex determination in common Indian mugger (*Crocodylus palustris*)

A Thesis submitted for the degree of

DOCTOR OF PHILOSOPHY

To

Jawaharlal Nehru University, New Delhi

by

AMIT ANAND

CCMB

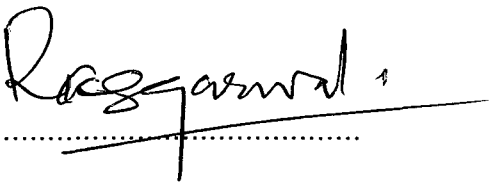
CENTRE FOR CELLULAR AND MOLECULAR BIOLOGY

HYDERABAD - 500007, INDIA

January 2009

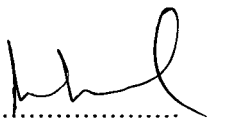
CERTIFICATE

The Research work presented in this thesis was carried out at the Centre for Cellular and Molecular Biology, Hyderabad. This work is original and has not been submitted in part or full to any other university for any other degree and diploma.



Dr. Ramesh K. Aggarwal

Research Supervisor



Amit Anand

Research Fellow

Acknowledgements

I begin with expressing my gratitude to my thesis supervisor Dr. Ramesh K. Aggarwal, for his immense support and guidance in course of my PhD. His understanding and accomplishment have provided a good basis for the present thesis. His eye for detail and perfection are exemplary and will remain an inspiration.

My sincere thanks to Dr. Lalji Singh for making CCMB an excellent place to work. As a Director, he maintained high standards in CCMB and also established new facilities.

Thanks to Dr A.J. Rachel and Dr K Thangaraj for their help and encouragement.

Thanks to Nandini for her expert help in Confocal microscopy, thanks to Nagesh and Mehar for oligonucleotide synthesis and DNA sequencing. Thanks to all the members of Instrumentation Group, Fine Chemicals, Digital Imaging and Xerox section.

I also express my gratitude to the authorities of Nehru Zoological park, especially to Ramulu and co-workers for making crocodile egg collection safe and easy.

I also take this opportunity to express my thanks to previous and present members of my lab for making lab a wonderful place. My thanks to Prasad, Raman, Krishnakumar, Albert, Ajeet, Kathiresan, Velavan, Uday, Rama, Prasanna, Rajkumar, Rajendrakumar, Annapurna, Indranil, Jamila, Phanindra, Srikanth, Paridhi, Janani, Sabeera, Richa, Megha, Upasana, Maninder, Prasoona, Madhubala and Minu.

Members of the West Wing Third Floor who have made my long tryst an unforgettable pleasant experience, from helping me in my experiments, to those great outings, cooking and eating together, and reviewing movies and accompanying me. My sincere thanks to Anjali, Arun, Umaji, Hemakumar, Amritha, Avinash, Jojo, Rajeev, Niraj, Nizam, Aditya, Ashwin, Vanni, Sreekar, Sreedhar, Joshi.

Thanks to summer trainees and M.Sc. dissertation students in the lab.

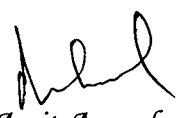
Thanks to my batchmates, Pantnagar group at Hyderabad and friends on whom I relied for every little scientific and personal help. I express my gratitude to Shweta, Brajesh, Pratibha, KT, Deepankar, Arvind, Jayendra, Pankaj, Anoop, Praveen, Ram, Vasanthi, Pallavi, Sumit, Soji, Guru, Tirumal, Shoeb, Rajkumar, Krishnan, Dev, Ashish.

Sincere thanks to Suryanarayana garu for keeping the lab in great condition and to Hanif, Visvanath, and Bhaskar for helping out whenever required.

My Wife Priyanka, whom I cannot thank enough for her unconditional support, companionship and encouraging me during those critical moments when it was needed most.

Words can express too little of my sense of gratitude towards my parents. I always look at them in the moments when I need inspiration, support, patience and guidance. I dedicate this thesis to my parents.

Financial support from the CSIR and CCMB is duly acknowledged.



Amit Anand

Table of contents

List of figures

List of Tables

Abbreviations

µCi	micro Curies
µg	microgram
µl	microlitre
µm	micrometer
bp	base pair
BSA	bovine serum albumin
cDNA	complementary DNA
DAPI	4'6-diamino-2-phenyl indole
dATP	deoxy adenosine 5' triphosphate
DNA	Deoxyribonucleic acid
dNTP	deoxy nucleotide 5' triphosphate
EDTA	Ethylene diamine tetra acetic acid ethylene diamine
EST	Expressed Sequence Tags
FITC	Fluorescein isothiocyanate
GSP	Gene Specific Primer
IPTG	isopropyl thio-b-D-galactoside
kb	kilo base
LB	Luria Bertani
M	molar
mg	milligram
ml	millilitre
mM	millimolar
MMLV	Moloney Murine Leukemia Virus
MOPS	3-(n-morpholino) propane sulphonic acid
mRNA	messenger RNA
ng	nanogram
nm	nanometre
OD	optical density
ORF	open reading frame
PBS	phosphate buffer saline
PCR	polymerase chain reaction
PI	propidium iodide
pk	phospho kinase
RACE	rapid amplification of cDNA ends
RNA	ribonucleic acid
rpm	rounds per minute
RT-PCR	reverse transcription-PCR
SDS	Sodium dodecyl sulphate

SSC	Sodium saline citrate
SSPE	Sodium saline phosphate EDTA
TAE	Tris-acetate-EDTA
TBE	Tris-borate-EDTA
TE	Tris-EDTA
U	unit
UTR	Un-translated region
UV	ultra violet
X-gal	5-bromo-4-chloro-3-indolyl-b-D-galactoside

Abbreviation of genes

Amh	Anti-Mullerian hormone
Arx	Aristaless-related homeobox gene
Clk1/4	*Clk2* like protein kinase1/4
COUP	Chicken ovalbumin upstream transcription TF-II
Dax1	Dosage-sensitive sex reversal gene
Dhh	Desert hedgehog
Dmrt1	Doublesex mab-3 related transcription factor1
Emx2	Empty spiracles homologue 2
Fgf9	Fibroblast growth factor 9
Fog2	Friend of GATA; zinc finger protein, multitype2
FoxL2	Forkhead helix family transcription factor L2
GAPDH	glyceraldehydes 3-phosphate dehydrogenase
GATA4	GATA binding protein4
H19	H19, imprinted maternally expressed transcript
HSF1	Heat shock factor-1
Igf1r	Insulin like growth factor 1 receptor
IME4	Initiator of meiosis 4
Lhx9	LIM homeobox protein 9
M33	Mouse homologue of polycomb genes
Pdgfrα	Platelet derived growth factor-alpha
SLF35F5	solute carrier family 35, member F5
Sox9	SRY-box containing gene 9
Sf1	Steroidogenic factor-1
Sry	Sex-determining Region Y
Unr	Upstream to n-RAS
Wt1	Wilms tumor protein-1
Wnt4	Wingless-type MMTV integration site family, member 4
Xist	X (inactive)-specific transcript

Chapter 1

Introduction

Sex determination is a fundamental biological process of great importance, whereby bipotential gonads develop into a testis or into an ovary, essential for the propagation of the species. Nature has evolved an astonishing variety of genetic and epigenetic sex determining systems which all achieve the same result, the generation of two sexes, a male or a female. In the animal kingdom, a large number of sex determination mechanisms are seen. In vertebrates, this process is accomplished by highly specialized sex chromosomes (GSD - *G*enetic *S*ex *D*etermination) to simple environmental cues (known as ESD - *E*nvironmental *S*ex *D*etermination). Overall, GSD is the major mode of sex determination, seen in all five phyla of vertebrates including mammals, and is much better understood. In GSD species, the specialized sex chromosomes may differ: morphologically (heteromorhic or homomorphic), in numbers (from one pair to many pairs), and even in their effect due to the presence of several minor sex-determining genes (Bull 1983; Solari 1994; Grutzner *et al.,* 2004). In comparison to GSD, very little is known regarding ESD and its underlying mechanism. The ESD species, in general, seem to lack apparent sex chromosomes or sex specific genes that could consistently produce a genetic difference among zygotes, and their sex appears to be mainly decided post fertilization by the simple cues like temperature, pH etc. of the environmental milieu of the developing embryos. Among different environmental cues seen in ESD species, the incubation temperature as the sex-determining factor is the most important and widespread. The latter is called *T*emperature-dependent *S*ex *D*etermination (TSD) which is chiefly found in the reptilians; it occurs in all crocodilians, tuartaras studied till to date, and is common among turtles (Valenzuela & Lance 2004).

Sex determination is exclusively temperature dependent in all crocodilians. The particular period of embryonic development during which temperature appears to be critical for the sex determination is called *T*emperature *S*ensitive *P*eriod (TSP) (Bull 1983; Solari 1994). In TSD species, the sex ratios can be easily manipulated by changing the incubation temperature during TSP. The Indian mugger (*Crocodylus palustris*) is one such species in which exclusively males or females can be produced by manipulating the incubation temperature of fertilized eggs. It is demonstrated that the incubation temperature at and below 31.0°C and above 33.0°C results exclusively in female development whereas varying proportion of both males and females are produced only between the narrow range of 31.0°C-33.0°C with 32.5°C supporting exclusive male development (Lang *et al.,* 1989). The embryonic phase wherein the sex determination process is sensitive to environmental temperature has been broadly delineated to developmental stages 21st to 25th (spanning approximately 21 - 25 days of

development) by temperature shift incubation studies (Lang & Anderson 1994). All the field studies and universal absence of sex chromosomes in crocodilians suggest that gene(s) responsible for primary sex determination are present in all embryos, but are differentially and selectively expressed at specific time points during development as a function of temperature, leading to the development of either males or females. The pattern of sex ratios vis-à-vis critical temperature range, suggests the female development being the default mechanism, but male development needing specific activation of factor(s) in response to temperature. Because of these unique features, *Crocodylus palustris* provides an ideal and attractive system to: a) gain insight into the molecular basis of TSD; and b) for identification and characterization of temperature sensitive biological control amenable to fine tuning by temperature. The possibility of such temperature sensitive biological control would be of great significance in both basic and applied biotechnology research.

Several studies done in the past (mostly in mammalian species) have helped in detailing the developmental/cellular events, as well as, identification of many of the candidate genes involved in- maintenance of bipotential gonad primordia, sex determination and early differentiation of gonads (Merchant-Larios & Moreno-Mendoza 2001; Pannetier *et al.,* 2004; Swain & Lovell-Badge 1999; Lovell-Badge *et al.,* 2002; Wilhelm *et al.,* 2007). Major morphological events occurring during the embryonic gonad development are broadly conserved among vertebrates *viz.*, formation of testicular cords immediately after sex-determination/decision, strengthening of cortex in developing female gonads and regression of Wolffian ducts (Morrish & Sinclair 2002; Yao & Capel 2005; Smith & Sinclair 2004). A similar broad conservation has also been evident even for the major sex determining genes. The *Sry* (Sex determining region of Y chromosome) is revealed as a key testis-determining gene in mammals. In addition, a number of other genes, *Wt1, Sf1, Sox9, Dax1, Lim1, Emx2, Lhx9, Wnt4, Fgf9 GATA4, Amh* and *Dmrt1* have been identified and implicated either in the maintenance of the bipotential gonadal primordia and/or in the early differentiation of gonads after sex determination (Swain & Lovell-Badge 1999; Wilhelm *et al.,* 2007). These genes generally show dimorphic expression in males and females. Genes like *Sf1, Wt1, M33, Lim1, Emx2* etc. are shown to express early with a role in growth and survival of bipotential gonad, while *Sox9*, *Dax1, Fgf9, Amh* and other genes are involved in the differentiation and development of the gonads. Many of these genes appear to be conserved through evolution at least in their organization, although, the same may not be true for their role in sex-determination. It is now well documented that there exists conspicuous differences in the spatio-temporal expression of many such sex-related genes in different systems, *e.g.*, in *Gallus* (having

ZZ/ZW sex determining system), *Sox9* expresses after *Amh*, a situation just reverse to that of mammals wherein *Amh* is the transcriptional target of Sox9 (de Santa *et al.,* 1998). Similarly, *Sry* is absent in the *Gallus* and it is the *Dmrt1* gene (present on the Z chromosome) showing early expression in the bipotential gonads that is suggested to be the possible male sex deciding candidate gene (Smith *et al.,* 2003; Raymond *et al.,* 1999). Nevertheless, at this stage, our understanding of the putative genetic cascade underlying the sex determination mechanism remains poor, mainly because of lacking information regarding functional interplay between the known and many other unknown (yet to be identified) sex candidate gene(s). Filling these missing pieces is the challenge for future.

In contrast, relatively few studies have been carried out in TSD species that too in recent years and these are mainly limited to identification and expression analysis of conserved sex determining genes earlier identified in GSD species/mammals and hormonal effect on TSD. These studies, in general, have revealed 'hormonal interplay' as an important component of the underlying sex-determination mechanism(s), and homologues of many of the sex-related genes have been identified such as, *Sf1, Wt1, Sox9, Amh, Dmrt1,* and *Aromatase* gene of hormonal cascade (Shoemaker *et al.,* 2007; Sarre *et al.,* 2004; Western & Sinclair 2001). Presence of such genes even in TSD species exhibiting sex-specific quantitative variation in their expression during gonadal development (Morrish & Sinclair 2002; Sinclair *et al.,* 2002), suggests a broad evolutionary and functional conservation between the genetic and hormonal machinery underlying the GSD and TSD. However, none of these studies done to date provides any insight about the nature/mode of the molecular machinery in embryo that senses the external temperature and triggers the process of sex-determination. Our understanding of the factor(s) involved in TSD and modulated by the temperature trigger is completely wanting. TSD thus, remains an enigmatic problem of developmental biology.

Unlike the mammalian sex-determination, a master regulator like *Sry* has not been detected in TSD species. On the other hand, another gene *Sox9* (one of the Sry related HMG box carrying gene) that has been proposed to be the "master" testis-determining factor downstream of *Sry* in mammals, is highly conserved among vertebrates and unlike *Sry*, is also suggested to have an important role in sexual development even in birds and reptiles. The *Sox9* gene is necessary for male development in mammals and a fortuitous insertional mutation has shown that constitutive *Sox9* expression is sufficient to direct full male development in the mouse in the absence of *Sry* (Bishop *et al.,* 2000). Inactivation of *Sox9* in XY embryos leads to sex reversal of male to female, while exogenous

expression of *Sox9* in XX gonads can induce testis formation in the absence of *Sry* (Vidal *et al.,* 2001). Sox9 is also known to interact with Sf1 and activates transcription of Anti-Mullerian hormone gene *Amh* (de Santa *et al.,* 1998). All these studies suggest *Sox9* to be the "master" testis-determining factor downstream of *Sry* in mammals. However, few recent studies in other non-mammalian species showing confounding differential expression pattern of *Sox9*, appear to suggest otherwise; that *Sox9* role as "master" testis effecter gene may not be conserved in TSD and possibly other GSD species. In some studies in chicken (GSD species) and alligator and turtle (TSD species), *Sox9* expression is observed to lack male/testis-specificity until testicular structures are established (Oreal *et al.,* 1998; Western *et al.,* 1999), while contrary to it, *Sox9* expression is found to be male-specific preceding testis organization in another TSD sea turtle *Lepidochelys olivacea* (Torres Maldonado *et al.,* 2002; Moreno-Mendoza *et al.,* 2001) more closely resembling the pattern in mice. More recently, the work done in CCMB on detailed expression profiling of *cpSox9* in another TSD crocodilian i.e., Indian mugger, amply demonstrate that the above referred observations about lack of male-sex specific expression prior to testis organization, may be a fall-out of the constraints of experimental design, as discerning male-specific expression becomes difficult unless critically analyzed due to *Sox9* transcriptional diversity during gonadogenesis (Agrawal 2006). The in-house work at CCMB, further suggests an important role for *Sox9* in male sex development/differentiation, albeit not as the candidate gene for primary sex-determination in TSD.

Lately, other major genes (other than *Sry*, *Sox9*) originally identified and implicated in the GSD in mammals are also being explored for their presence and possible role in TSD. Some such studies involving preliminary spatio-temporal expression analysis have revealed sex determining genes like, *Wt1, Dmrt1, Sf1* and *Aromatase,* as possible candidates having important role in TSD (Yao & Capel 2005). Among these, *Dmrt1* is a highly conserved gene that is shown to play an important role in GSD from being a master regulator in the teleost Medeka, candidate master regulator for male gonad development in birds, to a downstream gene required for the male gonad development in mammals (Raymond *et al.,* 2000; Shan *et al.,* 2000; Matsuda *et al.,* 2002; Hodgkin 2002). *Dmrt1* is also shown to express in early stages of embryonic development including the indeterminate bipotential gonadal primordial tissue in some studies in TSD/HSD vertebrates (Ferguson-Smith 2007; Sreenivasulu *et al.,* 2002; Sreenivasulu *et al.,* 2002; Murdock & Wibbels 2006). Similar to *Dmrt1, Wt1* and *Sf1* are also largely conserved and are implicated for an early role in bipotential gonad maintenance and sex determination in the mammals and few other GSD species. An expression analysis study

in the turtle also indicates an early role for these genes (Yao & Capel 2005). Moreover, notable sex-/developmental stage-specific differences during gonadal development have been observed in expression pattern of such major sex determining genes in the few studies on different TSD species, which further suggests proper characterization of the transcriptional diversity and expression analysis of these genes in TSD species is necessary (Yao & Capel 2005; Schmahl *et al.,* 2003; Western *et al.,* 2000; Ramsey *et al.,* 2007).

In contrast, to the apparent differences in expression patterns of the major sex-candidate genes, effect(s)/interplay of hormonal environment on early gonadal differentiation is revealed to be more conserved and universal among TSD species. A number of studies have shown that embryonic sex can rather be easily manipulated by modifying the hormonal regimen, e.g., estrogens treatment can lead to female development at male promoting temperature (MPT), while male embryos can be induced at female promoting temperature (FPT) by treatment with Aromatase inhibitors. In this hormonal cascade, Aromatase (CYP19) is the core important enzyme that regulates the conversion of androgens to estrogens (Lance 1997), and thus can override the effect of temperature on sex-determination. Now, there is a general consensus that temperature exerts its influence in species with TSD by acting upon the genetic mechanisms that govern steroidogenic enzymes or steroid hormone receptors (Sarre *et al.,* 2004). Therefore, Aromatase, one core regulator of hormonal cascade, may be one important target of close scrutiny in our pursuit of unraveling the molecular machinery underlying the TSD which is strongly influenced by the hormones (Valenzuela & Lance 2004).

To get insight into the molecular basis of TSD, *C. palustris* (Indian mugger) provides a model system. The temperature sensitive period in this species has been broadly defined corresponding to the developmental stages 21st to 25th (comprising ~22 - 30 days of embryonic development in male and female embryos, respectively) by egg shift incubation studies (Lang & Anderson 1994). Understanding molecular basis of TSD using Indian mugger is one major program in the area of developmental biology at CCMB. To address this interesting but challenging problem multiple approaches are being pursued that broadly include: isolation and characterization of homologues of known sex-determination related candidate gene(s), search for new genes with putative role in TSD, and creating large biological resources for future functional genomic studies. In this background, the present study was planned to identify/isolate crocodilian homologues of three important candidate genes namely, *Dmrt1, Wt1* and *Aromatase* from *C. palustris* and to examine their spatio-temporal expression during gonad

development through TSP. The characterization of these candidate genes from Indian mugger was expected to unravel their functional role (if any) in TSD, and also to help ascertain the extent of their conservation between corresponding genetic elements involved in TSD/GSD. In addition, it was planned to explore their sex-specificity and potential use as biomarker for embryo sexing to delineate the determinate/indeterminate phases of TSD thus helping in narrowing down the long TSP which is an important need to undertake effective functional genomic studies. Simultaneously, the study also proposed to attempt search for novel differentially expressed gene(s) during early TSP that may be: involved in the developmentally important decision of sex determination and differentiation in case of TSD and/or immediate targets of the machinery that senses the temperature and regulates gene expression.

In specific, the main objectives of the study were:

- Isolation, characterization and expression analysis of crocodilian homologues of candidate genes: *Aromatase*, *Dmrt*1 and *Wt*1 from *Crocodylus palustris.*

- Search for the gene(s) differentially expressed during early temperature sensitive period (TSP) in developing gonads of putative male and female embryos, and their characterization.

Chapter 2

Review of Literature

Vertebrate Sex determination: An overview

2.1 Sex determination and differentiation

Sex determination is a fundamental biological process that is of profound importance for the development of individuals and the formation of sex ratios in natural population. It is among the most primary decisions made during embryonic phase of life, which results in to development of anatomically and physiologically distinct male and female organisms and allows sexual reproduction to occur. In this process, in metazoans undifferentiated "bipotential" embryonic gonads are committed to either testes or ovaries, is guided by genetic or environmental cues. In many organisms, the sex of an offspring is irreversibly determined by its sex chromosomes, or rather, a set of genes on the chromosomes, regardless of any environmental variation. This condition is known as "genotypic sex determination" (GSD) wherein the sex of an individual is fixed at the time of fertilization. However, in some organisms the immediate environment of the fertilized egg/embryo seemingly determines whether the offspring will become a male or female, a condition referred to as "environmental sex determination" (ESD). In cases, where the decisive environmental factor is temperature, it is referred as "temperature dependent sex determination" (TSD). Overall, GSD is the major mode of sex determination, seen in all five phyla of vertebrates including mammals, and is much better understood. In GSD species, the specialized sex chromosomes may differ: morphologically (heteromorhic or homomorphic), in numbers (from one pair to many pairs), and even in their effect due to the presence of several minor sex-determining genes (Grutzner *et al.,* 2004; Solari 1994; Bull 1983). In comparison to GSD, very little is known regarding ESD and its underlying mechanism. The ESD species, in general, seem to lack apparent sex chromosomes or sex specific genes that could consistently produce a genetic difference among zygotes, and their sex appears to be mainly decided post fertilization by the simple cues like temperature, pH etc. of the environmental milieu of the developing embryos. Among different environmental cues seen in ESD species, the incubation temperature as the sex-determining factor is the most important and widespread. The latter is called *T*emperature-dependent *S*ex *D*etermination (TSD) which is chiefly found in the reptilians; it occurs in all crocodilians, tuataras studied till to date, and is common among turtles (Valenzuela & Lance 2004).

After the commitment of gonad into a testis or ovary, the process of differentiation takes place in a sexual dimorphic fashion, which is a programmed cascade of genetic and

hormonal events. The founding cell populations in the bipotential gonad have the capacity to differentiate into ovarian or testicular cell types via several complex events like cell migration and cell fate determination (figure. 2.1). In addition to germ cells, there are three cell lineages present in the bipotential gonad of mammals. The supporting cell lineages differentiate into follicular cells in the ovary and Sertoli cells in the testis. The function of these cells is to provide an appropriate growth environment for the germ cells that are surrounded by them. Steroidogenic cell lineages, which differentiate into theca and granulosa cells in females and Leydig cells in males, produce sex hormones. On the other hand, the connective cell lineage, comprising peritubular myoid and stroma cells participate in the formation of the organ as a whole. During the indifferent stage, the superficial coelomic mesothelium of the genital ridge continues to proliferate into loose connective mesenchymal tissue and forms the cortical cords (the primary sex cords). In females, the cortical sex cords split into clusters and differentiate into granulosa cells without penetrating deeply into the mesenchyme, while the mesenchymal cells differentiate into theca cells. Together, the theca and granulosa cells form the follicles that envelope the germ cells and secrete steroid hormones via a two-cell system. Theca cells first produce androgen, which is then diffused into the granulosa cells and get converted to estrogens. Trophic hormones, luteinizing hormone (LH) and follicle-stimulating hormone (FSH) regulates this system (reviewed by (Wilhelm *et al.,* 2007; McLaren 2000; Swain & Lovell-Badge 1999; Patek *et al.,* 1991).

In males the testicular cords contain Sertoli cells, which surround the primordial germ cells. Sertoli cells are the first cell type known to differentiate and are believed to act as the organizing centre of the male gonad as well as to orchestrate the differentiation of all other cell types. In fact, Sertoli cells are believed to induce peritubular myoid cells to come in place to form testicular cords. Sertoli cells also support the germ cell development in the male specific manner. The deeper layer of the mesenchyme forms the interstitial mesenchymal cells of the testes, further differentiating into testosterone-producing Leydig cells. In addition to the Sertoli and Leydig cells, there are peritubular myoid cells and endothelial cells along with other uncharacterized cell types in the interstitial space of the early testis. These cells move into the gonad from the mesonephros (Wilhelm *et al.,* 2007; Yao & Capel 2005; Yao *et al.,* 2004; Lovell-Badge *et al.,* 2002; Swain & Lovell-Badge 1999; Pieau *et al.,* 1999; Smith & Joss 1994b).

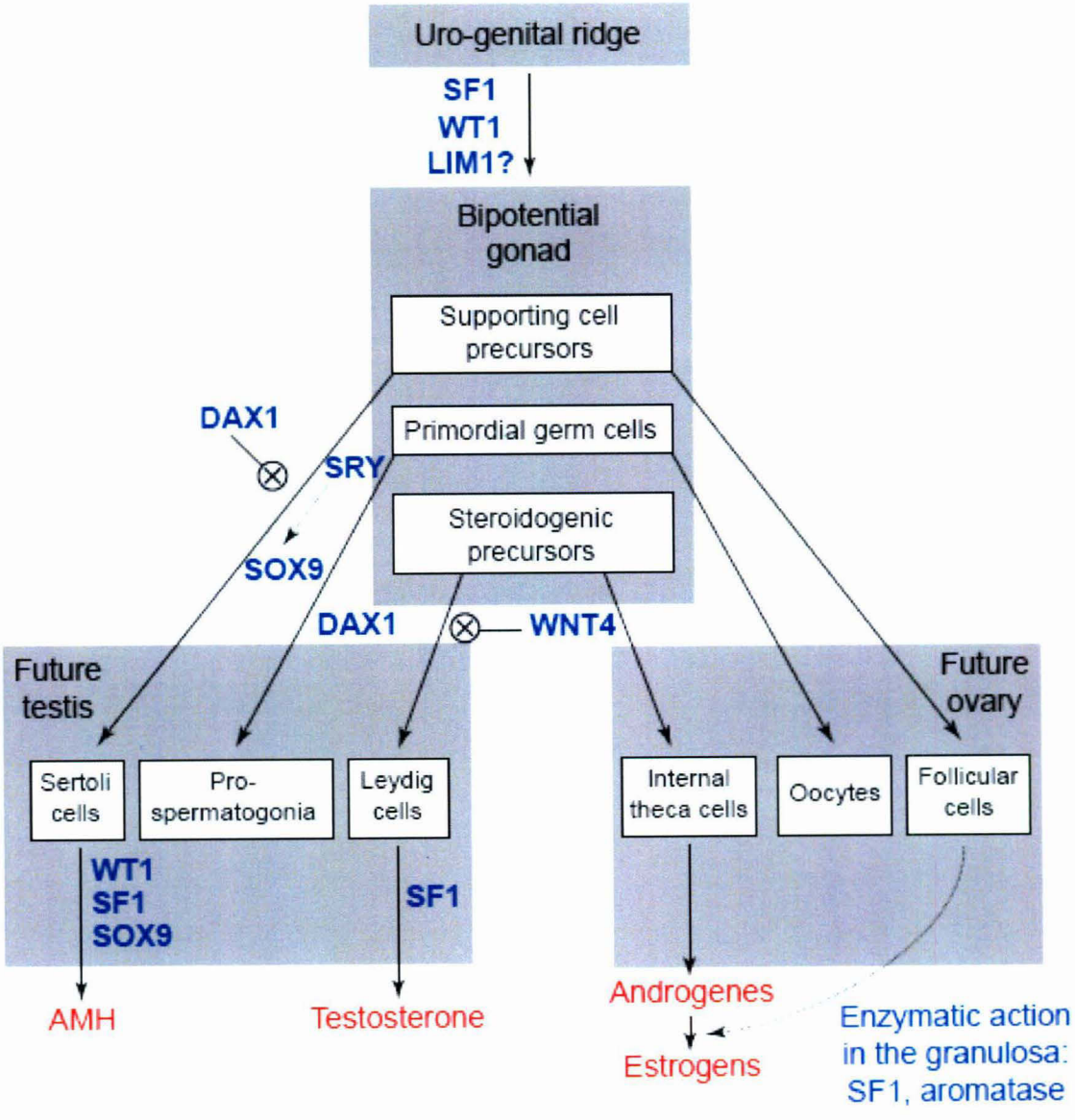


Figure 2.1: Differentiation scheme of the bipotential gonad into ovary and testis. (Adapted from Viaman and Pailhoux 2000).

2.2 Vertebrate Sex determination

Sex determining mechanism varies considerably across vertebrates. During early embryonic stages gonads appear as bipotential primordia. The undifferentiated "bipotential" embryonic gonads are committed to either testes or ovaries, is guided by genetic or environmental cues in metazoans (Sekido & Lovell-Badge 2007; Wilhelm 2007; Yao & Capel 2005). This process occurs both in GSD and TSD species.

2.2.1 Early events during gonad development

Once committed the bipotential gonad differentiate into either testis or ovary following complex cellular events programmed by an intricate regulatory network of genetic factors. Sex determining mechanisms and gonad development are briefly reviewed in the following sections.

2.2.1.1 Events at cellular level during early gonad development

Gonad development has been well understood in mammals wherein, mouse by far the most studied model organism. Events in other mammalian models follow the same basic pattern notwithstanding the differences in their anatomy and timing of the events.

Gonad development during embryonic stages can be divided into two phases. The initial phase is characterized by the development of gonadal ridge, which is bipotential and identical in male and female (Kim & Capel 2006). The second phase involves the development of testis or ovary from the bipotential gonad by the timely expression of some of the genes causing onset of sex specific cellular migration/organization (Wilhelm *et al.,* 2007). Though the sexual fate of the mammalian embryo is cast at fertilization, still the bipotential gonad develop and retains capacity to acquire testicular or ovarian fate until the expression of *Sry* (Ottolenghi *et al.,* 2007). The bipotential or the primordial gonads arise with in the immediate mesoderm, which comprises of three segments i.e., pronephros, mesonephros and metanephros, which develop into adrenal primordium, gonad and kidney respectively. The indifferent gonads grow as thickening on the ventromedial surface of the mesonephros, which is supplemented by the cells that migrate to the site of the developing gonads from the coelomic epithelium and from the adjacent mesonephros. The early mammalian gonad is composed of bipotential precursor cells including precursor cells for supporting and steroid cells. These Precursor supporting cells have an ability to differentiate into testis specific Sertoli cells or ovary specific follicle (granulose) cells (Palmer & Burgoyne 1991a; Palmer & Burgoyne 1991b)

and precursor steroid cells differentiate into steroid secreting cells (Leydig cells in males and theca cells in females) (Wilhelm *et al.,* 2007).

2.2.1.2 Development of ductal systems in early gonads

Male and female gonads are made up of the ductal systems; the primordia of these ductal systems are present in the developing mesonephros. The Male duct system is called Wolffian duct while the female duct system is called Mullerian duct. During the initial phase of the gonad development, both types of duct (Mullerian and Wolffian) run together in the mesonephros. However, only one of them develops and other regresses as the bipotential gonads progress to become testis or ovary (Warne & Kanumakala 2002).

2.2.1.3 Migration of the primordial germ cells to the genital ridge

Interestingly, the primordial germ cells (PGCs) do not arise at the spot where the genital ridge develop, but they migrate from the posterior end of the primitive streak (Lawson & Hage 1994). Migration starts from the 9.0 dpc (days post coitum) in mouse. By the time migration starts the posterior end of the primitive streak give rise to hind gut region. The PGCs travel through the hind gut and dorsal mesentery to reach to the site of developing gonads. During their migration, PGCs continue to divide and by the time these reach to the developing gonads their number increases to approximately 3000 cells. Origin and migration of PGCs can be tracked down by following expression of alkaline phosphatase. Survival of the PGCs during migration is dependent on the expression of pluripotency marker c-KIT (a tyrosine kinase receptor on the surface of the germ cells) and secretion of the stem cell factor (SCF) by the surrounding tissue (del-Stenzel *et al.,* 1998). Migration of PGCs is over by 11.5 dpc (del-Stenzel *et al.,* 1998). After coming in to genital ridge, the germ cells lose their motility and form an aggregate; however, they continue to proliferate. Till 13 dpc the germ cells maintain their bipotential state.

In case of male gonad development, the germ cells become completely encircled by the developing testicular cords by 13 dpc; hereafter they enter into meiosis and remain arrested as T1 prospermatogonia (in G_0/G_1). On the other hand, in developing female gonads, the germ cells proliferate for a longer duration (till 13.5 dpc) before entering into the meiosis, and remain arrested in the first meiotic prophase (Hilscher *et al.,* 1974). These PGCs can differentiate in to either prospermatogonia or meiotic oocytes, and importantly this differentiation is not dependent on the chromosomal sex of the PGCs themselves, but it is dependent on the chromosomal sex of the somatic cells in the

vicinity and their (PGCs) position in developing gonads. This implies that XY PGCs can grow as oocytes in the female embryos and XX PGCs can develop in to prospermatogonia in the male embryos (del-Stenzel *et al.,* 1998; Anderson *et al.,* 1998; Palmer & Burgoyne 1991a).

2.2.2 Development/differentiation of the gonads

2.2.2.1 Testicular development

Start of testicular differentiation in mouse is marked with *Sry* expression from 10.5 to 12.5 dpc, which peaks at 11.5 dpc. However, these are the pre-Sertoli cells which express *Sry* or in other words the cells expressing *Sry* differentiate in to Sertoli cells (Swain & Lovell-Badge 1999). This fact indicates the importance of this cell lineage in male gonad development. Sertoli cells are considered to be organizing centre of the male gonads and coordinate development/organization of the other cell types in the developing testis (Wilhelm *et al.,* 2007). Later on, Sertoli cells are found associated with the germ cells and nurture their development in to sperms.

The migration of the cells from coelomic epithelium and adjacent mesonephros starts from 9 dpc onwards (figure 2.2); these cells are progenitors of the Sertoli cells and Leydig cells in males and granulosa cells and theca cells in female (Albrecht & Eicher 2001). A subset of the *Sry* expressing cells become pre-Sertoli cells that delaminate from coelomic epithelium and migrate to the site of developing gonads (Capel *et al.,* 1999; Karl & Capel 1998). The Sertoli cells are the first cell type that differentiates in the gonads and therefore are the first indicator that gonad has passed from indifferent stage to testis development. *Sry* expression starts in the cells that have already delaminated from the coelomic epithelium and *Sry* expression is not the cause of their delamination from the coelomic epithelium. Later when *Sry* expressing cells start co-expressing *Sox9* (Wilhelm *et al.,* 2005), the pre-Sertoli cells become Sertoli cells. The Sertoli cells are polarized when they form epithelial aggregates to make testicular cords that occurs after 12.5 dpc (Wilhelm *et al.,* 2007). One remarkable feature about the Sertoli cells is their strong bias towards the presence of Y chromosome, an observation from a chimeric mouse of XX and XY genotype wherein the testis had 50% XX cells and 50% XY cells but 90% of the Sertoli cells were XY (Palmer & Burgoyne 1991a). These experiments indicate that Sertoli cells are the only cells that require cell autonomous expression of *Sry*. However, few recent experiments also suggest that the XX Sertoli cells which do not express *Sry* can be recruited as Sertoli cells when mixed with other prostaglandin D2

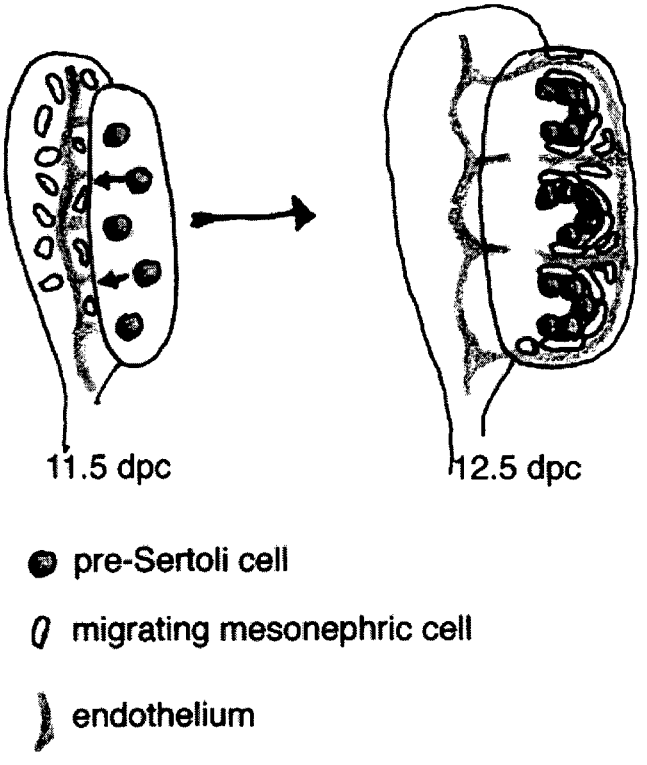


Figure 2.2: Schematic representation of male-specific cell migration in the mammalian gonad. At 11.5 dpc, pre-Sertoli cells of the XY gonad (solid arrows) signal endothelial cells and other cells in the mesonephros to begin migrating into the gonad. By 12.5 dpc, interactions between migrating mesonephric cells and pre-Sertoli cells establish testis cords. Migrating cells are represented in at least three lineages: peritubular myoid cells, endothelial cells, and cells associated with the endothelium (adapted from Tilmann and Capel 2002).

secreting Sertoli cells (Wilhelm *et al.,* 2005). It is now believed that the paracrine signaling *via* prostaglandin D2 may act as a backup mechanism in case of Sry dysfunction or it probably is a feedback mechanism for *Sox9* expression. It has been shown that at least 20% of the precursor cells should be XY in case of chimeric XX-XY mouse to make a male gonad (Patek *et al.,* 1991). This implies that 20% of the potential *Sry* expressing cells can induce 80% non Sry expressing precursor cells to develop in to Sertoli cells. All these observations strongly suggest that Sertoli cells have a very important role in the male gonad development.

While Sertoli cells develop, many other cells migrate from the adjacent mesonephros making gonads to grow bigger in size. However this migration is dependent on the Sry expression (Capel *et al.,* 1999). These immigrating cells from mesonephros give rise to peritubular myoid cells that enclose the Sertoli cells and complete formation of testicular cords, along with endothelial cells (that form male specific vasculature) and the steroidogenic Leydig cells. However, the genetic/cellular factors, responsible for inducing above migrations, are yet to be characterized.

Peritubular myoid cells (PM cells) form a layer of flattened cells surrounding the Sertoli cells demarcating the testicular cords. They strengthen the testicular cords and facilitate movement of the mature sperm through the semniferous tubules in adult testis to the seminal vesicle. Migration of the PM cells to the site of developing gonads marks the formation of the testicular cords by 12.5 dpc. Sertoli cells first differentiate and then aggregate around the germ cells (Jost 1972). Importantly, the formation of testicular cords is dependent on the PM cells migrating from the mesonephros as gonad explant cultures lacking mesonephros show disrupted testicular cords (Buehr *et al.,* 1993), but is not affected by the absence of the germ cells (McLaren 1991; Merchant 1975).

During later stages of testicular development, the testicular cords get separated from each other and give rise to interstitium, which harbour steroidogenic, Leydig cells. Leydig cells often lie in clusters close to the blood vessels. The Leydig cells develop twice in the life of a mammal, first during embryonic stages where androgen secretion causes foetal masculinization, then adult Leydig cells which develop after birth.

2.2.2.1.1 Primordial Germ cell development

PGCs in male gonad remain arrested as T1 prospermatogonia in G_0/G_1 stage till birth (Taketo *et al.,* 1984). However, germ cells play a more active role in Ovary development. Ovary does not show primary signs of differentiation in absence of germ cells. Germ

cells in the male gonads are not committed to spermatogenesis fate till 12.5 dpc, till the cells continue expressing *oct4*, a pluripotency marker that ceases to express after commitment of PGCs (Bowles *et al.,* 2006; McLaren & Southee 1997).

2.2.2.2. Embryonic Development of female gonads

Unlike rapid development of male gonads between 10.5 to 12.5 dpc, the female gonad development remains slow. Though expression of female-specific genes start from the 11.5 dpc onwards (Nef *et al.,* 2005), early signs of female gonad differentiation at cellular level is only between 14.5 dpc to 18 dpc. During this period germ cells come close together and cyst forms around them by epithelial flat layer of squamous pregranulosa cells. These cysts remain interconnected through cytoplasmic bridges. Development of primordial, primary and antral follicle occurs after birth (Wilhelm *et al.,* 2007; Swain & Lovell-Badge 1999).

2.2.3 Cellular events during early embryonic gonadal development in reptiles

Remarkable similarities are seen in the structure and function of mature gonads among vertebrates. In general, vertebrate gonads arise as bipotential organs, initially indistinguishable between male and females. However, studies have shown differences between mammalian and reptile gonad development. In contrast to mammalian gonads, where germ cells and somatic cells are mixed during bipotential stages itself, reptilian gonads have sex cords since beginning, which are made up of somatic cells distributed all over and germ cells remain restricted to the coelomic epithelium (Yao *et al.,* 2004; Wibbels *et al.,* 1991). The presence of loosely arranged undifferentiated sex cords in the bipotential stages in reptiles is therefore, a distinct feature from mammals and is probably related to how germ cells arrive and populate the gonad (Yao & Capel 2005). Germ cells are present in the interior of gonads in mouse, while in turtle they are present on the coelomic surface. This probably can be correlated to migration of the germ cells to gonads which is well studied in mouse (mammals) and chicken but remains to be worked out in reptiles. In mammals, germ cells migrate through the hindgut mesentery traversing the mesonephros to the interior of the gonads (del-Stenzel *et al.,* 1998), while in chicken, germ cells circulate through the vascular system, extravagate from the small vessels and then enter the neighboring thickened coelomic epithelium, which is connected to the gonads (Ukeshima *et al.,* 1987).

However, gonadal morphogenesis after germ cell migration to genital ridge is almost comparable between mouse and turtle. At MPT in turtle, the sex cords are strengthened and enclose germ cells into it and at FPT these sex cords are retracted back to medulla and are regressed (Yao *et al.,* 2004). In mouse, proliferation in gonads increases after Sry expression, which is also seen in gonads of TSD species where cellular proliferation increases at MPT (Schmahl *et al.,* 2003). Sex cords in mouse are made by the cellular migration from the coelomic epithelium and the adjacent mesonephros; however in turtle, lineage tracking experiments indicate that primitive sex cords formation occurs only by invagination of the coelomic epithelium (Yao & Capel 2005), and unlike mouse studies in turtle to track mesonephric cell migration were not successful. On the other hand, in a separate experiment where gonads were separated from the mesonephros during bipotential stages, sex cords were formed at MPT, establishing that mesonephric cell migration is not important for sex cord formation in the turtle (Moreno-Mendoza *et al.,* 2001). This study also suggested that it is only gonadal tissue that probably senses the temperature stimulus, and not by any other embryonic tissue like brain that was suspected to be involved.

2.3 Genotypic Sex Determination (GSD)

GSD is seen in all five phyla of vertebrates. However, most of our present knowledge has come from studies done on mammals, which have male heterogametic (XX-XY) system. In the mammals, sexual fate is determined genetically by the presence of the Sry gene on the Y chromosome (Sex determining Region of Y-chromosome). In Contrast, female heterogamety is found in birds and snakes having ZZ male and ZW female sex chromosomes. To date, it is not clear whether the sex-determining mechanism in ZW systems is a dominant ovary-determining gene residing on the W chromosome or whether a double dose of a Z-linked gene triggers testis determination (Clinton 1998).

Over the years, significant progress has been made in our understanding of the underlying mechanisms of GSD. The H-Y antigen (a minor histocompatibility antigen) was proposed as a tool to identify the heterogametic sex (Engel & Schmid 1981; Engel *et al.,* 1981). However, it gave conflicting results *e.g.* in case of a GSD exhibiting turtle species (*Siebernockiella crassicollis*) where the female was found to be H-Y positive despite cytogenetic studies confirming male as the heterogametic sex (Carr & Bickham

1981). During the last two decades efforts have been made to establish heterogamety by searching for sex specific DNA. Bkm satellite DNA characterized by repetitive GATA sequences was isolated from W-chromosome of the banded krait snake *Bungarus faciatus*. The Bkm-related sequences were shown to be linked to the W or Y-chromosome in many species, and using these DNA as probe led to elegant demonstration of sex reversal phenomenon in mice due to translocation of Y DNA on the X chromosome (Jones & Singh 1985; Singh & Jones 1982; Jones & Singh 1981). The sex specificity of Bkm-related satellite DNA was also suggested in few species of sea turtle and fishes (Nanda *et al.,* 1992; Demas *et al.,* 1990). Notwithstanding these positive studies, a general unequivocal role of Bkm sequences in GSD is yet to be demonstrated. Studies on XXY/XO and other sex chromosome abnormal individuals established that sex in human is not decided by the X chromosomes and/or autosomes ratio, and that Y chromosome contains a dominant allele for sex determination. Later it became obvious that development of testis is associated with presence of a Y linked locus called TDF (testis determining factor) in human or TDY (testis determining region on Y) in mouse. This was adjacent to the PAR (pseudo autosomal region), which is the only homologous region between X and Y chromosomes. As an abnormality, the pairing between PAR and X-chromosome may extend to the non-homologus regions and could result into transfer of Y specific DNA to the X chromosome. Study of four such XX males who had a 60 kb of Y DNA (on one of the two X-chromosomes) led to mapping of the testis determining gene SRY (sex determining region on Y) (Sinclair *et al.,* 1990), which was later confirmed to be major testis determining factor by studies on additional sex-reversal cases (Berta *et al.,* 1990). Similarly, counterpart of Sry in mouse as testis determining factor was also proven using loss and gain of function studies in different genetic backgrounds (Koopman *et al.,* 1991; Koopman *et al.,* 1990).

In humans, SRY mutations were found in patients with XY gonadal dysgenesis (Berta *et al.,* 1990). It was also demonstrated that XX mice transgenic for *Sry* develop as males with normal testis (Koopman *et al.,* 1991). *Sry* encodes a putative transcription factor that contains a high mobility group (HMG) box DNA binding protein. The timing of *Sry* expression in the mouse is consistent with a role in sex determination. *Sry* expression is detected in the genital ridge at 10.5 dpc; it peaks at 11.5 dpc and by 12.5 dpc its expression becomes feeble(Koopman *et al.,* 1990). *Sry* expression correlates with the proliferation of bipotential supporting cells, some of which are the precursors of Sertoli cells in the XY gonad (Schmahl *et al.,* 2000). Thus, there are many evidences to support SRY as a key factor necessary for sex determination in mammals. However, very little molecular information is available about the regulation of Sry. WT1 (Wilms' tumor 1

gene) +KTS, transcription factor GATA4 and FOG2 have been proposed in the transcriptional or post transcriptional regulation of the *Sry* (Hammes *et al.,* 2001)}.

Apart from SRY, autosomal gene SOX9 (Sry related HMG box) has been implicated in sex determination. SOX9 was originally mapped in campomelic dysplasia (CD) patients, which showed testicular dysgenesis leading to sex reversal in up to 75 % of XY individuals (Foster *et al.,* 1994; Wagner *et al.,* 1994). Human XX males were also identified with SOX9 duplication (Huang *et al.,* 1999). Studies on transgenic mice determined the precise function of SOX9 in testis development. Ectopically expressed *Sox9* driven by *Wt1* promoter in transgenic XX mice showed testes development consisting of apparently normal Sertoli and Leydig cells (Vidal *et al.,* 2001). Recently, it has been shown that XX mice, which lack estrogen receptors ERα and ERβ, express *Sox9* in granulosa cells of the ovaries just prior to their trans-differentiation into Sertoli cells (Dupont *et al.,* 2003). Thus SOX9 remains one potential candidate as a sex-determining gene. In Mice, *Sox9* is expressed at very low levels in the genital ridge of both sexes at 10.5 dpc, but soon after *Sry* expression, strong upregulation in *Sox9* expression is observed in male gonads while it is down regulated in female (Morais da *et al.,* 1996; Kent *et al.,* 1996). Recently, *Sox9* upregulation has been demonstrated by transient expression of Sry in Sertoli cells precursors (Sekido *et al.,* 2004). These observations suggest *Sox9* as a target for Sry, however, this possibility was finally proven where SRY binding on the Sox9 promoter is essential for *Sox9* expression (Sekido & Lovell-Badge 2008).

2.4 Genetic factors in GSD

To date, a number of genes have been implicated in the process of genetic sex determination in mammals. Potential role of some of such genes during gonadal development is now better understood by the genetic manipulation studies using mice (figure. 2.3). As indicated earlier, in GSD sex of the embryo is decided at the time of fertilization however manifestation of sex determination occurs later. Before sex determination of the developing gonad a bipotential gonad forms, which has no morphological difference between male and female. Several of these genes are required for the maintenance of the bipotential primordium (*Wt1, Sf1, Emx2, M33, Igf1r* and *Lhx9*) (Wilhelm *et al.,* 2007). Role of all these candidate genes are briefly discussed below.

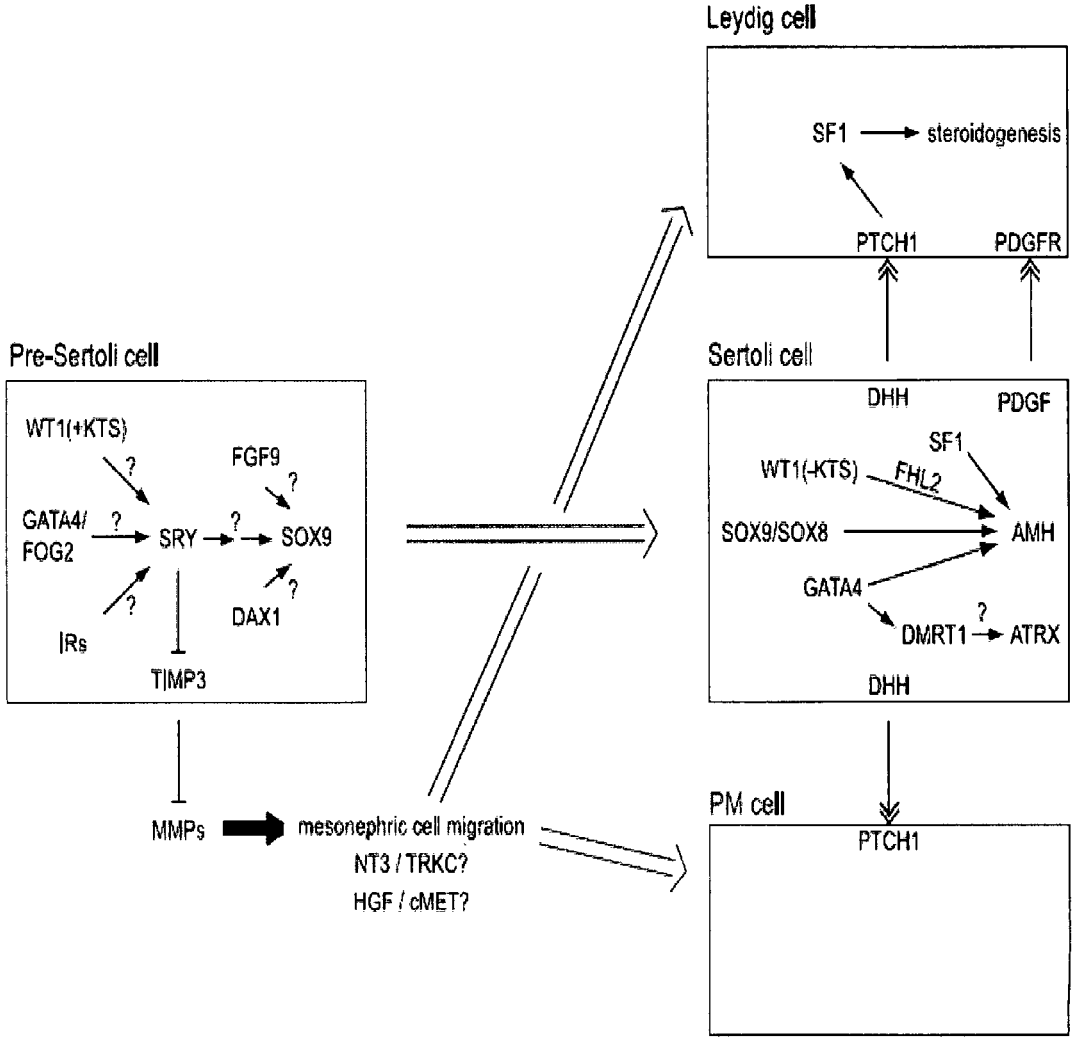


Figure 2.3: Postulated interaction of molecular players involved in early testicular development. Double-headed arrows indicate binding to a receptor while coloured arrows (blue, red, green) indicate differentiation of precursor cells into testis-specific cell types; black, bold arrow show gene important for cellular process (adapted from Wilhelm *et al.*, 2007).

2.4.1 Genes involved in early growth and survival of the gonad

2.4.1.1 *Emx2* (empty spiracles homologue 2)

Emx2 (empty spiracles homologue 2) is a homeobox gene. Its absence also does not affect the migration of PGCs. Morphologically, the emx -/- gonads lack the thickening of the coelomic epithelium that is the first sign of gonadal development; and Mullerian and Wolffian ducts never form. These mutants lack the gonads like *Sf1* and *Lhx9* null mutants, and no development of the genital tracts is seen (Miyamoto *et al.,* 1997). However nothing is known about the regulation and downstream targets of *emx2.*

2.4.1.2 *Lhx9* (LIM homeobox protein 9)

Lhx9 (LIM homeobox protein 9) is a homeobox gene, which is expressed in urogenital ridges of mice at 9.5 dpc much earlier to sexual differentiation. After morphological differentiation of gonads, transcripts of *Lhx9* localizes in the interstitial region. In *Lhx9*$^{-/-}$ XX and XY mice, genital ridge fails to proliferate and a discrete gonad is not formed. Moreover *Sf1* (described below) expression was detected at minimum level in Lhx9-deficient genital ridge, indicating that it may have a role preceding *Sf1* in developmental cascade (Birk *et al.,* 2000). Lhx9 acts upstream to *Sf1* and binds directly to the *Sf1* promoter and has additive effect to the Wt1 induced activation in vitro (Wagner *et al.,* 2002).

2.4.1.3 *Igf1r* (insulin like growth factor 1 receptor), *Irr* (insulin receptor-related receptor) or *Ir* (insulin receptor)

The role of insulin receptor tyrosine kinase family members *Igf1r, Irr* and *Ir* has been shown in the male sexual differentiation. Triple knock out for all the three receptors led to a completely female phenotype in XY mice. Reduced expression of both Sry and Sox9 in these mice indicates that insulin signaling pathway is essential for male sex determination (Nef *et al.,* 2003).

2.4.1.4 *M33*

M33 is a mouse homologue of polycomb genes in drosophila. Disruption of carboxy-terminal region of M33 in mice ($M33^{cterm}$/ $M33^{cterm}$) showed a high rate of male to female sex reversal. The defect of gonadal growth was observed near the timing of *Sry*

expression indicating the role of *M33* upstream to *Sry* (Katoh-Fukui *et al.,* 1998). Recently M33 was found to regulate expression of *Sf1* in the adrenal gland and spleen (Katoh-Fukui *et al.,* 2005). It is expected to play a similar role during the bipotential gonad development; however, its extracellular roles and molecular functions are yet to be studied.

2.4.1.5 *Sf1* (Steroidogenic factor 1)

Loss of *Sf1* leads to complete failure of adrenal and gonad development, obesity and abnormalities of ventromedial pituitary and hypothalamus gonadotropes (Shinoda *et al.,* 1995; Luo *et al.,* 1994) The gonads of *Sf1* mutant embryo do not develop beyond the early indifferent stage. SF1 is shown to interact with *Sox9* promoter along with SRY and it also interacts with Sox9 and is proposed to be involved in Sertoli cell specific expression of AMH in the XY gonad (Sekido & Lovell-Badge 2008; de Santa *et al.,* 1998). *Sf1-/-* mutants show XY sex reversal wherein Mullerian ducts develop into uteri, oviducts, and upper vagina. All these studies suggest an important role of *Sf1* in the formation of the bipotential gonads and later male sexual development (Wilhelm *et al.,* 2007).

.2.4.1.6 *Wt1* (Wilms Tumor Suppressor)

Wt1 is a zinc finger gene that was initially identified as a gene involved in pediatric kidney cancer. *Wt1* is expressed in the embryonic kidneys, urogenital ridge, developing gonads, adrenal glands heart and brain etc. It is very important for the early embryonic development of the Kidney, gonads and adrenal, as *Wt1* homozygous knockouts mice are embryonic lethal (Kreidberg *et al.,* 1993), lack kidneys, gonads, adrenal glands and defects in heart formation. Gonads in the Wt1 knockout mutants developed only till 10.5 dpc (bipotential state) and did not develop after this point. However, the Wt1 mutation does not affect the germ cell migration (Moore *et al.,* 1999; Kreidberg *et al.,* 1993). In mouse *Wt1* is known to produce a total of 24 isoforms generated by combination of alternative splicing, alternative translation start site and RNA editing. However, two isoforms are of particular importance in the sex determination, which on the basis of presence/absence of codons for three amino acids (KTS) between exon 8 and exon 9, are called +KTS and –KTS isoforms. Wt1 in mammalian gonads normally remain localized to the coelomic epithelial cell layers, and developing Sertoli cells in males or granulose cells in females (Mundlos *et al.,* 1993). Wt1 encodes a protein that acts as a transcription activator, as well as, transcriptional suppressor depending upon the cell

type and the promoter context. Wt1 has been shown to be a potential activator of the Sf1, and thus was considered to be important for the maintenance of the bipotential primordium (Wilhelm & Englert 2002). The two *Wt1* isoforms perform different functions during the gonad and kidney development. The -KTS isoform is more active transcriptional regulator, while the +KTS isoform is shown to localize with the splicing speckles (Ladomery *et al.,* 1999; Larsson *et al.,* 1995). Biochemical studies show that the two Wt1 isoforms have different interacting partners: *e.g.*, the +KTS isoform shows stronger affinity for the splicing factor U2AF65 (Davies *et al.,* 1998), whereas -KTS isoform binds preferentially to steroidogenic factor Sf1 (Nachtigal *et al.,* 1998). Subsequently, NMR relaxation experiments demonstrated that the insertion of the KTS sequence leads to an increased linker flexibility and loss of DNA binding of zinc finger 4, which results in the differential affinity of +KTS and -KTS variants to DNA (Laity *et al.,* 2000).

Detailed studies of mouse knockouts for the +KTS and –KTS isoforms revealed that both *Wt1* isoforms play an important role in kidney and gonad development and homozygous deletion of either of the isoforms was proven to be lethal after birth due to severe kidney abnormalities (Hammes *et al.,* 2001), whereas heterozygous deletion provides better understanding of the function of individual isoforms. In the +KTS heterozygous deletion mutants the ratio of +/- KTS isoforms was shifted towards the – KTS isoform, and in – KTS isoform deletion mutants the ratio was shifted towards the +.KTS isoform. These mutants mimicked the Denys-Drash Syndrome and Frasier Syndrome respectively. The +KTS ablated mice showed increase in the stromal tissue and decrease in the tubular epithelium. Immunolocalization of Pax2 and 2α integrin in the +KTS deficient mice showed malformation of the glomerular structure with the red blood cell leaking into Bowmen's space, indicating an overall defect in the formation of glomerular tuft. On the other hand in –KTS ablated mice, kidney development was severely impaired with marked reduction in size accompanied with the increase in the stromal tissues, decrease in the nephrogenic zone and macroscopic hemorrhages in the glomerular zone. However, analysis of gonads in these homozygous mutant mice, each of the isoforms showed their sex-specific role. The –KTS isoform knockout showed streaked gonads without any differentiated tissues, defects in the genital duct formation and relatively low expression levels of *Dax1*. In contrast, +KTS ablation in XX or XY mouse led to development of female gonads with no apparent differences in the expression levels of early sex determining genes like *Dax1, Sf1* and *Lhx9* with wild type mouse. However, localization of Dax1 in the +KTS ablated XY mice was female specific suggesting sex reversal at molecular level. Similarly *Sox9* localization was also found to be female

specific in the +KTS ablated mice, with its expression dropped to a level which is sufficient to bring about sex reversal in humans (Foster *et al.,* 1994; Wagner *et al.,* 1994). Moreover *Sry*, levels was also significantly dropped in the ~ 25% +KTS ablated animals. No such effects were seen in –KTS ablated animals. The results thus suggest that +KTS isoform of *Wt1* has an important role in the male sex determination, operating even upstream of *Sry*; while –KTS isoform was found to be generally required for the cell survival (Hammes *et al.,* 2001).

2.4.1.7 SRY (Sex determining region on Y)

It is now well evident that a number of genes are involved in the formation and maintenance of the bipotential primordium both in male and female; and a subset of genes are required for the early differentiation of gonads once the sex is determined. Sry is found to be one master regulator gene in mammals, expression of which is necessary to decide the testicular fate of the developing bipotential gonad. *Sry* expression is observed in a wave like manner, which starts from 10.5 dpc, peaks by 11.5 dpc and vanishes by 12.5 dpc in male gonads (Morais da *et al.,* 1996). It has a characteristic DNA binding High mobility group (HMG) domain, which binds to DNA in a sequence specific manner and induces a sharp bend of 60-85^0. Biochemical and mutation analysis have revealed that DNA binding and bending are integral parts of *Sry* function. All the mutations on *Sry* leading to XY sex reversal are seen in the HMG box region of Sry (Harley *et al.,* 2003). In order to understand the *Sry* function at molecular level efforts have been made to identify interacting proteins as well as target genes. To date, only two interacting partners of SRY have been identified: a novel protein containing only a Kruppel associated box (KRAB) domain (Oh *et al.,* 2005; Thevenet *et al.,* 2005) and SIP1 (SRY interacting protein 1). KRAB-O protein interacts with HP1 (heterochromation Protein 1) thus, leading to gene silencing (Oh *et al.,* 2005). However function/role of these interactions in sex determination is yet to be determined. Apart from interacting partners, our understanding about the Sry targets was generally wanting; although *Sox9* was presumed its target. This was recently proven by a study from Lowell-Badge and colleagues that Sry binds to *Sox9* promoter along with Sf1 and both were required for *Sox9* expression.

2.4.2 Genes involved in gonadal differentiation

Organogenesis of testis and ovary must require the integration of many signalling molecules and network of genes, most of which remain unidentified. As described above,

many key genes implicated in sex determination pathways, were identified by analysis of phenotype and mapping sex-reversing mutations in the human population. Fortuitous discoveries in targeted mouse mutants have also led to identification of many new genes and their respective functions. Characterization of morphological events in testis organogenesis has helped in prediction of several signalling pathways by analogy to similar developmental systems. Search for the novel members in the pathway has also been initiated by adopting several expression screens that compare XX and XY gonads during development. These strategies have led to the identification of many dimorphically expressed genes. However, their functional attributes in gonad development remains a challenge (Menke & Page 2002; Koopman *et al.,* 2002). Early gonad differentiation is largely conserved among vertebrates; particularly among mammals a common pathway downstream to *Sry* is seen. However, extent of this conservation is still to be defined. Some of the important genes involved in the early gonad differentiation are discussed below.

2.4.2.1 *ARX* (aristaless-related homeobox gene)

Knockout mouse for X-linked aristaless-related homeobox gene (Arx) is shown to have abnormal testicular Leydig cell differentiation in developing testis. The gene is thus thought to be involved in the pathway that promotes Leydig cell differentiation (Kitamura *et al.,* 2002).

2.4.2.2 *Dax1* (dosage-sensitive sex reversal gene)

Dax1 is located on the X-chromosome; it is a member of orphan nuclear receptor super family and is also called NR0B1. DAX1 interactions has been demonstrated with Sf1 (Suzuki *et al.,* 2003; Ikeda *et al.,* 1996), estrogen receptors α and β (Iyer *et al.,* 2006; Zhang *et al.,* 2000), androgen receptor and progesterone receptor (Agoulnik *et al.,* 2003) for recruiting co-repressors to the above transcriptional units. *Dax1* expression is found in developing adrenal glands, gonads, hypothalamus and the pituitary (Hanley *et al.,* 2000). In the developing testis, *Dax1* expression starts in the somatic cells before and shows a strong upregulation in Sertoli cells till it ceases by 12.5 dpc. Subsequently, a second phase of upregulation is seen in interstitial cells between 13.5 to 17.5 dpc. In contrast, in the developing ovary *Dax1* is expressed till 14.5 dpc (Ikeda *et al.,* 1996).

Dax1 was originally thought to be an 'anti-testis' factor. Duplication of *Dax1* resulted into XY sex reversal (Swain *et al.,* 1996; Bardoni *et al.,* 1994). Overexpression of *Dax1* in mouse, showed male to female sex reversal, in which *Sry* expression was delayed or

weakened (Swain *et al.,* 1998). Loss of function study on the other hand, showed no effect on the female development, but a complete impairment of testicular development in XY mice (Meeks *et al.,* 2003b). It is now known that Dax1 acts downstream or in a parallel pathway with *Sry* and promotes organization of testis cords (Meeks *et al.,* 2003a; Meeks *et al.,* 2003b). Recent evidences also suggest that Dax1 acts along with Sry in the up-regulation of *Sox9* (Bouma *et al.,* 2005).

2.4.2.3 *Fgf9* (Fibroblast growth factor 9)

Fibroblast growth factor 9 (*Fgf9*) is widely expressed in mouse embryo (Colvin *et al.,* 1999). It is a secreted signaling molecule implicated in the Sertoli cell proliferation and differentiation. *Fgf9* is member of fibroblast growth factors family which is responsible for embryogenesis of several organs like lungs, limbs and anterior pituitary (Ornitz & Itoh 2001). *Fgf9* expression is found in both male and female developing gonads during the bipotential stage; however, during later stages its expression becomes restricted to the testicular cords in the male gonads (Schmahl *et al.,* 2004). *Fgf9*$^{-/-}$ XY mice showed a phenotype ranging from testicular hypoplasia to complete sex reversal. Expression of male markers downstream of *Sry* was not observed in these mice. Since *Fgf9* expression begins shortly after *Sry*, it is believed to have a role in Sertoli cell proliferation and differentiation, however XX gonads develop normally (Schmahl *et al.,* 2004; Colvin *et al.,* 2001) A recent study now conforms that *Fgf9* expression is required for the cellular proliferation (in the developing male gonads) as well as Sertoli cell differentiation, which implies that it plays a dual role during the male gonad development (Kim *et al.,* 2007).

2.4.2.4 Gata4 (GATA binding protein4)/Fog2 (Friend of GATA; zinc finger protein, multitype2)

The phenotype of *Fog2*$^{-/-}$ *Gata4*$^{ki/ki}$ (Gata4^{ki}, a V217G amino acid substitution) XY mice showed reduced expression of *Sry*. Three genes crucial for normal Sertoli cell function (*Sox9, Mis* and *Dhh*) and three Leydig cell steroid biosynthetic enzymes (p450scc, 3βHSD and p450c17) were not expressing in these mice. *Wnt4*, a gene required for normal ovarian development was expressed ectopically. *Wt1* and *Sf1*, which are express prior to *Sry,* were found expressing in these gonads. Based on these observations, GATA4/FOG2 have been implicated in promoting *Sry* expression and differentiation of Sertoli cell lineage (Tevosian *et al.,* 2002).

2.4.2.5 *Sox9*

Sox9 is an important gene implicated for its role in sex determination and early gonad differentiation *Sox9* is expressed in sexually dimorphic manner, upregulated during Sertoli cells differentiation and shows highly conserved protein structure among all vertebrates. Moreover, it appears that *Sox9* role in sex determination is also conserved through vertebrates evolution despite differences in switch mechanisms underlying sex determination *e.g.* SRY in mammals except for mole vole (Just *et al.,* 1995), ZW chromosome gene(s) in bird (Oreal *et al.,* 1998) and incubation temperature of egg in turtles and alligator (Western *et al.,* 1999; Moreno-Mendoza *et al.,* 1999). These observations together support the view that *Sox9* is a critical sex-determining gene in vertebrates regardless of the initial switch being genetic or environmental. *In vitro* studies suggest that pig *Sox9* cDNA transactivates *Sry* promoter but pig SRY did not. Potential binding site for Sox9 in *Sry* promoter is conserved in human and pig but not in mouse. When this site was mutated in pig, the Sox9 transactivation effect was reduced to 70 %. In the light of these results, feedback transcription of pig *Sry* was proposed for Sox9 (Daneau *et al.,* 2002). In addition, Sox9 is also shown to regulate AMH activity. It is shown that SOX9 binds to AMH promoter and cooperates with SF1 to activate transcription of AMH (de Santa *et al.,* 1998). In mice, males with homozygous mutation in Sox9 binding site of *Amh* promoter do not transcribe *Amh* resulting in pseudohermaphrodites. These *in vivo* observations suggest that one of the downstream targets of SOX9 may include Amh (Arango *et al.,* 1999).

2.4.2.6 AMH (Anti-Mullerian hormone; also referred as Mullerian inhibiting substance, MIS)

AMH is a member of beta superfamily of growth and differentiation factors. The principal function of AMH is proposed to induce regression of the Mullerian duct (the anlagen of female reproductive tract) during male sex differentiation ((Behringer *et al.,* 1994). *Amh*$^{-/-}$ XY gonads showed normal Sertoli cell differentiation (Ross *et al.,* 2003). Despite the observation that *Amh* is expressed in early testis development, its specific role in male gonad development is yet to be identified.

2.4.2.7 *Dmrt1* (doublesex mab-3 related transcription factor1)

The *doublesex* and *Mab-3* related transcription factor 1 (*Dmrt1*) is a member of protein family that shares a DNA binding motif called DM domain. The *doublesex* in drosophila and *Mab-3* in *Caenorhabditis elegans* are involved in sex determination. *Dmrt1* is one of

the most conserved sex determining gene that is involved in process of sex determination and differentiation in all vertebrates from fish to mammal (Swain 2002; Ottolenghi *et al.,* 2002). It is demonstrated to be a master regulator for sex determination in teleost *Medaka* (Matsuda *et al.,* 2002), and in *Gallus* where its presence on the Z chromosome makes it an important candidate for the master regulator (Raymond *et al.,* 1999). In all the vertebrates examined till date (TSD or GSD), *Dmrt1* expresses at higher levels in the developing male gonads (Wilhelm *et al.,* 2007; Hodgkin 2002). However in mammals *Dmrt1* does not seem to be involved in sex determination but plays important role in testis differentiation. A knock out study shows that Dmrt1 is required for the postnatal testis differentiation and is completely dispensable for the ovarian development (Raymond *et al.,* 2000). In comparison, *Dmrt1* seems to play a more early and important role in the non-mammalian species where *Sry* is absent.

2.4.2.8 *Wnt4*

Wnt4 is a member of Wnt family of secreted molecules, which was originally identified as mammalian homologues of the drosophila wingless gene. *Wnt4*$^{-/-}$ XX gonads showed masculinized characters demonstrated by the absence of Mullerian duct (presumptive female reproductive tract) and the presence of Wolffian duct (presumptive male reproductive tract). In addition, expression of steroidogenic enzymes required for testosterone production was also seen in *Wnt4*$^{-/-}$ females (Vainio *et al.,* 1999). This gene, therefore, seemingly plays a crucial role in female development pathway.

2.4.2.9 *FoxL2* (forkhead helix family transcription factor L2)

FOXL2 is a member of forkhead/winged helix family of transcription factors, which play an important role in the early development (Kaufmann *et al.,* 1996). *FoxL2* is expressed in female specific manner after 12.5 dpc. Early female specific expression of *FoxL2* is conserved among several vertebrate species (Cocquet *et al.,* 2002). It is expressed in mesenchymal pre-granulosa and granulosa cells and its expression ceases after birth (Schmidt *et al.,* 2004). Knockout studies in mouse have showed that FOXL2 is essential for the pre-granulosa cells maturation thus for the maintenance of ovary development (Ottolenghi *et al.,* 2005; Uda *et al.,* 2004). FOXL2 deletion causes an early folliculogenesis thus leads to ovarian failure. Despite its important role *FoxL2* is not considered to be candidate for the ovarian determining factor.

2.4.2.10 Ovary determining genes

Ovary determining genes are yet to be identified. There is also lack of information about the genes expressed since early stages of female gonad determination/differentiation. However, recent large scale transcriptional analysis of the early gonads indicates, a robust female specific genetic program far earlier than expected (Bowles *et al.,* 2006; Nef *et al.,* 2005; Bouma *et al.,* 2005). Further study of these genes is expected to elucidate the regulatory network for the determination and early differentiation of female gonad.

2.4.3 Cross Talks between Sertoli cells and other cell types during gonadogenesis

As discussed in the earlier section, the Sertoli cells are organizing centre of the male gonad development. At cellular level, there seems to be a lot of cross talks between somatic cells, which is important for the testicular development. The most studied molecule secreted by Sertoli cells is Amh (Anti Mullerian Hormone). Amh is a protein of TGF-β superfamily. The Amh receptor is expressed in mesenchyme surrounding the Mullerian duct in both male and female gonads, and in accessory cells like Sertoli and granulosa cells. However a loss of function study for the AMH has shown intact Mullerian ducts but no effect on the testis development, therefore is was concluded that AMH does not play a very important role in testis development (Behringer *et al.,* 1994).

Dhh and Pdgrα are among the other important molecules secreted by Sertoli cells. However, genes involved in the Sertoli cell mediated function which induce mesonephric cell migration, cell proliferation and the mitotic arrest of the germ cells by 12.0 dpc are still to be explored

2.4.3.1 *Dhh* (Desert Hedgehog)

Desert Hedgehog (Dhh) is a signaling protein, which is produced by Sertoli cells. *Dhh* expression starts from 11.5 dpc in somatic cells and thereafter in Sertoli cells. *Dhh*$^{-/-}$ XY gonad shows defective differentiation of fetal Leydig cells and testis cord formation. The role of *Dhh* is proposed in maintaining high levels of *Sf1* in pre Leydig cells (Yao *et al.,* 2002).

2.4.3.2 *Pdgfra* (platelet derived growth factor-alpha)

Pdgfr-alpha is a member of platelet derived growth factor receptor family proteins. *Pdgfra*$^{-/-}$ XY gonads display disruption in the organization of interstitial cells and testis cord. These mice also show defect in the development of Leydig cells. Accordingly, the gene has been implicated downstream of Sry in the cord formation and Leydig-cell differentiation (Brennan *et al.,* 2003).

2.5 Environmental/Temperature dependent Sex determination

ESD signifies that the environment experienced during early development determines the sex of an individual irreversibly. The existence of ESD in a species can be experimentally established when sex ratio vary significantly according to the environment in which the developing embryos are reared. A wide range of environmental cues has been shown to influence the sex determination process. For instance, factors such as osmotic stress, level of oxygen and carbon dioxide may also affect sex determination in turtles (Gutzke & Paukstis 1983; Ackerman 1981a; Ackerman 1981b). Some studies have also indicated effects of pH and social conditions on sex determination in some fish species ((Holmgren & Mosegaard 1996; Romer & Beisenhertz 1996; Francis & Barlow 1993). A classical example of social condition based ESD is seen in the worm *Bonellia viridis* (Echiura: Bonelliidae). In this marine invertebrate species, vast majority of undifferentiated larvae metamorphose into males when exposed to females but differentiate into females in the absence of females (Jaccarini *et al.,* 1983). Among various environmental factors, influence of temperature on sex determination (TSD) has been most widely reported in a number of vertebrate species; especially a number of reptilian species have been shown to employ temperature dependent sex determination (TSD). In addition, few species of birds, fishes and amphibians have also been reported to have TSD (Valenzuela & Lance 2004). It was in 1966, when the first report suggesting that temperature could influence the sex ratio was published from observations in a lizard from West Africa, Agama agama (Charnier 1966). The second report documenting TSD in a turtle species (*Chelydra sepentina*) was published a few years later, which greatly stimulated interest in this topic (Yntema 1976). By the end of 1970s a number of publications from different group showed that TSD is widespread in reptiles and it occurs in natural as well as under artificial conditions. It was found to occur

in all crocodilians and tuataras, and in many turtle species, (Valenzuela & Lance 2004). However, a vast majority of reptilians are yet to be ascertained for the type of sex determination they undergo (Janzen & Paukstis 1991).

In TSD it is the temperature of the surroundings that act as trigger for the sex determination during later stages of embryonic stages. The specific period of development, when temperature trigger is taken, is called temperature sensitive period (TSP). Temperature acts as trigger in at least three patterns that are characterized by the sex ratios produced as function of temperature. In two lizards (*Agama agama*, *Eublepharis macularius*) and in alligators, originally it was described that eggs incubated at low temperature give rise to 100% females, and eggs incubated at high temperature give rise to 100% males (F→M). In many turtles it is the other way round: 100% males at low temperature and 100 % females at high temperature (M→F). In other turtles and crocodilian species including Indian mugger, incubation at specific intermediate temperature leads to 100% males, whereas both low and high temperatures lead to only females (F→M→F) (Lang *et al.,* 1989; Bull 1983). Interestingly, more recent elaborate studies (Valenzuela & Lance 2004) based on larger ranges of temperature have also revealed a F→M→F pattern of sex determination even in *Eublepharis macularius* and *Alligator mississippiensis*, unlike previous reports. It is possible that similar studies with warmer temperatures than those used in earlier studies, might well yield the similar pattern in *A. agama*. Therefore, most of the lab/field based egg incubation studies at specified temperature range suggest two overwhelming patterns in TSD species (figure. 2.4). These as a consensus have been classified as TSD-Ia, TDS-Ib and TSD-II. TSD-Ia is called as male-female (MF) pattern wherein lower temperature favours males development and higher temperature favours female development, while in TSD-Ib the opposite is true i.e. lower temperature favours female development and higher temperature favours male development and it is called female–male (FM) pattern. In contrast, TSD-II is called female-male-female pattern where lower and higher temperatures favour female development and it is only the intermediate temperature that favours the male development. The incubation temperature that yields 1:1 sex ratio is called pivotal (threshold) temperature. The range of temperature at which both male and female embryos are hatched is called Transitional range (TR) (Mrosovsky & Pieau 1991). Accordingly, TSD-Ia/TSD-Ib have one pivotal temperature, while TSDII has two pivotal temperatures. The individual TSD species broadly follow one of the above patterns, however range of male and female promoting temperature, sex ratios pivotal temperatures etc. vary between species (Lance 2003).

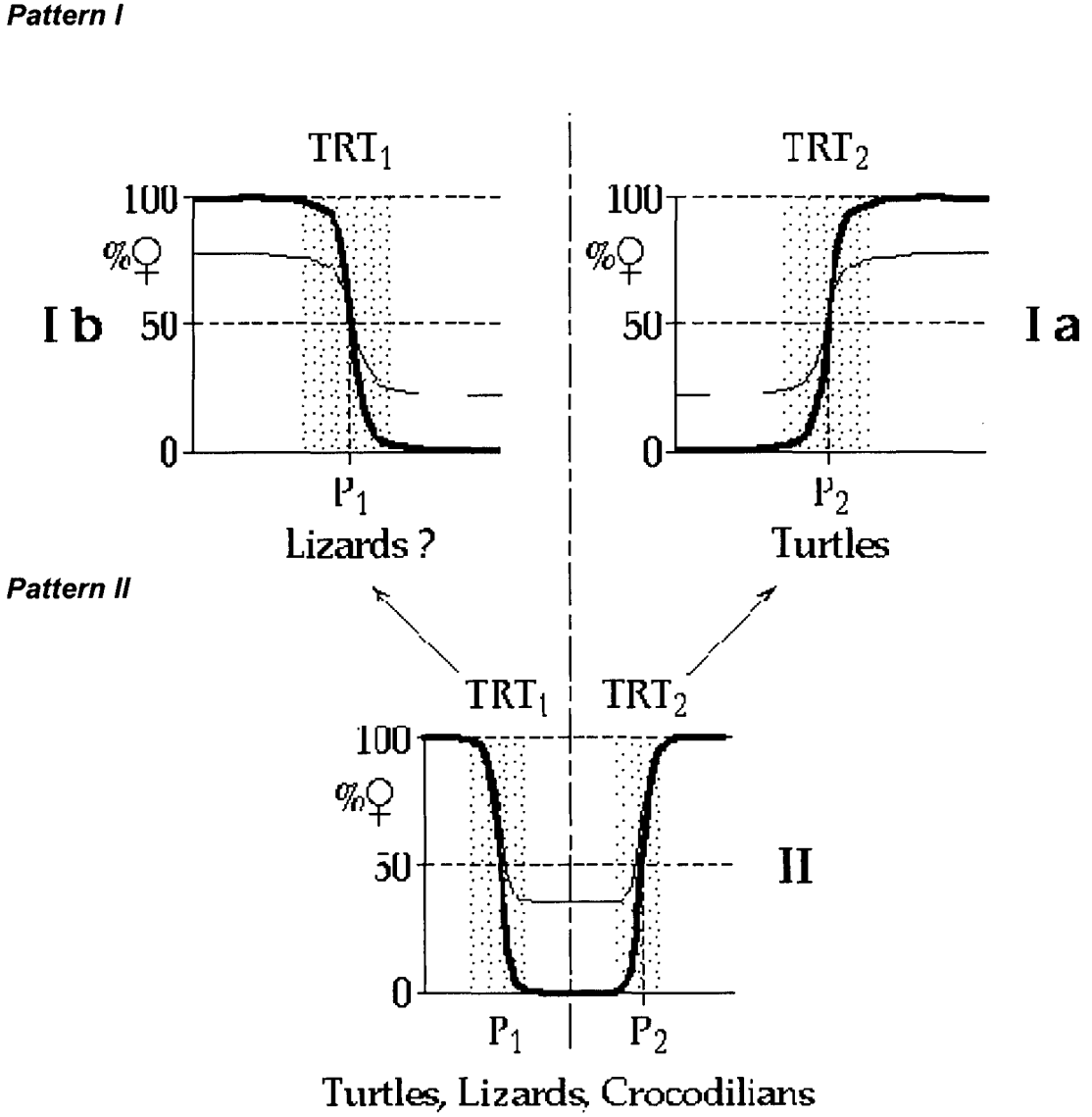


Figure 2.4: **Two most common patterns of TSD in reptiles:** In pattern I, there is only one transition zone or pivotal temperature (represented by TRT) that corresponds to the temperature supporting to a population wide 1:1 ratios of sexes. Pattern I can be classically divided into Ia (MF; Male - Female) and exactly opposite 1b (FM; Female -Male. In contrast, two pivotal temperatures are observed in pattern II. In the graph, temperature increases from left to right (adapted from Valenzuela and Lance 2004).

Most of the members of three reptilian groups (crocodilians, lizard, turtles) that have been known to exhibit TSD follow TSD-II pattern with the exception of sea turtles and in some of the emydid pond turtles (*e.g. Chrysemys, Emys, Trachemys*), wherein pattern TSD-Ia has largely been documented. Furthermore, in most of the F→M→F cases (TSD-II), there is only a very narrow temperature range at which males or both sexes are produced. Transition zone can be defined as intermediate pivotal temperature at which 50% of each male and female individual can be obtained. As mentioned above this transitional range of temperature may differ in different TSD species from being less than 1°C (marine turtle *Dermochelys coriacea*) to about 4°C (in *Chelonia mydas*, another marine turtle) (Pieau 1996). Moreover, transition zones or threshold temperatures may also vary sometimes intraspecifically. For example, in one study on the turtle *Chelydra sepentina*, 88 % males were produced at 20°C, but three subsequent clutches produced 100 % males at the same temperature (Janzen & Paukstis 1991).

2.5.1 Mixed sex determination

Sometimes a combination of environmental and genotypic sex determination, with major genetic factors, can be present within the individuals. Examples are the fish species *Menidia menidia* (Conover & Heins 1987) and *Limanda yokohamae* (Goto *et al.,* 2000), and the turtle *Emys orbicularis* (Girondot *et al.,* 1994; Zaborski 1986). There are some reptiles with ESD, which also show signs of heterogamety (Ewert *et al.,* 1990; Demas *et al.,* 1990; Nakamura *et al.,* 1987; Wellins 1987). Similarly, a form of TSD has also been reported in poultry (Ferguson 1994a; Ferguson 1994b), while all birds are known to have ZZ/ZW sex chromosomes.

2.5.2 Transition between GSD and TSD

Evolutionarily, no distinct line can be drawn between TSD and GSD. We are yet to find answers to questions like which mechanism of sex determination was primitive and how they evolved. Even the prediction of predominant direction of change can be very difficult with the fact that distribution of TSD is haphazard across the whole reptilian phylogeny. At the level of order, crocodilians have TSD, whereas their most closely related taxon, the birds do not. At the level of suborder, snakes have GSD with a ZW chromosome sex determining system whereas lizards display both GSD and TSD. Likewise at family and genus level too, the variation in sex determining mode is observed in many turtle species. Thus, at this point it would appear that both GSD and TSD have evolved a number of times independently in the reptilian species. Although many hypotheses have

been proposed on the evolutionary advantages of TSD in reptiles, its adaptive significance remains controversial (Sarre *et al.*, 2004; Shine 1999). Some of the hypotheses fall into a few categories (Janzen & Phillips 2006).

Phylogenetic inertia: TSD is an ancestral form of sex determination with no current adaptive significance. However, biggest problem before this hypothesis is why TSD is selected among a huge number of species.

Inbreeding avoidance: by resulting in single-sex clutches from natural nests, TSD reduces deleterious effects of inbreeding among siblings.

Differential fitness: The most plausible theoretical models for the evolution of TSD rely upon differential fitness of male and female offspring incubated under different thermal conditions (Charnov & Bull 1977). TSD can enhance maternal fitness by enabling the embryo to develop as the sex best suited to those incubation conditions. Many experiments have supported this hypothesis including the a recent work (Warner & Shine 2008)

In contrast, the possibility that GSD has evolved to ensure a balanced male/female ratio as the global temperature change can skew the sex ratio of TSD animals has also been examined recently. The study surveys the prevalence of GSD/TSD through vertebrate evolution (figure 2.5) and suggests the possible role of TSD mechanism in the demise of long extinct dinosaurs (Miller *et al.*, 2004).

2.5.3 TSD in other vertebrates

2.5.3.1 TSD in turtles

TSD is very frequently observed in 64 species of turtle; whereas, remaining 70% of the 257-280 turtle species remain unexamined for their sex determining mechanisms. Homomorphic sex chromosomes seem to be a necessary prerequisite for TSD, but GSD in turtles may occur under homo- or heteromorphic chromosomes (Bull 1980). Chelonians (turtles and tortoise) exhibit large variation in their sex determining mechanisms. With the exception of the genera *Platemys* (XY, Chelidae), *Staurotypus* (XY, Kinosternidae), *Siebenrockiella* (XY) and *Kachuga smithii* (ZW, Bataguridae), other turtle species, in general, lack heteromorphic sex chromosomes (either XY male or ZW female heterogamety). Among chelonians, member species of families' chelidae, trionychidae and staurotypidae exhibit only GSD. Within the Emididae, family members

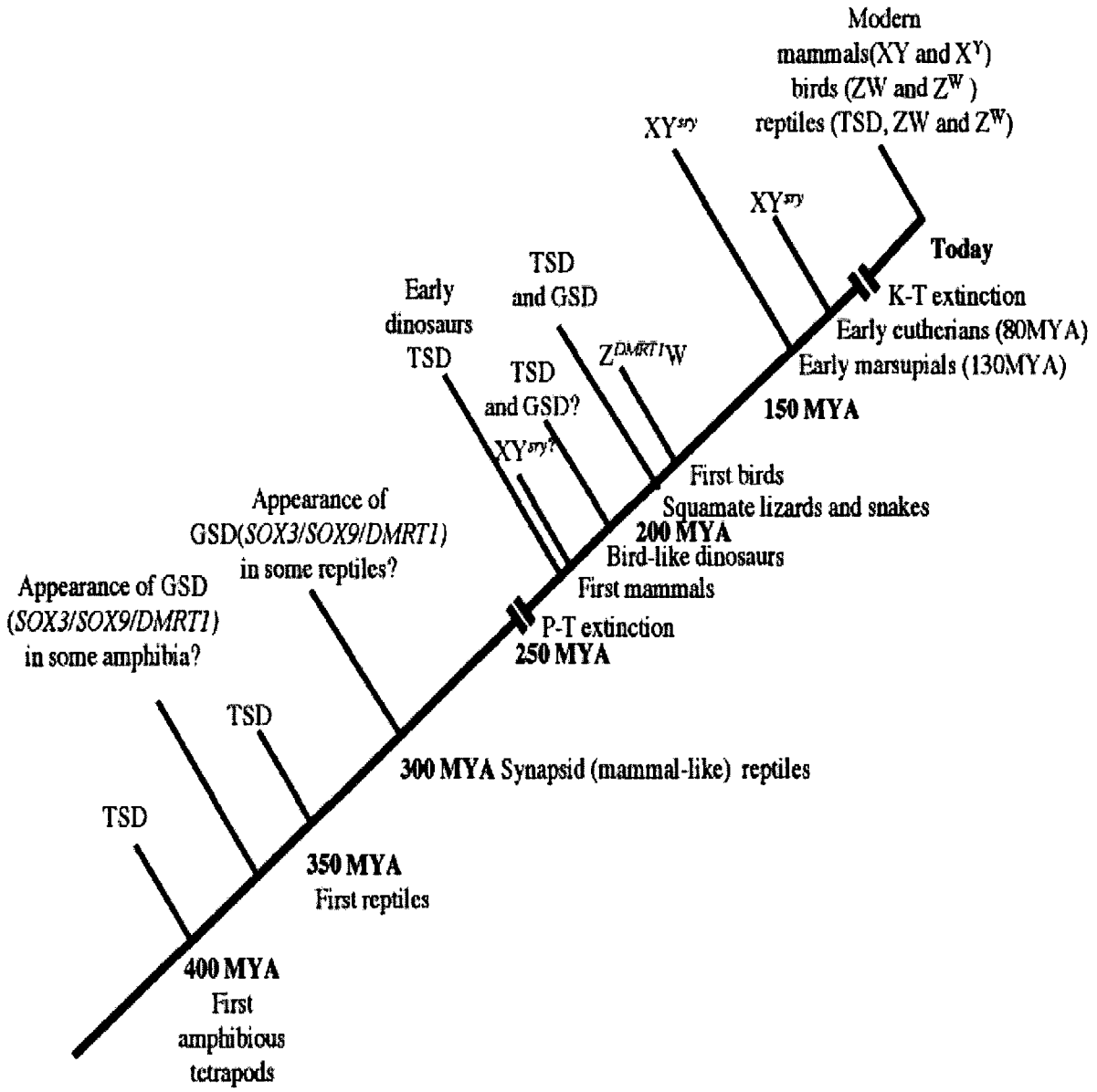


Figure 2.5: **Chronology of sex determination during vertebrate evolution.** TSD is assumed to have been the dominant form of sex determination for much of Paleozoic era from 500 to 250 MYA (million years ago). GSD is suggested to have evolved before the Permian extinction (P-T) event around 254 MYA (adapted from Miller *et al.*, 2004).

of batagurinae and emidinae exhibit both GSD and TSD, though later is much more common. In contrast, TSD is seen as prevalent sex determining mechanism in all other families that have been studied to date. In addition, there are few species, wherein it is still uncertain whether they employ GSD or TSD, *e.g., Macroclemys temminckii,* a species of Chelydridae, was first reported to exhibit TSD, but later was shown that incubation temperature does not affect the sex of offsprings (reviewed by Ciofi and Swingland 1997).

Pattern TSD-I is predominantly found in some Bataguridae and the Carettochelyidae, Cheloniidae, Dermochelyidae, Emydidae and Testudinidae whereas TSD-II is prevalent in Pelomedusidae, Kinosternidae, *Macroclemys temminckii* (Chelydridae), and a few Bataguridae. TSD-I is mainly found in the turtles in which adult females are bigger in body size than adult males while TSD-II is common in turtles with smaller female body size or in which body size is not dimorphic (for more Information http://www.eti.uva.nl/turtles/Turtles2a.html).

2.5.3.2 TSD in lizards

In lizards, iguanids and scincids both demonstrate GSD, where male heterogamety appeared to be ubiquitous. Both GSD and TSD have been described in Agamidae, Eublepharidae, Gekkonidae and Lacertidae families of lizard. However, the coexistence of both of these mechanisms is still uncertain for Iguanidae and Vaanidae (Viets *et al.,* 1994). Occurrence of viviparity (in almost 20 % of lizards) and parthenogenesis (*e.g.* Geckos, Agamids and chameleons) has been suggested for the recurrence of GSD in these taxa, whose reproductive mechanisms presumably require genetic input for sex determination (Ciofi & Swingland 1997; Bull 1980). Recent evidences from wide range of temperature effects on sex ratio of offspring, suggest that pattern TSD-II is prevalent in lizards (Pieau 1996). A well studied system among lizards is *Calotes versicolor,* wherein recent studies have ruled out the presence of TSD; and further suggest its sex determination to be rather under the hormonal control (Ganesh *et al.,* 1999; Ganesh *et al.,* 1997).

2.5.3.3 TSD in Fishes

First report of TSD in fishes came from a study on normally hermaphroditic *Rivulus marmoratus,* wherein low temperature was found to favour male differentiation (Harrington 1967). Subsequently, TSD has been reported in seven orders of fishes with the exception of *Menidia.* In Fishes, other form of labile sex determination is also

frequently reported, like sex change and frequent sex reversal. In many fishes with TSD, the gender is seemingly determined by a genotype by environmental (GXE) interaction.

Despite the reports of TSD in fishes, it has not been studied in detail. Moreover insight of available literature on the ecology and evolution of sex change/ TSD in reptiles, the nature of TSD seen in fishes is rather questionable, and there is need to confirm whether the observed TSD is true or just an anomaly of GSD that occurs at extreme environmental conditions/ ecologically irrelevant temperatures (Conover 2004).

2.5.3.4 TSD in crocodilians

TSD in crocodilians was first reported in *Alligator mississippiensis* (Ferguson & Joanen 1982). Subsequently, TSD has been documented in 11 of the 22 extant species of Crocodylia, which represent both the crocodilian lineages (*i.e.*, alligators/caimans and crocodiles/gharials) (Lang & Anderson 1994). The widespread demonstration of TSD and the universal absence of heteromorphic chromosomes in this group raised the possibility that TSD is a characteristic of crocodilians. This is in contrast with the diversity of sex determining modes in turtle, lizard, and fishes examined so far. Earliest reports suggested a F→M types of TSD in alligator with low temperature supporting female, and high temperature male development (Deeming & Ferguson 1988). However, more recent detailed examination of five species and the preliminary data on another four species of crocodilians (Lang and Andrews 1994) showed that all of them follow pattern TSD-II (i.e. FMF). The relationship between temperature and sex ratio is also remarkably similar among the species. In all crocodilian species studied in detail, exclusively females are produced at or below 31.0°C, with the FM transition zone being between 31.0°C and 32.0°C. On the other hand, maximum male producing temperatures ranges from 32.0 to 33.0°C (Lang & Anderson 1994). In alligator, exclusively males are produced at 32.5 to 33.0°C. Before and after this range, there is a pivotal temperature where a mixed ratio is obtained followed by female producing temperature. In *Crocodylus palustris*, temperature at or below 31.0°C produce 100 % females, but temperatures between 31.0 and 33.0°C produced varying proportions of males. A temperature of 32.5°C produces100% males (Lang *et al.,* 1989); although Lang and Anderson (1994) reported a value of 93% males (figure 2.6), but it is unclear whether this value represented additional information. Hence, *Crocodylus palustris* exhibits the FMF pattern, but at any particular male producing temperature there were significant clutch effects on the proportion of male hatchling.

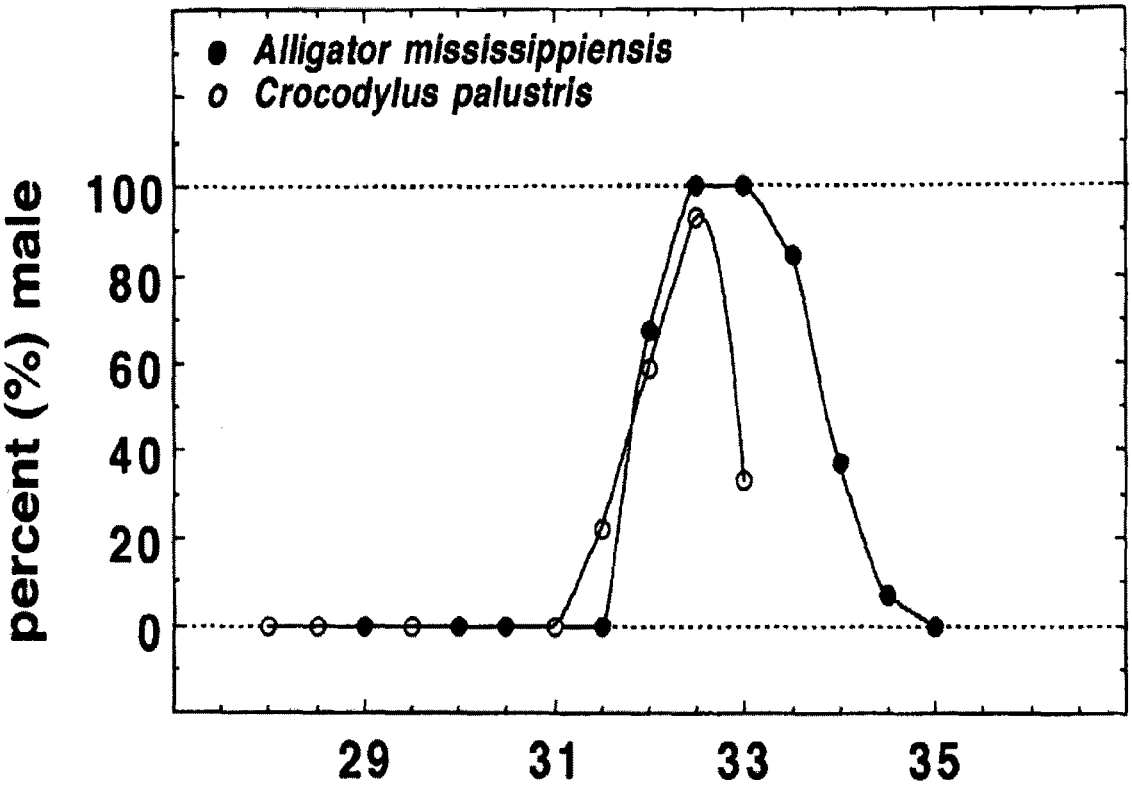


Figure 2.6: Sex ratios expressed as % males *vs.* constant incubation temperature (^{0}C) for the American alligator *(Aliigator mississippiensis)* and the Indian mugger, (*C.rocodylus palustris).* Constant temperature of 32.5^{0}C is defined as a maximum (93%) male producing temperature in *C. palustris* (Lang and Andrews 1994).

2.5.4 Temperature sensitive period or window (TSP/TSW)

Period of embryonic development during which temperature stimulus seems to be pivotal for determination of sex of embryos prior to gonadal differentiation is generally referred as temperature sensitive window (TSW/TSP). The first classical attempt to determine TSP involved temperature-shift experiments performed by (Pieau & Dorizzi 1981) in *Emys orbicularis*. The eggs were first incubated at 25°C (100 % male producing temperature; MPT) and then transferred to 30°C (100 % female producing temperature; FPT) or *vice versa* at different stages of embryonic development. Permanent changes in the sexual differentiation of gonads were examined during various stages and the TSP was defined from stage 16 to 21 at 25°C, and stage 16 to 22 for 30°C. In addition, it was shown that before and after TSP, incubation temperature did not influence the sex of offspring. Similar egg shift experiments, in *Alligator mississippiensis* demonstrate, sex is determined earlier at female producing temperatures (FPT) compared to that at male producing temperatures (MPT). The TSP at female producing temperature (31°C) was estimated during stage 20.5 to 24.5, whereas at male producing temperature (33°C), it was found spanning stages 21.5 to 25.5 (Lang & Anderson 1994). In alligator, gonadal differentiation had also been described at the histological and ultra-structural levels (Smith & Joss 1994b; Smith & Joss 1993). At 33°C, the onset of testis differentiation is signaled by the formation of pre-Sertoli cells during stages 21-22 and these cells differentiate into Sertoli cells during stage 23. Similarly ovarian differentiation is characterized by the proliferation of cortical germ cells during stages 22 - 23 at 30°C. Therefore it was considered that the critical period for gonadal differentiation spans stages 21 - 24 in alligator.

In *Crocodylus palustris* preliminary data from switch-once experiments indicated that femaleness is irreversibly set at about stage 20 (30 days) at 30°C, whereas maleness is set irreversibly at about stage 25 at 32 °C (Lang et al. 1989). However, detailed studies later on demonstrated that at 32°C (MPT) TSP broadly spans the stages 21.5 to 24.5 (approximately 30 to 45 days post egg laying), while the TSP for female sex determination (at 30°C) was identified spanning between stages 20 to 23 or 35 to 45 days post egg laying (Lang & Anderson 1994). It should be noted that the temperature sensitive period appears to increase as incubation temperature increases (Webb *et al.*, 1987). Earlier egg shift assays in *Crocodylus palustris* were carried out at 32°C, which does not represent the maximum male producing temperature. Therefore, at 32.5 °C (maximum male producing temperature), TSP can still be longer than suggested earlier and can be broadly defined corresponding to the developmental stages 21 to 25. Proper

staging of *Crocodylus palustris* embryos during development at 30.0°C (FPT) and 32.5°C (MPT) could be carried out according to (Ferguson 1987), earlier in the laboratory. The morphological or molecular markers for identification of gonads during sex determination are yet to be established in this species, except for *cpSox9* that is indicated to be a potential marker for the identification/onset of male gonad development from 23rd stage onward (Agrawal 2006).

2.5.5 Effects of steroids on gonadal differentiation of TSD species

The role of hormonal interplay during early sex determination/differentiation in TSD is among one of the major differences between TSD and GSD. In mammals aromatase expression is postnatal. The effect of synthetic estrogens on early gonadal development was quite obvious even before the TSD was first observed in turtles. Embryonic or larval treatment with estrogen could feminize genetic male individuals up to sex-reversal in birds (in few species), reptiles (lizards), amphibians and fish (Burns 1961). For the first time absolute feminization of the turtle eggs kept at MPT was demonstrated by Pieau (1974). The development of ovotestes had been observed in marsupials after early estrogenic treatment of young genetic males in the pouch. However, androgenic treatments were successful only in some fish and amphibian species rendering genetic female to male (Pieau *et al.,* 1994). Effect of estrogen (estradiol benzoate) and androgen (testosterone propionate) were also examined during the early gonad development of few TSD species *e.g. Emys orbicularis* and *Testudo graeca.* Estradiol was found to induce gonadal feminization, from ovotestis to ovary, depending on the time and dose during embryonic stage at MPT. However, testosterone did not reverse sexual phenotype at FPT. Indeed, if testosterone is administered to eggs at MPT, about half would hatch as females. This effect is probably due to the aromatization of the testosterone to estradiol, as administration of both testosterone and aromatase inhibitor to eggs produce only males at MPT. If a non-aromatizable androgen such as dihydrotestosterone (DHT) is administered to eggs at FPT, there is no discernable effect on the ovarian nature of gonad. However, if DHT is administered to eggs incubating at an intermediate incubation temperature that produces a mixed sex ratio, most or all of the hatchlings will be male (Pieau *et al.,* 1999; Crews *et al.,* 1994).

Another important discovery was the detection of higher levels of enzyme Aromatase during later stages of TSP in several TSD species. Later on, a number of studies were focused to investigate precise role of oestrogen content, aromatase activity and aromatase gene expression in the GAM (genital ridge-adrenal-mesonephros) complexes

during the TSP in turtles and crocodilians. *In vitro* secretion of steroids hormones (including testosterone, estradiol) by adrenal-kidney-gonad complexes (AKGs) and the stimulation of these hormones in response to gonadotropin (FSH), adrenocorticotropin (ACTH) and growth hormone (GH) was analyzed during TSP of *Trachemys scripta* (White & Thomas 1992a; White & Thomas 1992b). Increased activity of aromatase was detected after TSP in the AKGs of both male and female in *Trachemys scripta*. However, brain of putative female exhibited a significantly higher level of aromatase activity compare to male at the beginning of TSP (Willingham *et al.,* 2000). In contrast, in *Malaclemys terrapin*, high expression of aromatase mRNA was detected in the embryonic brain at the beginning of TSP whereas ovarian expression was detected later in the TSP (Jeyasuria & Place 1998; Jeyasuria & Place 1997). Similar to turtle species, increased aromatase enzyme activity was detected only in GAM tissues of developing embryos incubated at FPT in *Crocodylus porosus* and *Alligator mississippiensis* and not in these at MPT (Smith *et al.,* 1995; Smith & Joss 1994b). Expression of aromatase mRNA in *Alligator mississippiensis* was also detected in brain of putative male and female embryos at early TSP, whereas female-specific expression in GAM tissues was only detected after the TSP (Gabriel *et al.,* 2001). Moreover the similar pattern of aromatase expression (Jeyasuria & Place 1998) was reported in *Trachemys scripta* (Murdock & Wibbels 2003a). These studies in several TSD species indicated that no significant difference existed in levels of aromatase and oestrogens between the developing male and female gonads during TSP in the AKG/GAM complexes. However, in almost all studies some differences were found either at the end and/or after the TSP, but essentially not during TSP. Therefore, initially it was thought that oestogens are not involved in the early steps of ovarian differentiation (Murdock & Wibbels 2003a; Smith *et al.,* 1995; Smith & Joss 1994a). It was also hypothesized that aromatase activity in the brain might be responsible for determining gonadal sex in TSD reptiles despite contradictory evidences from different species (Gabriel *et al.,* 2001; Jeyasuria & Place 1998); (Merchant-Larios & Villalpando 1990).

2.5.6 Aromatase activity *vis-a-vis* incubation temperature

The above-mentioned view was proposed based on the studies carried out on AKG/GAM complexes. However, subsequent studies on gonads alone and the results obtained from freshwater and marine turtles indicated that aromatase activity in the gonads depends on the incubation temperature. Direct correlation in the aromatase activity and the incubation temperature was observed in the turtle gonads, wherein increased aromatase activity was observed at high temperature of incubation (FPT) during TSP. Egg shifts

experiments performed during the TSP showed expected increase or decrease in female embryo-specific aromatase activity in gonads, whereas shifts after the TSP were found ineffective (Pieau *et al.,* 1998; Desvages *et al.,* 1993b; Desvages & Pieau 1992). Similar results were obtained in the in the salt water crocodile (Smith & Joss 1994a). Therefore, the contrasting observation of lack of differences in aromatase/oestogens in AKG/GAM complexes at FPT and MPT in *Crocodylus porosus, Alligator mississippiensis* and *Trachemys scripta* (Murdock & Wibbels 2003a; Gabriel *et al.,* 2001) might be due to the masking effect of adrenal/mesonephros activity over gonadal tissue. Indeed, during TSP, gonad represents a very small portion of the GAM complex and aromatase activity is also present in the adrenal/mesonephros. For the same reason, using RT-PCR, difference in aromatase mRNA levels between male and female GAMs were not seen during TSP in *Alligator mississippiensis* and *Trachemys scripta* (Pieau & Dorizzi 2004).

2.5.7 Gonads synthesize oestrogens

Studies on hormonal metabolism during embryogenesis reveal that all active enzymes involved in steroidogenesis are present in the gonads during the TSP of *Emys orbicularis* (Desvages & Pieau 1991). These studies also showed high oestrogens level in the separated gonads in the beginning of TSP at FPT in *Emys orbicularis* and *Dermocheys coriacea* (Desvages *et al.,* 1993a; Dorizzi *et al.,* 1991). These studies, thus, suggest that oestrogens are synthesized by the gonads during the TSP. Based on these studies it was proposed that gonadal endogenous oestrogens and not the extragonadal oestrogens are involved in TSD. Important role for endogenous aromatase in the female gonad development, is further strengthened by observations that almost in all TSD species treatment of aromatase inhibitors cause sex reversal or formation of ovotestis (Piferrer *et al.,* 2005; Pieau *et al.,* 1999; Lance 1997; Wennstrom & Crews 1995).

2.5.8 Role of oestrogens during TSP

The implication of oestrogens in ovary differentiation has been irrefutably confirmed by the results of treatments with antioestrogens and aromatase inhibitors prior to and /or during TSP. The time-period during which the treatment is efficient coincides with TSP (Gutzke & Chymiy 1988). During this period, binding of tritiated oestradiol by the gonads was shown to occur at both FPT and MPT (Smith & Joss 1994a). Moreover, oestrogen receptors mRNA were detected in the putative female and male at the beginning of TSP (Bergeron *et al.,* 1998). In *Trachemys scripta*, shift from MPT to FPT as well as oestrogenic treatment at the same embryonic stages during the TSP had similar effects

on gonadal structure, i.e. inhibition of testicular development preceding the proliferation of germ cells in the cortex (Wibbels et al. 1993). Taken together these results suggested that, at FPT, oestrogen synthesized by the gonads themselves direct their differentiation into ovaries.

Similarly, the experiments and assays carried out on the excised gonads of turtle embryos further provided evidence that oestrogens play an important role in ovarian differentiation during the early TSP (figure 2.7). Similar to reptilians, in fish and amphibians also, oestrogens and aromatase have been shown to be involved in the ovarian differentiation. Moreover, the involvement of oestrogen is also shown in birds that all display GSD (Elbrecht & Smith 1992). The upregulation of aromatase expression in fish and chicken have been demonstrated in the early female gonad development (Oreal *et al.,* 2002; D'Cotta *et al.,* 2001) (Nishikimi *et al.,* 2000);.

Based on these results, oestrogens and aromatase have been implicated in the early steps of ovarian differentiation. Nonetheless, future studies are needed to examine more precise regulation/role of aromatase gene expression in sex determination/differentiation in species with strict GSD and TSD; and also to identify the downstream molecular and cellular targets of oestrogens in the cortical and medullary parts of the gonads during sex differentiation (Pieau & Dorizzi 2004). Furthermore, the search for upstream effectors of *Aromatase* transcription during TSP in response to temperature may probably be the first step to unravel the molecular mechanism *vis-à-vis* hormonal interplay underlying the TSD.

2.6 Expression pattern of conserved sex-determining genes in TSD

Near similar effects of hormonal modulation in the GSD and TSD species suggest existence and expression of relatively conserved sex related genes during gonadal development. Indeed, despite highly varied "triggering" mechanism(s), a variety of vertebrates appear to share relatively similar, highly conserved downstream genes that operate in gonadal sex differentiation.

Several important genes, which have a role in GSD, are largely shown to be conserved among the vertebrates. Moreover, relatively similar hormonal interplay in GSD and TSD species also suggest that triggering mechanism(s) for the sex determination may vary,

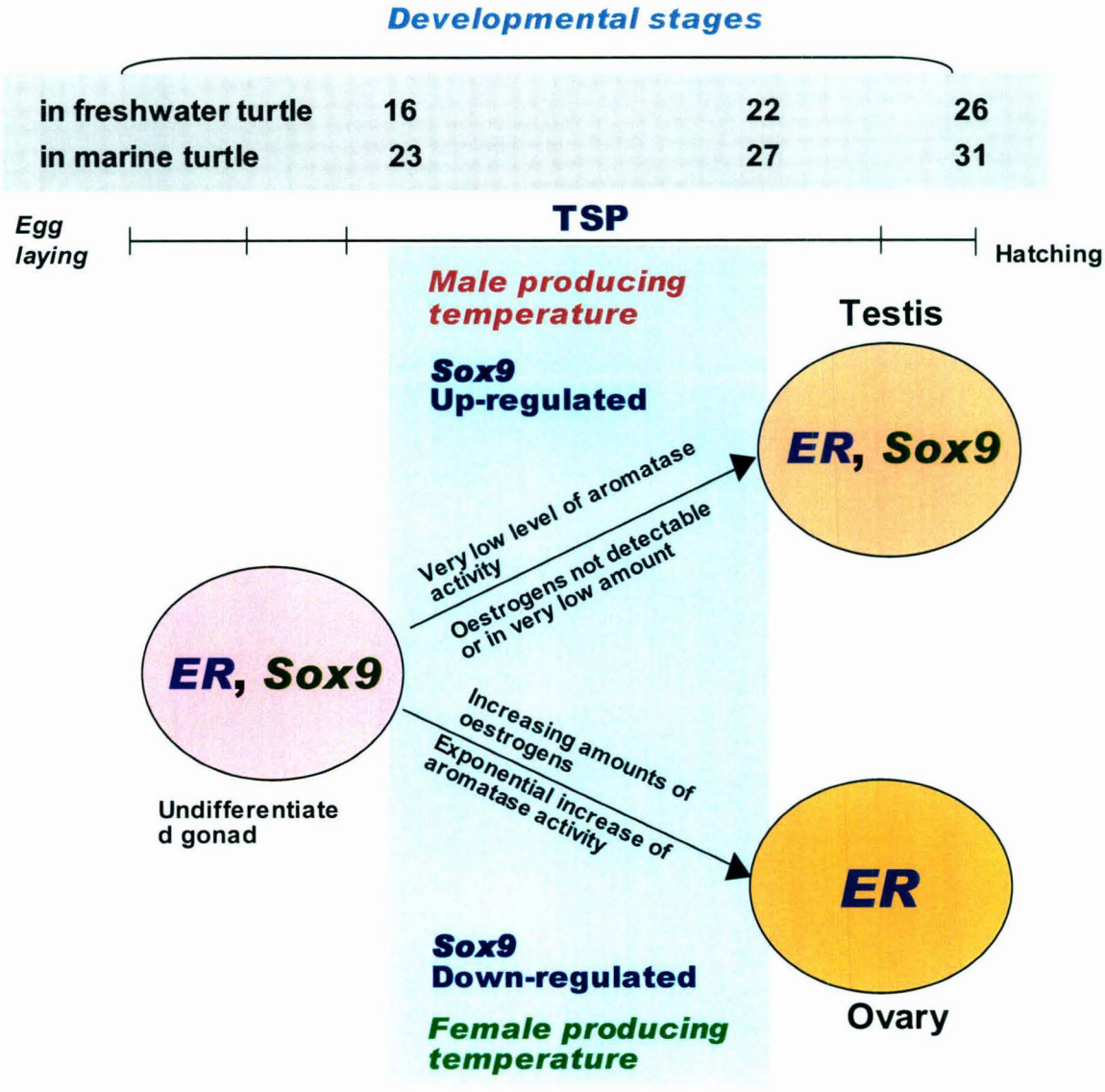


Figure 2.7: Molecular and physiological events in the gonads consistent with the involvement of endogenous oestrogens in the first steps of ovarian differentiation in freshwater and marine turtles. (ER- oestrogen receptor). (adapted from Pieau and Dorizzi 2004).

but the downstream processes in early gonad differentiation may be same (Morrish & Sinclair 2002).

Increasing number of molecular studies undertaken in recent years to investigate the genetic component(s) of sex determination/differentiation in TSD species largely support the above with the notable exception of SRY. Some of the sex-related candidate genes involved in mammalian GSD that have also been indicated during gonadogenesis in TSD reptilian species (mainly turtles) include: *Sox9, Dmrt1, Sf1, Dax1, Amh,* and *Wt1* etc (figure 2.8, 2.9). These studies also suggested that despite a broader conservation evident in sex-related genes across GSD/TSD species, there might exist some fundamental differences in their role in the sex determination/differentiation process.

2.6.1 *Wt1* (Wilms Tumor protein-1)

Wt1 has been identified and studied in several TSD species of turtles and alligator (Western *et al.,* 2000; Spotila *et al.,* 1998). Its expression starts very early in development much before TSP however, there is no clue about the possible role it has in TSD species. Moreover, no differences are seen in its expression in developing gonads at MPT and FPT (Western *et al.,* 2000; Spotila *et al.,* 1998). *Wt1* in mammals undergoes extensive alternative splicing; however, in TSD species isoforms of *Wt1* have not been reported except in *Trachemys scripta* (Spotila & Hall 1998). WT1 amino acid sequences of TSD species show nearly 85% similar with that of mammals, however, important proline rich domain at the N-terminus is absent. N-terminal domain of the mammalian WT1 is larger than their reptilian counterparts, which indicates possible different role of WT1 in the TSD species.

Localization of *Wt1* in mammals has been well studied. *Wt1* in female gonads is expressed in cells of coelomic epithelial origin, very feeble in medulla and later in granulose cells during different stages of development. On the other hand in embryonic male gonads *Wt1* expression was restricted to testicular cords and the tunica albugenica, and later to Sertoli cells in the mature gonads (Mundlos *et al.,* 1993). In a study done in developing gonads of *Trachemys scripta, Wt1* localization tracks the development of the male and female gonads at MPT and FPT *i.e.* during formation of testicular cords from the sex cords at MPT and thickening of the cortex at the FPT (Schmahl *et al.,* 2003). All these studies amply suggest that it is an important early expressing sex related gene; and thus, there is a need to explore its function/splice

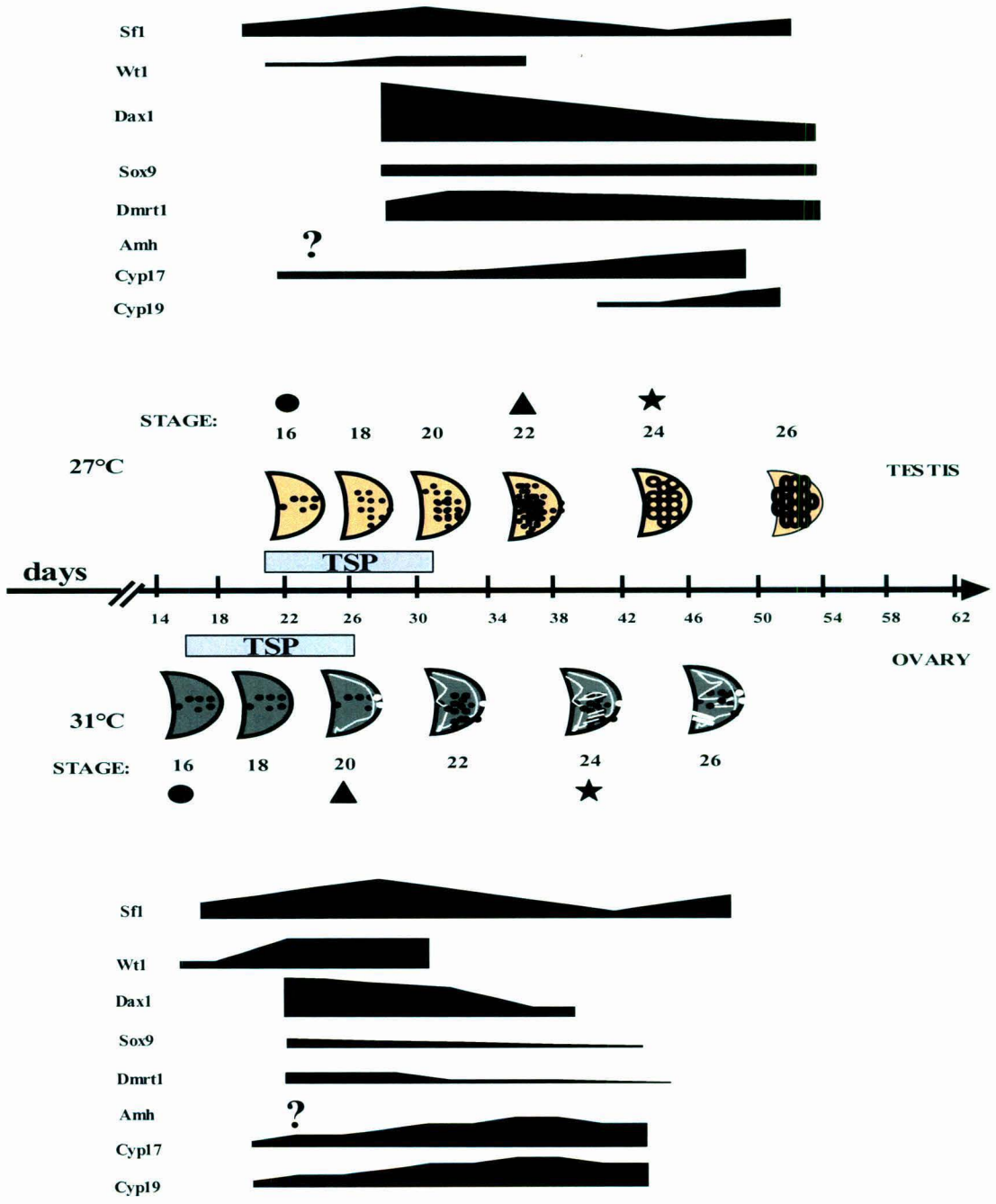


Figure 2.8: Sexually dimorphic gene expression during gonadal development in the turtle embryo. For each of the gene, line thickness reflects relative level of expression compared to that in opposite sex (but not between genes within a sex). Symbols: ● Bipotential gonad; ▲ cord formation begins; ★ cord formation complete (adapted from Morrish and Sinclair 2002).

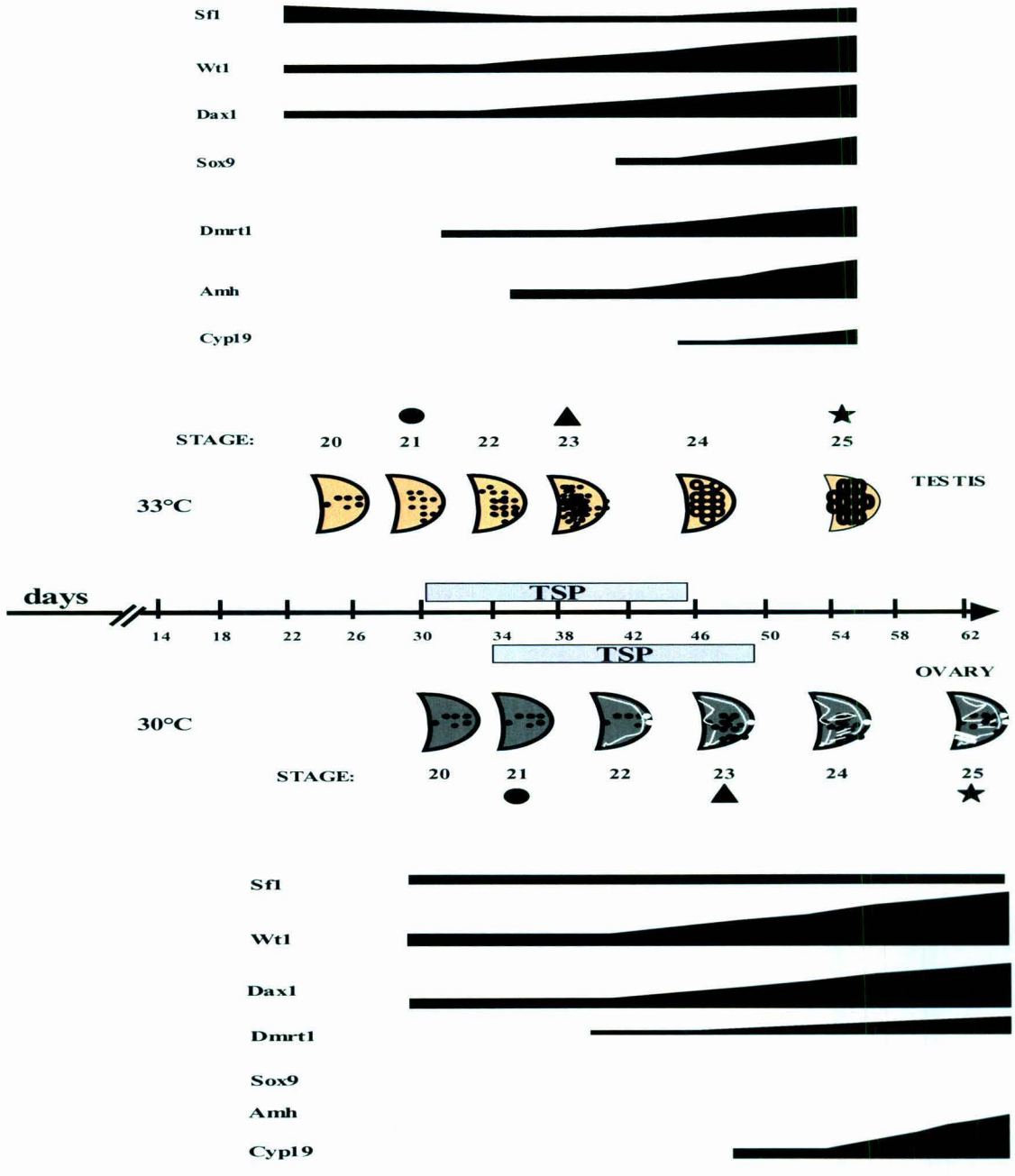


Figure 2.9: Sexually dimorphic gene expression during gonadal development in the alligator embryo. For each gene line thickness reflects relative level of expression compared to that in opposite sex (but not between genes within a sex). Symbols: ● Bipotential gonad; ▲ cord formation begins; ★ cord formation complete (adapted from Valenzuela and Lance 2004).

isoforms (if any) in much detail in TSD species, as in mammals Wt1 isoforms play a crucial role (Hammes *et al.,* 2001).

2.6.2 *Sf1* (Steroidogenic factor 1)

Sf1 is an early gene involved in the maintenance of the bipotential primordium in the GSD species and is shown to be largely conserved in several TSD species (Ramsey *et al.,* 2007; Valenzuela *et al.,* 2006; Crews *et al.,* 2001; Western *et al.,* 2000; Fleming *et al.,* 1999). In GSD species *Sf1* is not upregulated in the bipotential primordium of the XY embryos, though its expression increases in the differentiating male gonad. *Sf1* is a central player in the regulation of the steroidogenic enzymes in the gonads, SF1 is required for the *Amh* expression in the male gonad in the mammals (de Santa *et al.,* 1998); however, its relation with aromatase expression is not known (as in the GSD species aromatase expression can be not be detected during embryonic stages)

Contrary to above, in all the TSD species, *Sf1* expression is reported high at MPT since bipotential stages of gonad development, moreover, increase in the estrogen concentration or shifting eggs from MPT to FPT reduces the *Sf1* expression (Crews *et al.,* 2001). These studies show that higher estrogen or aromatase concentration antagonizes *Sf1* expression, which is unique in TSD species and nothing is known about the underlying mechanism. These features of *Sf1* i.e. early upregulation of *Sf1* at the MPT and its antagonistic relation with the estrogen producing machinery makes it an important gene for study in TSD.

2.6.3 *Sox9*

Besides its role in mammalian sex determination and differentiation, *Sox9* expression has been studied during early gonadal development of TSD species. In marine turtle olive ridley (*Lepidochelys olivacea*) TSP covers stages 21 to 26. In this species, *Sox9* expression by RT-PCR was detected to be ubiquitous both at MPT (26°C) and FPT (33°C) throughout TSP; whereas down regulation was observed only from stage 26 onward at FPT. The immuno-histochemical analysis also showed the same expression pattern, indicating that *Sox9* is expressed both in male and female gonads during early TSP, but only in male after TSP (Torres Maldonado *et al.,* 2002). Not a similar situation was evident in another turtle species *Trachemys scripta*, wherein *Sox9* expression (detected by *in situ* hybridization) was seen throughout the undifferentiated gonad spanning stages 15 to 20 of TSP, followed by preferential expression in the testis only from 20th stage onward (Spotila *et al.,* 1998). However, somewhat contrary to turtles, in

alligator, the male-specific expression of *Sox9* was also evident during later part of TSP itself. Nonetheless, since the upregulation of *Sox9* occurred during the structural organization of testis, it was postulated that Sox9 might be involved in the testis differentiation but not in determination (Western et al. 1999). Whatever is the pattern of expression of *Sox9* in the developing gonads of turtles and alligator, its differential expression correlates well with the morphological sign of testis, suggesting: a) it plays a role in the testis differentiation in these species; though b) the underlying mechanism(s) may differ for male-specific *Sox9* upregulation in *Alligator mississippiensis* and *Trachemys scripta* or female-specific downregulation in *Lepidochelys olivacea*.

Recently, *cpSox9* (crocodilian homologue of *Sox9*) has been isolated and characterized in our lab (Agrawal 2006). Very interestingly, the detailed expression profiling studies revealed extensive transcriptional diversity of *cpSox9* in GAM and other embryonic tissues, and most notably, a male GAM specific upregulation of one such *cpSox9* isoform. This elegant study also indicated extensive splicing of *Sox9* in mammalian model mouse. Thus, the major findings of this study were that *cpSox9*: a) expression is dimorphic and male specific; and, b) has important role in testis differentiation but not in sex determination (TSD). The study further suggested that apparent lack of *Sox9* sex specificity during TSP as noted in the earlier studies (some referred above) was most probably the outcome of poor experimental design, and/or masking effect of surrounding tissue GAM.

2.6.4 *Dax1* (dosage-sensitive sex reversal gene)

During the gonadogenesis of *Alligator mississippiensis*, and *Lepidochelys olivacea*, *Dax1* expression was seen throughout the TSP at both FPT and MPT. However, in *Lepidochelys olivacea*, increased level of expression was detected in the ovaries at stage 27, once they were morphologically distinct. The results from both these TSD species suggest that *Dax1* may not be involved in TSD (Torres Maldonado *et al.,* 2002; Western *et al.,* 2000).

2.6.5 *Amh* (Anti Mullerian Hormone)

Though *Amh* is transcriptional target of the Sox9 in mammals this does not seem to be the case in many TSD species wherein *Amh* expression either precedes *Sox9* or its expression starts with *Sox9.* Male-specific expression of *Amh* is observed during the TSP of *Alligator mississippiensis* and *Trachemys Scripta*. In alligator, Amh is upregulated in the gonads at MPT preceding male-specific *Sox9* expression, whereas in turtle, the

timing of upregulation of Amh in male gonads was found at stage 17 onward and its correlation with *Sox9* expression is yet to be precisely established (Western *et al.,* 1999).

2.6.6 *Dmrt1* (*Doublesex mab-3 related transcription factor1*)

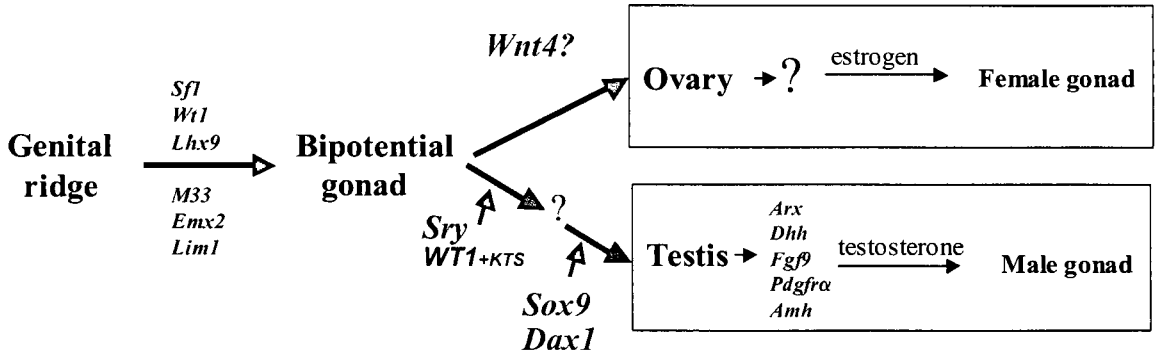
Dmrt1 is one of the most conserved sex determining gene that also has invertebrate counterparts *i.e. doublesex* in drosophila and *mab-3* in *Caenorhabditis elegans (Raymond et al., 2000).* *Dmrt1* is an important gene required for the later stages of testis differentiation in mammals. However, *Dmrt1* plays a wide variety of roles in sex determination/differentiation across vertebrates (Hodgkin 2002). In teleost, *Medaka Dmrt1* acts as master regulator for the male sex determination similar to what Sry does in mammals (Schartl 2004; Matsuda *et al.,* 2002). In *Gallus, Dmrt1* presence on the W chromosome and its early male specific expression makes it a candidate for the master regulator (Smith *et al.,* 2003). In several TSD species, *Dmrt1* has been identified and shown to be upregulated at the MPT since early stages of TSP indicating an early role for *Dmrt1* in TSD.(Shoemaker *et al.,* 2007; Murdock & Wibbels 2006; Murdock & Wibbels 2003b; Torres Maldonado *et al.,* 2002; Shibata *et al.,* 2002; Kettlewell *et al.,* 2000). Furthermore, it was noteworthy that *Dmrt1* is upregulated in developing gonad of embryo that are shifted to MPT from FPT and is downregulated in the embryos treated with estrogen and/or shifted to FPT from MPT.

Major differences in the expression pattern of conserved sex determining genes between mouse and turtle are shown in the figure 2.9 (taken from Yao and Capel 2005)

2.7 Missing links: Why TSD needs a special attention

Vertebrates have a wide variety of the sex determining mechanisms and till very recent it was believed that downstream processes to the sex determination including early differentiation of gonads is conserved among vertebrates (Sarre *et al.,* 2004; Morrish & Sinclair 2002). In species with GSD, sex is determined by genetic factors that operate largely independently of environment, whereas in reptiles with TSD, sex is determined after fertilization by the environmental conditions that prevail during embryonic incubation, and largely independent of direct genetic influences. The conventional view, which emerged from the early work on sex determination in reptiles, is that these two mechanisms are mutually exclusive and can therefore be viewed as discrete and fundamentally different. However, there is great conservation of the genes involved in sexual differentiation and the hormonal cascades that govern differentiation. The mechanisms by which genetic and environmental factors determine sex can be expected

Genotypic sex determination



Temperature-dependent sex determination

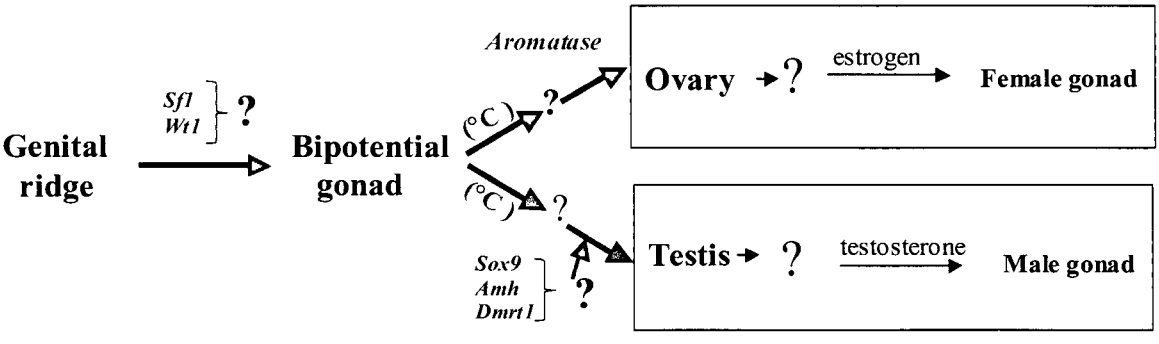


Figure 2.10: Current paradigms for Genotypic and Temperature dependent sex determination.

to draw upon elements of this common machinery. Now there are evidences that sex determining pathways have evolved over time. Cellular migration is very important for the embryonic gonad development. In mouse, PM cells including steroidogenic cells migrate from mesonephric mesenchyme and Sertoli cells from the coelomic epithelium (Wilhelm *et al.,* 2007). However in *Gallus,* mesonephric mesenchyme contributes both to steroidogenic cells and Sertoli cells (Sekido & Lovell-Badge 2007). Similarly, migration of the Germ cells to site of the gonad differentiation is different between the mammals and non-mammals (in Gallus and fish Medeka) (Nakamura *et al.,* 2007; Kurokawa *et al.,* 2006). On the other hand there are also few studies addressing germ cell and the accessory cell migration in the TSD species. In TSD species, cellular migration is expected to be more similar to that seen in non-mammalian vertebrates; in case of turtle, dye tracking experiments suggest otherwise i.e. cell migration pattern was more similar to mammals (Yao *et al.,* 2004). Similarly, recent analysis of conserved sex determining genes and cellular events reveal some fundamental differences between mammalian and non-mammalian TSD species (Yao & Capel 2005). In another study, cultured gonads detached from the mesonephros showed normal development of the male gonads (Moreno-Mendoza *et al.,* 2001). All these studies amply demonstrate that mechanisms of sex determination/differentiation among vertebrates have evolved over time, despite an apparent broad conservation of underlying major elements/events; and thus remains an enigma for geneticists, biologists and evolutionarists.

Our existing knowledge of initial trigger such as *Sry* in mammal or temperature in species with TSD is only the beginning of an expanding field of investigations. A number of questions remain unsolved. What about the regulatory genes that respond to these signals? In more specific terms, whether these signals are activator or repressor of downstream genes in the male pathway? How do they interact to complete the cascade of molecular events? What are the other players? The downstream pathways during differentiation might be conserved in the species with GSD and TSD, but the complete profile of the genetic factors involved in this process has not yet been completely charted out (figure. 2.10). Our present understanding of downstream genes involved in TSD has, by and large, been dependent on the identification and expression of homologues of genes that have been shown, or implicated in gonad determination and/or differentiation in mammals such as, *Sox9, Wt1, Dmrt1, Dax1, Amh*, and *Sf1*. The timing and differential expression of some of these genes can also vary among different TSD species during the critical sex-determining period; for example Sox9 is expressed only in the developing male gonad of alligator, but in turtle, initially it is expressed in both male and female gonads and then down-regulated in female during the differentiation, however, proper

characterization for *Sox9* was not done in above studies. Therefore expression analysis of such candidate genes in various species and identification of new candidate genes during the gonad determination/differentiation period is expected to provide better understanding of this complicated pathway. It is hoped that with continuous advancements in genomic tools and increasing interest and fervor of biologist (around the globe) the underlying mechanism of TSD and SD at large would soon be resolved, and that this study is step to achieve this cherished goal.

Chapter 3

Materials and Methods

In the present study, an attempt has been made to study the underlying molecular basis of TSD using Indian mugger as a model system. In specific, the study aimed to identify/characterize crocodile homologues of *Aromatase*, *Wt1* and *Dmrt1*, the three important sex-related genes, and also attempted to search for novel differentially expressed genetic factors during TSP. Briefly, experimental plan involved collection of crocodile eggs on day zero of egg laying, followed by incubation at FPT (30.0°C) and MPT (32.5°C) and collection of various embryonic tissues from 21st to 25th developmental stages that span TSP. The embryonic tissues were used for DNA/RNA isolation and preparation of the tissue sections, for the identification/isolation, characterization and expression profiling of genes.

Crocodile homologues of *Wt1, Aromatase and Dmrt1* are isolated from *C. palustris* (Indian mugger), using a number of PCR-based approaches, such as, RT-PCR, 5′ RACE, 3′ RACE, Genomic PCR and Genome Walking; and using both cDNA as well as genomic DNA from GAM tissues of developing embryos. The spatio-temporal expression analysis of above genes was performed using semi-quantitative RT-PCR, real-time RT-PCR, Northern hybridization, *in-situ* hybridization and immunostaining. The study interestingly revealed multiple transcribed isoforms for all the above genes that were later cloned and sequenced; and some were also followed for their putative functional significance by transfection studies in GAM-derived embryonic primary cell lines of crocodile.

In parallel, attempts were made to search for novel gene candidates differentially expressed during TSP in the developing GAM tissues of putative male crocodile embryos. This was done using suppressive subtractive hybridization (SSH) approach, which provides a highly efficient improved method of cDNA subtraction with relatively less number of false positives. SSH was carried out with the RNA pool isolated from GAM of male and female embryos representing developmental stages 21+22.

3.1 Living materials

3.1.1 Crocodile eggs/Tissues

Indian mugger, *Crocodylus palustris* is a schedule IV endangered species protected under the Wild Life Act of Govt. of India. Hence, requisite permissions were obtained

from the appropriate Government authorities in the beginning of the project and a memorandum of understanding was signed between the CCMB and the Nehru Zoological Park (NZP), Hyderabad, India, for supply of necessary materials needed for the work. Egg clutches of *C. palustris* were collected from NZP, on zero day of egg laying during February-March (the reproductive season) of every year and incubated in the laboratory at constant temperatures of 30.0 ± 0.1 °C (FPT) and 32.5 ± 0.1 °C (MPT). Developing embryos were sacrificed after fixed incubation periods, corresponding to the developmental stages 21 to 25 that define the TSP in Indian mugger (Table 3.1; Lang *et al.* 1989). Identification of developmental stages and their correspondence with incubation temperature/days were as per Ferguson 1987.

Five embryonic tissues viz., embryonic GAM (Genital ridge-Adrenal-Mesonephros complex), brain, kidney, heart and liver were dissected out from etherized embryos and snap frozen using liquid nitrogen. These were then stored frozen at –70°C until used for the subsequent isolation of RNA, Protein and DNA. Some of the samples were directly treated with 4% formaldehyde for tissue fixation and later sections have been prepared to study localization of genes.

Table 3.1. Days of Incubation at FTP and MTP corresponding to different developmental stages through TSP, of embryos of Indian mugger.

Incubation temperature	**Developmental Stage**				
	21	**22**	**23**	**24**	**25**
30.0°C (FTP)	36	40	45	51	60
32.5°C (MTP)	31	36	40	45	51

3.1.2 Bacterial strains

E. coli DH5α strain F^-, *E.coli* DH10α strain F^-, *rec* A1 *end* A1 *hsd* R17 ($r_k m_k^-$) d(*lac* ZYA-*arg* F) U169 (f80 lacZ dM 15) *sup* E44 *thi* – 1 *gyr* A96 *rel* A1 (BRL, 1986) were used for all transformation and cloning of recombinant plasmids.

3.1.3 Embryonic Cell lines of Crocodile

Primary cell lines derived from the embryonic brain and GAM tissues were used for transfection studies. These cell lines were established from the tissues taken from developing Crocodile embryos at MPT and FPT in a separate study. In brief, the embryonic tissues were treated with trypsin for 10 minutes and were carefully homogenized and were supplemented with DMEM medium containing 20% serum. A day after, medium was changed and the growing cells were transferred to another flask containing the same medium.

3.2 Chemicals/Bio-chemicals

3.2.1 Common Chemicals/plastic ware

All commonly used chemicals such as chloroform, isopropanol, methanol, formamide, glacial acetic acid, acetonitrile, glycerol, ethidium bromide, β-mercaptoethanol, SDS, acrylamide, bis-acrylamide, agarose, Ammonium per sulphate, TEMED, urea, tris base, boric acid etc. were obtained from SIGMA chemical company, MERCK and Qualigens. All restriction enzymes, T4 DNA polymerase, DNA molecular size markers were from New England Biolab. Ampli Taq Gold for all routine PCR was from Perkin Elmer. The plastic ware such as PCR tubes, centrifuge tube, micropipette tips, petri plates etc. were obtained from Axygen Scientific, Eppendorf, Tarson and Laxbro.

3.2.2 Radioactivity

The α-P^{32} dATP, α- P^{32} dCTP and α- P^{32} UTP were from Board of Radiation and Isotope Technology (BRIT) and Perkin Elmer.

3.2.3 Reagents and Kits

TRIzol reagent, cDNA synthesis kit (Invitrogen, Life Technologies)

SMART™ RACE cDNA Amplification Kit, Advantage 2 PCR Kit, SMART™ cDNA library construction Kit, Universal GenomeWalker Kit (BD Biosciences)

pMOS Blue Blunt Ended Cloning Kit, InsTA cloning kit (Fermentas), pEGFPN1 vector (clontech)

GFX PCR DNA and Gel Band purification Kit (Amersham)

Prime-It(R) Fluor Fluorescence Labeling Kit,

Message Clean Kit (Gene Hunter), RNA Later (Ambion), Massage Amplification Kit (ambion) QIAprep(R) Spin Miniprep Kit, MinElute Gel Extraction Kit (QIAGEN)

Random priming kit and Nick translation kit (BRIT-India)

ABI Prism(TM) Big Dye(TM) Terminator Cycle Sequencing ready reaction Kit (Applied Biosystem)

3.2.4 Bacterial media and commonly used solutions/buffers

A large number of working solutions/media were prepared in the lab following standard practices; some of the important ones are detailed below:

10 X TAE	0.9 M tris base, 10 mM EDTA, pH adjusted to 8.0 with glacial acetic acid.
10 X TBE	0.9 M tris base, 10 mM EDTA, pH adjusted to 8.0 with boric acid.
LB	1 % bactotryptone, 1 % NaCl, 0.5 % yeast extract, pH 7.0.
LB agar	LB containing 1.5 % bactoagar.
Ampicillin	100 mg/ml stock in sterile water
IPTG	100 mg/ml stock of isopropyl thio-β-D-galactoside in sterile double distilled water.
X-gal	20 mg/ml of X-gal in dimethyl formamide.
PBS	130 mM NaCl, 2.5 mM KCl, 5.3 mM Na_2HPO_4, 1.7 mM KH_2PO_4, pH adjusted to 7.4 with HCl.
10X MOPS Buffer	0.2 M MOPS, 20 mM CH_3COONa, 10 mM EDTA; adjust pH to 7.0 with 2 M NaOH

Phosphate Buffer	Na_2HPO_4 (1 M, 684 mL), NaH_2PO_4 (1 M, 316 mL) for 1 L of 1 M Phosphate Buffer, pH 7.2
Church Buffer	0.25 M Phosphate Buffer, pH 7.2, 0.2 mM EDTA pH 8.0, 1% BSA, 7% SDS
SSC (20 X)	3 M NaCl, 0.7 M Sodium citrate, pH adjusted to 7 with 10 N NaOH
SSPE (20X)	3.6 M NaCl, 0.2 M NaH2PO4 and 0.02 M EDTA, pH 7.4
TE	10 mM Tris-HCl (pH-7.5-8.0), 1 mM EDTA
SM	0.58 % NaCl, 0.2 % MgSO4, 50 mM tris-HCl (pH-7.5), and 0.01 mg gelatin (in 100ml)
SOB	2 % Bacto-tryptone, 0.5 % yeast extract, .05 % NaCl, 2.5 M KCl.

3.2.5 Oligonucleotides/Primers

A large number of primers were used in the study. These were, in general, custom synthesized for us by M/S Bioserve India Pvt. Ltd., Hyderabad. The primer details (ID, nucleotide sequence, purpose and source) are given in Tables 3.2 to 3.6.

3.3 General methods

3.3.1 Isolation of DNA and RNA

3.3.1.1 Genomic DNA isolation

In general, 500 mg tissue was homogenized in 1 ml of extraction buffer (100 mM NaCl, 10 mM EDTA, 50 mM Tris-HCl pH 8.0). Proteinase K and SDS were added to the final concentration of 100 µg/ml and 2% respectively. The homogenates were then incubated at 37°C for 4-6 hours. The proteins were extracted sequentially with organic solvents: phenol (Tris-saturated pH 8.0), phenol:chloroform (1:1 v/v) and chloroform:isoamyl alcohol (24:1 v/v), each time separating the aqueous phase (after thorough mixing with organic solvent for ~15 minutes) by centrifugation at 10,000 rpm for 15 minutes at 15^0C. After the third organic extraction, the separated aqueous phase was used to precipitate the genomic DNA by adding 1/10 volume of 3 M sodium acetate (pH 5.2) and equal volume of ice-chilled isopropanol. The gelatinous precipitate of high molecular weight

Table 3.2. Primers used in the study to isolate/charaterize the cpWt1 locus and cpWt1 transcribed isoforms

S.No.	Primer name	Sequence (5'-3')	Source	Purpose
1	Wt1F1	AAGGGCCATGCAACAGGATA	Designed from conserved domain of *Wt* 1 sequences of *Trachemys, Alligator, Taeniopygia, Xenopus, Gallus* and Human (Accession numbers **AF019779, X85730, U42011, AB033633, NM_000378** and **NM_024425;** primers designed from Alligator coordinates 310 to 841 bp)	To identify/obtain the initial fragment of the crocodile homologue of *Wt1*
2	Wt1R1	GCAGTTTAGTCATGTTCCTC		
3	WT15'	GCACTCGCCGGACATCTTGTATTCC	New, present study; designed from the initial 500+ bp amplified/transcribed *cpWt1* fragment	Primer designed for the 5' RACE
4	WT15'N	GGATATGAAAATGAGAACCACACTGCGC		Nested primer to the WT15'
5	AWT13'	GTGGTAATTCTGAGCTGACCAAGCTG		Primer designed for the 3' RACE
6	AWT13'N	GGAGGCAGAACCTTCCTTCAGTTGTT		Nested primer to the AWT13'
7	GWTR	CCATGGGTGTGTATTCTGTATTGG	New, present study, primers designed from 5' region	Genome walking to amplify *cpWt1* upstream region
8	WT1GW5'	GCGCAGTGTGGTTCTCATTTTCATATCC		
9	WTBIGF2	AAACGACACCAAAGACGACACACAGG	New, present study, primers designed from unique region of *cpWt1* **a** isoform	Forward primer for real-time PCR for isoform *cpWt1* **a**
10	WT1bigR4	GGCTTTGACATGTACACATACCCAACACA		
11	WT1SMF	TGTAGAGGAAGAAACAAACTTACCG	New, present study, primers designed from unique region of *cpWt1* **b** isoform	Forward primer for real-time PCR for isoform *cpWt1* **b**
12	Wt1SMR	CCAGGTGTCAACTGAACTACATACA		Reverse primer for real-time PCR for isoform *cpWt1* **b**
13	Wt1_BGWF1	GAAAATGAGAACCACACTGCGCCTATG	New, present study, primers designed common 625 bp region present in both isoform	Genome walking; from commong region to unique small isoforms coding region
14	Wt1_bigf1	GCCCAATACAGAATACACACCCATG		Nested primer to Wt1_BGWF1

Table 3.2. Primers used in the study to isolate/charaterize the cpWt1 locus and cpWt1 transcribed isoforms

15	Wt1_BGWR1	CTGATCGGACAATAGTTGGAGCTACTC	New, present study, primers designed from exon-2 region coding *cpWt1* **a**	Genome walking; from exon-2 to unique small isoforms coding region
16	WT1bigR	TGCACACATGAAGGGACGTTTTT		Nested primer to Wt1BGWR1
17	WT1GW_FN	CTGCAGCTTGGAATTGTGTATGTAG	New, present study, primers designed from intron-1 region to identify exon-2	Genome walking; from Intorn-1 to exon-2
18	WT1GWBF1	CCTATGCAGGTTCAGAAGGATGTTTTACCA		Nested to WT1GW_FN
19	WT1INT2_1	GATTCTGTTGCTTTTGCTTTGTCA	New, present study, primers designed from intron-2 region to identify exon-3	Genome walking; from Intorn-2 to exon-3
20	WT1INT2_2	CACGGCTGCTGTAAATTGGTATA		Nested to WT1INT2_2
21	Wt1_exon4F	TGAAGACTCACACCAGGACTCATACAGG	New, present study, primers designed from exon-3 region and extreme 3' end of *cpWt1* **a**	Genome walking; to amplify intron-3 region
22	Wt1_exon5F	TCATCACAACATGCACCAGAGGAACA		
23	Wt1BFN1	aagcttATGCACCTAAGACAGTTCTACTCC	Primers designed to amplify coding region of *cpWt1* **a** with HindIII and BamH1 sites	Forward primer with HindIII site
24	Wt1BRN1	GgatccAAGGGCCAACTGCAGTTTAGTC		Reverse primer with BamHI site
25	Wt1BFN1	aagcttATGCACCTAAGACAGTTCTACTCC	Primers designed to amplify coding region of *cpWt1* **b** with HindIII and BamH1 sites	Forward primer with HindIII site
26	Wt1RSN1	ggatccTAAACGATCGGTAAGTTTGTTTC		Reverse primer with BamHI site
27	WT1SCF	ATGCACCTAAGACAGTTCTACTCC	Primers designed to amplify coding region of *cpWt1* **b** without any restriction sites	Forward and reverse primers to amplify coding region of *cpWt1* **b**
28	WT1SCR	TTATAAACGATCGGTAAGTTTGTTTC		
29	WT1SCF	ATGCACCTAAGACAGTTCTACTCC	Primers designed to amplify coding region of *cpWt1* **a** without any restriction sites	Forward and reverse primers to amplify coding region of *cpWt1* **a**
30	WT1SBR	AAGGGCCAACTGCAGTTTAGTC		

Table 3.3. Primers used in the study to isolate/charaterize the cpDmrt1 locus and cpDmrt1 isoforms

S.No.	Primer Name	Sequence (5'- to 3'-end)	Source	Purpose
1	Dmrt1F1	GGAACTAGGGATCAGCCACCCAGTC	Designed from conserved domain of *Dmrt1* sequences of Human, Mouse, *Gallus, Tranchemys, Pelodiscus* (Accession numbers **NM_021951, NM_015826, AF211349, AY316537, AB179697**) primers designed from *Trachemys* coordinates 15-535 bp)	To identify/obtain the initial fragment of the crocodile homologue of *Dmrt1*
2	Dmrt1R1	CCAGGTAGGGAGCAGGTGGGTAGTAAG		
3	DM5'	GTCCTGAATAAGCATCCGTCCCTCTG	New, present study; designed from the initial 500+ bp amplified/transcribed *cpDmrt1* fragment	Primer designed for the 5' RACE
4	DM5'N	GCTTGTTGAGTGAGTTGGGCTGCTGCT		Nested primer to the DM5'
5	DM3'	GCCGAGAGGCAGCGGGTGATG		Primer designed for the 3' RACE
6	DM3'N	AAGGAAAACAGTGGTGGAAGCTCCTGT		Nested primer to the DM3'
7	DMF1	GGAACTAGGGATCAGCCACCCAGTC	New, present study; Primers designed from commmon region present in all isoforms	To measure overall *cpDmrt1* levels by Real-Time RT-PCR
8	DMR1	TCTGATGGCGAGGTGGACACTGTAG		
9	GDMF	GAGGGACTGCCAGTGCAAAAAGT	New, present study; Primers designed from the conserved DM-Ser region present in all	To amplify a 322 bp probe sequence for Northern analysis of *cpDmrt1* /isoforms
10	GDMR	CAAGTCAGACGTGCTCTCCAGATG		
11	Dmrt1_4	CAAGCCAGTCAGGAAACCAGTG	Isoform specific, present study	Forward primer for real-time PCR for isoform *cpDmrt1* **e**
12	Dmrt1eR	TCAGCTCTTTTGTACATGGATTAT	Isoform specific, present study	Reverse primer for real-time PCR for isoform *cpDmrt1* **e**
13	Dmrt1_5	CCGCCACGTATATGTCAAGCCAGTCAGGA	Isoform specific, present study	Forward primer for real-time PCR for isoform *cpDmrt1* **d**
14	Dmrt1dR	GAGATACTGGATCCCACCCCTGAACCT	Isoform specific, present study	Reverse primer for real-time PCR for isoform *cpDmrt1* **d**
15	DMisoa1b1F	CTTCCCTGGTGTTACAGAAACCAACCTG	Isoform specific, present study	Forward primer for real-time PCR for isoforms *cpDmrt1* **a1/b1** fdesigned from the 3'UTR region

Table 3.3. Primers used in the study to isolate/charaterize the cpDmrt1 locus and cpDmrt1 isoforms

16	DMisoa1b1R	ACACCAGGGAAGTACACCATAAGCAA	Isoform specific, present study	Reverse primer for real-time PCR for isoform *cpDmrt1* **a1** and **b1** fdesigned from the 3'UTR
17	DMGW1	GCAGCCCAACTCACTCAACAAGCACAG	Present study	Primer for genome walking from exon-1 to intron-1
18	DMGW2	CAAGTCAGACGTGCTCTCCAGATG	Present study	Nested primer to the DMGW1
19	DMGW3	ATGGCATGACGGTTCTCGGAGCTCTTTAC	Present study	Primer for genome walking from exon-2 to intron-1
20	DMGW4	CTCCTGCATGTGTTCCTCAGATCTTCA	Present study	Nested primer to the DMGW3
21	DMGW5	TCACACTCCAGTACTCCTTTAGCGCTGTC	Present study	Primer for genome walking from exon-3 to intron-2
22	DMGW6	CTGATAGGAGAGCTGCTGGAAAGGGACAC	Present study	Nested primer to the DMGW5
23	DMGW7	CTCCTGCATGTGTTCCTCAGATCTTCA	Present study	Primer for genome walking from exon-2 to intron-2
24	DMGW8	GGTAGGTGCATGGTAACCCAAGGTTC	Present study	nested primer to DMGW7
25	GAPDHF	GAGCTCAATGGGAAACTTACTGG	In-house GAPDH sequence (from crocodile EST library/unpublished)	GAPDH primers for internal control
26	GAPDHR	AACCTGGTCCTCTGTGTATGCC		
27	UPM (Universal Primer mix Long primer)	CTAATACGACTCACTATAGGGCAAGCAGTGG TATCAACGCAGAGT	Clontech' SMART-RACE cDNA synthesis Kit primer	Primer against the adapter used in the primary PCR with the gene specific primer in both 5' and 3' RACE
28	Short primer	CTAATACGACTCACTATAGGGC		
29	NUP (Nested Universal Pimer)	AAGCAGTGGTATCAACGCAGAGT		Nested to Universal primer mix used for secondary PCR with the nested GSP
30	AP1	GTAATACGACTCACTATAGGGC	Clontech; Genomewalker Kit primer	Primer against genome walking adapter for use with GSP
31	AP2	ACTATAGGGCACGCGTGGT		Nested primer to AP1, used with nested GSP in secondary PCR

Table 3.4. Primers used in the study to isolate/charaterize the cpAromatase locus and cpAromatase isoforms

S.No.	Primer Name	Sequence (5'- to 3'-end)	Source	Purpose
1	AROF1	ATAGAGTGTCAGGGGCAGCGATCA	aromatase sequence from *Alligator, Trachemys, Gallus, Xenopus, Taeniopygia*, and *Homo sapiens* (Accession numbers **AY029233, AF178949, NM_001001761.1, NM_001076691** and **X13589**) AroF1, AroR1 and AroF2 and AroR2 spanned region of 577 to 1495 and 787 to 1270 bp respectively on the alligator aromatase mRNA	To identify/obtain the initial fragment of the crocodile homologue of *Aromatase*, subsequent RT-PCT and semiquantitative RT-PCR experiments
2	AROR1	GGCAATAAACTTCCCTACAC		
3	AROF2	CTGGCAAGCACTTTTACTGA		
4	AROR2	CACAACTGGCTGGTATCTCA		
5	Aro5'	CCACGTGCTCATCCAGTTTTTCAAC	New, present study; designed from the initial 800+ bp amplified/transcribed *cpAromatase* fragment	Primer designed for the 5' RACE
6	Aro5'N	TTCAGTGGTCACCTCTTGCAGTTTGTC		Nested primer to the Aro5'
7	Aro3'	TGGCTGTACAAGAAATATGAGGAGTC		Primer designed for the 3' RACE
8	Aro3'N	TGATGAAGGAAATTGAAACTGTTATG		Nested primer to the Aro3'
9	AGW2	TGGGCACGGTGGCAGCTGGCATGGC	New, present study; designed from, designed from the 5' end of *cpAromatase*	Primer designed for genome walking to explore upstream region
10	AGW1	TGGAAGCAATGGGGAAAGAGGCAGAC		Nested primer to AGW2
11	Aro_bigF	TGGGAGCAAGCTTGGGTTACAGTG	Designed from unique region of bigger isoform	In combination with AroF2 to amplify the unique region of the bigger isoform

Table 3.5. Primers used in the study to isolate/charaterize the novel SLF35F5 like ncRNA of crocodile

S.No.	Primer Name	Sequence (5'- to 3'-end)	Source	Purpose
1	SC35F	GGGAAACTCAAAGGCCTCAAATCCA	Present study; desigend from clone A4r of subtractive library	Amplification/expression analysis of subtracted EST clone A4r by RT-PCR, semiquantitaive RT-PCR and real-time RT-PCR
2	SC35R	CCCTTGAAAGCATTTGCACTGTAAGAA		
3	SC35F2	TAAGGCCATTAATGACAATGCACATCC	Present study; desigend from clone A4r of subtractive library	For the 5' RACE
4	SC35F3	GGAAACTCAAAGGCCTCAAATCCAGTA		Nested primer for SC35F2
5	SC35R	CCCTTGAAAGCATTTGCACTGTAAGAA		For the 3' RACE
6	SC35R2	TACTGGATTTGAGGCCTTTGAGTTTCC		Nested primer for SC35F2
7	SCINTF	TGGTCTACATTAAGGTTTTGCAAGTCT	Present study; designed from 2552 bp SLF35F5 like crocodile mRNA	For amplfication of 1700 bp of SLF35F5 like crocodile mRNA; also for 5' and 3' RACE
8	SCINTR	AGCCAAGGCATGAAAACCACTATAG		
9	SC35FN	AAGGCAGTGAGTTAGGGAAAGTC		Nested primer for SCINTR for 5' and 3' RACE
10	SC35RN	AAGCAATAATTCTGACTCCATGG		Nested primer for SCINTF for 5' and 3' RACE
11	T7SC35F	TAATACGACTCACTATAGGTGGTCTACA TTAAGGTTTTGCAAGTCT		Forward/reverse primers for preparation of *in-vitro* transcribed probe
12	T7SC35R	TAATACGACTCACTATAGGAAGGCAGTG AGTTAGGGAAAGTC		
13	SenseF	GGACTCCAGACATAATAGGGTTGGCTA	Present study; designed from the non-overlapping 3'-end of the sense transcript of novel ncRNA	Forward/revesre primers for the real-time RT-PCR to analyze expression of sense transcript
14	SCINT1R	GTTTCCCAGACAATGTGAACGTGTGTA		
15	antisense F	GTCTTCTCAAGTACCCCCAGGGTATG	Present study; designed from the non-overlapping 3'-end of the anti-sense transcript of novel ncRNA	Forward/revesre primers for the real-time RT-PCR to analyze expression of anti-sense transcript
16	SCINT1F	AATCTCTGGGTTTCGTCTTGTGACTTG		
17	FE_ANSEN	CAATAGAGCCTTGTTAATCTGCAGAG	Present study; designed from ends of anti-sense transcript of novel ncRNA	Forward/revesre primers for amplification of complete antisense transcript of novel ncRNA
18	RE_ANSEN	GCTAACATTGTGTGGCCTTCACAGC		
19	FE_SEN	CAGGGAGGAGGTAGTGCACATC	Present study; designed from ends of sense transcript of novel ncRNA	Forward/revesre primers for amplification of complete sense transcript of novel ncRNA
20	RE_SEN	GTGTACACGGCCTCAGCTGCATTAT		

Table 3.6. Miscellaneous primers used for: RACE, Genome-walking, and Real-time PCR for few sex determining genes

S.No.	Primer Name	Sequence (5'- to 3'-end)	Source	Purpose
1	sox9gdf	CAGCACAAGAAGGACCACCCCGACT	New, previous study in lab, primers designed from *cpSox9* conserved region	Forward and reverse primers for real-time RT-PCT for expression analysis of *cpSox9*
2	sox9_3'	ACTACAAGTACCAGCCCCGGAGGAGAA		
3	E3F2	CCGAACTGGAAAACCCATTGCTGTTA	New, present study, primers designed from the E3 clone of subtractive library (matches to *cpUNR*)	Forward and reverse primers for real-time RT-PCR and semi-quantitative RT-PCR for expression analysis of *cpUNR*
4	E3R2	TCCCCATTACGTTCGTAGCATACACT		
5	CAMHRACEF	GGAGCGGCTGGTGCGGTCGGAGGAG	primers from a separate study in the lab	Forward and reverse primers for real-time RT-PCR for expression analysis of *cpAmh*
6	CAMHRACER	GCAGCACCCTCCCTTCCAGCCGCCAG		
7	Srcf	GCACATGCACTCCCTCCAGGTCGAC	primers from a separate study in the lab	Forward and reverse primers for real-time RT-PCR for expression analysis of *cpSf1*
8	Sf1raf	CTCCCCGCTCAGGTGCTTGTGGTAC		
9	CC1F1	GCGCAGTCTGATCTTTCCAGCTCTTTA	New, present study, primers designed from subtractive library clones matching (j5) to CC1.3	Forward and reverse primers for real-time RT-PCT and semi-quantitative RT-PCR for expression analysis of crocodilain homologue of *CC1.3*
10	CC1R1	TTGGCTTGGAAGTCTGGCATTTATTTG		
11	Clk1F1	GACCCTCCTCATCATCCTCTACACT	New, present study, primers designed from subtractive library clone (c2) matching to Clk1	Forward and reverse primers for real-time RT-PCT and semi-quantitative RT-PCR for expression analysis of crocodilain homologue of *Clk1*
12	Clk1R1	TCTGTTCCTTGCCATGTTGTCTCCTG		

Table 3.6. Miscellaneous primers used for: RACE, Genome-walking, and Real-time PCR for few sex determining genes

13	A16F1	TGTGGAGTAAAACTGACCATTCCTTATCA	New, present study, primers designed from theA16 clone of subtractive library (matches to cpUNR)	*General primers for RT-PCR for cpUNR*
14	A16R1	TGTGGAGTAAAACTGACCATTCCTTATCA		
15	GAPDHF	GAGCTCAATGGGAAACTTACTGG	In-house GAPDH sequence (from crocodile EST library/unpublished)	GAPDH primers for internal control
16	GAPDHR	AACCTGGTCCTCTGTGTATGCC		
17	GAPDH_SMF	GAGCTCAATGGGAAACTTACTGG	In-house GAPDH sequence (from crocodile EST library/unpublished)	GAPDH primers for internal control
18	GAPDH_SMR	AACCTGGTCCTCTGTGTATGCC		
19	UPM (Universal Primer mix Long primer)	CTAATACGACTCACTATAGGGCAAGCAGTG GTATCAACGCAGAGT	Clontech' SMART-RACE cDNA synthesis Kit primer	Primer against the adapter used in the primary PCR with the gene specific primer in both 5' and 3' RACE
20	Short primer	CTAATACGACTCACTATAGGGC		
21	NUP (Nested Universal Pimer)	AAGCAGTGGTATCAACGCAGAGT		Nested to Universal primer mix used for secondary PCR with the nested GSP
22	AP1	GTAATACGACTCACTATAGGGC	Clontech; Genomewalker Kit primer	Primer against genome walking adapter for use with GSP
23	AP2	ACTATAGGGCACGCGTGGT		Nested primer to AP1, used with nested GSP in secondary PCR

DNA was spooled and washed twice with 70 % ethanol. The DNA was then dried and dissolved in TE (pH 8.0).

3.3.1.2 Isolation of RNA from tissues and Cells

RNA was isolated from soft tissue samples as well as from cultured cells using TRIzol reagent. Approximately, 20 to 50 mg of tissue samples were homogenized in 1 ml of TRIzol reagent and incubated for 5 minutes at ice. The homogenates were extracted for proteins using 200 µl of chloroform (thorough mixing for 15 seconds and incubation for 5 minutes) and then centrifuged at 13,000 rpm for 15 minutes at 4°C. The upper phase was collected, to which 500 µl of isopropanol was added and incubated at –20°C for 15 minutes. RNA was precipitated by centrifugation at 13,000 rpm for 15 minutes at 4°C. The RNA pellet was washed with 80 % alcohol, air-dried and dissolved in sterile water and stored at –70°C.

To isolate RNA from the cells, approximately 20,000 to 50,000 cells were pelleted down by centrifugation and washed twice with PBS to remove any traces of media. Pelleted cells were then suspended in Trizol for homogenization. After this, same steps as described above were followed for the RNA isolation.

3.3.1.3 Quantitation of RNA and DNA

RNA and DNA samples were quantified by spectroscopic method i.e. measuring O.D. at 260 and 280 nm wavelengths, using ND1000 Nanodrop (Thermo Scientific, USA).

3.3.2 Preparation of Ultra competent cells

Competent cells of *E. coli* DH5α strain were made following standard protocol. For the purpose, ~250 ml of LB medium was inoculated with 1% overnight culture and grown at 18°C till it reached an OD of 0.6 at 600 nm. The culture was then kept standing on ice for 10 - 15 minutes. The cells were pelleted at 2500g for 10 minutes at 4°C, resuspended in 80 ml of TB buffer (10 mM PIPES, 15 mM $CaCl_2$, 250 mM KCl, 55 mM $MnCl_2$ pH 6.7 with 1 N KOH) and incubated on ice for 10 minutes. The cells were again pelleted as before, suspended in 20 ml of TB buffer containing 7% DMSO and incubated on ice for 10 minutes. These cells were then aliquoted in ~200 µl volume, the aliquots were snap frozen in liquid nitrogen and stored at –70°C.

3.3.3 Transformation

The frozen competent cells were thawed slowly on ice just before use. Transforming DNA (50 - 100 ng) in a maximum volume of 10 µl was mixed with competent cells and incubated on ice for 20 - 30 minutes. The cells were subjected to heat shock for 90 seconds at 42°C followed by rapid chilling on ice. Approximately 800 µl of LB medium was added to the cells, which were then incubated for 1 hour at 37°C with gentle shaking. Transformed cells were pelleted and resuspended in 50 - 100 µl LB media and were then plated on LB-agar plates containing 100 µg/ml of ampicillin, 10 mM IPTG and 20 µg/ml X-gal. The plates were incubated at 37°C for 14 - 16 hours and the recombinant (white) colonies were picked up for subsequent plasmid isolation.

3.3.4 Plasmid DNA Isolation

The protocol of Alkaline lysis method as described by Sambrook et al. (1989) was used with minor modification for small scale isolation of plasmid DNA. Approximately 1.5 ml of bacterial culture was used. Cells grown overnight at 37°C (along with appropriate antibiotic) were pelleted at 10,000 rpm for 5 minutes. The bacterial pellet was resuspended in 100 µl of ice-cold solution I (50 mM Glucose, 25 mM Tris.HCl, and 1 mM EDTA, pH 8.0) by vigorous shaking. Afterwards 200 µl of freshly prepared solution II (0.2 N NaOH, 1% SDS) was added and mixed by inverting the tube gently for five times. Then 150 µl of solution III (3 M potassium acetate, pH adjusted to 4.8 with glacial acetic acid) was added and mixed by vortexing the tube gently in inverted position. The tube was incubated on ice for 5 minutes and centrifuged at 12,000 rpm for 10 minutes at 4°C. The aqueous phase was collected in a separate sterile tube and DNA was precipitated with 0.6 volume of isopropanol followed by washing with 80% alcohol. The plasmid DNA was air dried and dissolved in 50 µl of TE containing 20 µg/ml of RNase.

However, when large-scale plasmid DNA was required, plasmid clones were processed using Qiagen midiprep kit following the manufacturer's instructions.

3.3.5 Polymerase Chain reaction (PCR)

All routine PCR reactions were set in a 20 µl volume containing 10 - 25 ng of DNA, 1X PCR buffer, 150 µM dNTPs, 1.5 mM $MgCl_2$, 2 pmol of each primer and 1 U of Ampli Taq

Gold (Perkin Elmer). Thermal conditions for PCR were determined on the basis of primers used and /or information available in the published literature.

3.3.5.1 DNA sequencing

For sequencing, 50 - 100 ng of DNA (Plasmid and/or PCR product,) was mixed with 1 pmol of primer and 1.8 µl of BigDye Terminator v3.1 Ready Reaction Mix (Applied Biosystems, USA) in 5 µl reaction volume. The sequencing PCR was carried out in GeneAmp 9700 (Perkin Elmer) thermocycler using the profile: 30 cycles of denaturation at 96 °C for 10 seconds, annealing at 50 °C for 5 seconds, and extension at 60 °C for 4 minutes. The DNA was then precipitated with 50 µl of 95 % ethanol and 2 µl of 3 M sodium acetate followed by centrifugation at maximum speed for 20 minutes, washed twice with 80 % ethanol, air dried and dissolved in 10 µl of sequencing grade HiDi formamide (Applied Biosystem, USA). Finally the processed samples were analyzed on ABI 3700 or ABI 3730 automated DNA sequencers (Applied Biosystem, USA).

3.3.6 RT-PCR, 5′ and 3′ RACE (Rapid Amplification of cDNA Ends)

3.3.6.1 Reverse transcription and RT-PCR

For RT-PCR (conversion of RNA to a DNA copy) reactions, 2 - 3 µg of total RNA was reverse transcribed using the cDNA synthesis kit (Invitrogen) as per the manufacturer's instructions. The RNA was mixed with 1 µl of 10 mM dNTPs, 250 ng of Oligo-dT and/or 25 ng of random hexamer in a 10 µl reaction volume. The sample was incubated at 65°C for 5 minutes, quick chilled on ice for 2 minutes and then following components were added to the reaction: 2 µl of 10X RT buffer, 4 µl of 25 mM $MgCl_2$, 2µl of 0.1 M DTT and 1 µl of RNase Inhibitor. Subsequently, the reaction tube was incubated at 42°C for 2 minutes and 1 µl of Reverse Transcriptase was added to the reaction followed by incubation at 42°C for 90 minutes and inactivation of enzyme at 70°C for 15 minutes. The resulting first strand cDNA was diluted 5 times and 1-2 µl of diluted cDNA was used as template for subsequent RT-PCR reactions.

3.3.6.2 5′ and 3′ RACE

RACE is a PCR based technique for the amplification of 5′ and 3′ ends of mRNA transcript. This technique is employed when a part of gene is known and complete gene (including 5′ and 3′ ends) has to be identified. This involves PCR amplification using

gene-specific primer (GSP) and the primer specific to adapter sequence ligated to cDNA at either 5′ or 3′ end.

In the present study, RACE was done for several genes using Clontech's SMART RACE cDNA amplification kit. Sex and tissue specific 5′-/3′- RACE ready cDNA were prepared following the manufacturer's instructions. Briefly, 3-4 µg of total RNA was subjected to reverse transcription, for each RACE ready cDNA preparation. RNA was mixed with 1 µl of SMARTIIA oligo and 1 µl of 5′CDS primer in a 5 µl reaction volume for preparation of 5′ RACE ready cDNA and RNA sample was mixed with 1 µl of 3′CDS primer in 5 µl reaction volume for preparation of 3′ RACE ready cDNA. Sample was incubated at 70° C for 2 minutes and immediately chilled on ice for 2 minutes. This was followed by addition of 2 µl of 5 X First-Strand buffer, 1 µl of DTT (20 mM), 1 µl dNTP Mix (10 mM) and 1 µl PowerScript Reverse Transcriptase to each sample, tubes were then spinned briefly and incubated at 42°C for 1.5 hours. Reverse transcriptase was inactivated by heating the samples at 72°C for 15 minutes. Resulting first strand cDNA products were diluted 7 fold with sterile water and 1 µl of the diluted sample was used as template in each RACE PCR experiments for amplification of cDNA ends using gene-specific and adapter-specific primers. All RACE-PCR reactions involved two rounds of amplification steps. Gene-specific primer drives the specific reaction, whereas adapter-specific primer is common to all cDNAs, therefore, first amplification often resulted in to either a smear or multiple amplified products. Second round was performed with the GSP2 primer (nested primer to gene specific primer used in first round) and nested adapter primer to bring in more specificity, using the diluted (100 times) amplified product from first round. Details of specific RACE reactions used for different genes are described in the corresponding sections.

3.3.7 Northern Analysis:

About 15-20 µg RNA was electrophoresed in a 1 % denaturing formaldeyde agarose gel. The agarose gel was prepared in 1X MOPS buffer and samples were electrophoresed at 60 V for 4-5 hours. The eletrophoresed RNA samples were then capillary transferred to Hybond N^+ membrane (Amersham Biosciences) in 10X SSPE for 12-15 hours, after transfer blots were subjected to UV crosslinking. The RNA blot thus prepared was hybridized using radiolabeled probe (10^7- 10^8 cpm) at 65°C for 16 hours (hybridizations were done in Church Buffer). The blot was washed twice (10 minutes each) at 65°C with 2X SSC, once each with 1X SSC and 0.1X SSC, all washing solutions contained 0.1%

SDS. Finally the blot was drained and analyzed for hybridization signals in phosphor imager (Fuji-Film).

3.3.8 Preparation of radiolabeled probe for Northern hybridization

The cDNA, which was to be used as probe, was radiolabeled using Random Priming Kit (Jonaki) in 50 µl reaction volume that comprised 100 - 300 ng of cDNA, random 9-mer primers, dNTPs (0.8 mM each), 3 µCi of α-P^{32} dATP and 2 units of Klenow enzyme. The labeling reaction was carried out at 37°C for 2 hours.

In some cases, probe for Northern hybridization was also radiolabeled by PCR amplification. For the purpose, PCR was done in 50 µl reaction volume comprising primers for the target region, approximately 20-50 ng of cDNA as template, 5 µl 10X PCR buffer, 4 µl 25mM $MgCl_2$, 100 µM dCTP, dTTP, dGTP and 50 µM of αP^{32} dATP with 10 µM dATP was also added with 2U of Taq DNA polymerase. Cycling conditions were kept as follow: 94°C initial denaturation, 35 cycles of 94°C for 45 seconds, 62°C for 45 seconds and 72°C for 45 seconds, final extension at 72°C for 5 minutes.

In both cases, unincorporated radioactive dNTP were removed by passing PCR product through a sephadex G50 column. Northern hybridization was done as described before.

3.3.9 Real-time RT-PCR

Real-time RT-PCR was done using ABI 2X SYBR green PCR mix. Reactions were carried out in 10 µl reaction volume. Reactions were set in triplicates using 5 µl of SYBR green mix and 2.0 pM of primers and appropriate quantity of cDNA. GAPDH was used as internal control using GAPHSF and GAPDHSR primers (Table 3.5). As a control to test PCR efficiency of each gene, standard curves were made for each primer pair with different dilutions of cDNA (any sex or stage). Cycling conditions comprised: 55°C for 2 minutes, 95°C for 10 minutes for initial denaturation, 40 cycles for 95°C for 15 seconds 62°C for 15 seconds and 72°C for 30 seconds, followed by a dissociation curve analysis 95°C for 15 seconds 65°C for 15 seconds and 95°C for 15 seconds. Experiments were done mainly for $\Delta\Delta$Ct values to analyze comparative expression levels. Annealing temperature was standardized separately for every primer set and dissociation curve was examined for the presence of additional peaks, Only those primer pairs were selected which yielded a single peak in dissociation curve analysis.

All PCR reactions were done in ABI 7900HT real-time PCR (Applied Biosystem, USA). Data analysis was done using ABI's SDS2.1 software package (Applied Biosystem, USA).

3.3.10 Genome walking

Genome walking is a strategy to explore any genomic region for which some sequence information is already available *i.e.* either RNA sequence arising from the locus or a part of locus for which flanking region has to be explored. In this study, genome walker kit from Clontech was used for this purpose. Briefly genomic DNA was digested with the four blunt end hexacutter restriction enzymes separately (*Dra*I, *EcoR*V, *Stu*I and *Pvu*II). Each digested DNA pools was then purified and ligated with the Genome walker adapter in a 10 µl reaction volume. Ligation reaction comprised.

Digested and purified DNA (500 ng) + water	6.1 µl
Genome walker adapter (25 µM)	1.9 µl
10X T4 DNA ligase buffer	1.0 µl
T4 DNA ligase	1.o µl (10U)

Ligation was carried out at 16^{0}C for 14-16 hours in a thermocycler after which the reaction was stopped by heating at 70^{0}C for 10 minutes. Finally each digested and ligated pool was then diluted to 80 µl by adding sterile water to obtain PCR ready genome walker libraries.

The above libraries were used as template to carryout genome walking reactions using GSPs. Each genome walking PCR (performed in two rounds of amplification) comprised: 1 µl of diluted genome walking library, 1µl of the genome walking adapter primer 1 (AP1), 5 pM of gene specific primer, 2 µl of the 10X Advantage PCR buffer, 150 µM dNTPs and 0.4 µl of 50X Advantage polymerase mix in a 20µl reaction volume. Cycling conditions were: 95^{0}C for 2 minutes, followed by 22 cycles of 95^{0}C for 15 seconds, 68^{0}C for 3 minutes and final extension of 72^{0}C for 5 minutes. Secondary PCR reactions were then set in 20 µl reaction volumes; each reaction contained 1 µl of 50 fold diluted primary PCR product as template, 1 µl nested adapter primer (AP2), 5 pM of nested gene specific primer, 2 µl of 10X Advantage PCR buffer, 150 µM dNTPs and 0.4 µl of the 50X Advantage Polymerase mix. Cycling conditions were same as above except the number of cycles was reduced to 14 from 22. PCR products examined for successful amplification were purified (using Qiagen's PCR purification columns) and cloned in the

T/A cloning vector (T/A cloning kit, invitrogen). Plasmids were sequenced using M13 universal primers to obtain the derived genomic sequence.

3.3.11 Fluorescence *In situ* Hybridization

3.3.11.1 Tissue Sectioning

Immediately after dissection, embryonic tissues were fixed in 4% formaldehyde and 0.5 M phosphate buffer. The tissues were then washed in a graded series of ethanol baths: 30 %, 50%, 70%, 90% and 100% for 30 minutes each at room temperature and then embedded in paraffin wax blocks. These blocks were cut into 10 µm thick sections on a Leica microtome followed by mounting onto slides. The slides were then incubated at 37°C overnight on a hot plate before storage at room temperature till further use.

3.3.11.2 Probe Preparation

Generally, cDNA was used as probe to study localization of target transcript(s). For the purpose, purified cDNA was labeled with FITC using random priming kit from Strategene (Prim –It Fluor Fluorescence Labeling Kit). Approximately 200 - 250 ng of target gene-specific DNA (PCR product) was used for each labeling reaction in 50 µl volume as per the manufacturer's instructions. The probe was then ethanol precipitated, air dried and dissolved in 15 µl of sterile water.

3.3.11.3 Deparaffinization of tissue sections

Deparaffinization of stored tissue sections was carried out as follows: washing three times for 5 minutes in xylene; two times for 5 minutes in 100% ethanol; two times for 5 minutes in 95% ethanol; and once for five minutes in 80% ethanol. The sections were then placed in endogenous blocking solution (methanol + 2% Hydrogen peroxide) for 20 minutes at room temperature followed by rinsing twice for 5 minutes each in deionized water. Rehydration of sections was then carried out with 1 X phosphate buffer saline (PBS, pH 7.4)

3.3.11.4 *In situ* hybridization

Deparaffinized sections were treated with 5 µl of proteinase K (20 mg/ml) diluted in 100 µl of 1 X PBS for 3 minutes. The sections were washed with 1 X PBS for 5 minutes and air-dried. Fluorescent labeled cDNA probe (2 µl) was mixed with 27 µl of hybridization mix (see below) and denatured by keeping in boiling water for ~5 minutes followed by

quick chilling on ice. The sections were then hybridized, covered with cover slip and sealed with rubber solution.

Composition of Hybridization mix (1ml) containing 50 % formamide

Component	*Stock/amount*	*Final Concentration*
Formamide (Sigma)	500 µl	50 %
20 X SET	200 µl	4 X
Dextran sulfate	100 mg	10 %
Sheared *E. coli* DNA (1 µg/µl)	20 µl	20 µg/ml
Heparin (5 mg/ml)	100 µl	500 µg/ml
H_2O	up to final volume 1 ml.	

The slides were incubated at 37°C in humid-chamber for 12 - 16 hours. After hybridization, the tissue sections were washed as follows:

Washing solution	*Temperature*	*Time/purpose*	*Washes*
2 X SSC	RT	*to remove cover slip*	
2 X SSC	45 °C	15 minutes -	1 wash
50 % formamide + 2 X SSC	45 °C	10 minutes -	3 washes
1 X SSC	45 °C	15 minutes -	2 washes
2 X SSC	45 °C	15 minutes -	4 washes
2 X SSC	4 °C	overnight -	2 changes

After washing, excess 2 X SSC was drained out and 20 µl of mountant with counter stain (Vecta-sheild mounting medium with PI or DAPI) was added to each slide. The slides were covered with coverslip.

3.3.11.5 Immunohistochemistry

Specific polyclonal antibody (Aromatase) was applied as primary antiserum. The deparaffinized sections were incubated for 30 minutes at room temperature in a humid chamber and then washed twice with 1 X PBS. The FITC-labeled anti IgG against primary antibody, at a 1:200 dilution was then added as a secondary antibody and incubated at room temperature for 30 minutes in a humid chamber protected from light to

avoid quenching of fluorescence. The propidium iodide (PI) was used as the counter stain. The sections were washed with 1 X PBS and mounted with vectashield mounting medium. The staining patterns were recorded with an excitation wavelength of 488 nm using Axioplan-2 microscope (Ziess)/LSM 510 Meta Confocal microscope (Zeiss).

3.4 Isolation/characterization of homologues of conserved sex-determining genes from *Crocodylus palustris* (Indian mugger)

(Information about all the primers mentioned in this section is given in Tables 3.2 - 3.6).

3.4.1 Isolation and expression analysis of *cpWt1* (*Wt1* homologue of *Crocodylus palustris*)

3.4.1.1 Primer design and initial identification of *cpWt1*

Wt1 sequences from the non-mammalian and mammalian vertebrates species were retrieved from NCBI, were aligned and primers Wt1F1 and Wt1R1 were designed to amplify a 490 bp conserved region. The PCR product obtained using these primers was sequenced for confirmation.

3.4.1.2 Expression analysis

The sex-/stage- specificity of *cpWt1* expression in GAM tissue was ascertained by semi-quantitative RT-PCR. For the purpose WT1F1 and WT1R1 primers were used along with cDNA (as template) prepared from the RNA isolated from GAM tissues of the embryos kept at MPT and FPT, for all five developmental stages (21-25) representing TSP. PCR was performed in a 20 µl reaction volume containing 1 µl cDNA as template, 2 µl 10X PCR buffer, 1 µl of WT1F1 (5 pM), 1 µl of WT1R1 (5 pM), 150 µM dNTPs, 1.5 mM $MgCl_2$, 1U Taq DNA polymerase. GAPDH was used as internal control using GAPDHF and GAPDHR primers. Cycling conditions used for PCR comprised: initial denaturation at 94^0C for 5 minutes followed by 30 cycles of 94^0C for 45 seconds, 58^0C for 30 seconds and 72^0C for 45 seconds and final extension of 5 minutes at 72^0C.

3.4.1.3 RACE analysis

Primers for 5′ and 3′ RACE were designed from the 490 bp long *cpWT1* region amplified by the WT1F1 and WT1R1 primers. Primers WT15′ and WT15′N were designed for 5′ RACE, that was carried out using both male and female RACE ready cDNA. RACE reactions were done in two rounds as described in section 3.3.6.2. In the first round primer WT15′ was used with the universal primer mix. Primary PCR products were diluted 30 times to generate template for secondary PCR that was carried out using WT1N5′ and nested universal primer. Primers for the 3′RACE were designed after sequencing of the 5′ RACE products. 3′ RACE was also done in two rounds with both male and female RACE ready cDNA, first round of amplification was done using primer AWT15′with Universal primer mix and 1 µl of 30 times diluted primary PCR products were used for the secondary PCR with AWT15′N and nested universal primers. RACE analysis revealed two isoforms of *cpWt1* differing significantly in their size.

3.4.1.4 Expression analysis of bigger and smaller *cpWT1* isoforms

Sex-/stage-specific expressions of the two identified *cpWT1* isoforms were ascertained by semi-quantitative RT-PCR using isoform specific primers (primers WT1F1 and WT1R1 for isoform *cpWt1***a** and primers WT1SMF and WT1SMR for isoform *cpWt1***b**: Table 3.2). PCR and cycling conditions were same as described in section 3.4.3.2.

Isoforms *cpWt1***a** and *cpWt1***b** expression in GAM tissues of embryos kept at MPT and FPT through TSP was also analyzed by Real-time RT-PCR. Reaction conditions/cycling parameters and procedures for data analysis were as described in section 3.3.9. Primer pairs used for *cpWt1***a** and *cpWt1***b** were: WT1exon 3F/WT1exon5R and WT1SMF/Wsmr, respectively.

3.4.1.5 *In-situ* Hybridization

Localization of the Isoforms *cpWt1***a** and *cpWt1***b** was studied in the developing gonads and the mesonephros by *in-situ* hybridization using isoform specific cDNA probes. A part of the unique region of each of the two isoforms was used separately as probe to examine expression of individual isoforms. For bigger isoform, primers WT1exon3f and WT1bigR4 (Table 3.2) amplified a 632 bp region which was not present in the smaller isoform. Similarly, primers WT1SMF and WT1SMR amplified a region of 356 bp in the smaller isoform that was not present in the bigger isoform. Probes were prepared using Prime–It Fluor Fluorescence Labeling Kit as described in section 3.3.11.2.

Ten micron thick de-paraffinized sections of GAM tissue, representing developmental stages from 21 to 25 of embryos kept at MPT and FPT were hybridized separately with isoform specific fluorescent labeled probes. In-situ hybridization was done as described in section 3.3.11. Sections were counter stained with Propidium iodide. Tissue sections hybridized with hybridization buffer without probe and RNase treated sections hybridized with the labeled probes served as control for the autofluorescence and non specific hybridization. Slides after hybridization were visualized in the LSM 510 confocal microscope.

3.4.1.6 Genome walking for *cpWT1*

Genomewalker libraries prepared earlier (3.3.10) were used for the genome walking experiments to explore *cpWt1* locus and its putative upstream region. To study genomic organization of *cpWt1,* genome walking was performed from the 5′ end of the *cpWt1* transcript in order to maximize chances of not missing any intron, Primers AWT5′N and its nested primer WT1BGWF1 (Table 3.2) were used for this purpose. Reaction conditions and cycling conditions were same as described (section 3.3.10). Products from the primary and secondary PCR were cloned and sequenced. Similarly to explore the upstream region of *cpWt1,* genome walking was attempted using primers GWTR and nested primer WT1GW5′N for the primary and secondary PCR respectively. PCR products were cloned and sequenced.

3.4.1.7 Cloning of *cpWT1* isoforms in the EGFPN1 vector

In an attempt to study function of the *cpWt1* isoforms, coding region of both *cpWt1* isoforms were cloned in the EGFPN1 vector and transfected in the primary cell lines derived from the embryonic GAM and Brain tissues. EGFPN1 vector codes for enhanced Green Fluorescent Protein (EGFP) and any reading frame can be cloned to the N-terminal of the EGFP. Primers were designed to clone the entire reading frame of both *cpWt1* isoforms (i.e. from start codon to last amino acids excluding stop codon). Forward primer WT1BFN1 (starts from ATG) was same for both isoforms and has a *Hind*III site at its 5′ end while the reverse primers were isoform specific i.e. WT1BRN1 and WT1RSN1 for the isoforms *cpWt1***a** and *cpWt1***b,** respectively. Both reverse primers had *Bam*HI sites at their 5′ end. Male GAM cDNA pooled over TSP was used as template to amplify both *cpWt1* isoforms. Reactions were set in 20 µl volumes using 1 µl cDNA as template, 5 pM of each forward and reverse primers, 150 µM dNTPs, 2 µl 10X PCR buffer and 1U of AmpliTaq gold. Cycling conditions were: initial denaturation at 95^0C for 10 minutes

followed by 35 cycles of 95^{0}C for 30 seconds 60^{0}C for 30 seconds and 72^{0}C for 90 seconds and final extension for 5 minutes, PCR products were cloned in the InsTAClone vector. Plasmids were isolated by alkali lysis method (section 3.3.4). Recombinant plasmids were digested with the *Hind*III and *Bam*HI enzymes in 25 µL reaction volume containing 2.5 µL NEB buffer-2, 2 µg plasmid DNA and 5 U of each *Hind*III and *Bam*HI enzymes. After 6 hours of incubation, complete reaction was loaded on agarose gel to separate and isolate the isoform specific insert having sticky ends generated by *Hind*III at 5′ end and *BamH*I at 3′ end. The released insert was eluted from agarose gel using Gel Elution Kit (Qiagen). Vector pEGFPN1 (2.0 µg) was also digested in the same manner as described above and purified after digestion using Qiagen's PCR purification kit following manufacturer's instructions. Both eluted bigger and shorter isoform fragments were ligated to the digested/purified pEGFPN1 separately as described below.

Linearized pEGFPN1 vector (40 ng)	1 µl
cpWT1(bigger/smaller) (120ng)	2 µl
T4 DNA ligase (2U)	1 µl
10X T4 DNA ligase buffer	1 µl
Sterile water	5 µl

Reactions were incubated for 8 hours at 22^{0}C. The ligation mix was transformed in the *E. coli* DH5α competent cells, Plasmids were isolated and sequenced to confirm proper in-frame insertion of *cpWt1* isoforms. The resultant constructs containing, *cpWt1***a** and *cpWt1***b** isoforms was referred as cpWt1aEGFPN1 and cpWt1bEGFPN1.

3.4.1.8 Transient transfection studies

The constructs cpWt1aEGFPN1 and cpWt1bEGFPN1 were transfected separately to the primary cell lines derived from the embryonic GAM and Brain tissues of *C. plaustris* (available in the lab at CCMB). Transfection was done using liposome mediated DNA delivery methods. Cells were subcultured and allowed to grow at 26^{0}C. Transfections were done in 25 mL flask or on glass coverslips. Normally at the time of transfection, cell number should reach around 2.5–7.5 x 10^6 in a 25 mL flask and 0.5–2.0 x 10^5 on the

coverslips. For transfection in the flasks, cells were first washed with the PBS and 3 mL of fresh media was added. DNA-effectene complex was prepared by mixing 2 µg of DNA (EGFP1/cpWt1aEGFPN1/cpWt1bEGFPN1) with the 16 µl of enhancer and 24 µl of effectene reagent following the manufacturer's instructions. DNA with effectene was incubated for 15 minutes and then mixed with 1 mL of media (DMEM with 20% serum) and the media containing effectene DNA complex was added to flask. For transfection on the coverslips, 0.5 µg DNA was mixed with the 0.8µl of the enhancer and 6µl of the Effectene reagent, this reaction mix was allowed to stay for 15 minutes before adding 100 µl of the DMEM (with 20% serum). Cells on the coversplips were also washed with PBS gently and were added with fresh media. Media containing the transfection complex was added to the cells gently drop by drop in both experiments and gentle swirling was done to ensure proper mixing and the cells were allowed to grow for 24-48 hours.

3.4.1.8.1 Localization study

After 24 hours of transfection, cells growing on coverslip were washed with PBS, fixed in 4% formaldehyde and counterstained with DAPI. Fixed cells transfected separately with EGFP1/cpWt1aEGFPN1/cpWt1bEGFPN1 were examined under Axioplane 2.0 microscope and LSM 510 confocal microscope to detect GFP localization in the cells.

3.4.1.8.2 RNA isolation from the cell and reverse transcription

After 36-48 hours of growth, transfected cells were scraped to bring them in suspension. Detached cells were centrifuged at 2,000 rpm and total RNA was isolated from pelleted cells using Trizol method as described in section 3.3.1.2. One µg of RNA was taken for reverse transcription using Superstrand first strand cDNA synthesis system (Invitrogen Corporation) as described in section 3.3.6.1.

3.4.1.8.3 Real-time RT-PCR

RT products prepared above from the transfected cells (having EGFP1 or cpWt1aEGFPN1 or cpWt1bEGFPN1 construct) were subjected to real-time RT-PCR for measuring expression levels of few sex determining genes such as *cpSox9*, *cpSf1*, *cpDmrt1*, *cpUnr* (Upstream to N-ras), *cpAMH*, SC35F5 like nc RNA, *cpWt1***a**, *cpWt1***b**, crocodilian homologue of *CC1.3* and GAPDH as internal control. Primers used in this study are mentioned in Table 3.2, 3.3 and 3.5. Real-time RT-PCR was carried out as mentioned in section 3.3.9 and data was analyzed using SDS2.1 software. Each

transfection experiment was repeated thrice, followed by real-time analysis of expression levels of the above mentioned genes to ensure reproducibility.

3.4.2 Isolation and expression analysis of *cpDmrt1* (*Dmrt1* homologue of *Crocodylus palustris*)

3.4.2.1 Primer design for initial identification of *cpDmrt1*

Dmrt1 sequences from other vertebrate species especially reptiles were aligned and primers (Dmrt1F1 and Dmrt1R1) (Table 3.3) were designed from the conserved region spanning 512 bp. Amplified product was cloned and sequenced for confirmation of *cpDmrt1*

3.4.2.2 Expression analysis

Primers used above (Dmrt1F1 and Dmrt1R1) were used also to analyze sex-/stage-specific *cpDmrt1* expression analysis for the *cpDmrt1* in the GAM tissues of *C. palustris* kept at MPT and FPT through TSP. All these PCR reactions were done for 32 cycles with annealing temperature of 58°C. GAPDH primers were used as internal control.

3.4.2.3 RACE Analysis

RACE was done using 5′ and 3′ male and female RACE ready cDNA. Primers were designed from the 512 bp *cpDmrt1* sequence obtained in the first step (3.4.2.1). Primers DM5′ and DM5′N were used for the 5′ RACE. Primary PCR was done using DM5′ and universal primer mix. 5′ RACE ready cDNA from male and female GAM were used as template in separate reactions. PCR reactions were set in a 20 µl reaction volume containing 1 µl diluted 5' RACE ready cDNA as template, 1 µl universal primer mix, 1 µl DM5′N, 2 µl of 10X Advantage PCR buffer, 150 µM dNTPs and 0.4 µl of 50X advantage taq polymerase mix. PCR conditions used were: initial denaturation at 95^0C for 2 minutes followed by 28 cycles of 95^0C for 15 seconds, 68^0C for 30 seconds and 72^0C for 2 minutes and a final extension step of 72^0C for 5 minutes. Primary PCR products were diluted 30 times and 1 µl of diluted sample was used as template for the secondary PCR. Secondary PCR reaction was done with DM5′N and nested universal primers. Cycling conditions were same as above except for the number of cycles that was reduced to 16. 3′ RACE reactions were done using same conditions as used for 5′ RACE except that in these reactions, 3' RACE ready cDNA was used as template, primer DM3′ was used for primary PCR and primer DM3′N for the secondary PCR.

3.4.2.4 Real-time PCR for expression analysis of *cpDmrt1*

Real-time PCR was done as described in section 3.3.9. Overall *cpDmrt1* expression levels were analyzed using primers designed from the region conserved in all *cpDmrt1* isoforms, namely GDMR1 and DMSLGW2 (Table 3.3). Similarly, expression pattern of four of the isoforms; *cpDmrt1***a1**, *cpDmrt1***b1** (combined expression), *cpDmrt1***d** and *cpDmrt1***e** were analyzed using isoform specific primers. *cpDmrt1***a1** and *cpDmrt1***b1** isoforms were amplified using primers Dmrt1a1b1F and Dmrt1a1b1R, isoform *cpDmrt1***d** with primers Dmrt1dR and Dmrt1_4, and isoform *cpDmrt1***e** with primers Dmrt1eR and Dmrt1_5. Ct values were determined after subtracting standard deviation from the average Ct value. ΔΔCt values were then calculated to compare the expression levels. Data was analyzed using SDS2.1 software.

3.4.2.5 Genome walking for the *cpDmrt1* intronic sequence

Nested primers DMGW1 and DMGW2 were used to amplify the region from exon-1 to intron-1 and primers DMGW3 and DMGW4 for amplification of intron-2 from the exon-3. Similarly, primer DMGW4 amplified the intronic region after exon-3 and primers DMGW5 and DMGW6 amplified intron-3 from the exon-4. Genome walking PCRs were carried out as described in section 3.3.10. In each case, amplified products were cloned and sequenced to obtain sequence information.

3.4.3 Isolation and expression analysis of *cpAromatase* (Aromatase homologue of *Crocodylus palustris*)

3.4.3.1 Primer Design and identification of *cpAromatase*

Aromatase sequences of reptilians and other non-mammalian vertebrates were aligned and primers were designed from conserved region. Two sets of primers were designed namely Aromatase F1, R1 and aromatase F2, R2 (table 3.4) and were subjected to PCR with the template first strand cDNA (prepared from pooled male and female GAM RNA). Cycling conditions were: initial denaturation at 94°C for 10 minutes, followed by 32 three step cycles (94°C for 30 seconds 62°C for 30 seconds and 72°C for 1 minute) and final

extension step of 72°C for 5 minutes. Fragments were cloned in T/A cloning vector and positive clones were isolated and confirmed by sequencing.

3.4.3.2 RACE for *cpAromatase*

Primers for 5′ and 3′ RACE were designed from the 900 bp fragment of *cpAromatase* obtained in the first step (section 3.4.3.1). Primers Aro5′ and Aro5′N were designed for 5′ RACE and Aro3′ and Aro3′N for 3′ RACE. As described in section 3.3.6.2, RACE was done in 2 rounds of PCR amplification. Primary PCR was carried out in a 20 µl reaction volume using 1 µl RACE ready cDNA as template with 1 µl universal primer mix, 1 µl gene specific primer (Aro5′ with 5′ RACE ready cDNA as template and Aro3′ with 3′ RACE ready cDNA as template), 2 µl of 10X Advantage PCR buffer, 150 µM dNTPs and 0.4 µl of 50X Advantage Taq polymerase mix. PCR conditions were: 94°C for 2 minutes for initial denaturation followed by 25 cycles of 94°C for 30 seconds 65°C 15 seconds 72°C for 90 seconds and a final extension step of 72°C for 5 minutes. Primary PCR product was diluted 30 times and 1 µl of diluted sample was used as template for secondary PCR. Secondary PCR was done in a 20 µl reaction volume with nested universal primer and nested gene specific primer (Aro 5′N for 5′ RACE and Aro3′N for 3′ RACE). Cycling conditions were kept same as above except the number of cycles was reduced to 15. PCR products were cloned in T/A cloning vector and positive clones were sequenced.

3.4.3.3 Expression analysis by semi-quantitative RT-PCR

Total RNA isolated from GAM tissues of embryos kept at MPT and FPT, respectively, representing all the five developmental stages (21-25) spanning TSP, were reverse transcribed. The resulting single stranded cDNA were then subjected to semi-quantitative RT-PCR to analyze expression pattern of *cpAromatase* at MPT and FPT through TSP using primers AroF2 and AroR2. PCR profile was kept same as described in section 3.4.1.1 except it had 28 cycles instead of 32 cycles. RT-PCR products were electrophoresed in 1.5 % agarose gels, stained in ethidium bromide and visualized in UV. Expression analysis of *cpAromatase* was also done, in the brain tissue of the embryos at MPT and FPT through TSP, in the same manner as described above.

3.4.3.4 Isoform specific RT-PCR

Two Aromatase isoforms were detected after 5′ RACE. The smaller isoform lacked approximately 150 bp region near its 5′ end. Primer ArobigF was designed

corresponding to region absent in the smaller *cpAromatase***b** isoform. This was used with primer AroF2 to examine tissues/stage/sex specific expression of *cpAromatase* isoforms.

3.4.3.5 Northern Hybridization

A 500 bp region of cpA*romatase* cDNA amplified by primers AroF2 and R2 was used as probe in Northern hybridization. The radiolabeling of probe was achieved using PCR amplification by including α-P^{32} dATP in the PCR reaction mix (see section 3.3.8). Northern hybridization was done as described before (section 3.3.7).

3.4.3.6 Immunostaining

Aromatase localization was studied using anti Aromatase antibody raised in goat (obtained from Abnova, USA). Immunostaining was done as described in section 3.3.11. FITC conjugated goat antirabbit antibody was used as secondary antibody.

3.5 Search for differentially expressed genes during early TSP

An attempt was made to identify gene(s) showing differential expression in the GAM tissues of the embryos kept at MPT and FPT especially during early TSP (developmental stages 21-22). This was done using highly efficient approach of the SSH (suppressive subtractive hybridization). The schematic of SSH is described in figure 3.1.

3.5.1 RNA isolation and double stranded cDNA synthesis

RNA was isolated from the embryonic GAM tissues from embryos of 21^{st} and 22^{nd} stages kept at MPT and FPT. RNA was isolated using Trizol as described in section 3.3.1.2. Approximately 2 µg of RNA (pooled separately for GAM tissues at MPT and FPT of stages 21 and 22) was subjected to cDNA synthesis using Clontech SMART PCR cDNA synthesis kit. Double stranded cDNA was prepared as per manufacturer's instructions. Briefly, RNA was mixed with CDS primer and BD SMARTII oligonucleotide to make a total volume of 5 µl. The mix was denatured by heating at 65^0C for 2 minutes followed by snap cooling for 2 minutes. The mix was incubated at 42^0C for 90 minutes after addition of: 5 X first strand cDNA synthesis buffer (2 µl), 100mM DTT (1 µl), dNTPs (1 µl) and MMLV RT (1 µl). Reverse transcriptase was inactivated by heating at 72^0C for 5 minutes. Three µl from the resulting first strand cDNA was then subjected LD-PCR for generating

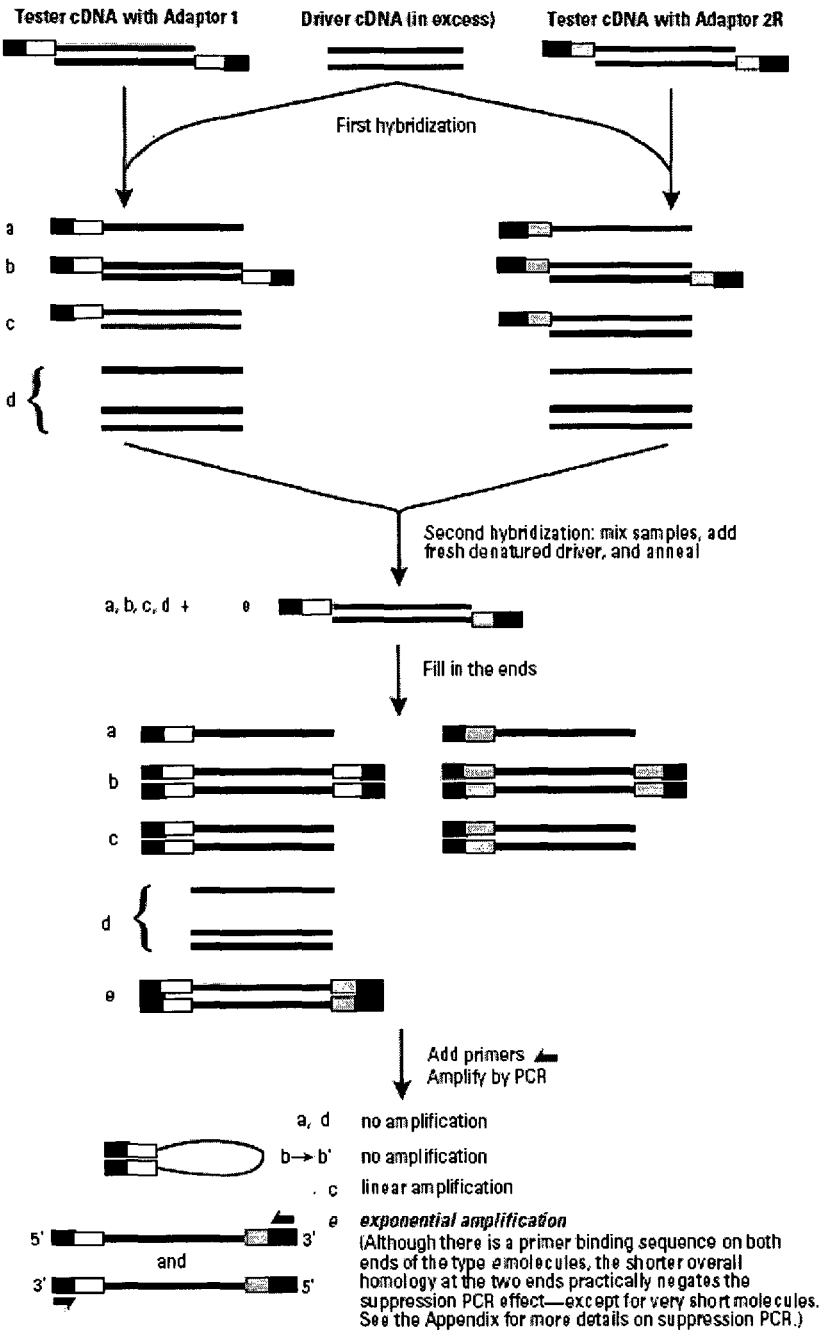


Figure 3.1: **The schematic representation of the suppressive subtraction hybridization strategy** (adapted from Clontech PCR select kit manual) used in the present study to identify differentially expressed genes in the GAM tissues at MPT.

double strand cDNA using 5′ PCR primer IIA which has complementary sites to CDS primer that tags 3′ end with its linker and SMARTII oligo that tags 5′ end of the transcript. The 100 µl LD-PCR reaction comprised: 3 µl of RT product, 10 µl of 10X Advantage 2 PCR buffer, 2 µl of 50X dNTP mix, 2 µl of 5′ PCR primer IIA, 2 µl of 50X Advantage polymerase mix and 71 µl of deionized water. The cycling conditions for PCR were: 94^{0}C for 1 minute (initial denaturation) followed by 16 cycles of 94^{0}C for 15 seconds, 65^{0}C for 30 seconds and 68^{0}C for 6 minutes. The resulting double stranded cDNA was purified by alcohol precipitation. The total LD-PCR product (100 µl) was mixed with 5 µl of 3M sodium acetate and 100 µL isopropanol followed by incubation at -20^{0}C for 4 to 5 hours. cDNA were precipitated by centrifugation at 14000 rpm for 20 minutes. cDNA pellets were washed in 80% ethanol and air dried. Finally, each cDNA prep was dissolved in 8.0 µL of sterile water and quantitated using spectrophotometer.

3.5.2 Restriction digestion of cDNA

Each cDNA prep (made from female and male GAM tissue RNA) was separately digested with RsaI (blunt end tetra cutter restriction enzyme). Care was taken that cDNA quantity was at least 2 µg for the tester cDNA and 4 µg for the driver cDNA for a single subtraction experiment (involving both forward and reverse subtraction). Restriction digestion was carried out in 50 µl volume reaction containing 2-4 µg of cDNA, 5 µl of 10X NEB buffer 2 and 10 to 20 units of *Rsa*I, at 37^{0}C for 4 hours followed by purification as described above. The purified, digested DNA was finally dissolved in 6 µl of sterile water.

3.5.3 Ligation of adapter to digested cDNA

The tester cDNA was divided in two parts, and ligated with two separate adapters (adapter 1 and adapter 2R). Adapter ligation reaction was set in 10 µl volume containing: 1 to 1.5 µg of Tester cDNA, 100µM of adapter (1 of 2R), 1µl T4 DNA ligase and 1µl of 10X T4 DNA ligase buffer. In another reaction, 50 ng tester cDNA were ligated with both adapters 1 and 2R (to serve as unsubtracted tester control). T4 DNA ligase was heat inactivated by keeping at 65^{0}C for 10 minutes after reaction. The adapter ligated cDNA was directly used for hybridization. Efficiency of ligation was checked by GAPDH primers and adapter specific primers. The digested non-ligated cDNA was (control to check efficiency of ligation) amplified with forward and reverse GAPDH primers whereas ligated cDNA was amplified with one of the GAPDH primer in combination with adapter specific primer. Relative intensity of amplified bands from ligated and unligated cDNA indicate the efficiency of ligation

3.5.4 Subtractive Hybridization

Step I: In the first step, hybridization was separately performed for two types of tester cDNA (ligated to different adapters, tester population 1 and 2) with digested Driver cDNA (in 2-3 fold molar excess). The hybridizations were performed in 4X hybridization buffer supplied in the subtraction kit. After mixing tester, driver cDNA and hybridization buffer, samples were denatured at 98^0C and hybridization was performed at 68^0C for 6 to 8 hours, but not more than 10 hours in any case.

Step II: Second round of hybridization ensures further enrichment of the differentially expressing genes and normalization of rare and abundant massages. For second round of hybridization, both pools (i.e. tester populations ligated to adapter 1 and 2 and hybridized separately with the driver) from first round of hybridization are mixed together and freshly denatured driver is added again. This hybridization was also carried out at 68^0C for 10 to 12 hours. After completion, the hybridized cDNA was diluted 50 times and subjected to PCR for amplification of differentially expressed genes.

3.5.5 PCR amplification and cloning of subtracted cDNA

This SSH approach used, does not involve physical separation of unhybridized cDNA at any step because it offers efficient enrichment of differentials and normalization of massages added with suppression PCR effect. Therefore, as per manufacturer's instructions two rounds of PCR amplification were performed. In the first round of PCR, diluted subtracted cDNA and unsubtracted tester control were amplified with primer 1 which has binding site on both the adapters (i.e. adapter 1 and 2R). In the second round of PCR, 1 µl of 10 fold diluted primary PCR products were used as template along with primers specific to both adapters, which had a binding site downstream of primer 1. Afterwards, the PCR products were checked on agarose gel which generally reveals a smear along with few major bands in both subtracted and unsubtracted (controls). The success of the subtraction indicated by the difference between band patterns, which is observed because subtracted cDNA pool is expected to contain enriched, differentially expressed or significantly upregulated gene(s). The amplified products after secondary PCR of subtracted cDNA were cloned in pMOS vector following manufacturer's instruction. Plasmids were isolated from the recombinant clones by alkali lysis method (section 3.3.4). Inserts were sequenced with M13 universal primers from the flanking MCS region of inserts.

3.5.6 Validation of differential expression for subtractive library ESTs (fragments)

There are several methods that can be used to analyze differential expression like RT-PCR, Real time RT-PCR, Northern and Reverse Northern. Northern blotting requires amplification of individual insert and their use as probe, while for RT-PCR insert needs to be sequenced for design of primers. In comparison, Reverse Northern, where several clones can be validated for their differential expression in a single hybridization, by and then, is most commonly used method in such studies.

3.5.6.1 Reverse Northern

Reverse Northern is an easy but preliminary method for validation of differentially expressed candidate genes/ESTs. Technique is particularly useful, when the number of fragments to be validated is large and the basic source material (such as RNA) is limited to perform the conventional Northern analysis. In this method, multiple fragments either plasmids or their inserts amplified by PCR, are transferred to a nylon membrane, which are then hybridized with the radio labeled tester and driver cDNA separately to check for unique/upregulated/downregulated expression of the blotted fragments on the membrane (Zhang *et al.*, 1996). In this study, multiple cloned fragments selected from subtractive cDNA libraries were amplified (with universal M13 primers from the flanking MCS region), and denatured by adding 10 µl of 2 N NaOH and boiling for 5 minutes, followed by addition of 10 µl of 3 M Sodium acetate (pH 5.0) and diluted to a total volume to 110 µl with dH_2O. Each of these were then loaded in duplicate on N^+ nylon membranes (50 µl each) with the help of Slot-blot apparatus (Amersham) under vacuum. Nylon membranes (blots) after transfer were washed twice with 100 µl of neutralizing buffer (1.5 M NaCl, 0.5 M Tris, 1 mM EDTA) at room temperature for 10 minutes each; blots were UV cross-linked, rinsed in 3 X SSC and stored at RT till hybridization. The probe was prepared by subjecting ~5 µg total RNA (from GAM tissue) for cDNA synthesis using one of the dNTPs radiolabeled (^{32}P dATP). In addition, control reactions were set by blotting GAPDH fragment on the same blots to serve as loading control. All reactions were performed with Oligo-dT primer and 3 µCi of ^{32}P dATP using cDNA synthesis kit (Invitrogen life technologies). cDNAs were purified with G50 columns to remove unincorporated ^{32}P label and then eluted in 50 µl of sterile water. Blotted fragments on N^+ nylon membranes were hybridized with total cDNA probe at 65°C in 0.5 M phosphate buffer containing 7 % SDS overnight. The membranes were washed at 65°C with 2 X SSC for 15 minutes; 1 X SSC for 15 minutes; and 0.2 X SSC for 10

minutes. Hybridization signals were then analyzed in a phosphor-imager system (Fujifilm, FLA 3000).

3.5.6.2 Other methods employed to validate differential expression

The fragments narrowed down by Reverse Northern analysis were further validated by techniques like semi-quantitative RT-PCR, Real time RT-PCR and Northern Blotting following the standard methods as described in section 3.2.2. In all these cases, the candidates were validated for their differential expression in GAM tissue of the embryos kept at MPT/FPT through TSP. One of the selected candidates that could be validated for their differential expression was subjected to RACE analysis (as described in section 3.3.6.2) to pull out the complete gene. After that, genome walking was also done to identify/characterize the complete locus of the selected gene. Finally, Northern analysis was done using sense and antisense probe as we suspected the selected gene is being coded from both of its strands.

3.5.6.3 Preparation of sense and anti-sense probes

Sense and antisense probes were prepared for the cloned gene (identified from the cDNA subtraction) using T7 RNA polymerase. A modified approach was followed where promoter site (core region) of T7 RNA polymerase was added to the 5′ end of the primer. We designed both primers (forward and reverse) with core promoter at 5′ ends of the primers. Accordingly, two PCR reactions were set wherein; the tagged forward primer was used in combination with non-anchored reverse primer in one reaction and in another tagged reverse primer was used in combination with non-anchored forward primer. The PCRs yielded amplicons, one has T7 promoter site tagged to top strand and has T7 site attached to bottom strand. In one reaction, T7 Tagged primer T7SC35F was used with SC35FN (probe to detect sense transcript) and in another reaction T7 tagged primer T7SC35R was used with SCINTF (probe to detect antisense transcript) (Table 3.5). Two 50 µl PCR reactions with above mentioned primers sets comprised; 1 µl genomic DNA (20 ng) as template, 5 µl 10X PCR buffer, 3 µl of T7SC35F/T7SC35R (5 pM), 3 µl of SC35FN/SCINTF (5 pM), 150 µM dNTPs, 1.5 mM $MgCl_2$, 2U Taq DNA polymerase. PCR products were purified using Qiagen PCR purification kit following manufacturer's instructions. The purified PCR products thus obtained, one having T7 promoter tagged to the beginning of top and another having T7 promoter at the beginning of bottom strand, were subjected to in-vitro transcription. The 50 µl reactions comprised: 5 µl ATP, 5 µl CTP, 5 µl GTP, 5 µl T7 10X reaction buffer, 1 µl UTP, 4 µl

P^{32}UTP, 1 µl of T7 RNA polymerase and 400 ng purified PCR product. The reaction mix was incubated at 37^{0}C for 2 to 3 hours. Probe was purified of unincorporated radiolabeled nucleotide with G50 column. Integrity of the probe was checked by running it on the agarose gel. Specific activity of probe was monitored by liquid scintillation counter (Packard Scintillation analyzer).

3.5.6.4 Northern Hybridization

The blots having 10 µg of total RNA from GAM of male and female embryos of developmental stages 21+22 were prepared in duplicates and hybridized separately with probes to detect sense and antisense transcripts as described in section 3.3.7. Hybridization buffer composition was 50% formamide, 6X SSC, 1% SDS, 5X Denhardt's solution, 5% dextran sulphate and 100µg/mL sheared salmon sperm DNA. Hybridizations were done at 42^{0}C for 12 to 14 hours. After hybridization, Northern blots were washed once with 2X SSC, 1X SSC, 0.5X SSC and 0.1 X SSC for 15 minutes each at 55^{0}C, all the washing buffers contained 0.1% SDS. After washing blots were checked in phosphorimager for hybridization signals. Northern hybridization with GAPDH probe was also done for the loading control.

3.5.6.5 Strand Specific RT PCR

To ascertain presence of both sense and antisense transcripts, strand specific RT PCR was employed. The First strand cDNA was prepared separately by reverse transcription using gene specific forward and reverse primer. It is based on the assumption that only the reverse primer (primers annealing to the 3′-end and amplifying the 5′-region like a oligo dT primer) would find a template in the RT, if mRNA does not have an antisense counterpart, and in case, where both the sense and antisense transcripts are present, both the gene specific forward and reverse primers would find template during reverse transcription. Accordingly, in case of presence of only one strand (i.e. sense) in RT-PCR reaction, only cDNA primed by reverse GSP should give amplification, while in case where both sense and antisense transcripts are present, cDNA primed by both forward and reverse GSPs are expected to give amplification. For the purpose, male and female GAM RNA samples from 21st stage were subjected to reverse transcription with SC35INTF and SC35INTR, separately, after reverse transcription (section 3.3.6.1), PCR with both RT products was done using same primers or a nested set of primers (SC35F/SC35R). Finally RT products generated separately by *cpWt1***a** specific primers Wt1F1 and Wt1R1 were also subjected to PCR by same primers, to bring in control for

the transcription from only one strand. As a prerequisite, RNA samples were given DNAase treatment and was followed by reverse transcription experiment using oligo dT primers, DNA contamination was ruled out by PCR with primers Wt1BF1 and Wt1bigR flanking an intron in *cpWt1* gene (that amplifies a 918 bp fragment with genomic DNA but a 198 bp fragment with RNA. Further the PCR was carried out for 40 cycles so that even traces of DNA contamination could be detected.

3.5.6.6 Fluorescence *in-situ* Hybridization (FISH)

FISH studies were done as described in section 3.3.11 to further analyze the expression pattern of SC35F5 like non-coding RNA. For the purpose, GAM tissue sections from the embryos kept at MPT and FPT representing stages 21 to 25, were hybridized with FITC labeled probe. The region amplified by SC35FN and SCINT1F primers that mostly comprised the region which was unique to transcript specific, was labeled by FITC and used as probe (section 3.3.11)

3.6 *In-silico* analysis

Following software and web based programs were used in the study:

3.6.1 Primer designing:

Gene Tool Lite version 1.0 (http://www.biotools.com/products/genetool.html), currently withdrawn from the server.

Primer 3 (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi)

3.6.2 Sequence alignment:

Codon code aligner (Codon code Corporation, USA)

Auto assembler (Perkin-Elmer software for mac), Clustal-X 1.8 (Thompson et al. 1997), Bioedit (http://www.mbio.ncsu.edu/BioEdit/bioedit.html),

ClustalW (http://www.ebi.ac.uk/clustalw/)

3.6.3 Sequence homology search:

NCBI Blast search (http://www.ncbi.nlm.nih.gov/BLAST)

VecScreen (http://www.ncbi.nlm.nih.gov/VecScreen/VecScreen.html)

3.6.4 Gene finder, splicing:

GENSCAN (http://genes.mit.edu/GENSCAN.html)

NetGene2 Server (http://www.cbs.dtu.dk/services/NetGene2)

3.6.5 Protein translation and secondary structure prediction:

The Predict Protein server

(http://www.embl-heidelberg.de/predictprotein/submit_adv.html#top)

Prediction of conserved Domains and protein-protein interaction prediction

Expasy proteomics server, ELM search

3.6.6 Transcription factor binding prediction

TESS (Transcription element search system) (www.cbil.upenn.edu/tess)MatInspector (www.genomatix.de/products/MatInspector)

TFSEARCH (www.cbrc.jp/research/db/TFSEARCH.html)

AliBaba2 (www.gene-regulation.com/pub/programs/alibaba2/index.html)

3.6.7 *In-Silico* analysis for subtractive library sequences

After sequencing a large number of clones, library was analyzed for the Unigenes before proceeding to BLAST. Unigenes were analyzed from the library sequences by a perl program Clobb.pl. Databases were also analyzed against tissue specific databases which were downloaded from SANBI (South African National Bioinformatics Institute). The Subtractive library was also screened for noncoding RNA by using a pipe line which involved BLAST like BLASTn, BLASTx, tBLASTx etc. ORF prediction was done using fungal genome website. Several thresholds were kept for each kind of test, like threshold for e-value, ratio-test etc, to predict mRNA like non coding RNA candidates. All computational analysis was done on CCMB servers using standalone version on BLAST algorithm. RNA folding analysis was done by using MFOLD program v 2.3 developed by Martin Zucker, at the RNA Page server at IMB Jena.

Chapter 4

Results

The present study comprised an effort to have some newer insights of the underlying molecular basis of temperature dependent sex-determination using Indian mugger as a model system. In specific, the study was undertaken with two major aims *i.e.*,1) to isolate the crocodilian homologues of a few conserved sex determining genes, and ascertain their spatio-temproral expression in detail to understand their possible role in TSD, if any; and 2) simultaneously, to search for novel, candidate gene(s) having a role in TSD by carrying out global screens for differentially expressing genes in the developing gonadal tissues of putative male and female embryos during the temperature sensitive period. The results of experiments carried out to this end, are described in the following pages.

All primers referred in these sections are detailed in **Tables 3.2 - 3.6**.

4.1 Isolation and expression analysis of conserved sex-determining genes

In this study, efforts were made to isolate and characterize the crocodilian homologues of three conserved sex determining genes namely, *Wt1, Dmrt1* and *Aromatase* that have been implicated to have an important role in the GSD and also shown to be important in TSD, in some of the preliminary studies (see section2.6). *Wt1* has been shown to be a very important gene for the urogenital development in vertebrates, and most interestingly it has been shown to perform different functions through multiple isoforms. *Dmrt1* is one of the evolutionarily most conserved sex determining genes, and is demonstrated to have an important but different roles in male sex development in GSD vertebrates. On the other hand, *Aromatase* has already been indicated to have a very important role in TSD; it has been shown to express early in all TSD species and is associated with the female gonad development.

To isolate crocodilian homologues of the above three genes from Indian mugger, a number of PCR based approaches were used in this study; these invariably started with RT-PCR using heterologus primers designed from the conserved region(s) of the target gene sequences especially from non-mammalian vertebrates. The partial sequences obtained in the initial RT-PCR reactions were then used to isolate the complete cDNA/gene using RACE and genomic walking strategies. Subsequently, detailed spatio-temporal analysis, especially in GAM tissues of embryos from MPT and FPT through TSP was carried out (in each case), to have insights about the tissue-/stage-/sex-

specific expression (if any). Analysis revealed expression specificity to different levels for each of the three crocodilian homologues; and even more interestingly, revealed multiple transcripts (novel isoforms) for each gene. The latter observation necessitated efforts to isolate the coding region (in case of *cpDmrt1* and *cpWt1*) to have an idea about the possible origin of the multiple isoforms. Studies to this end provided newer insights about the organization of the coding regions of the two genes and also demonstrated that the multiple isoforms represented only the splice variants and are not the products from different genes.

Furthermore, for two genes (*cpWt1* and *cpAromatase*) localization in the developing GAM tissues and mesonephros was also studied, which could give an idea about the cell type specificity or association with development of a particular cell type(s) in the developing gonads.

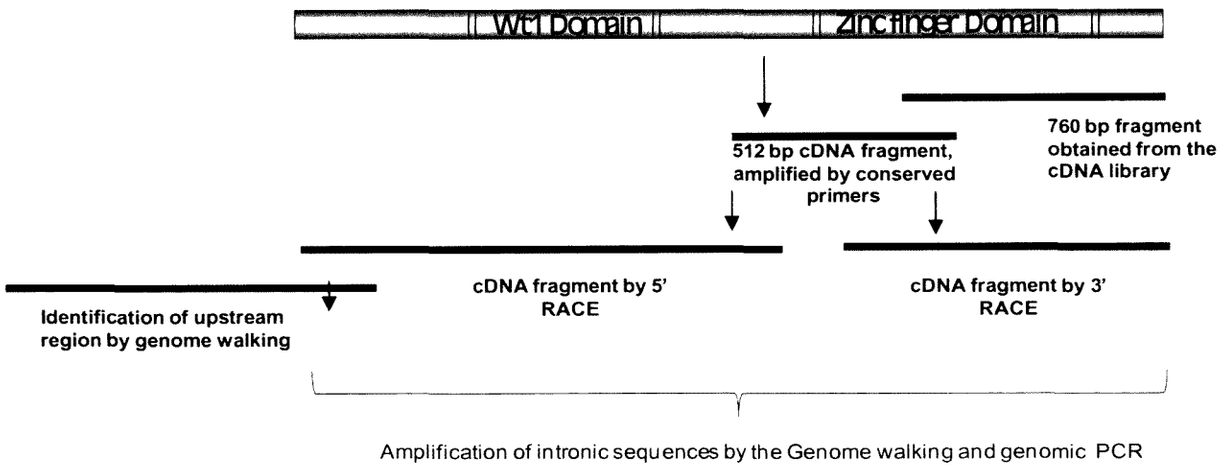
4.1.1 Isolation and characterization of *cpWt1* (*Wt1* homologue of *C. palustris)*

4.1.1.1 Isolation of partial *cpWt1*

Wt1 has been studied extensively in mammals and many other non-mammalian GSD vertebrates, considerable information is available about its transcriptional diversity and function during embryonic development. *Wt1* sequences from amphibian, bird and reptile were retrieved from NCBI and compared to identify the most conserved domain, which corresponded to 512-1041 bp of alligator *Wt1* sequence (X85730), and was used to design primers Wt1F1 and Wt1R1. The primers were used for RT-PCR to identify/isolate *cpWt1* sequence/transcript, using the pooled male and female GAM tissue cDNA as template. RT-PCR yielded a 530 bp amplicon, which was sequenced and confirmed to be *Wt1* sequence. This partial *cpWt1* transcript showed complete sequence homology to *Wt1* sequence of *Alligator* and a very high similarity (95%) to that of *Trachemys* and *Gallus*.

4.1.1.2 Isolation of complete *cpWt1*

The strategy used to pull out the complete *cpWt1* gene is shown in figure 4.1A. The initial 530 bp of the *cpWt1* obtained above was used for designing primer to carry out 3' and 5' RACE *i.e.* two primers Wt15' and nested Wt15'N for 5' RACE, and primers AWt13' and nested AWT13'N for 3' RACE. All RACE reactions were done as described in



B

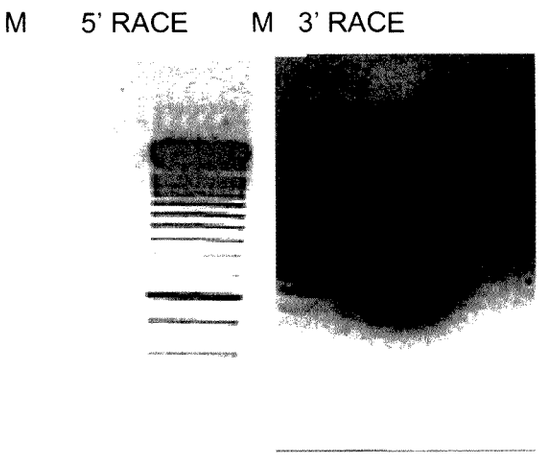


Figure 4.1: **Identification and Isolation of *cpWt1* (isoforms and locus). A)** Schematic representation of the approach used for isolation of *cpWT1*. In the 1st step, partial *cpWt1* was identified by RT-PCR with conserved primers, which was followed by 5'-/3' RACE to obtain complete transcribed cDNA, and finally genome walking was used to ascertain/isolate the possible intron(s), as well as, upstream region; and **B)** RACE reactions to isolate complete *cpWt1* transcript expressed in GAM tissue; the left panel represents 5' RACE that resulted in only a single band of approximately 800 bp, whereas right panel represents 3'RACE that resulted into two major bands of approximately 800 and 1200 bp. M represents DNA sizing ladder.

section 3.4.1.3. 5' RACE resulted into a single fragment but interestingly 3' RACE revealed two amplicons of approximately 800 and 1200 bp (figure 4.1B). Further, similar amplification profiles were observed using male and female RACE ready GAM cDNA. RACE products were purified and cloned in pCR 2.1 vector and in each case; at least 100 recombinant plasmids were sequenced. Sequence analysis revealed two isoforms of *cpWt1* of sizes 1560 bp and 1088 bp, which were named as *cpWt1***a** (bigger) and *cpWt1***b** (smaller) respectively. Both isoforms shared only 625 bp towards their 5' end, and the remaining 455 bp region of isoform *cpWt1***b** and 935 bp of *cpWt1***a,** at the 3' ends, were different (figure 4.2 A). BLAST homology based analysis indicated that the isoform *cpWt1***b** was novel and was not described in any other vertebrate.

4.1.1.3 Homology analysis of *cpWt1* isoforms

BLAST homology analysis (of both the isoforms) indicated a unique 5' end in both the *cpWt1* isoforms. Interestingly, all the evolutionarily relatives *i.e.* reptiles and birds have significant heterogeneity in their 5' region that corresponds to 1-308 bp of *Alligator* Wt1; 1-897 bp in *Trachemys Wt1* and 1- 138 bp in *Gallus Wt1* transcript. At nucleotide level no other vertebrate *Wt1* sequence showed homology with the 1-519 bp region of both *cpWt1* isoforms (figure 4.2B). Only exception was the Human *Wt1* transcript where 100 bp corresponding to exon 5 were showing similarity to the 42 to 141 bp region of *cpWT1.* Also the similarity search at genomic level revealed that 1-223 of the initial 519 bp of *cpWt1* matched with the intronic region between exons 2 and 3 of *Gallus Wt1* coding region and similarly to a part of human *Wt1* intronic sequence. However, a region from 225 to 497 bp does not match to any *Wt1* coding sequence in the vertebrate.

However after 519 bp, the rest of the *cpWt1***a** isoform was completely similar to the *Wt1* transcripts of other vertebrate. It was noteworthy that *cpWt1***a** had longer 3' end compared to those of *Alligator* and *Gallus Wt1* sequences *i.e. Gallus Wt1* terminates by 1094 bp of *cpWt1***a**, and Alligator *Wt1* terminates by 1480 bp of *cpWt1a* sequence. In comparison isoform *cpWt1***b** that also had a unique 519 bp on 5' end (similarity to other vertebrate *Wt1* transcripts has been described above), showed similarity to *Wt1* sequences of other vertebrate only for 100 bp *i.e.* 519 to 625 bp only. Its region after 625 bp (till 3' end) did not have any match in database and thus was indicated to be unique.

Isoforms *cpWt1***a** and *cpWt1***b** were predicted to code for proteins of 294 and 160 amino acids (figure 4.3). Isoform *cpWt1***a** has 195 bp 5' UTR, 885 bp coding region and 480 bp 3' UTR. Isoform *cpWt1***b** has 195 bp 5' UTR, 483 bp coding region and 410 bp 3' UTR.

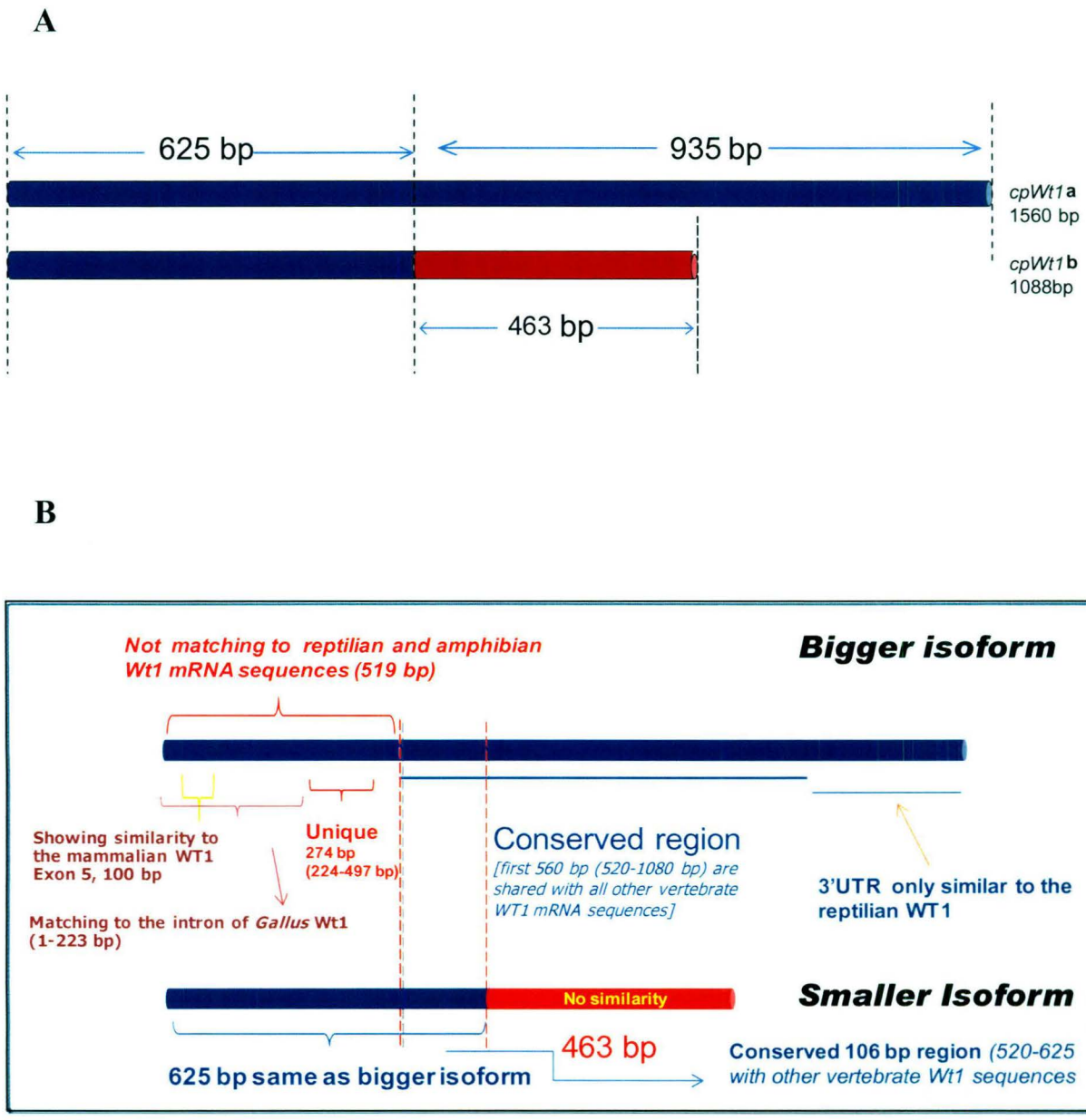


Figure 4.2: **Sequence comparison of the two *cpWt1* isoforms obtained in RACE from GAM tissue. A)** Schematic representation of alignment of two *cpWt1* isoforms; common region between both isoforms is represented in blue, while the non-matching region (which is not evolutionarily conserved in *cpWt1***b**) is represented in red; and **B)** *In-silico* sequence comparison of *cpWt1* isoforms. Note that only the middle region is relatively conserved across vertebrates while 5'-end is strikingly resemble to the *Wt1* intronic region of *Gallus* and to a part of *Wt1* exon in mammals.

The predicted proteins of the both isoforms shared 144 amino acids from their N-terminal end, and translated *cpWt1***b** had 17 unique amino acids at its C-terminal end (figure 4.3). Furthermore, 107 amino acids from their N-terminal do not match with any Wt1 protein sequence in the database. However, putative protein from *cpWt1***a** matched other WT1 proteins from 108 amino acids onwards till end; similarly protein from *cpWt1***b** matched other WT1 proteins from 108 to 143 amino acids and then carried a 17 amino acids long unique C-terminal end. The 108 to 158 amino acids of predicted cpWt1**a** protein resemble to conserved WT1 domain, while its 175 to 294 amino acids resemble to the conserved zinc finger motifs. It was also noteworthy that putative protein of *cpWt1***b** isoform lacked all the zinc finger motifs; whereas WT1 proteins of all other vertebrates, including that of *cpWT1***a** have four zinc finger motifs. A minimum evolution tree constructed using maximum Composite Likelihood model of evolution revealed that: *cpWt1***a** isoform was closest to the *Alligator Wt1* sequence and shared an overall high similarity to the reptilian and *Gallus Wt1* sequences, which together defines the archosuarial lineage of evolution (figure 4.4).

4.1.1.4 Genomic organization of *cpWt1* and isolation of *cpWt1* upstream region

Presence of a unique 5' region of *cpWt1* isoforms and a novel *cpWt1***b** isoform necessitated studies to understand organization of *cpWt1* locus *vis-a-vis* origin of both *cpWt1* isoforms. Accordingly, efforts were made to isolate the *cpWt1* locus (coding region) by genome walking approach.

4.1.1.4.1 Genome walking

Genome walking was started from 5' end; just before the 625 bp of the sharing region between the two *cpWt1* isoforms. For the purpose primers AWT13' and AWT3'N (used for 3' RACE) were used in primary and secondary genome walking PCR reactions to obtain the genomic region towards 3' coding region of the *cpWt1*; similar experiments were repeated with many primer combinations till intron-exon structure was revealed; for primer combination details refer to table 3.2. Genome walking efforts thus revealed a 3720 bp long cpWt1 genomic region that could explain the origin of the *cpWT1* isoforms by alternative splicing. Overall *cpWt1* locus (that could be isolated in this study) revealed to comprise four exons and three introns (figure 4.5A), though presence of additional regions towards the 3' end cannot be discounted. The length of four exons (exons-1 to exon-4) was: 625, 153, 198 and 584 bp respectively, and the three introns were of 788, 780 and 592 bp (partial intron-3). Analysis of the genomic sequence revealed that the

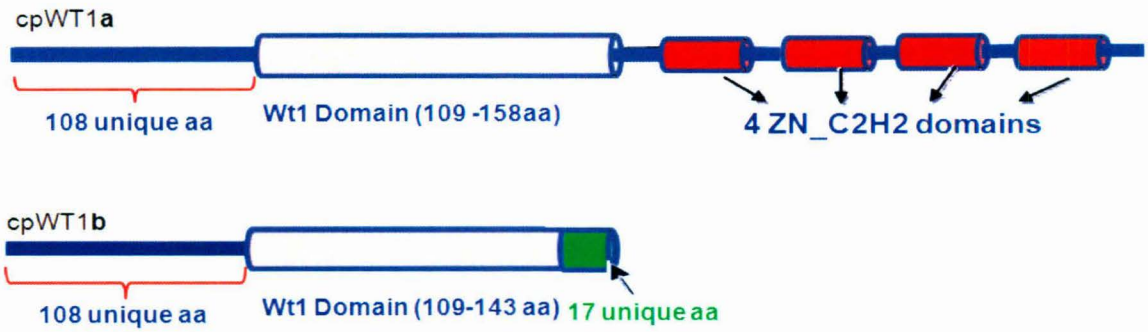


Figure 4.3: **Schematic representation of the domains present in the predicted cpWT1 proteins.** Note that both proteins have: a) 108 unique aa at N terminal region; b) a largely conserved Wt1 domain with 14 aa truncation and additional 17 unique aa at the C-terminal end in cpWt1**b;** and c) a conserved zinc finger domain only in the cpWT1**a** protein.

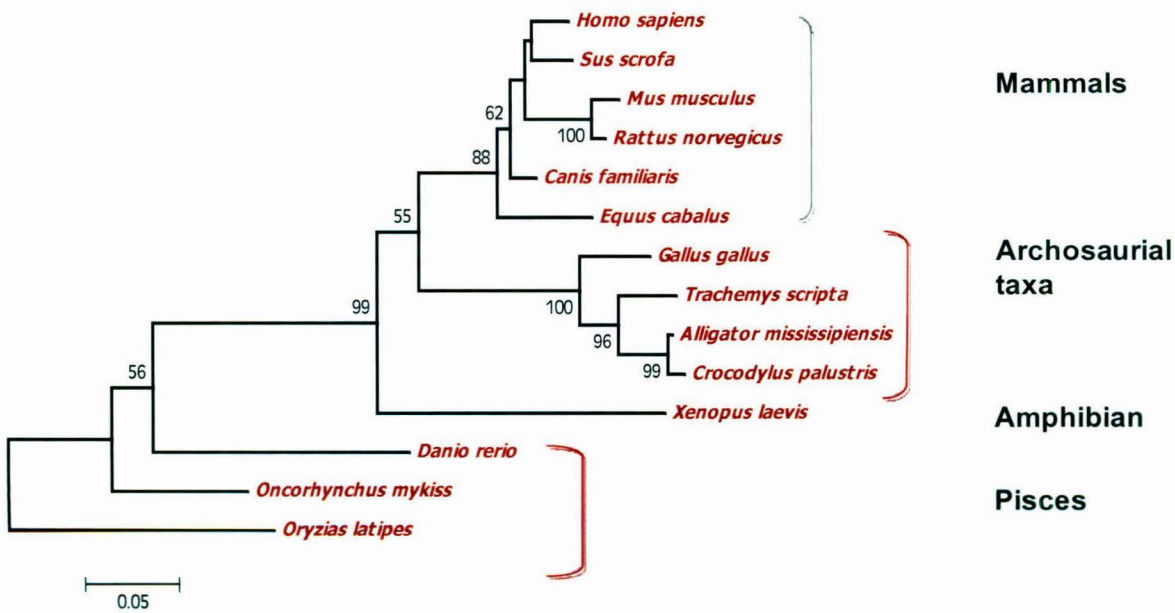


Figure 4.4: **Neighbour-joining tree showing the generic/phylogenetic relationship of *cpWT1*a with *Wt1* sequences of other vertebrates.** The tree is generated using the Kimura 2-parameter based genetic distance estimates and MEGA4.1 software. A total of 473 complete nucleotide sites and empirical nucleotide frequencies/ gamma-correction were considered in deriving genetic estimates. Analysis revealed *cpWt1* to be closest to the *Wt1* of *Alligator* and overall as part of the archosaurial lineage. Values at nodes represent bootstrap values.

*cpWt1***a** isoform was coded by four exons, while the smaller *cpWt1***b** isoform represented by a continuous stretch of DNA comprising 625 bp of exon-I and 463 bp of intron-I (figure 4.5A).

In addition to *cpWt1* coding locus, genome walking was carried out to identify/isolate the *cpWt1* upstream region using primers GWTR and WT1GW5', designed from 5' ends of the *cpWt1* transcripts. This experiment led to isolation of a 1121 bp long *cpWt1* upstream region (figure 4.5A).

4.1.1.4.2 *In-silico* analysis of *cpWt1* upstream region

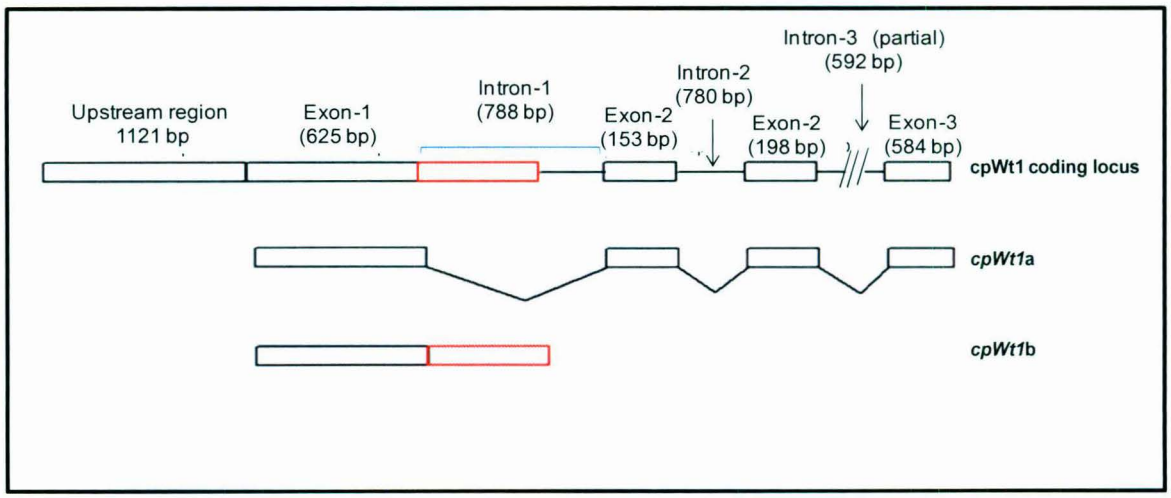
BLAST based homology analysis showed that the putative promoter region of *cpWt1* has no resemblance with *Wt1* promoters of other vertebrates; here it is important to mention that designated promoter region of *Wt1* in *Gallus* does not match to those of mammals and that similar sequence for other reptilian species are yet to be described. Upstream *cpWt1* region was also analyzed for binding site(s) of regulatory factors using several algorithms like, MatInspector, TESS, TFSEARCH etc. Invariably programs predicted that *cpWt1* upstream region has all the hallmarks of promoter region, like consensus TATA binding site (figure 4.5B). In addition, the upstream region had many HSF1, HSFTF binding sites, which suggest its possible importance in TSD. The region was also predicted to have binding sites for a number of regulatory factors like, AR/PR/ER (androgen receptor/progesterone receptor/estrogen receptor) (figure 4.5B).

4.1.1.5 Expression analysis of *cpWT1* isoforms

4.1.1.5.1 Tissue specific expression analysis of *cpWt1* isoforms

Expression of *cpWt1* isoforms was checked in five embryonic tissues *viz.* heart, liver, brain, kidney and GAM of *C. palustris* by RT-PCR using cDNA, synthesized for each tissue from pooled male and female RNA. Analysis using the isoform specific primers (Wt1F1/Wt1R1 for *cpWt1***a** and Wt1SMF/Wt1SMR for *cpWT1***b** isoform) revealed that both the *cpWt1* isoforms expressed in all tested embryonic tissues without any apparent tissue specificity (figure 4.6A).

A



B

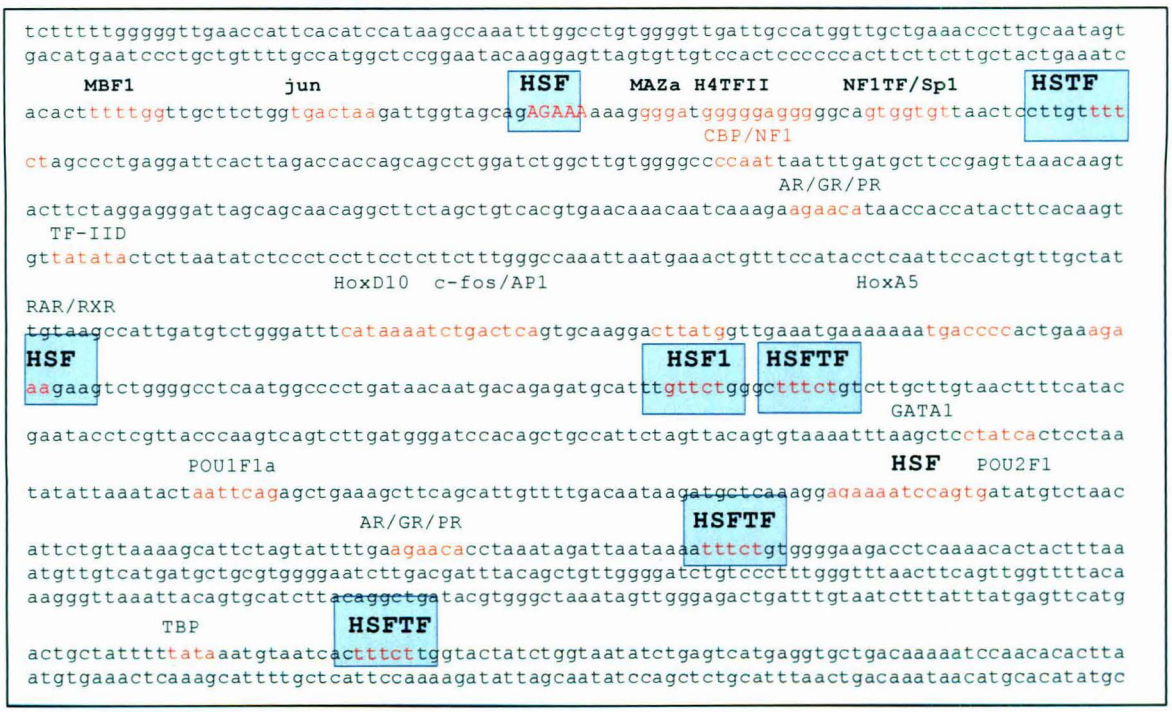


Figure 4.5: ***In-silico* characterization of the *cpWt1* genomic locus *via-a-vis* isoforms and its upstream region. A)** Schematic representation of the *cpWt1* genomic locus and the suggested alternative splicing leading to two isoforms. Boxes represent exons, whereas straight/'V' lines represent intronic regions. The region shown in red is intronic region transcribed in *cpWt1***b** isoform as novel 3' region; and **B)** Part of 1121 bp upstream region of *cpWt1* showing putative binding sites for a number of transcription factors and DNA binding proteins. Analysis revealed a number of binding sites for heat shock transcription factors (HSF, HSF1, HSTF and HSFTF).

4.1.1.5.2 Expression analysis in GAM tissue

Expression of *cpWt1***a** and *cpWt1***b** isoforms in GAM tissue of the embryos kept at MPT and FPT through TSP was followed using semi-quantitative RT-PCR with specific primers (see section 3.4.1.4). Analysis revealed a near comparable expression of *cpWt1***a** in GAM tissues irrespective of the putative sex/incubation temperature of the embryos and developmental stages (figure 4.6B). In contrast, expression of shorter isoform, *cpWt1***b** was found to be significantly higher in the GAM tissues of the embryos kept at MPT than those at FPT throughout the TSP (figure 4.6C). Expression levels of *cpWt1* isoforms was also monitored by the real-time RT-PCR using primers WT1BIGF2 and WT1BIGR4 for *cpWt1***a** and WT1SMF and WSMR for *cpWt1***b**. Real-time RT-PCR also revealed no significant difference in the expression levels of the *cpWt1***a** in the GAM tissues from MPT and FPT, or between 21st to 25th developmental stages (figure 4.7A). However, isoform *cpWt1***b** expression levels were revealed to be 40 folds higher at the MPT compared to that at FPT (figure 4.7B) all through stages 21 to 25 at MPT.

Subsequently, expression analysis of *cpWt1***b** isoform was also monitored in other embryonic tissues at MPT and FPT through TSP using real-time RT-PCR (*viz*, Heart, Brain, kidney and liver). The analysis revealed no significant difference in expression levels between MPT and FPT in these tissues (figure 4.8).

Thus the results conclusively show that one of the isoform of *cpWt1* (*cpWt1***b**) expresses at significantly higher levels only in the male gonads since the very beginning to the end of TSP.

4.1.1.6 Localization of *cpWT1* isoforms in developing GAM

Localization of *cpWt1* isoforms was followed by Fluorescent *in-situ* hybridization (FISH) in the GAM tissues (developing gonads and mesonephros) individually using cDNA probes specific to the unique regions of *cpWt1***a** and *cpWt1***b**. Details of the probe preparation and *in-situ* hybridization are outlined in section 3.4.1.5.

4.1.1.6.1 Localization of the bigger isoform (*cpWt1a*) in developing gonads

Putative male and female GAM tissue sections of all five stages representing TSP (i.e. stage 21-25), hybridized with probes specific to *cpWt1***a**, were examined separately

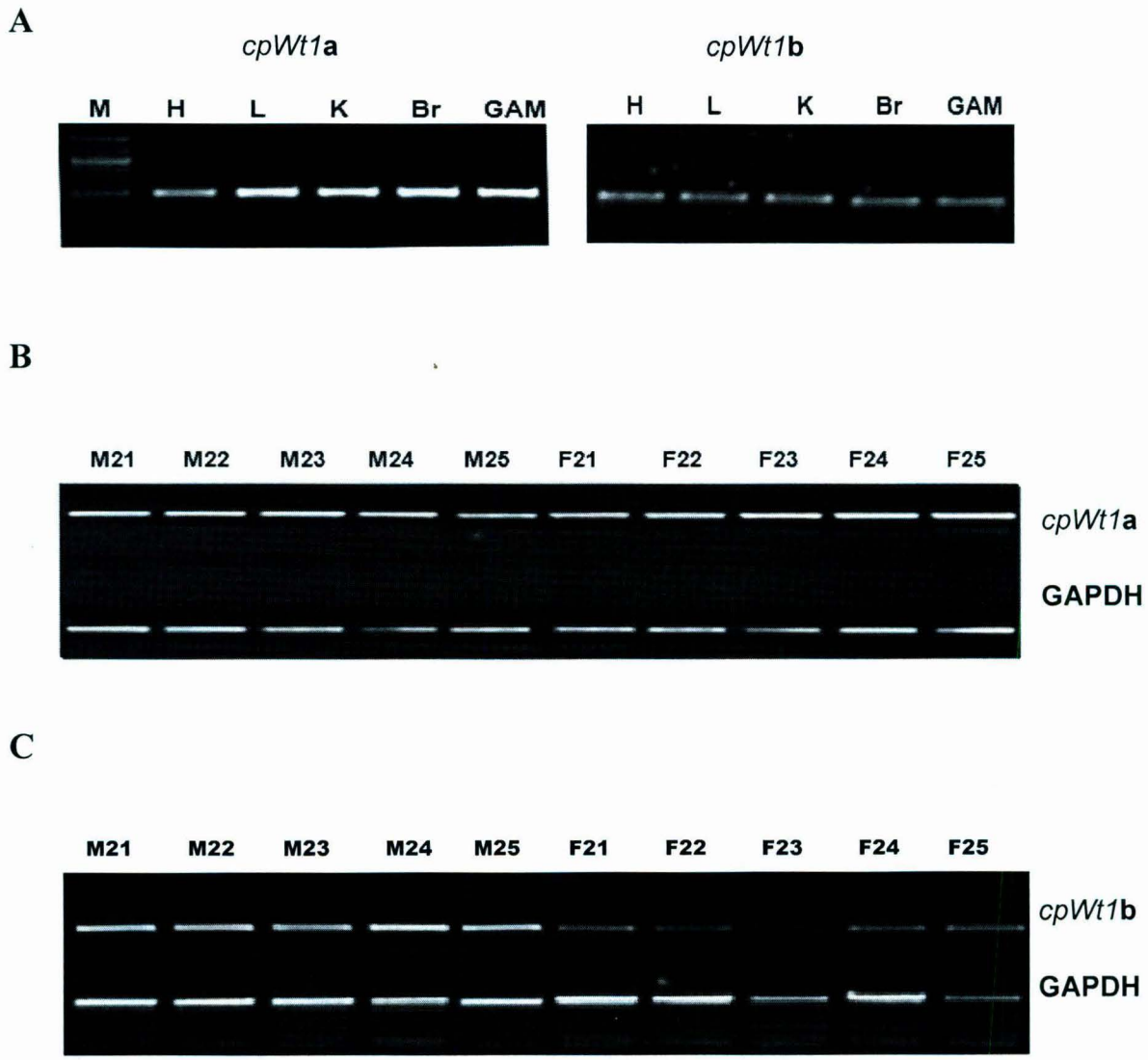


Figure 4.6: **Semi-quantitative RT-PCR based expression analysis of *cpWt1* isoforms in GAM and other embryonic tissues through TSP. A)** Expression of both isoforms in all five embryonic tissues; **B)** no quantitative variation *cpWt1***a** expression in GAM of male or female; and **C)** Many fold higher expression of isoform *cpWt1***b** in male GAM than that of female throughout TSP. Lanes: M: 100 bp marker ladder, H: Heart, L: Liver, K: Kidney, Br: Brain, GAM: Gonad-adrenal-mesonephros complex. M21-25 and F21-25 represent developmental stages 21 to 25 at MPT and FPT, respectively. *GAPDH* was used as internal control for expression analysis.

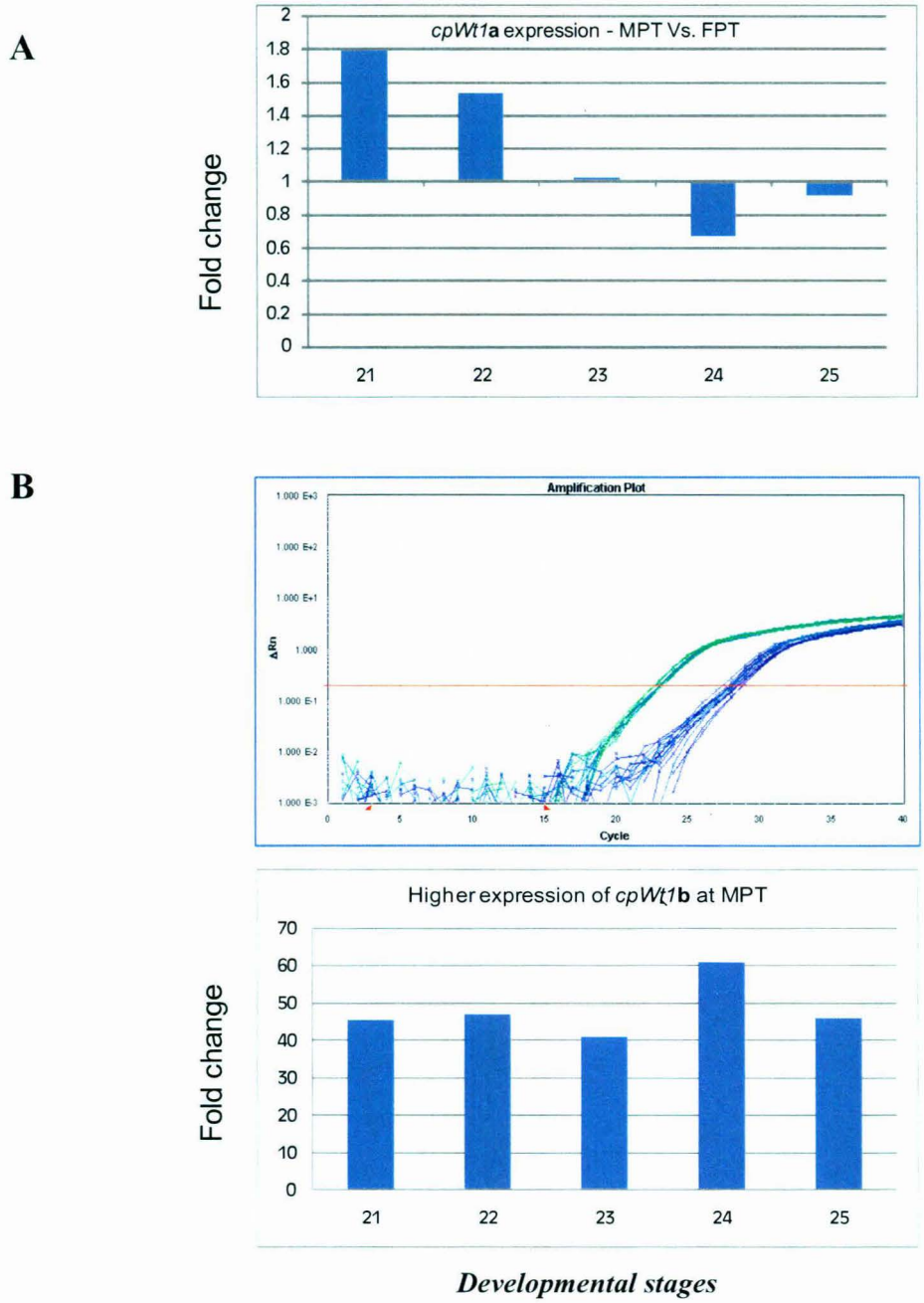


Figure 4.7: **Real-time RT-PCR analysis of *cpWt1* isoforms expression in GAM tissues of male and female embryos through TSP. A)** No significant difference is seen in expression of *cpWt1***a** in the GAM tissue of male and female embryos during TSP; and **B)** In comparison atleast 40-60 fold higher expression is seen for isoform *cpWt1***b** in the GAM of the embryos kept at MPT than those at FPT. Upper panel shows the amplification plot of *cpWt1***b** wherein male and female samples are represented in green and blue respectively; and the lower plot represents the difference in expression levels after normalization of data.

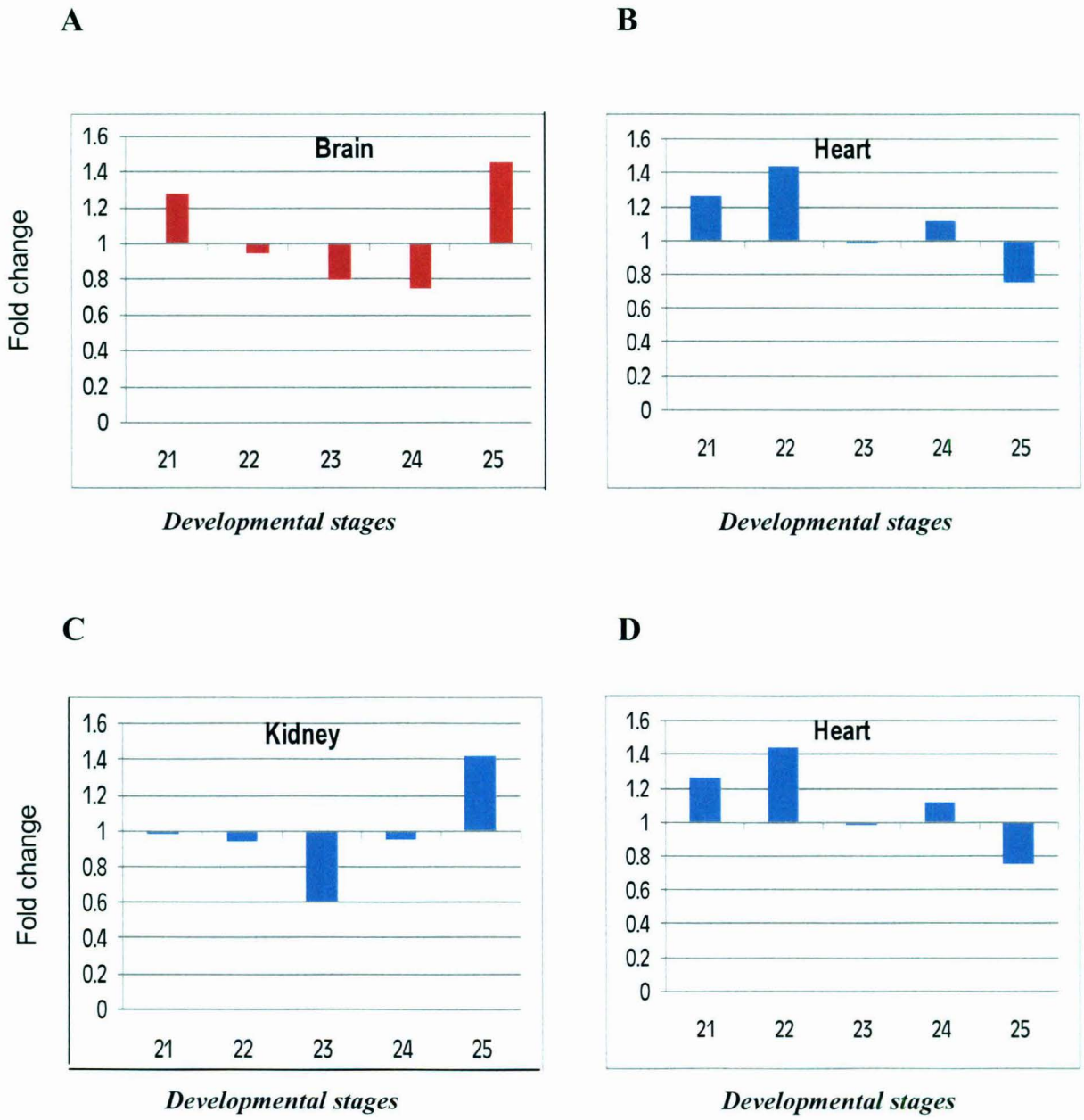


Figure 4.8: **Real-time RT-PCR analysis of *cpWt1*a isoform expression in brain, heart, kidney and liver tissues of embryos kept at MPT and FPT through TSP.** Comparative expression in: **A)** brain tissues; **B)** heart tissues; **C)** kidney tissues; and **D)** liver tissues. Note that expression in all the tissues is comparable with no significant sex specific bias.

across areas of developing gonads and mesonephros. The *cpWt1***a** expression (staining) in developing gonads did not show any cell type specific localization. Moreover, there was no difference in the signal intensity between the male and female developing gonads indicating equal level of expression in male and female developing gonads all through TSP. During 21st to 23rd stage, the male and female gonads have no difference in the cellular organization, however after 24th stage, development of testicular cords was seen which were properly visible by 25th stage in the male GAM. In contrast, in female gonads no obvious cellular organization was apparent except slight thickening of cortex towards the 25th stage. The localization pattern of *cpWt1***a** isoforms was thus in continuity of RT-PCR results that have revealed equal expression of *cpWt1***a** at MPT and FPT across TSP without any cell type specificity in developing gonads (figure 4.9).

4.1.1.6.2 Localization of the smaller isoform (*cpWt1*b) in developing gonads

Localization of *cpWt1***b** in the developing gonads revealed a very interesting localization/expression pattern in male gonads developing at MPT (figure 4.10). During initial stages (stage 21-22), *cpWT1***b** expressing cells were scattered and did not have any localization pattern in the GAM. However, 23rd stage onwards, *cpWT1***b** expression became prominent in the cells that appeared to arrange themselves in tubule like structures; which became more organized in cord like structures that defined the developing testicular cords in later stages. The development of testicular cords was evident in the sections studied by optical microscopy without any staining or monochrome images (figure 4.11). By the 24th stage, the *cpWt1***b** probe staining was seen only in the cells that formed testicular cords. Further, PI staining of the hybridized section reveals migration of cells inside the testicular cords of male GAM from 24th stage onwards. Interestingly, *cpWt1***b** probe did not stain the cells that migrate to the lumen of testicular cords or the interstitium cells between the adjacent testicular cords (figure 4.12). The testicular cords appeared most prominently by the 25th stage and were seen filled with many cells visible but not stained with *cpWt1***b** probe as seen first in stage 24 (figure 4.12). The results thus demonstrated a significant correspondence between the expression/localization of shorter *cpWt1* isoform and the cellular events and order of testicular cord development in vertebrates involving Sertoli cells.

In contrast, to the male GAM, expression of the *cpWt1***b** isoform in the GAM of female embryos (from FPT) was found to be invariably low and scattered. Staining did not give any indication, of cell type specificity or any cellular events throughout the developmental

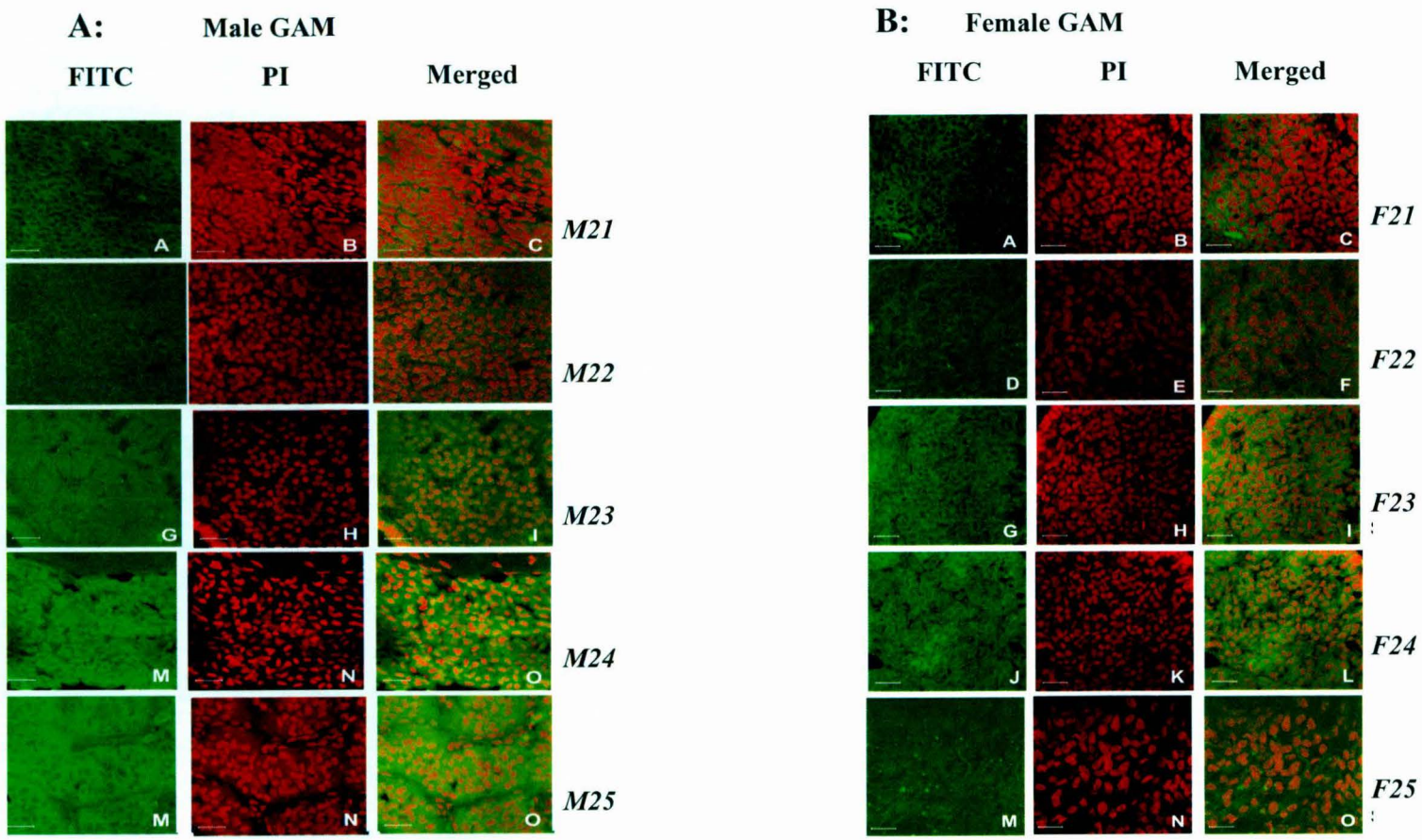


Figure 4.9: **Localization of isoform *cpWt1*a isoform in the developing GAM tissues of male and female embryos through TSP (21st to 25th stage).** *In-situ* hybridization was done using a probe specific to unique region of *cpWt1***a.** Probe was labelled with FITC (green) and cells were counter-stained with PI (red). Note similar localization pattern in male and female embryos with no sex or cell type specificity and similar expression levels through TSP. M21–M25 & F21-F25 represent 21st to 25th developmental stages of male/female embryos, respectively.

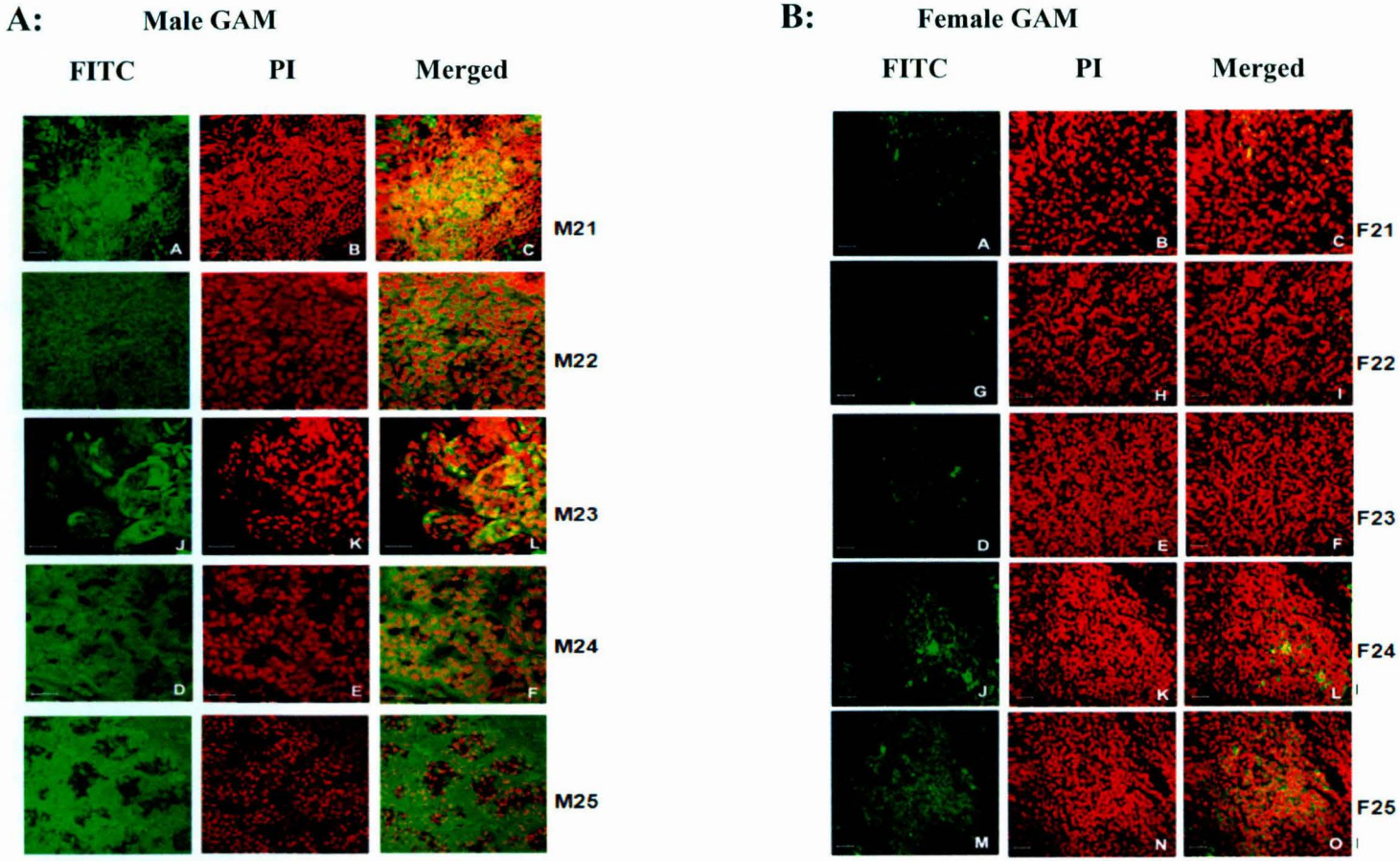


Figure 4.10 **Localization of *cpWt1*b isoform in the developing GAM tissues of male and female embryos through TSP (21st to 25th stage).** *In-situ* hybridization was done using a probe specific to unique region of ***cpWt1*b.** Probe was labelled with FITC (green) and cells were counter-stained with PI (red). Note sex-/tissue- specific localization and higher expression in male gonads, with specific expression in testicular cords through TSP. M21–M25 & F21-F25 represent 21st to 25th developmental stages of male/female embryos, respectively.

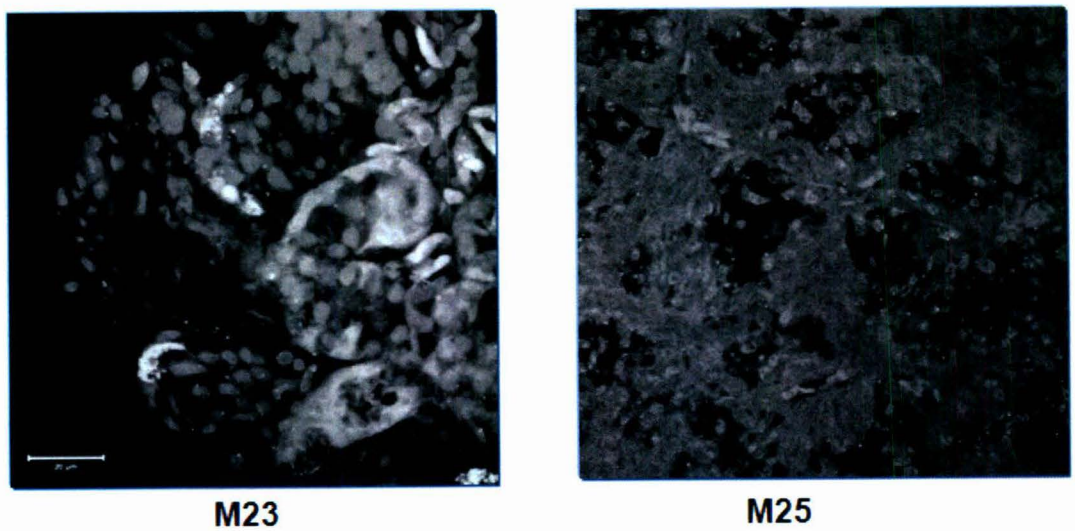


Figure 4.11: **Monochrome picture and unstained section of the developing gonads of embryos kept at MPT from developmental stages 23rd and 25th, respectively.** Note that during 23rd stage cord formation starts and by 25th they become conspicuous.

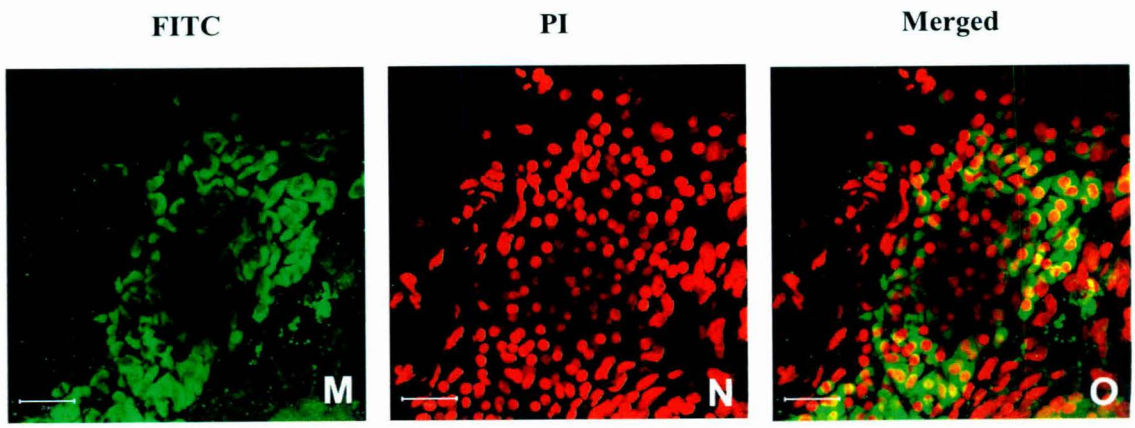


Figure 4.12: **Localization of isoform *cpWt1*b specifically in testicular cord cells of developing gonads from 24th stage embryo kept at MPT.** ***cpWt1*b** probe was labelled with FITC (green) and cells were counter stained with PI (red). Note ***cpWt1*b** expressing cells (in green) apparently form a cord like structure and the cells inside the cords (stained in red by PI) do not express ***cpWt1*b**.

stage 21 to 25 that defined TSP including thickening of cortex (figure 4.10). Which otherwise could be observed by PI staining by 25th stage.

4.1.1.6.3 Localization of smaller isoform (*cpWt1***b**) in developing mesonephros

Localization of smaller isoform was also studied in developing mesonephros region of GAM. Unlike the developing gonads, *cpWt1***b** expression in the mesonephros was comparable between male and female embryos. The labeled *cpWt1***b** probe stained the developing nephrons of male and female GAM tissues, equally without any difference in the intensity or any cell type specificity through all five developmental stages spanning TSP (figure 4.13).

4.1.1.7 Functional Studies of *cpWt1*

Expression profiling and localization studies (described above) amply suggested that the novel *cpWt1* isoform *i.e. cpWt1***b** (identified in this study) was male specific, and that it may have an important role in determination/differentiation of male gonads. To explore this aspect further, efforts were made to study the subcellular localization of both the *cpWt1* isoforms, and monitor their role (if any) in the regulation of transcription of other sex determining genes. For the purpose functional, studies were undertaken wherein labeled *cpWt1* isoform were transfected in the primary cell lines of the crocodile (previously established at CCMB) from the embryonic tissues.

4.1.1.7.1 Transcriptional profiling of the sex determining genes in the primary cell lines derived from the GAM tissues

Primary cell lines derived from the male, as well as, female GAM tissues (of crocodile embryos) were initially checked for differences (if any) in the expression levels of sex determining genes, akin to differences seen in tissues. For the purpose genes namely, *cpAmh, cpSf1, cpSox9, cpDmrt1, cpWt1b, cpUNR, CC1.3* homologue and *SLF35F5* like ncRNA were selected that have been indicated to have higher expression in the male GAM tissues in the present and earlier studies in the lab (Agrawal 2006); referred hereafter as male favoring genes. Expression levels of these male favoring genes, monitored using real-time RT-PCR, were found to be almost similar in different GAM derived cell lines, independent of the putative sex (male/female) and/or development stage (committed/non-committed) of the source embryos (figure 4.14A). Similarly, no differences in expression of these genes were observed when the cell lines were incubated at different temperatures i.e. one at 30^{0}C and other at 26^{0}C (figure 4.14B).

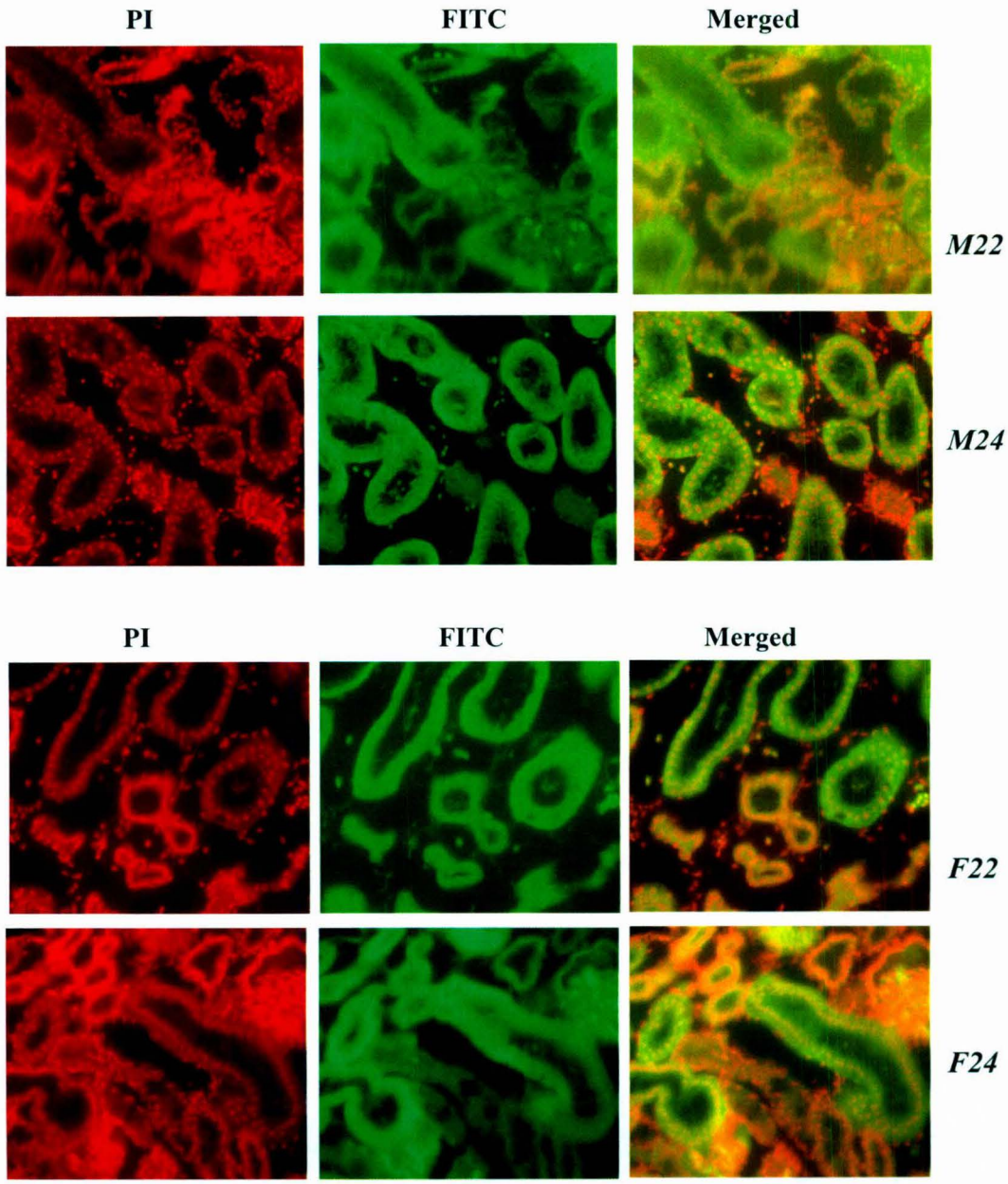
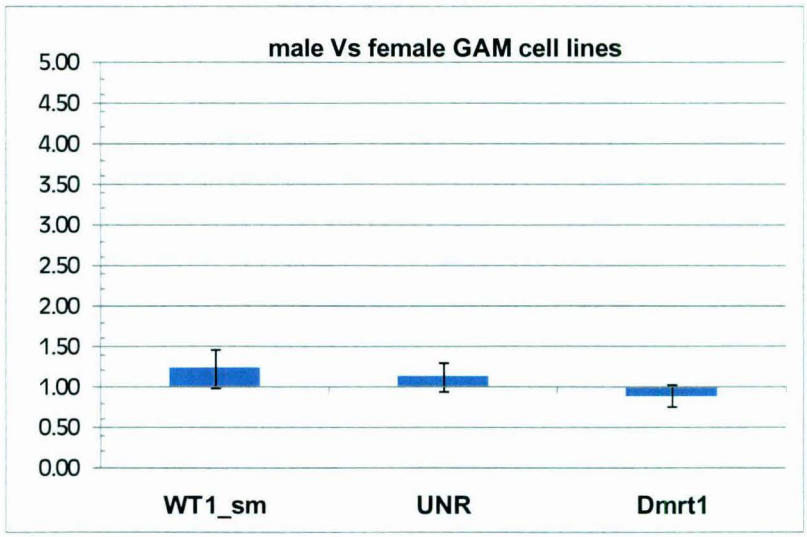


Figure 4.13: **Representative slides (from stages 22 and 24) showing similar pattern of localization of *cpWt1*b in the mesonephros tissue of male and female embryos through TSP.** *In-situ* hybridization was done using FITC labelled probe (green) and cells were counter-stained with PI (red). Note similar localization pattern in male and female embryos with no sex or cell type specificity and similar expression levels. **A)** Embryos kept at MPT; and **B)** embryos kept at FPT.

A



B

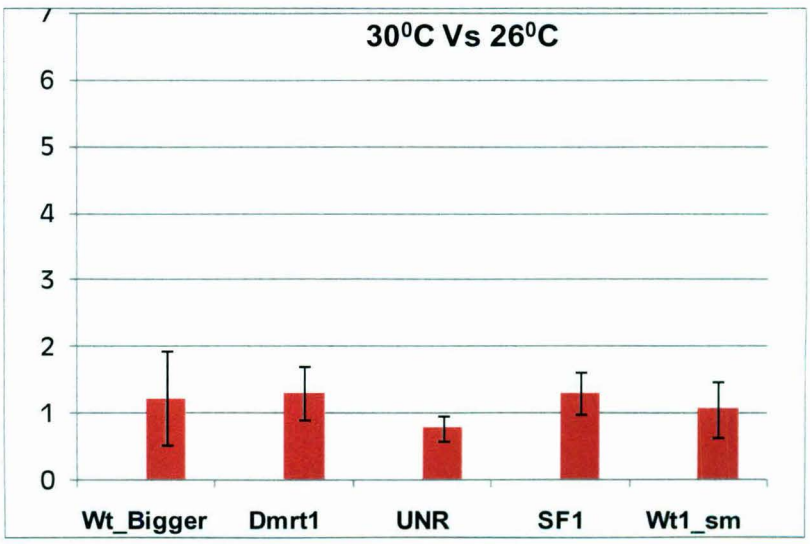


Figure 4.14: **Real-time PCR based expression analysis of few male gonad specific genes in the GAM derived primary cell lines.** Graph shows fold change in expression in cell lines derived from: **A)** embryos kept at MPT compared to those derived from embryos kept at FPT (cell lines were maintained at 26^0C); and, **B)** male GAM cell lines kept at 30^0C in comparison to the similar cell lines kept at 26^0C. Note almost similar expression of male specific genes independent of source of primary cell lines and/or temperature of incubation. The error bars represented the standard error values calculated for the 3 replicates.

4.1.1.7.2 Transfection of *cpWt1* isoforms in the primary GAM cell lines

Two *cpWt1* isoforms were cloned separately in the pEGFPN1 vector at the N-terminus of the EGFP protein to obtain the constructs cpWt1aEGFPN1 and cpWt1bEGFPN1for transfection studies (see section 3.4.1.8).

Transfection was achieved using effectene reagent as described in section 3.4.1.7. Primary cell lines derived from GAM tissue were transfected separately, with the two *cpWT1* constructs cpWt1aEGFPN1 and cpWt1bEGFPN1 and also with the vector pEGFPN1 as control.

4.1.1.7.3 Cellular Localization of *cpWt1* isoforms

Cells transfected with the pEGFPN1 (without insert) showed localization of GFP in the whole cell *i.e.* in nucleus, as well as, in cytoplasm. In comparison, the two isoforms showed different localization pattern when transfected. In the cells transfected with cpWt1aEGFPN1, EGFP mainly seen localized in the cytoplasm and very feebly in the nucleus (figure 4.15A). However, the EGFP localization in the cells transfected with the cpWt1bEGFPN1 was found to be strongly nuclear as well as in few cytoplasmic compartments (figure 4.15B). This different cytoplasmic localizations of two *cpWT1* isoforms was reproducible in three separate transfection experiments.

4.1.1.7.4 Expression analysis of male favoring genes in the cells transfected with the *cpWT1* isoforms

After 36 hours of transfection, RNA was isolated and expression analysis of the male favoring genes was carried out. All transfection experiments were repeated thrice for the expression analysis of sex determining genes by real-time RT-PCR. The increased (3-8 fold) expression level of transfected gene was indicator of successful transfection. It was notable that expression levels of *cpSox9* and *SC35F5* like ncRNA were undetectable in the non-transfected, as well as, vector only control transfected cells.

Expression levels of the male favoring genes in cells transfected with the vector carrying no insert (pEGFPN1), as well as, cells transfected with the cpWT1aEGFPN1 were found to be comparable. The analysis thus suggests that upregulation of *cpWt1***a** does not affect expression levels of male favoring genes (figure 4.16). Moreover, the expression of *cpSox9* and *SC35F5* like ncRNA remained undetectable after cpWT1aEGFPN1 transfection similar to the status of control cells. Likewise, no effect of *cpWt1***a** was seen

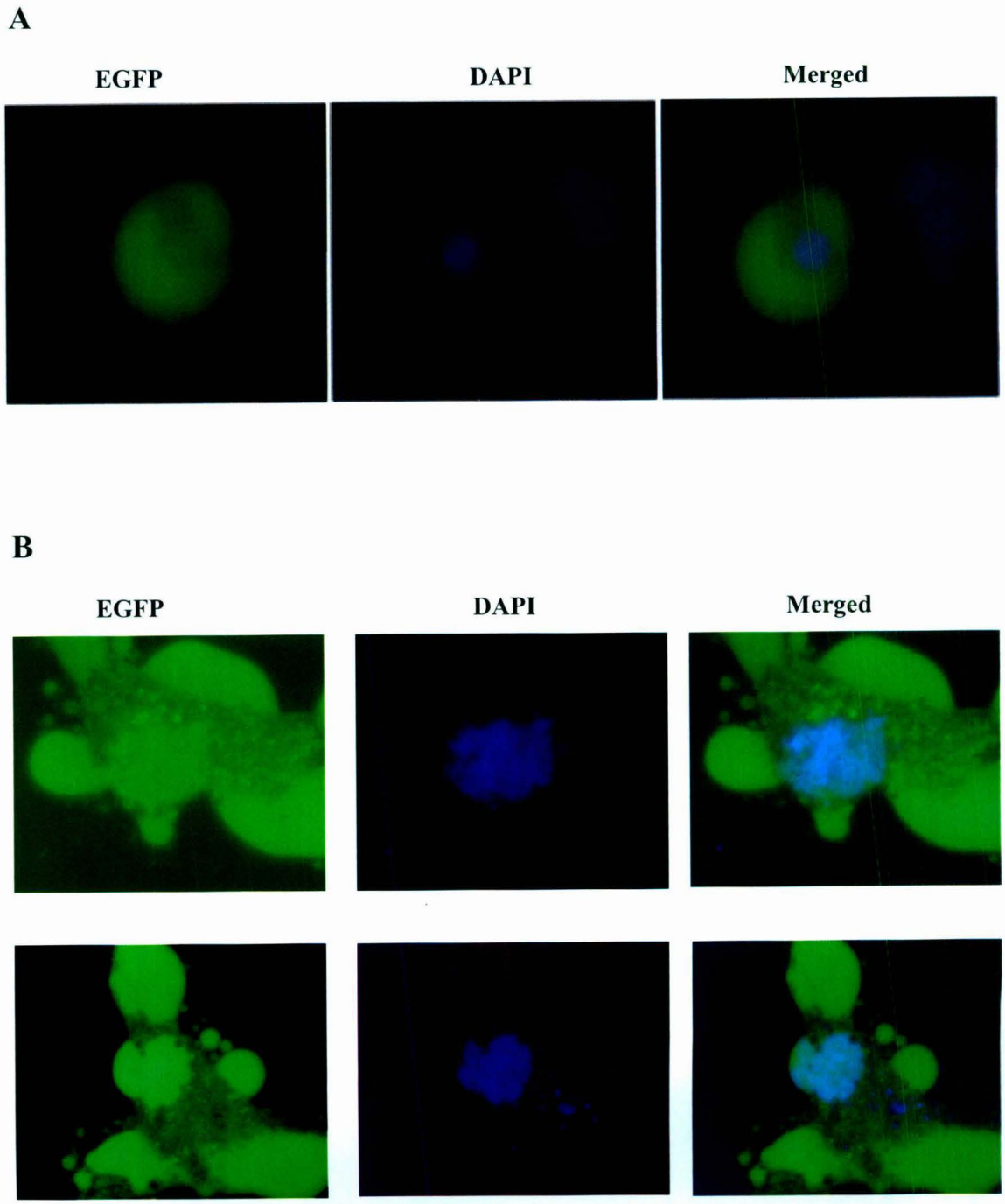
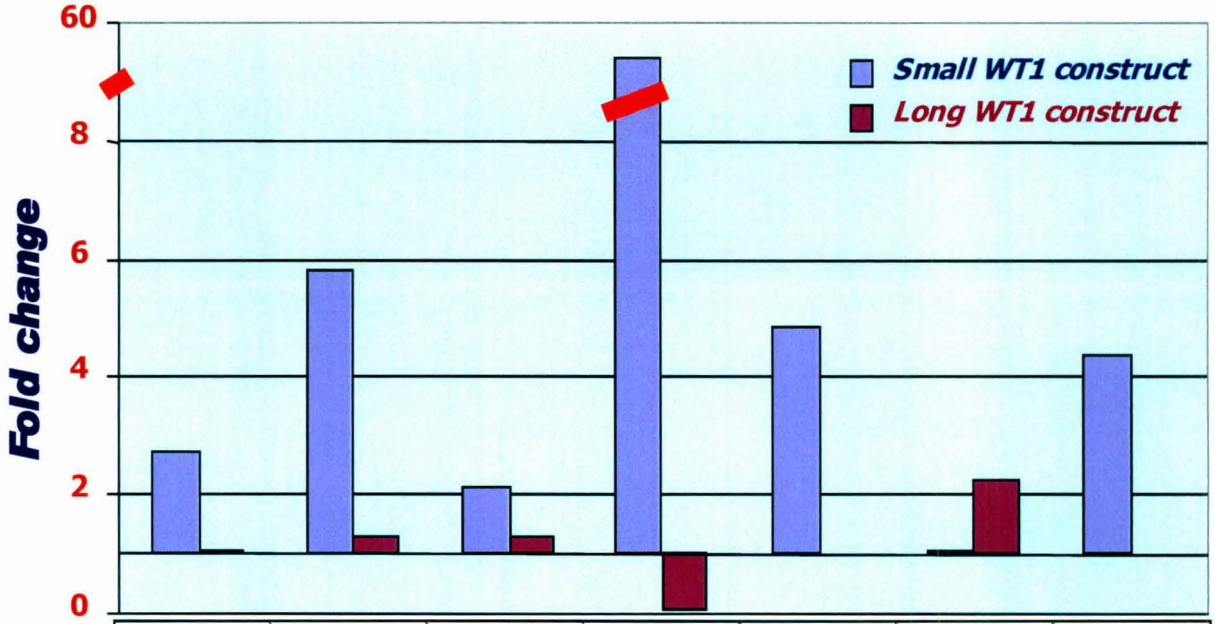


Figure 4.15: **Transfection studies of EGFP tagged *cpWt1* isoforms in GAM derived primary cell lines.** Note expression of: **A)** cpWT1aEGFPN1 mainly in the cytoplasm; and **B)** cpWT1bEGFPN1 mainly in nucleus along with some cytoplasmic compartments. Nucleus was stained with DAPI.



Genes	WT1_sm	Dmrt1	SF1	UNR	CC1.3	WT1_long	AMH
	2.73	5.77	2.15	54.26	4.87	1.09	4.39
	1.07	1.3	1.31	0.06		2.23	

Figure 4.16: **Real-time RT-PCR based expression analysis to compare change in expression levels of male gonad specific genes in the GAM derived primary cell lines transfected with the EGFP tagged isoforms cpWt1aEGFPN1 and cpWt1bEGFPN1.** Blue and purple bars indicate the fold-change over control cells (transfected with the vector only) after transfection with cpWt1bEGFPN1and cpWt1aEGFPN, respectively. Corresponding numerical values are tabulated below.

on expression levels of the male favoring genes even in experiments where expression was monitored after 48 hours of transfection.

In contrast to the above, cells transfected with the cpWt1bEGFPN1 construct revealed significant effect of shorter novel isoform (*cpWt1***b**) on all the tested male favoring genes (figure 4.16) In transfected cells, *cpWt1***b** expressed threefold compared the vector (control) transfected cells; most significantly this higher expression was accompanied with a concomitant significant increase in expression of all the male favoring genes, *i.e. cpAmh, cpCC1.3, cpDmrt1, cpSf1, cpUNR.* Increase in expression varied from 2 fold for *cpSf1,* 4-6 fold (for *cpDmrt1, cpAmh, cpCC1.3*) to a striking >50 fold for *cpUNR* (figure 4.16). Most importantly *cpWt1***b** transfection also induced the expression of important male favoring genes *cpSox9* and *SC35F5* like ncRNA (discussed in section 4.2.5), which were undetectable in the untransfected cells or cells transfected with the vector control (Figure 4.16). It was also interesting to note that the shorter isoform did not induce the expression of longer isoform *cpWt1***a**, which was comparable to the vector control. The results thus indicate a very important and early role of *cpWt1***b** in TSD.

4.1.2 Isolation and expression analysis of *cpDmrt1* (*Dmrt1* homologue of *C. palustris)*

The strategy used to isolate the *cpDmrt1* is depicted in figure 4.17. Briefly, a part of the *cpDmrt1* was obtained initially by RT-PCR using primers designed for the conserved DM domain of *Dmrt1* sequences of human, mouse, *Gallus*, Trachemys and Pelodiscus (NCBI accession numbers NM_021951, NM_015826, AF211349, AY316537, AB179697). This was followed by RACE reactions to identify/isolate full length cpDmrt1 transcript(s) that were later analysed for their expression using RT-PCR, quantitative Real-time PCR and Northern analysis. Finally, the *cpDmrt1* locus was isolated using the genome walking approach.

4.1.2.1 Initial identification of partial *cpDmrt1*

RT-PCR using primers Dmrt1F1/Dmrt1R1 using cDNA prepared from GAM tissues pooled over developmental stages 21-25, amplified a 517 bp fragment. The fragment was cloned and sequenced. Homology analysis confirmed it to be similar to the *Dmrt1* of vertebrates. The fragment shared 96% similarity with the *Dmrt1* sequences of *Pelodiscus* and *Trachemys*, 91% with that of *Gallus* and 88% with human *Dmrt1* sequence. This

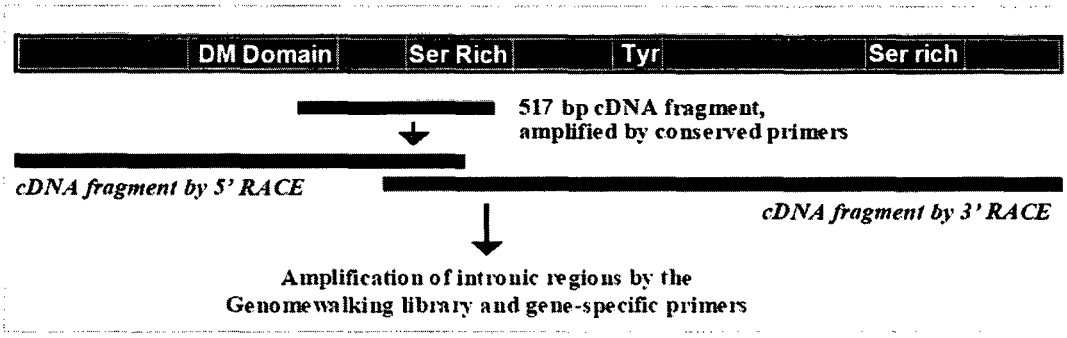


Figure 4.17: **Schematic representation of the approach used for identification and isolation of *cpDmrt1* (isoforms and locus).** In the 1st step, partial *cpDmrt1* was identified by RT-PCR with conserved primers, which was followed by 5'-/3' RACE to obtain different transcribed isoforms, and finally by genomewalking to ascertain/isolate the possible intron(s).

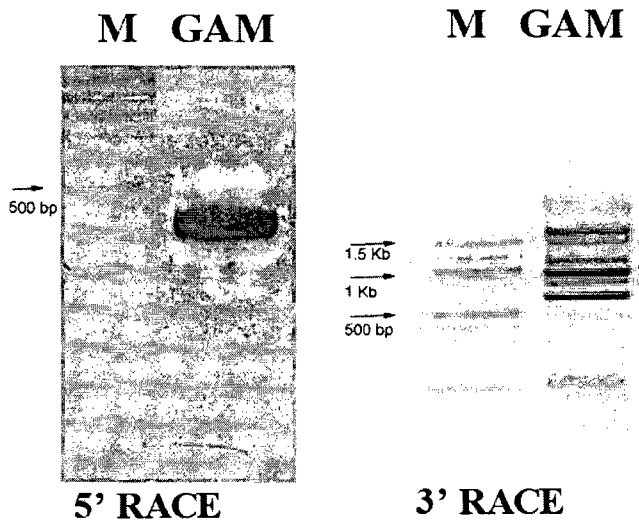


Figure 4.18: **End-specific RACE analysis suggesting the presence of multiple isoforms of *cpDmrt1* expressed in GAM tissue of developing crocodile embryos.** Note a single band in 5'-RACE (with primers DM5' and DM5'N), and multiple amplicons in 3'-RACE (with primers DM3' and DM3'N). Lanes: M—100 bp DNA ladder; GAM: Genital ridge-adrenal-mesonephros complex.

sequence was used to design gene-specific primers (DM5', DM5'N, DM3', DM3'N) for RACE to pull out the complete *cpDmrt1*.

4.1.2.2 Identification/isolation of multiple complete isoforms of *cpDmrt1*

RACE reactions were carried out to isolate the full length *cpDmrt1* cDNA/transcript using the RACE primers designed from the 517 bp *cpDmrt1* partial sequence (see above). Interestingly, 5'-RACE resulted in a single product of ~430 bp in size, while 3'-RACE resulted in several amplicons varying in size from ~300 – 1600 bp (figure 4.18) both with male and female GAM RACE ready cDNA suggesting the presence of multiple isoforms of *cpDmrt1*. Subsequent sequencing analysis of ~600 recombinant plasmids obtained from cloning of 5'-/3'-RACE products revealed a total of eight different isoforms of *cpDmrt1* ranging in size from 482 bp to 2060 bp. These isoforms are named as *cpDmrt1***a1** (2060 bp), *cpDmrt1***a2** (1119 bp), *cpDmrt1***b1** (1793 bp), *cpDmrt1***b2** (828 bp), *cpDmrt1***c** (945 bp), *cpDmrt1***d** (910 bp), *cpDmrt1***e** (1163 bp) and *cpDmrt1***f** (482 bp).

4.1.2.2.1 Diversity in *cpDmrt1* isoforms suggesting these to be splice variants

Sequence comparison of the eight *cpDmrt1* isoforms suggested these to be possibly splice variants differing from each other in: absence of a specific region, presence of a unique sequence(s) and/or in the length of their 3' UTR (figure 4.19). *In-silico* analysis further showed that while all isoforms had an identical short 5'-UTR, their coding regions and 3'-UTR varied significantly in size, ranging from 438 - 942 bp, and 19 – 1094 bp, respectively (Table 4.1).

Isoforms *cpDmrt1***a1** and *cpDmrt1***a2** are similar except that the *cpDmrt1***a1** has a very long 3' UTR (1092 bp for isoform *cpDmrt1***a1** while 140 bp in isoform *cpDmrt1***a2**). Similarly, the isoforms *cpDmrt1***b1** differs from *cpDmrt1***b2** only in its longer 3' UTR (1092 in *cpDmrt1***b1** and 140 bp in *cpDmrt1***b2**); and the two (*cpDmrt1***b1** and *cpDmrt1***b2**) lack a region of 284bp (corresponding to 364-647 bp) in comparison to the isoforms *cpDmrt1***a1** and **a2**, respectively. Isoform *cpDmrt1***c** shares complete similarity to *cpDmrt1***a2** except that it lacks 163 bp corresponding to 651-813 bp region of *cpDmrt1***a2**. Likewise, isoform *cpDmrt1***e** also shares complete similarity to the isoform *cpDmrt1***a2**, except for the presence of an additional 45 bases at 812 bp position. On the other hand, isoform *cpDmrt1***d** shares complete similarity to both *cpDmrt1***a1** and **a2** isoforms till its 812 bp, but hereafter has a unique 98 bp long 3' region that does not match even to any other *cpDmrt1* isoform. Shortest isoform *cpDmrt1***f** differs from all other isoforms in its unique 34 nucleotides long 3' end. Further, the analysis predicted

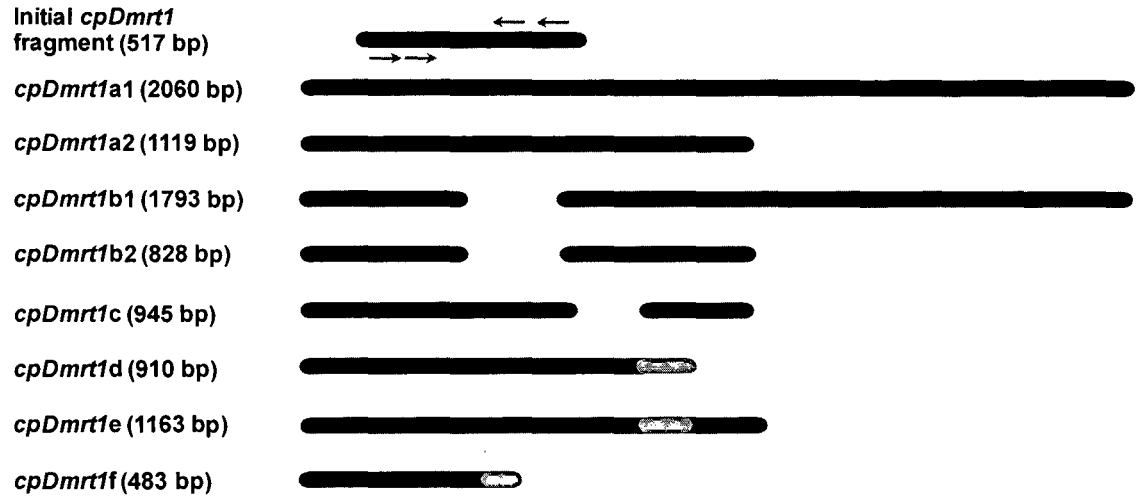


Figure 4.19: **Graphical representation showing the relative pair-wise homology between the eight *cpDmrt1* isoforms and the initial fragment of 517 bp revealed by RT-PCR.** The tentative positions of the primers used for initial RACE reactions are also indicated by solid arrows.

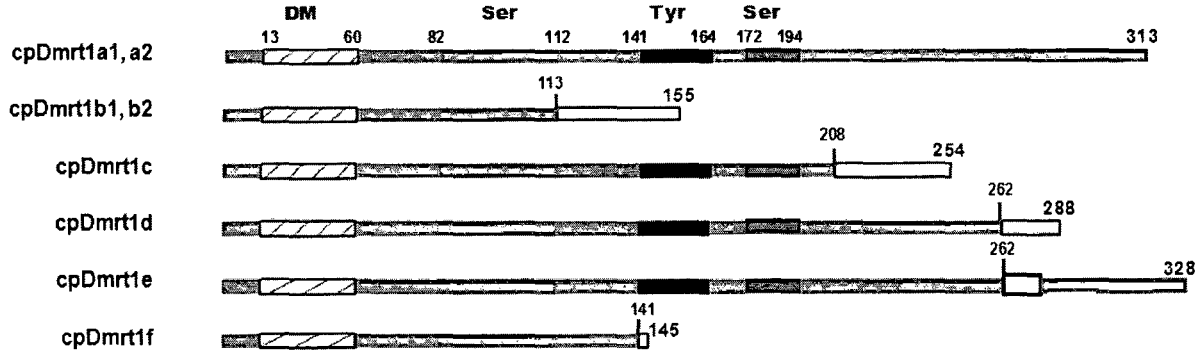


Figure 4.20: **Sequence alignment of predicted proteins of different *cpDmrt1* isoforms expressed in GAM during TSP.** All isoforms share the DM and first serine-rich domains, but differ significantly along their C-termini. White boxes represent the unique regions in the predicted proteins. Short vertical bars indicate the point of change in reading frames due to alternate splicing.

these isoforms to encode six proteins of 145, 155, 254, 288, 313 and 328 amino acids in size, all having the characteristic DM and serine-rich domains but unique C-termini owing to heterogeneity caused by possible splicing and/or shift in reading frame (figure 4.20; Table 4.1). The 313 aa long predicted protein CpDmrt1, coded by isoforms *cpDmrt1***a1** and *cpDmrt1***a2,** has a near conserved organization of the Dmrt1 predicted protein of vertebrates in having a DM domain (from 13 – 80 aa), two ***ser*** rich domains (from 82 - 112 aa, and 172 - 194 aa) and a tyrosine rich domain (from 141 to 184aa). Of all the six predicted proteins, this shows the highest resemblance to other vertebrate Dmrt1 counterparts (up to 95% in case of *Alligator* and 89% with other reptilian such as, *Trachemys* and *Pelodiscus*) with completely conserved DM domain, its flaking region and ser rich domains. Isoforms *cpDmrt1***b1/b2** are predicted to encode a shorter 155 aa protein of which first 113 aa are shared with the cpDmrt1 protein from *cpDmrt1***a1/a2,** and the remaining 42 aa are not conserved due to a shift in their reading frame expected to be caused by splicing. Similarly, the predicted proteins of *cpDmrt1***c** (254 aa) and *cpDmrt1***d** (288 aa) share 208 and 262 aa from their respective N-terminal with the protein coded by *cpDmrt1***a1/a2** isoforms, and have unique C-terminal ends of 46 and 26 aa, respectively, which do not match with any Dmrt1 protein sequences in public databases. Putative protein translation of *cpDmrt1***f** though shortest in length (145 aa) still retains the conserved DM domain and the 1st serine-rich domain. Interestingly, the putative protein of *cpDmrt1***e** has a unique 15 aa region (after the 2nd serine-rich domain near the C-terminal end) that is predicted to have a binding site for PDZ domain proteins and also a site for PIKK (Phosphoinositide-3 related protein kinase). Also, this is the only isoform for which no change is seen in the reading frame of its putative translation even after the inclusion of the additional bases splicing.

Table 4.1 *In-silico* predicted coding region coordinates of the eight *cpDmrt1* isoforms

Isoform	Length (bp)	Putative protein length (aa)/coding region (bp)	5' UTR length (bp)	3'UTR length (bp)
***cpDmrt1*a1**	2060	313 (942)	24	1094
***cpDmrt1*a2**	1119	313 (942)	24	153
***cpDmrt1*b1**	1793	155 (468)	24	1301
***cpDmrt1*b2**	828	155 (468)	24	336
***cpDmrt1*c**	945	254 (765)	24	156
***cpDmrt1*d**	910	288 (867)	24	19
***cpDmrt1*e**	1163	328 (987)	24	152
***cpDmrt1*f**	482	145 (438)	24	20

4.1.2.2.2 Homology analysis of *cpDmrt1* isoforms

BLAST based homology analysis revealed that at the nucleotide level, although all the isoforms were similar to the vertebrate *Dmrt1* sequences, it was *cpDmrt1***a2** that was the most conserved isoform showing a homology of 87, 85 and ~75% with *Dmrt1* sequences of *Trachemys*, *Gallus* and mammals, respectively, with completely conserved coding region and to some extent the 3'UTR.

In comparison, all the other isoforms had some unique regions (comprising, specific short unique regions and/or different lengths of UTRs) resulting from the alternative splicing, which did not match to other vertebrate *Dmrt1* sequences. Among these, the most notable regions were: 1) the unique long 3' UTR regions of *cpDmrt1***a1** and *cpDmrt1***b1**, approximately 700 bp of which (1154-1806 in case of *cpDmrt1***a1;** and 870 – 1657 bp of *cpDmrt1***b1**) did not match to any *Dmrt1* sequence in the database including *Dmrt1* coding regions of mammals, *Gallus* and *Xenopus*; 2) 98 bp 3' region of *cpDmrt1***d**; and 3) additional 45 bp near the 3' end of *cpDmrt1***e** that also did not match to any database sequence even when the search was extended to the genomes of *Gallus*, mammals and frog.

Phylogenetic analysis using *cpDmrt1**a2*** with *Dmrt1* sequences of other vertebrate species further confirms it to be a *Dmrt1* sequence most closely related to that of turtles followed by *Gallus* (figure 4.21). Moreover, the analysis also provides support to the archosaurial lineage of evolution.

4.1.2.2.3 5'-heterogeneity of *Dmrt1* in vertebrates

All the eight *cpDmrt1* isoforms had common 5' end; and homology analysis indicated the 5' end of the *cpDmrt1* to be 330 bp shorter than that of *Pelodiscus sinensis Dmrt1* sequence, which is the only reptilian *Dmrt1* sequence (known in the public domain) wherein 5' end is determined. However, 3' ends of *Dmrt1* sequences from reptiles were completely similar to the *cpDmrt1*, but were terminated at least 190 bp before 3' end of the *cpDmrt1***a2**.

Interestingly, sequence comparisons with the *Gallus Dmrt1* (which is also reported to undergo extensive alternative splicing) revealed that 5' ends of some of the *Gallus* isoforms was similar to the *cpDmrt1*; many other *Gallus* isoforms have shorter 5' ends. Likewise, heterogeneity in the 5' ends was also seen when *cpDmrt1* isoforms were

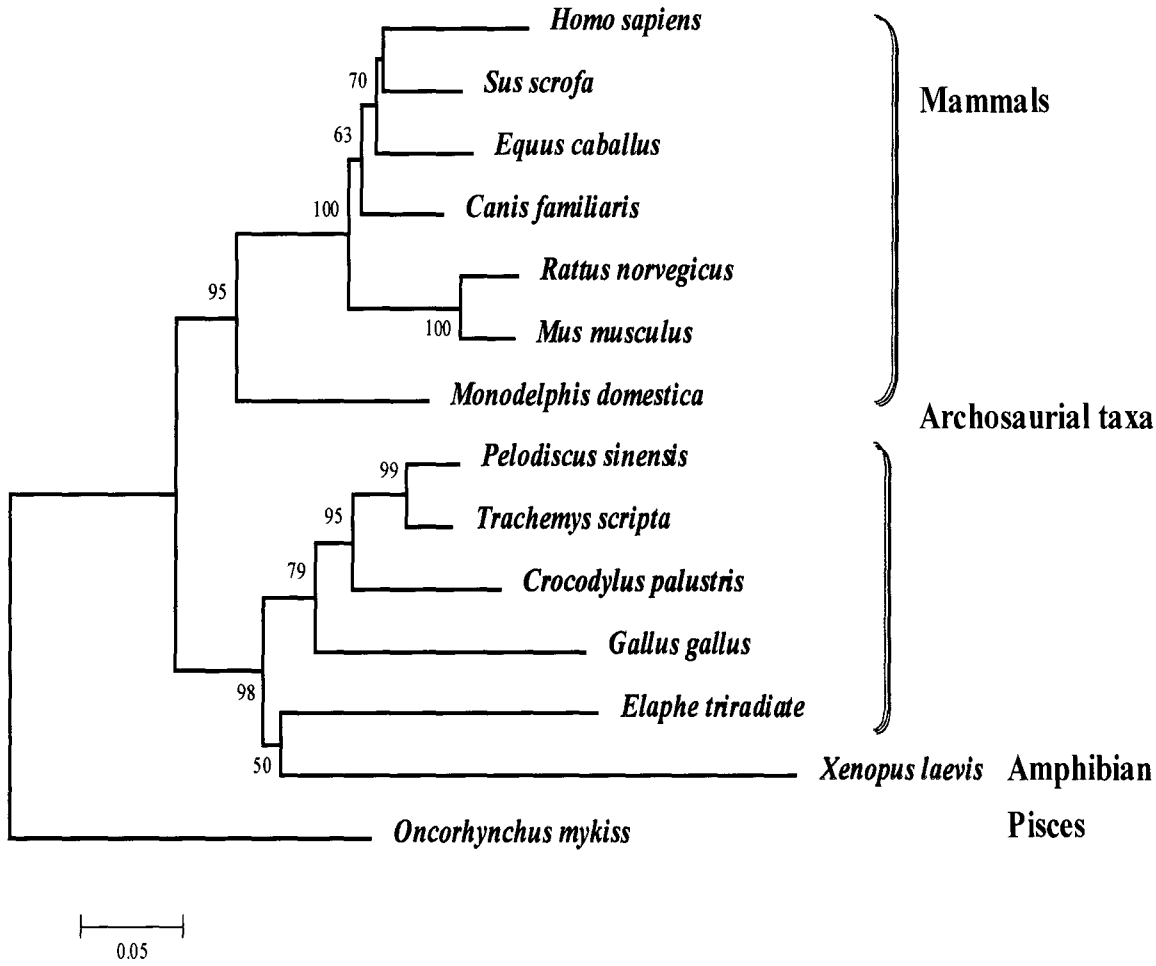


Figure 4.21: **Neighbor-joining tree showing the generic/phylogenetic relationship of *cpDmrt1*a2 with *Dmrt1* sequences of other vertebrates.** The tree is generated using the Kimura 2-parameter based genetic distance estimates and MEGA4.1 software. A total of 593 complete nucleotide sites and empirical nucleotide frequencies/gamma-correction were considered in deriving genetic estimates. All gaps/missing data were eliminated from the final dataset. The values on nodes are percent Bootstrap values (based on 1000 reiterations); sum of all branch lengths=1.53810388.

compared with the *Dmrt1* sequences from mammals, with latter having varyingly extended 5' ends (190 bp in *Canis*, 266 in pig, 166 in *Eqqus*, 493 in *Pan troglodytes* and 323 bp of the longest known isoform of human *Dmrt1*). Compared to the 5` end, the mammalian *Dmrt1* sequences exhibited a general similarity to the *cpDmrt1***a2** isoform till the end of the coding region; however these had completely dissimilar 3' UTRs.

A similar but clearer pattern was evidenced when the putative translations were compared. The putative protein from *cpDmrt1***a2** shared ~ 96% similarity to the Dmrt1 protein of *Pelodiscus sinensis* and 78-88 % with that of other vertebrates. Most of the observed variability was apparent towards the N-terminus. The cpDmrt1 protein predicted from *cpDmrt1***a2** was 55 aa shorter at its N-terminus but had 43 aa more at its C-terminal end when compared to predicted Dmrt1 protein of turtle *Pelodiscus sinensis*. *cpDmrt1***a2** translation was found to be 65-67 aa shorter at its N-terminal end compared to Dmrt1 proteins of many of the mammals; strikingly, its C-terminal end was revealed to be completely conserved in length and sequence with other vertebrate Dmrt1 proteins. On the other hand, *Gallus* Dmrt1 protein was one that shared similarity at both the N-terminal and C-terminal ends with the crocodilian Dmrt1; however the sequence similarity was not as high as shown with reptilian Dmrt1 sequence.

4.1.2.3 Genomic organization of *cpDmrt1* coding locus and alternate splicing

To understand the genomic organization and validate the origin of multiple transcribed isoforms, efforts were made to isolate the *cpDmrt1* coding locus by Genomic-PCR and Genome walking approaches. A number of primers (DMGW1, DMGW2, DMGW3, DMGW4, DMGW5, DMGW6, DMGW7, DMGW8) designed from the regions flanking the gaps in the alignment of the isoforms (figures 4.19) were used for genome walking (section 3.4.2.5). The *cpDmrt1* locus (figure 4.22), thus isolated, comprises three exons (exon-1, exon-2, and exon-3 of 650, 208 and 1344 bp length, respectively), and two introns (1311 bp long intron-1, and 252 bp of partially sequenced intron-2). Subsequently, comparisons of the genomic locus and expressed *cpDmrt1* sequences unequivocally demonstrated that all the eight isoforms were generated by alternative splicing (alternative 5' and 3' splice sites, exon-skipping, exonization of intronic sequences and alternative polyA sites) from the same genomic locus (figure 4.22). All the isoforms retain the part of exon-1 that codes for the conserved DM domain. It is noteworthy that contrary to the *Dmrt1* gene in many vertebrates, the *cpDmrt1* does not have any intron at the DM domain coding region. Isoforms *cpDmrt1***a2** and *cpDmrt1***b2** differ from isoforms *cpDmrt1***a1** and *cpDmrt1***b1** only in their 3'UTR (shorter by ~1.0 kb)

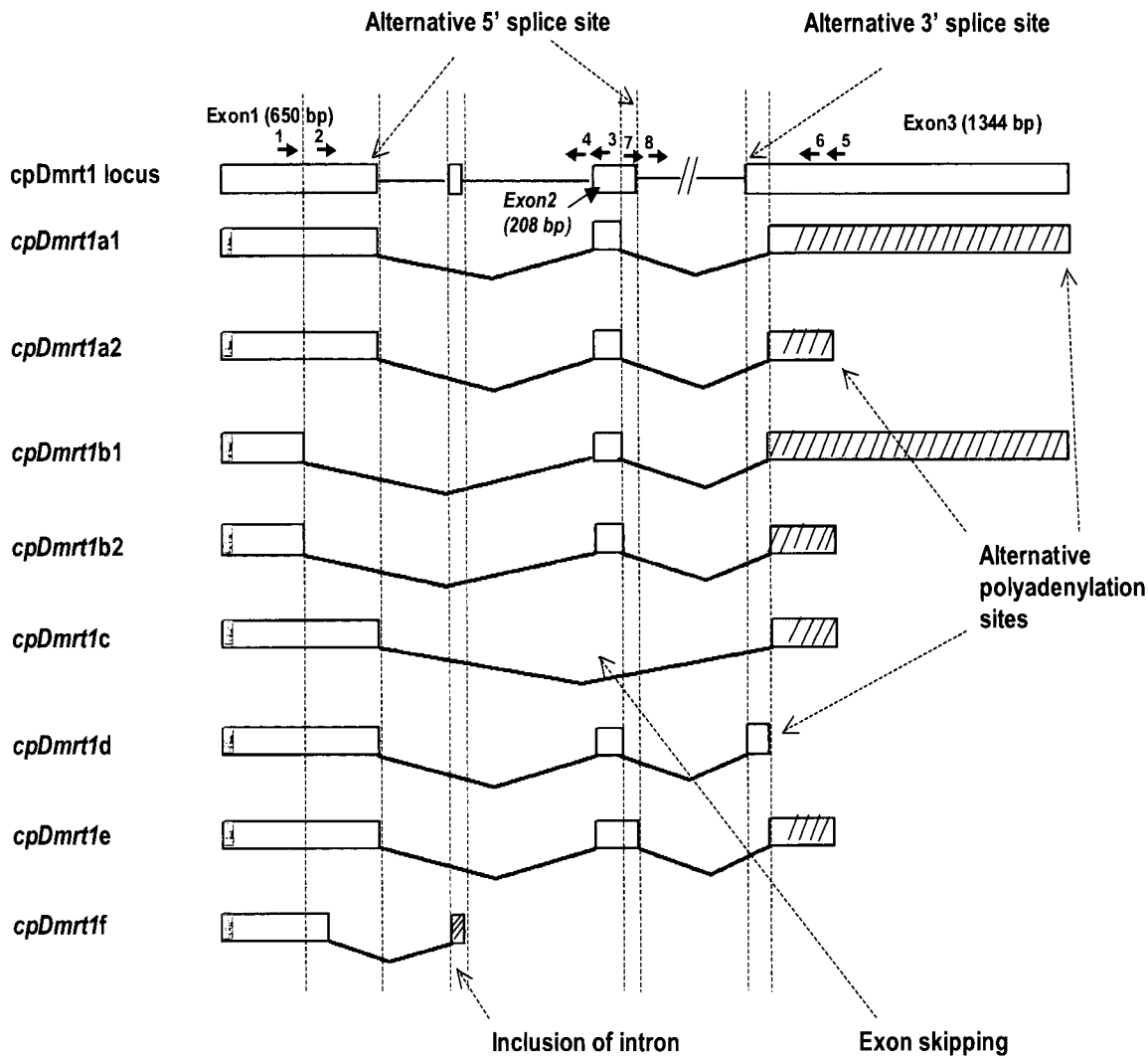


Figure 4.22: **Schematic representation of the *cpDmrt1* genomic locus and the alternative splicing leading to eight multiple isoforms.** White boxes represent three exons, whereas two intronic regions are shown as straight/'V' lines. Isoforms *cpDmrt1***b1**, *cpDmrt1***b2** and *cpDmrt1***f** lack part of exon-1, while isoform *cpDmrt1***e** has additional 45 bp of exon-2 due to alternative 5′ splice sites. Isoform *cpDmrt1***d** has a different 3′-end and isoforms *cpDmrt1***a1/b1** have longer 3′-UTRs because of alternative 3′-splice sites. Isoform *cpDmrt1***c** lacks exon-2 due to exon skipping. Shortest isoform *cpDmrt1***f** lacks part of exon-1, complete exon-2 and exon-3 but includes 38 bp region of intron-1 due to alternate splice sites. All isoforms other than *cpDmrt1***a1/b1** have anterior alternative polyadenylation signals and shorter 3′-UTRs. The solid black arrows indicate the primers and their locations, which were used in genome walking PCRs to identify the intronic sequences.

respectively, due to anterior alternative polyA signals. Isoforms *cpDmrt1***b1** and *cpDmrt1***b2** lack a part of exon-1 (284 bp) due to the use of an alternative 5'-splice site in exon-1, whereas isoform *cpDmrt1***e** has additional 45 bp due to a 5'-alternative splice site in exon-2. Isoform *cpDmrt1***c** lacks exon-2 (163 bp) and isoform *cpDmrt1***d** has a completely different 3'-UTR due to an alternative 3'-splice site in exon-3. The shortest isoform *cpDmrt1***f** was the most unique, comprising only part of exon-1 and exonization of a short sequence (34 nucleotides) of intron-1. Introns 1 and 2 don not have any significant evolutionary similarity.

4.1.2.4 Expression analysis of *cpDmrt1*

4.1.2.4.1 Tissue specific expression

cpDmrt1 expression was analyzed in the embryonic GAM, liver, heart, kidney and brain by RT-PCR using Dmrt1F1 and Dmrt1R1 primers. For each of the tested tissue, male and female RNA were pooled and subjected to reverse transcription, and the resulting cDNA products were used as tissue-specific templates in RT-PCR. The expression analysis, thus carried out, revealed that *cpDmrt1* is mainly expressed in the GAM and to a much lower level in the brain and kidney, but not in the heart and liver tissues.

4.1.2.4.2 Male-sex specific higher expression of *cpDmrt1* in GAM tissue

cpDmrt1 expression levels in the GAM tissues at male and female promoting temperature through TSP were analyzed using semi-quantitative and quantitative RT-PCR. RNA isolated from the male and female GAM tissues from stage 21 to 25 were reverse transcribed separately and used as template for PCR. A semi-quantitative RT-PCR using Dmrt1F and Dmrt1R (designed from common region present in all isoforms) revealed significantly higher expression in the GAM of male embryos (kept at MPT) in comparison to that of female embryos (kept at FPT) throughout TSP *i.e.* from stage 21-25 (figure 4.23A).

Similar results were obtained by real-time PCR that showed >21-fold higher expression of overall *cpDmrt1* in the earliest 21st stage followed by a similar 16-17 fold higher expression during stages 22 to 25 at MPT compared to FPT (figure 4.23B). However, no such sex-specific upregulation was seen in the brain and kidney during TSP (data not shown). Each QRT-PCR experiment was done in triplicate and was repeated twice.

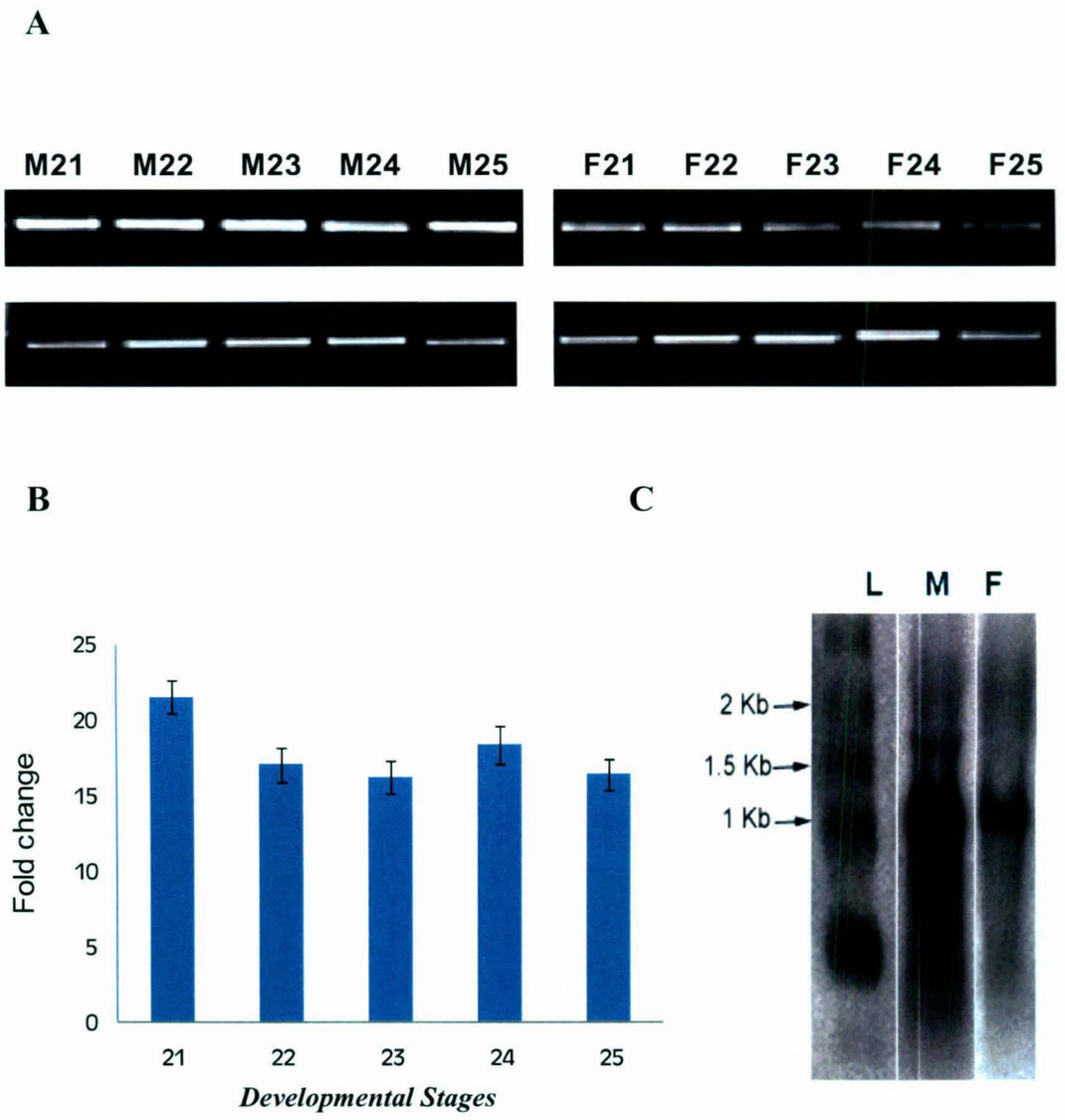


Figure 4.23: **Expression analysis of *cpDmrt1* through TSP. A)** Semi-quantitative RT-PCR, using primer Dmrt1F1 and R1 shows higher expression of *cpDmrt1* in the male GAM compared to female GAM tissue; **B)** Real-time RT-PCR shows significant upregulation of *cpDmrt1* in the male GAM tissues (M21–M25) compared to that from females (F21–F25) through TSP, using primer pair DMF1 and DMR1; GAPDH was used as control; and **C)** Northern analysis showing presence of several isoforms and significantly higher expression of *cpDmrt1* in male GAM; Lanes: L, M, F represent RiboRuler RNA ladder (Fermentas, USA), male and female GAM RNA, respectively. Error bars in 'B' represent standard error calculated over three replicates.

Northern analysis further confirmed higher expression of *cpDmrt1* at the MPT. Northern hybridization was done using a 322 bp region amplified by primers GDMF and GDMR primers (designed from common region present in all isoforms). Hybridization was carried out to detect *cpDmrt1* in male and female GAM tissue RNA (each was pooled over 21st to 25th stages). In addition to confirming the higher expression of *Dmrt1* at the MPT, Northern analysis also revealed presence of several isoforms in the GAM tissue (figure 4.23C). It was noteworthy that analysis suggested a differential expression of multiple isoforms with maximum expression for an isoform of >1 kb size (figure 4.23C), which corroborates with data from expression pattern of individual isoforms (see below).

4.1.2.4.3 Expression pattern of individual *cpDmrt1* isoforms

Expression of isoforms *cpDmrt1***a1/b1** (combined), and *cpDmrt1***d** and *cpDmrt1***e** (separately) was further analyzed using semi-quantitative, as well as, real-time RT-PCR in the GAM tissues at MPT and FPT during TSP (figure 4.24A, figure 4.25). The results revealed that *cpDmrt1***e** was the most abundant isoform followed by c*pDmrt1***a1/b1** (combined expression) contributing ~73% and 26%, respectively to the overall *cpDmrt1* levels at MPT; and ~60% and 23%, respectively to the overall *cpDmrt1* levels at FPT. Further, *cpDmrt1***e** expression peaked at 21st developmental stage, while *cpDmrt1***a1/b1** showed highest upregulation at 22nd stage. In contrast, the expression of *cpDmrt1***d** was <1% of the total *cpDmrt1* levels at both MPT and FPT, and its highest expression was observed at the 21st developmental stage. Overall, the analysis showed that all the tested isoforms were upregulated at MPT compared to FPT through TSP (figure 4.24A), but to different extents and at different time points (figure 4.24B). It was noteworthy that at its peak, isoform *cpDmrt1***e** expression was about 32 fold higher in male GAM compared to female GAM; and that this most abundant *cpDmrt1* isoform differs by only a 45 bp insertion in the coding region with the most conserved isoforms (i.e., *cpDmrt1***a2** or *cpDmrt1***a1**).

4.1.3 Isolation and expression analysis *cpAromatase* (*Aromatase* homologue of *C. palustris*)

Aromatase is a highly conserved gene among the vertebrates, which makes PCR based identification of this gene a convenient method. The strategy used for *cpAromatase*

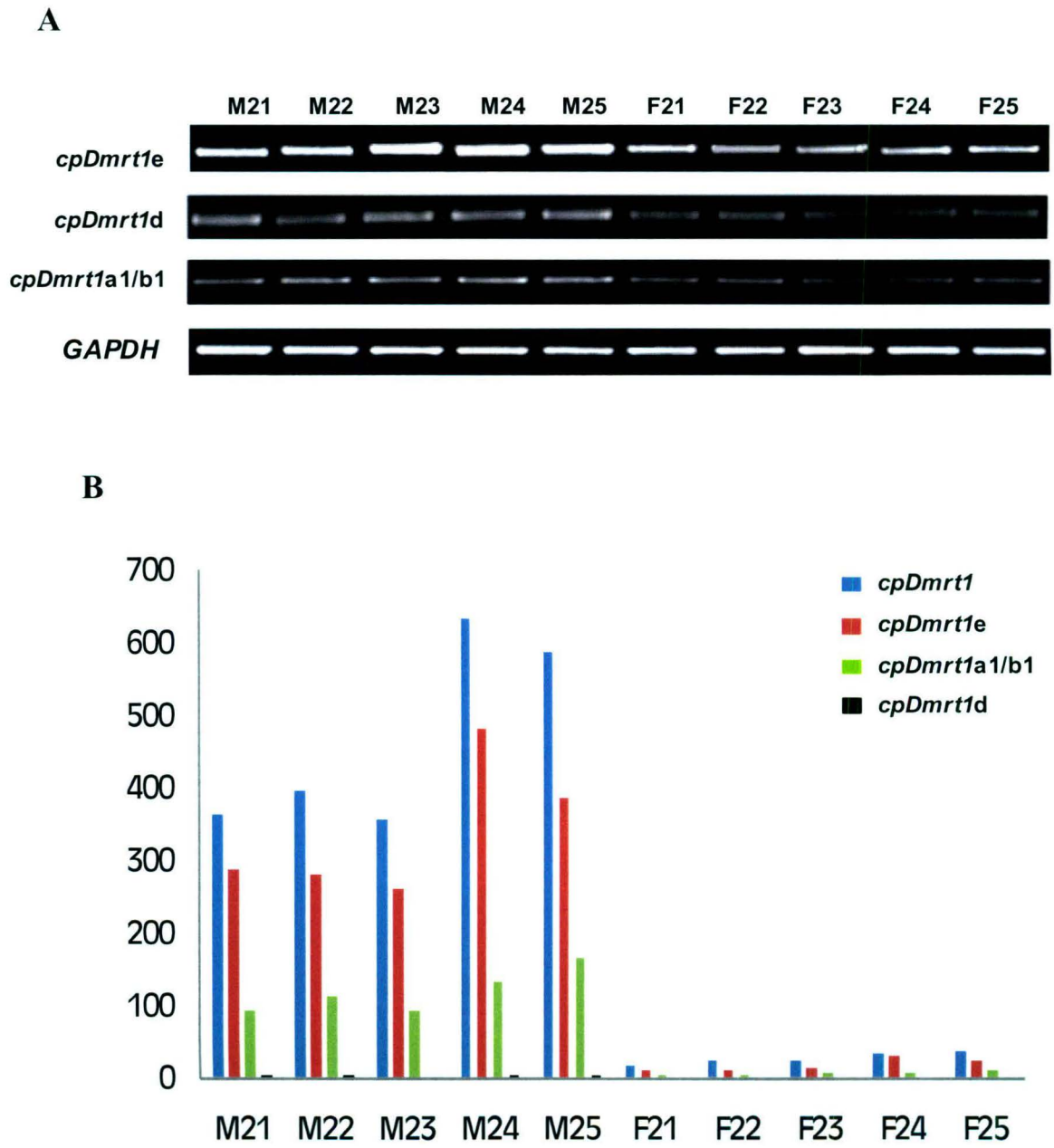


Figure 4.24: **Expression analysis of different *cpDmrt1* isoforms in GAM tissues of male/female embryos through TSP. A)** Semi-quantitative RT-PCR; and **B)** real-time quantitative RT-PCR. Note**:** significantly higher expression in male GAM (M21–M25) through TSP compared to female GAM (F21–F25); and *cpDmrt1***e** as most highly expressed isoform followed by isoforms *cpDmrt1***a1**/**b1**; and much lower expression of other isoforms.

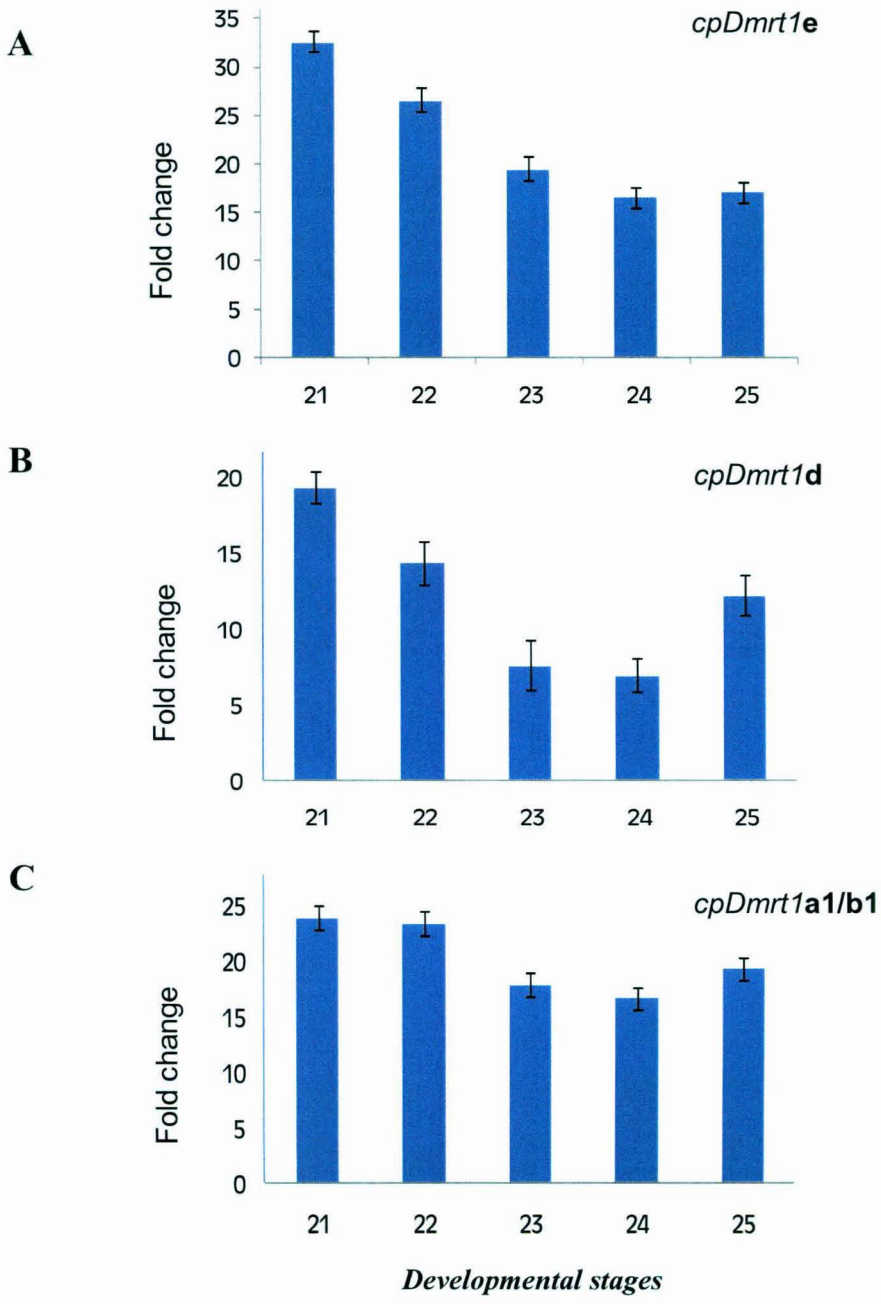


Figure 4.25: **Real-time quantitative RT-PCR based expression analysis showing significant higher expression of four different *cpDmrt1* isoforms in GAM tissue of male embryos (versus female embryos) through TSP (developmental stages 21 to 25). A)** *cpDmrt1***e** expression using primers Dmrt1eR and Dmrt1_4; **B)** *cpDmrt1***d** expression using primers DmrtdR and Dmrt1_5; and **C)** higher expression of *cpDmrt1* isoforms **a1** and **b1** monitored using long 3' UTR derived primers DMisoa1b1F/R**.** In each case, template RNA was normalized with *GAPDH*. Error bars in represent standard error calculated over three replicates.

Isolation is described in figure 4.26A. Complete *cpAromatase* isolation was achieved by 5' and 3' RACE.

4.1.3.1 Identification and isolation of partial *cpAromatase*

As described in section 3.4.3.1, non-mammalian *Aromatase* sequences from *Alligator, Trachemys, Gallus, Xenopus, Taeniopygia, and Homo sapiens* (Accession numbers AY029233, AF178949, NM_001001761.1, NM_001076691 and X13589, respectively) were retrieved from NCBI database. These sequences were aligned and two sets of primers AroF1/AroR1 and AroF2/AroR2 were designed spanning the overlapping conserved region corresponding 577 to 1495 bp and 787 to 1270 bp respectively, of the alligator *Aromatase* mRNA (Accession number AY029233). These primers were used to amplify *cpAromatase* using RT products made from the pooled male and female GAM RNA, as template. Primers AroF1 and R1 yielded a 917 bp product while primers AroF2 and AroR2 yielded a 483 bp product (figure 4.26B). Both PCR products were sequenced and homology analysis revealed these to share 98% similarity with the Alligator *Aromatase* and 94% and 90% homology to that of turtle and *Gallus aromatase* sequences, respectively. The 917 bp amplified region of *cpAromatase* corresponded to the Cyp450 domain of Aromatase that is responsible for the enzymatic action *i.e.* Aromatization of C19 steroids.

4.1.3.2 Identification/isolation of complete *cpAromatase* transcript(s)

Complete sequence of *cpAromatase* transcript(s) was isolated by 5', 3' RACE using primers designed from the 917 bp region of *cpAromatase* obtained in the first step (see above). 5' RACE done using primers Aro5' and Aro5'N (nested to former), resulted in two amplicons of sizes 700 bp and 900 bp, of which the smaller amplicon was poorly amplified (figure 4.26C). Similarly, 3' RACE carried out with Aro3' and Aro3'N primers, yielded a major amplicon of approximately 1 kb (figure 4.26C), amplicons were purified using Qiagen PCR purification column. Purified RACE products were cloned in the pCR2.1 vector and about 100 recombinant plasmids were sequenced. The analysis revealed two isoforms of *cpAromatase*; a bigger isoform of 2133 bp and smaller isoform of 1977 bp (figure 4.26D). The two isoforms are referred hereafter, as *cpAromatase***a** and *cpAromatase***b**. The isoform *cpAromatase***b** lacks 156 bp part which is present in the *cpAromatase***a** towards its 5' end (corresponding to 427 to 523 bp), and represents a novel isoform that has not been reported in any other vertebrate. Both *cpAromatase* isoforms have a 129 bp 5' UTR and 489 bp 3' UTR. Coding region of the bigger isoform,

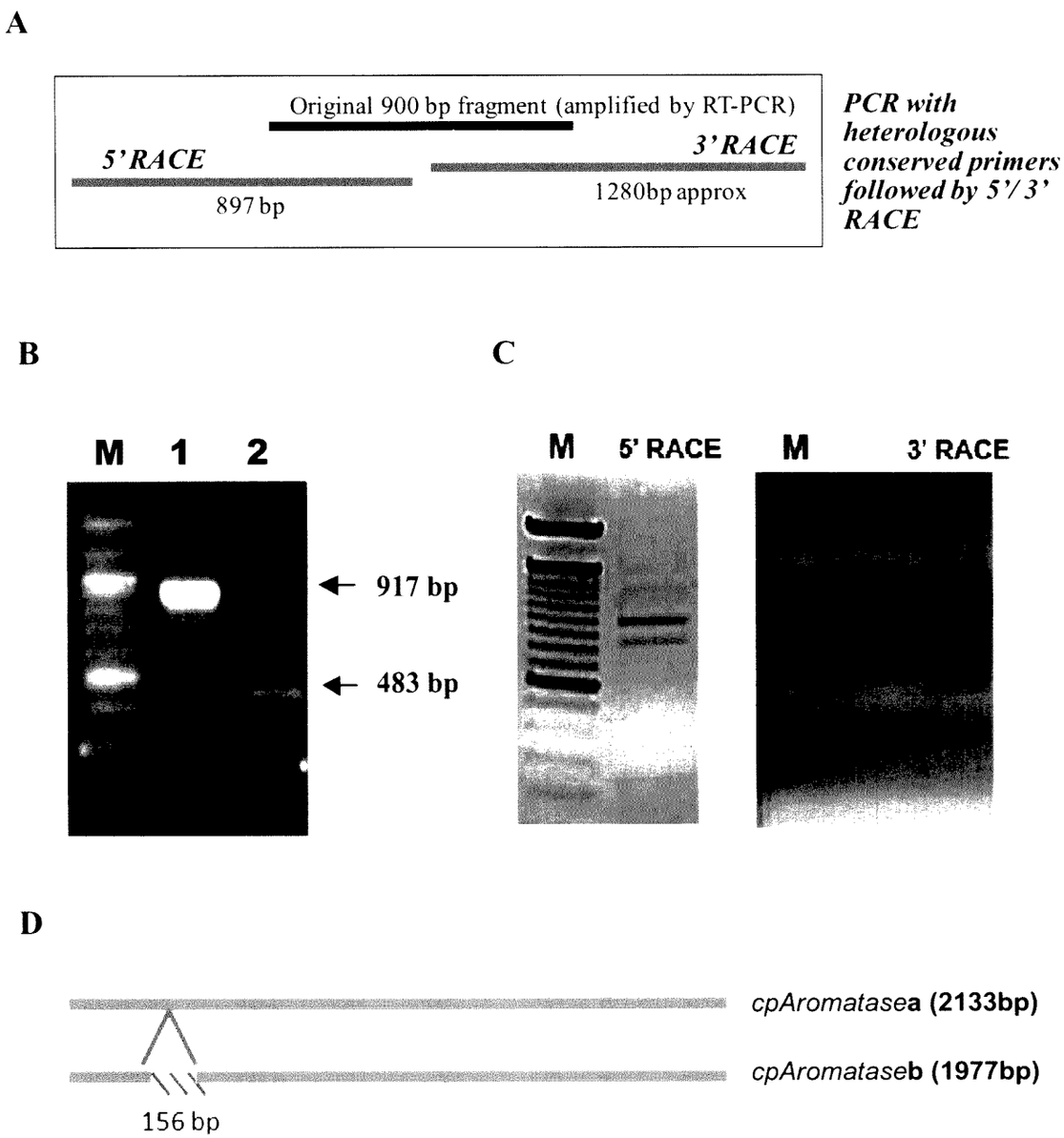


Figure 4.26: **Isolation of full-length crocodilian homologue of Aromatase gene (*cpAromatase*). A)** Schematic representation of strategy used to pull out the full length *cpAromatase*; **B)** Initial Identification of partial aromatase transcript of *Crocodylus palustris* by RT-PCR using pooled RNA from male and female GAM tissues; the primer combinations AroF1, AroR1 (lane 1) and AroF2, AroR2 (lane 2) amplified 917 and 483 bp fragments, respectively; **C)** 5'-/3'- RACE to pull out ends of *cpAromatase*. Note the two amplified bands in 5'-RACE; and **D)** Schematic representation of alignment of two *cpAromatase* isoforms. Note *cpAromatase***b** lacks a 156 bp region that falls in the coding region. M represents the 100 bp ladder lane.

*cpAromatase***a** is of 1515 bp while that of smaller isoform *cpAromatase***b** is 1359 bp. Putative translation of *cpAromatase***a** and *cpAromatase***b** isoforms results in to 504 and 452 amino acids long proteins, respectively. The putative translated protein of *cpAromatase***b** lacks 52 amino acids towards its N-terminal corresponding to 100-151 amino acids protein translation of *cpAromatase***a.** These 52 amino acids present only in the *cpAromatase***a** isoform are predicted to have myristylation and phosphorylation sites, which are required for trafficking of Aromatase protein. Except this domain the remaining putative proteins of the two *cpAromatase* isoforms were identical.

4.1.3.3 Homology analysis of *cpAromatase* isoforms and their putative proteins

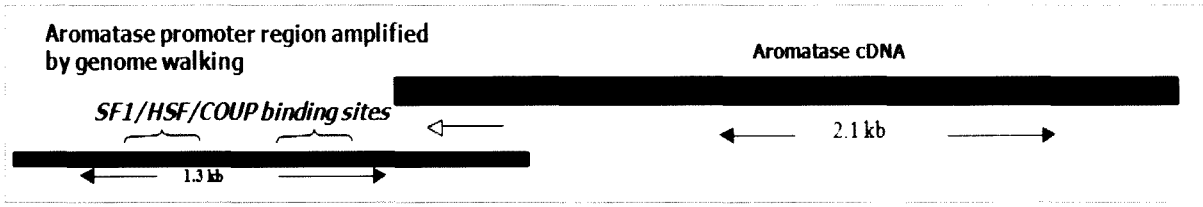
Sequence comparison at nucleotide level, revealed *cpAromatase* to be most similar (95%) to that of *Alligator Aromatase,* followed by approximately 85% similarity to *Aromatase* mRNA sequences of *Trachemys* and *Gallus,* and 70-75% sequence similarity to that of mammals. However, interestingly the observed sequence similarity was mainly restricted to coding region, while significant variation in the homology was visible for the 5'-/3'- UTR sequences, across the compared vertebrates. Strikingly *cpAromatase* shared a very high similarity with alligator *Aromatase* even for both 5'- & 3'- UTR regions; but in contrast it did not share similarity with the 3'-UTR region in case of turtle/*Gallus* sequences, and with none of the 5'- & 3'-UTRs of mammalian *Aromatase.*

Homology comparison at the protein levels showed that except first 42 amino acids, putative protein of both *cpAromatase***a** and *cpAromatase***b** isoforms belong to conserved p450 domain of vertebrate *Aromatase.* At the amino acid level, the cpAromatase predicted proteins share approximately 98% similarity to that of alligator, and approximately 88% similarity to that of other non-mammalian vertebrates like *Gallus.* Furthermore, it was noteworthy that most of variation revealed in these comparisons, were mainly restricted to the C-terminal of Aromatase proteins.

4.1.3.4 Isolation of *cpAromatase* upstream region

cpAromatase upstream region was isolated using genome walking strategy and primers designed from the 5' end region of the *cpAromatase* transcript. Accordingly, using the primer AGW2 in the primary PCR and Primer AGW1 in the secondary PCR rounds of genome walking, followed by cloning and sequencing, a 1405 bp upstream region of *cpAromatase* was obtained (figure 4.27A). The initial BLAST based homology analysis revealed that the 300 bp region of the *cpAromatase* upstream region shared similarity to the *Aromatase* promoter regions described in the Zebra finch, *Gallus*, *Courtnix* and

A



B

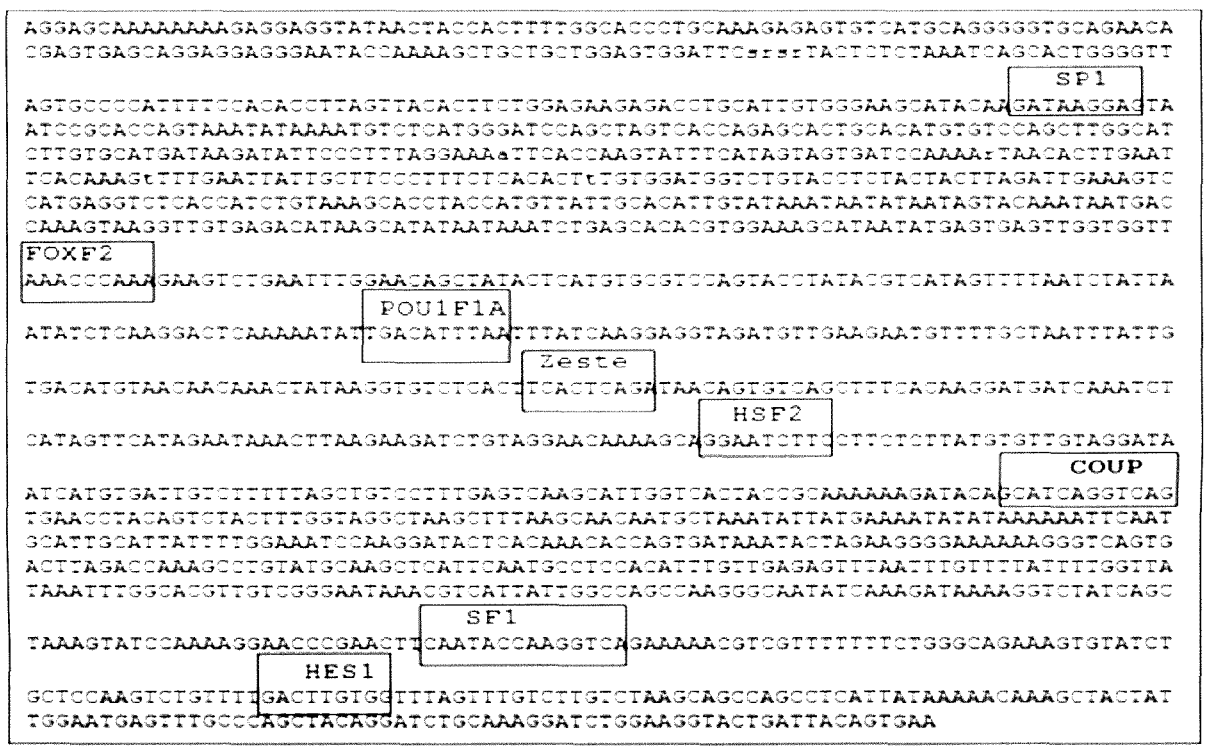


Figure 4.27: **Isolation and *in-silico* analysis of upstream region of *cpAromatase* coding region. A)** Schematic diagram showing the genome walking strategy used to identify/isolate the upstream region of the *cpAromatase* coding region; and **B)** 1405 bp *cpAromatase* upstream sequence showing the predicted binding sites for various transcription factors, some of which are indicated in the boxes.

Human; however, region after the first 300 bp does not resemble to the promoter sequences described in the above organisms. Here it may be noted that except *Gallus Aromatase* promoter region in any reptile or bird has not been explored.

4.1.3.4.1 *In-silico* analysis of the *cpAromatase* upstream region for regulatory elements

The 1405 bp upstream region of the *cpAromatase* was explored to find out whether it harbours any conserved binding site(s) for the important transcription or regulatory factor(s). The analysis was done using web based search engines like MatInspector and Transcriptional Element Search System TESS, revealed presence of TATA box at -29 position and many other conserved regulatory elements. The most important binding sites predicted were for: SF1 (steroidogenic factor1) at -138 bp and NR2F2 (Nuclear receptor subfamily 2 group F member 2) or COUP (chicken ovalbumin upstream transcription factor) at -470 bp (figure 4.27B).

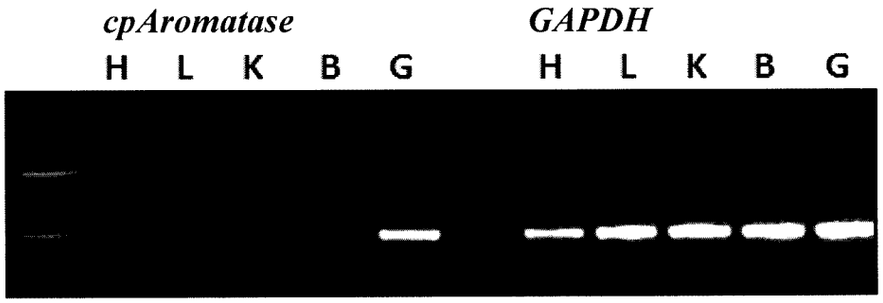
4.1.3.5 Expression analysis of *cpAromatase*

First of all, *cpAromatase* expression was examined to ascertain the tissue specificity if any. RT-PCR analysis with tissue specific pooled male and female cDNA (from embryonic GAM, Heart, Brain, Kidney and Liver) revealed *cpAromatase* expression only in the brain and GAM tissues (figure 4.28A). Subsequently, sex-/stage- specificity of *cpAromatase* expression was examined by monitoring the expression in the GAM and brain tissues from embryos reared at MPT and FPT. Semi-quantitative RT-PCR done to this end demonstrated that *cpAromatase* expression in GAM tissue was female specific, detectable from 23rd stage onwards at FPT, was also upregulated with the development with highest expression at stage 25th (figure 4.28B). On the other hand no expression was detectable in the male GAM tissues at MPT (figure 4.28B). Compared to GAM, no sex-specificity was seen for *cpAromatase* expression in embryonic brain tissues; comparable expression was seen in brain tissues from both MPT and FPT. Increasing expression was detectable from 23rd stage onwards in both cases with highest expression at stage 25 (figure 4.28C).

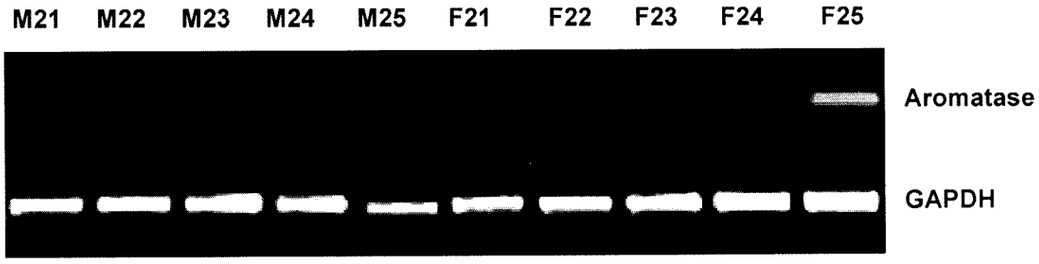
4.1.3.5.1 Tissue specific expression analysis of *cpAromatase* Isoforms

5' RACE analysis provided first indication that the two *cpAromatase* isoforms are probably expressed in the tissue specific manner. 5' RACE using RACE ready cDNA prepared from the pooled male and female brain RNA from 24th and 25th stage amplified

A



B **GAM tissue**



C **Brain Tissue**

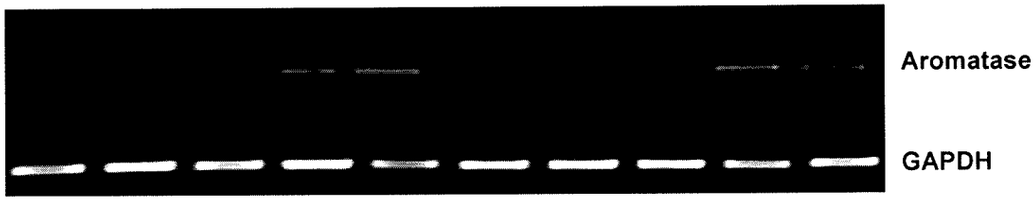


Figure 4.28: Expression analysis of *cpAromatase* by semi-quantitative RT-PCR in different embryonic tissues during TSP. The *cpAromatase* expression was only observed in: **A)** GAM and Brain tissues (for each tissue template cDNA comprised RNA samples pooled over all stages of TSP from the embryos kept at MPT and FPT; **B)** GAM tissues of embryos kept at FPT during 24^{th}-25^{th} developmental stages; and **C)** brain tissue at both male and female promoting temperature without a significant difference in expression levels. Lanes: H, L, K, B, G denote the five tissues- Heart, Liver, Kidney, Brain and GAM, respectively; M21-M25 and F21-F25 represent the five developmental stages 21 to 25 at MPT and FPT, respectively. *GAPDH* was used as internal control.

only one fragment of approximately 700 bp; while a similar RACE reaction with the RNA from GAM tissues revealed two clear amplicons; bigger (>900 bp) and smaller amplicon of approximately 700bp (figure 4.29A and B). The analysis thus suggested that only the smaller isoform *cpAromatase***b** expressed in brain tissue. Absence of longer *cpAromatase***a** isoform in brain was further confirmed by RT-PCR analysis. For this purpose, a primer AroBigF (designed from 156 bp region present only in cpAromatase**a** isoform) was used in combination with AroF2, as these could identify presence/absence of isoform *cpAromatse***a** expression. RT-PCR, done using these primers with male/female pooled cDNA of GAM and brain as template, revealed expression of *cpAromatase***a** only in GAM and not in brain tissue (figure 4.29C). Thus above two experiments, suggest that smaller *cpAromatase***b** isoform expresses only in brain.

Expression of more than one isoforms of *cpAromatase* was also confirmed by Northern hybridization. Total RNA pooled from female GAM tissues of 24th and 25th stages was transferred to a nylon membrane and was hybridized with a radio-labelled 500 bp long *cpAromatase* probe amplified with primers AroF2 and AroR2 that was common to both the isoforms. Northern hybridization revealed presence of two transcripts of about 2 kb size, confirming the presence of two *cpAromatase* isoforms in GAM tissues of *C. palustris* (figure 4.29D).

4.1.3.5.2 Immunolocalization of cpAromatase

Immunolocalization was done using a commercially available rabbit Anti-Aromatase antibody (raised with a peptide conserved in the Aromatase proteins of other mammals, *Gallus* and reptiles) and GAM tissues sections from the embryos at MPT and FPT representing all stages encompassing TSP. Immunolocalization revealed no *cpAromatase* expression in the GAM tissues from the embryos kept at MPT. On the other hand, *cpAromatase* expression was well evident in the GAM tissues from the embryos kept at FPT from stage 23rd onwards, which increased with the progression of development from 23rd to 25th stage resulting in highest expression in the 25th stage of female GAM tissue. Furthermore, these localization studies revealed no trace of the cell type specificity of *cpAromatase* expression otherwise confirms sex specific expression in female GAM. This was well evident as the anti-aromatase antibody stained all the cells of the developing gonads rather evenly till the later stages at FPT (figure 4.30). The Immunolocalization of *cpAromatase* thus confirms the RT-PCR results that have revealed a rapid increase in the *cpAromatase* levels from 23rd stage onwards in the female GAM tissues.

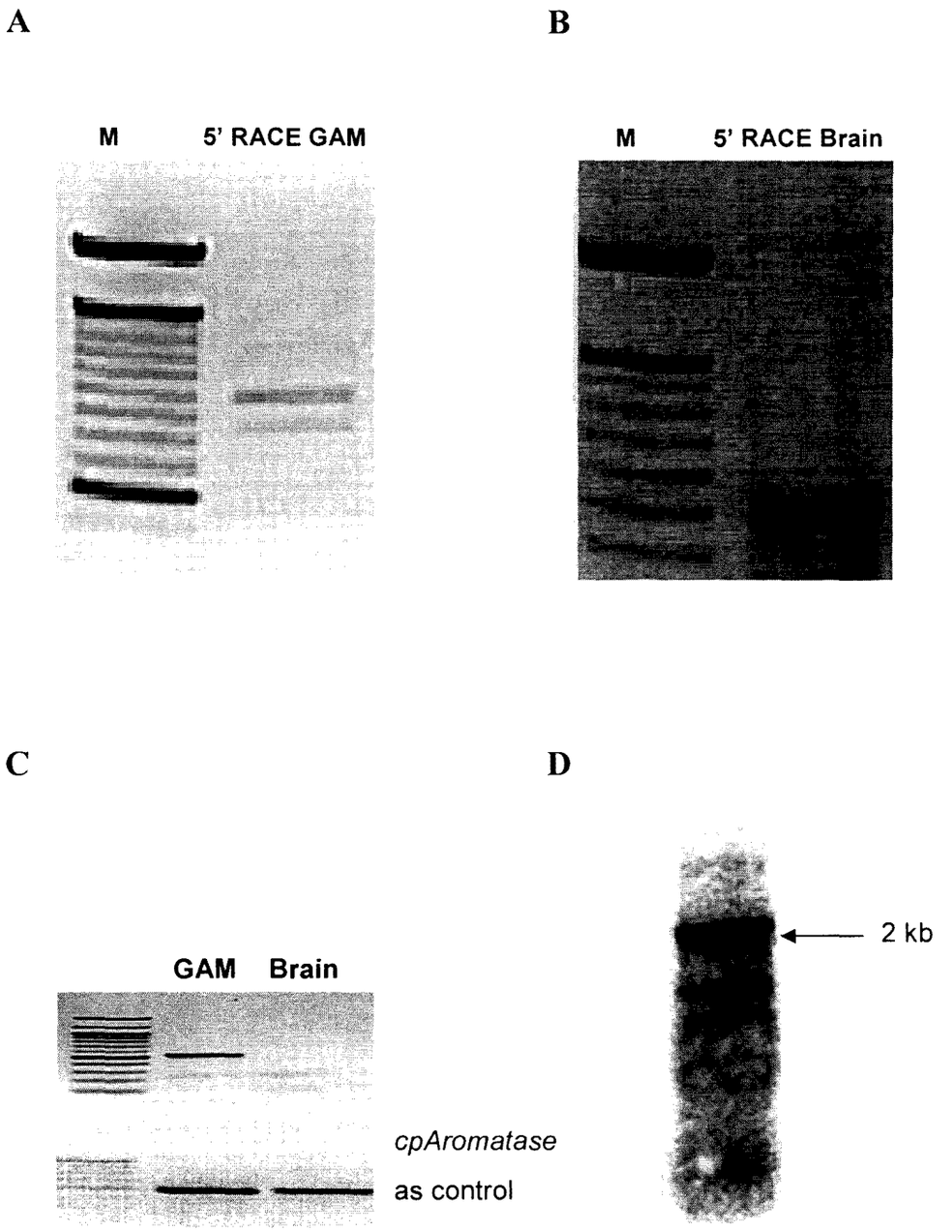


Figure 4.29: 5' RACE and expression analysis of *cpAromatase* in brain and GAM tissues indicating two isoforms and tissue specificity. 5'-RACE analysis indicating: **A)** two isoforms in GAM tissue; **B)** only one isoform in brain similar to the lower sized isoform seen in the GAM; **C)** RT-PCR analysis using primer designed from the region absent in the *cpAromatsae***b** (ArobigF and AroF2) shows amplification only in GAM, indicating *cpAromatsae***a** is expressed only in GAM; and **D)** Northern analysis with the pooled RNA from 24th and 25th stage female GAM tissue indicating presence of two *cpAromatase* isoforms.

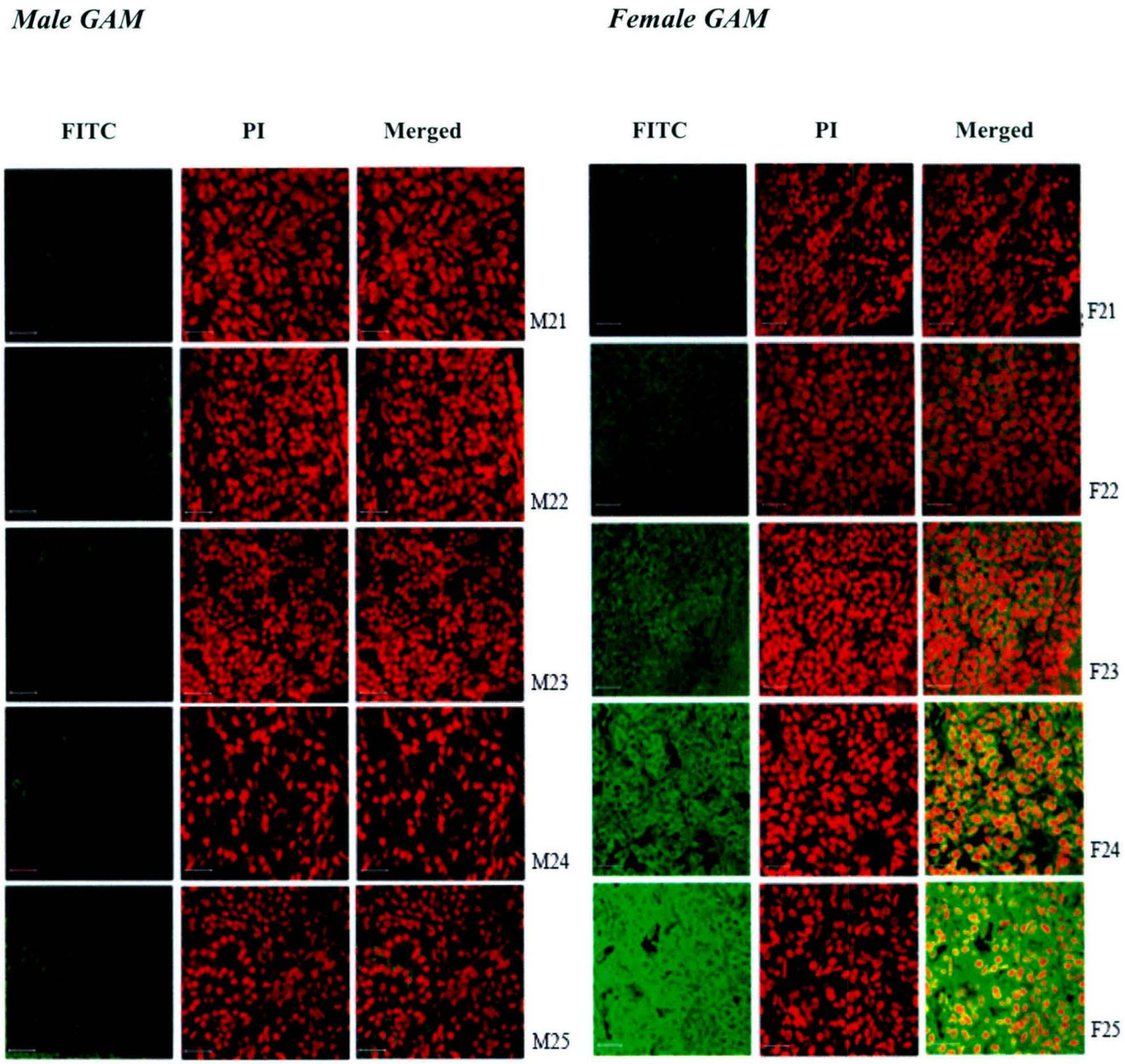


Figure 4.30: **Immnuostaining with anti-Aromatase antibody showing sex specific expression of *cpAromatase* in the developing gonads of embryos kept at MPT and FPT during TSP**. Note *cpAromatase* expression in the gonads at FPT (stage 23rd onwards). In contrast, no expression was seen in the developing gonads of embryos kept at MPT. FITC tagged anti-rabbit secondary antibody was used to follow *cpAromatse* expression and counter staining was done using PI.

4.2 Search for differentially expressed genes having putative role in TSD

As mentioned earlier, the master regulator(s)/genetic machinery that senses the temperature stimuli and decides the fate of a bipotential gonad in TSD remains to be discovered, necessitating search for novel differentially expressed genetic factor(s) involved in TSD. In recent years, a number of diverse techniques have been developed to isolate differentially expressed genes, among which PCR based SSH (Suppressive subtractive hybridization) is one suitable approach that can be efficiently used with meager amounts of target samples (RNA pool), which has been one major consideration in the present study. Therefore, SSH was performed using cDNA from male GAM tissues as 'Tester' and from female GAM as 'Driver'. The analysis revealed a number of putative differentially expressed cDNA fragments, some of which were tested for ascertaining their sex/stage-specific differential expression. One of these putative candidates was studied in detail, and very interestingly it was found to be a novel mRNA like non-coding RNA (mRNA like ncRNA) having an apparent role in male gonad development.

4.2.1 cDNA Synthesis/ cDNA subtraction using SSH approach

RNA was isolated from pooled 21+22 stage GAM tissues of embryos kept at MPT and FPT, separately. The two RNA pools were subjected to PCR based double strand cDNA (see section 3.5.1). In each case about 3-4 µg of double stranded cDNA was obtained that was used for cDNA subtraction.

Double stranded cDNA obtained above was column purified, and about 4 µg cDNA each from driver (female GAM cDNA) and tester (male GAM cDNA) was digested to completion using restriction enzyme RsaI. Adapter 1 and 2R were ligated separately to the digested tester cDNA (however both adapters were ligated together with 100 ng of tester cDNA to make unsubtracted tester control). Subtracted tester cDNA after two rounds of subtraction when compared with control unsubtracted tester cDNA (amplified by adapter specific primers), revealed major difference in band pattern on the gel, which indicated efficient enrichment of differentially expressed genes (figure 4.31).

4.2.2 Cloning, sequencing and BLAST based characterization of subtracted tester (male GAM cDNA) library

Secondary PCR products from subtracted tester cDNA were cloned in pMOS vector; approximately 700 plasmids were isolated from the resulting subtractive cDNA library

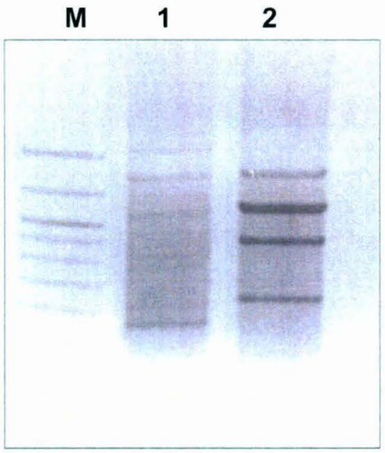


Figure 4.31: **PCR analysis to test the success of the cDNA subtraction.** The unsubtracted tester control and the hybridized cDNA after second round were subjected to PCR with adapter 1 and 2R specific primers (lanes 1, 2 respectively); show a successful enrichment of differentially expressed cDNA in the tester population. Lane M represents 100 bp DNA ladder.

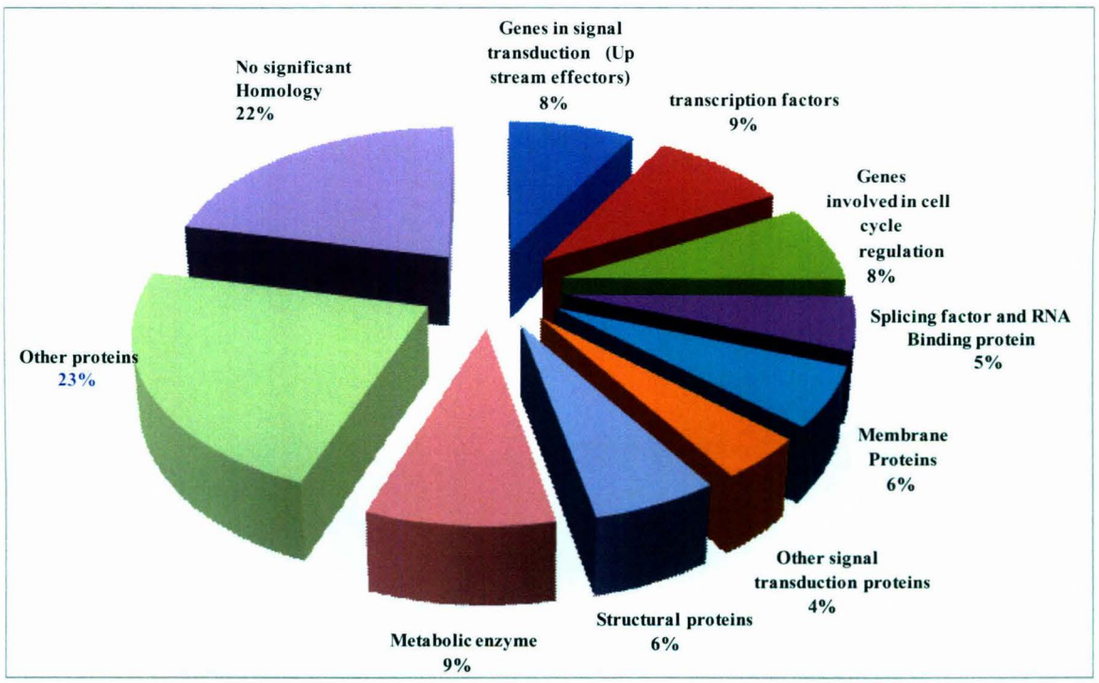


Figure 4.32: **Pie chart showing the functional classification of 167 non-redundant ESTs,** identified from the 700 clones representing the subtractive library generated using SSH approach from the 21st and 22nd stage GAM cDNA of male embryos as tester and female as driver. Note that a sizable fraction (22%) of ESTs does not have any significant similarity with the database sequences.

were sequenced. The insert sequences, edited for the vector and adapter motifs were tested in CLOBB program, which identifies the redundant sequences (makes a cluster of them). A manual check of resulting cluster revealed four clusters comprising splice variants. Analysis revealed a total of 167 unigenes (non-redundant fragments) in the subtractive library, which were checked individually by BLAST for their possible identity to the homologues in the public domain databases. A standalone version of BLAST was used for BLASTn, BLASTx, and tBLSATx against 'nt' and EST database. The analysis revealed a total of 131 (78.4%) sequences (from 167 uniqueness) showed complete or partial similarity to the sequences in the database, while 36 sequences appeared to be novel to crocodile (*i.e.* did not have a BLAST hit). Furthermore, the 131 cDNA sequences that showed similarity to database sequences, comprised functional diversity expected of a developing/differentiation tissue (table 4.2, figure 4.32) with a large proportion of regulatory genes belonging to cell cycle regulators, upstream effectors of signal transduction pathways, splicing factors, RNA binding proteins etc.

4.2.3 Validation of differential expression of selected cDNA clones

The subtracted cDNA clones obtained above, were expected to show a putative male GAM specific expression. Therefore, a preliminary effort was made to validate if the selected cDNAs are differentially expressed in the male GAM. For the purpose, 23 of the 167 non-redundant cDNAs were randomly selected and analyzed by Reverse Northern hybridization (see section 3.5.6.1). These included homologues of genes involved in regulation of gene expression, upstream effectors, important transcription factors, sex related genes and other important genes. Analysis revealed only one cDNA (clone a4r) to be expressed at considerable higher levels in the GAM tissues at MPT (figure 4.33). Reverse Northern was repeated to conform the result. Apart from the clone a4r, slight differences in hybridization signal/expression levels were apparent for few other fragments but were too subtle for any consideration.

Expression levels of 23 non-redundant unique clones of subtractive library were also analyzed by semi-quantitative RT-PCR for their differential expression. The analysis revealed three more cDNA clones (a16, d4 and 2n11), in addition of a4r cDNA (which was identified even in Reverse Northern, figure 4.33), to be differentially expressed *i.e.* having higher expression levels in the GAM of embryos kept at MPT, throughout TSP (figure 4.34). The validation studies thus demonstrated the success of SSH approach used in the study to isolate differentially expressed tester specific cDNAs with a relatively high efficiency of approximately 15-18% (4 of 23) tested clones. Also, the analysis

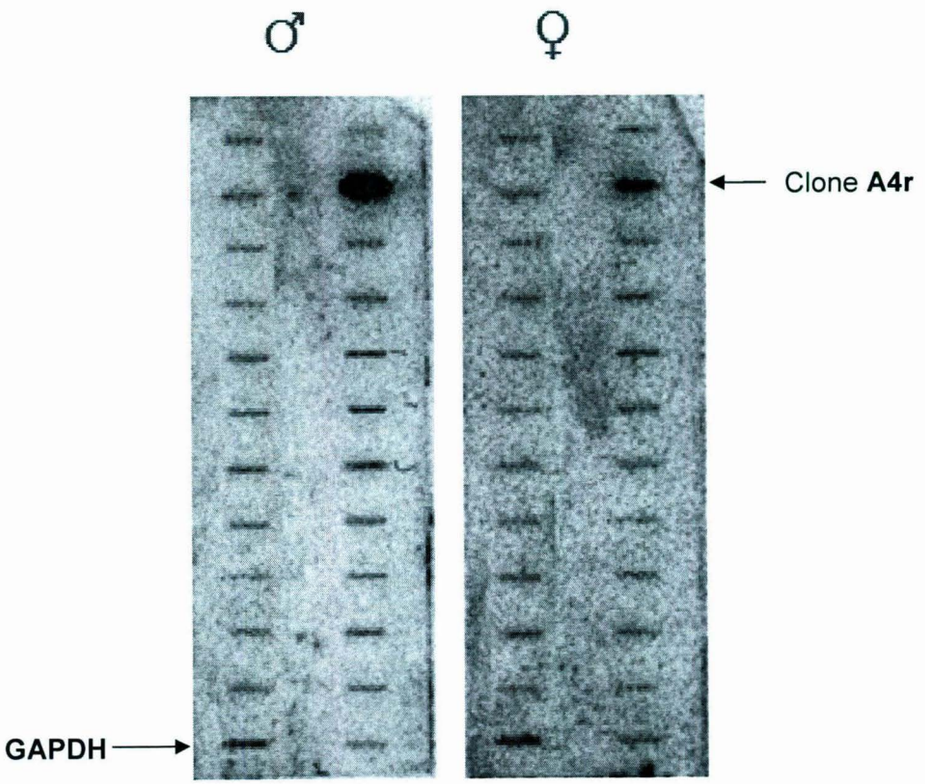


Figure 4.33: **Reverse Northern analysis of 23 randomly picked unique ESTs clones from subtractive library for validation of their sex-specific differential expression.** The DNA from each clone was slot blotted in duplicate and challenged separately with radio-labelled 21st + 22nd pooled cDNA from male and female GAM. Note significant difference in hybridization signal of clone A4r indicating higher/differential expression. *GAPDH* DNA was used as internal control for blotted DNA.

Protein family	Few important indicated gene(s)	Number of clones
Genes in signal transduction (Up stream effectors)	NRAS, ARP1, activin receptor type IIB, NFIA, ras rab proteins	14
transcription factors	NR2F2 (COUP), CAPER, HMGN3, HBOX13	15
Genes involved in cell cycle regulation	Clk1,Clk4, STY protein kinase , Ser-thr kinases	13
Splicing Factors and RNA binding Protein,	CCI.3, CCI.4, RNP1, RRM, RBM2 (Gene found on Y hromosome)	9
Membrane Proteins	Solute carrier family 35 no.35, Solute carrier family 1 no.5,	10
Other signal transduction proteins	Wsb, Notch-regulated ankyrin repeat protein, similar to RAB5 interacting protein 2 of *Gallus gallus*, Tribbles homolog 2 (TRB2) mRNA of *Gallus gallus*	7
Structural proteins	Albumin, Collagen	10
Metabolic enzyme	metR, Tyrosine monooxygenase	15
Other proteins	Gag viral proteins , hypothetical proteins etc.	38
No significant homology		36
	Total	167

Table 4.2: Functional classification of the 167 non-redundant ESTs obtained in the subtractive library generated using male GAM as tester and Female GAM RNA as driver.

revealed that Reverse Northern might not be helpful to highlight genes that may have subtle differences in expression levels.

The four cDNA that could be validated for their differential expression in the male GAM tissue (see above, figure 4.34) comprised; 1) Clone a16: 600 bp in size, showing near complete match to the upstream to nRAS gene of other vertebrates; *nRAS* codes for a protein having multiple cold shock domains. 2) Clone d4 approximately 700 bp in length, completely matching to the RBM (RNA binding motif) group of genes that act as splice factors. 3) Clone 2n11: approximately 600 bp in length, was completely matching to the *clk1* (cdk2 like kinase1) gene of other vertebrates; and 4) Clone a4r: 339 bp in size showing partial homology to the *SLF35F5* (solute carrier 35 family F member 5) gene. This clone was selected as a putative new candidate TSD related gene, and thus studied in detail in the present study (see section 4.2.5).

4.2.4 *In-silico* analysis of the subtracted cDNA clones to predict sex-/tissue-specific expression

In the present study, we could experimentally test only 23 cDNA clones (see above), as expression profiling based validation of each of subtracted cDNA clone for its differential expression was practically not feasible due to severe limitation of the source material (crocodile embryos/ tissues at MPT and FPT). Therefore, an *in-silico* approach was tried to have some idea about the putative expression specificity of the subtracted cDNA clones by sequence comparison with the known tissue-/sex-/stage- specific databases (downloaded from SANBI, South African national Bioinformatics institute) for various tissues namely, Heart, Liver, Gonads, adipose, Brain and Reproductive tissues. The preliminary analysis revealed many subtracted cDNA clones (of crocodile from present study) to share homology with genes having tissue-/sex-/stage- expression. Interestingly, a good proportion of such cDNAs were indicated to show similarity with the genes having brain and gonad specific expression. Furthermore, some of the subtracted cDNA clones obtained in this study were found to be complex in their homology, as these matched to two different genes from different databases or part of two different genes from different databases, indicating possible domain sharing. Three interesting examples of the subtracted cDNA clones are briefed below.

4.2.4.1 cDNA clone A3r

This 280 bp long crocodile cDNA clone shows homology with activin receptor type II B in BLASTn along its 115-209 (94 bp) bp region. Interestingly, this coordinate is seen in the

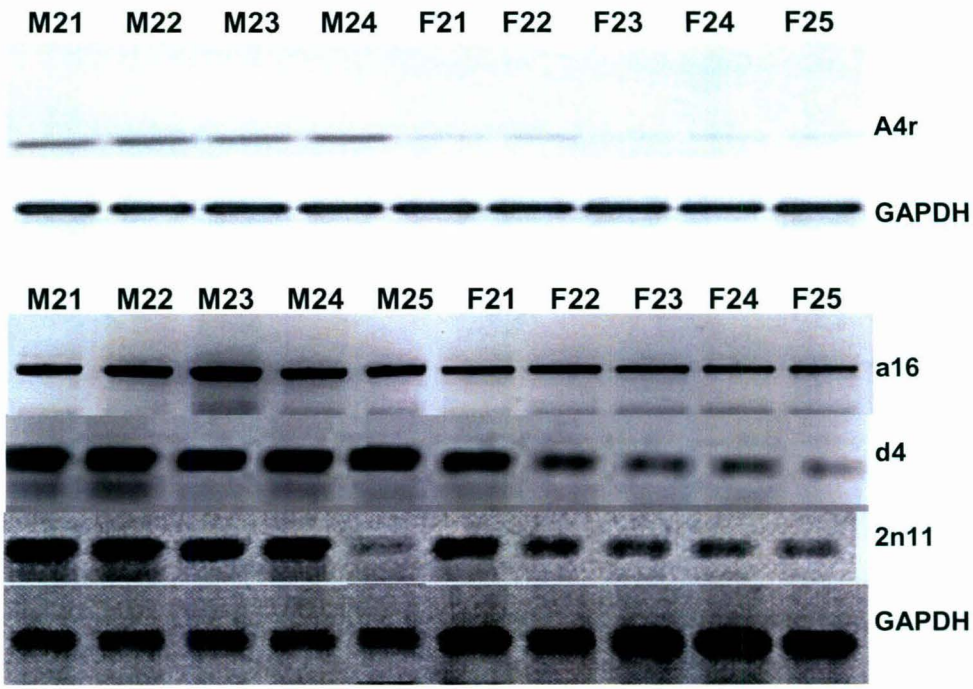


Figure 4.34: **Semi-quantitative RT-PCR based validation of few of the ESTs identified from the subtractive library (also screened in the reverse Northern analysis) for their differential expression in male and female GAM tissues through TSP.** Note higher expression in for EST clones A4r, d4 and 2n11, and to a lower extent for a16 male GAM through TSP. M21 - M25 and F21 to F25 represent 21st to 25th developmental stages at MPT and FPT respectively. *GAPDH* was used as internal control.

Figure 4.35: **Schematic representation of BLAST based homology of a4r EST clone.** Note that only 160 bp of the 339 bp long clone shows homology to the solute carrier family 35 group F member 5 (*SLF35F5*) gene and the flanking regions did not show any homology to the database sequence.

indicated homologus genes of the reproductive database and did not show any similarity to sequences from other tissue specific databases.

4.2.4.2 cDNA clone f15

The cDNA 2f15 (435 bp) showed homology to a Tribble like protein of drosophila and *Gallus* along its 196-402 bp region in BLASTn. Furthermore, the same region showed similarity to Tribble like protein present in the reproductive database only and not with sequences from other databases. Tribble like protein is a regulatory protein of MAP kinase pathway, and is reported to have a role in early embryonic development.

4.2.4.3 cDNA clone B5f

A stretch of 201 bp (120-320bp) of subtractive library clone b5f (312 bp) showed homology to NR2F2 (nuclear receptor subfamily 2, group F, member 2) gene of human and mouse. Same co-ordinates also showed similarity to the brain database gene sequence FLN29 in addition to NR2F2, and even more surprisingly, similarity with PBX/knotted 1 homeobox 2 (PKNOX2) gene from reproductive database. Expression of NR2F2 gene has not been reported in gonad of any species. It is possible that the cDNA may represent the part of crocodilian homologue of FLN29/NR2F2 with expression in GAM, or may be a novel gene just having a conserved domain shared by these genes. Moreover, the matching cDNA region defines the ligand binding domain of FLN29/NR2F2, that act in response to small steroid molecules like thyroid hormone, sex steroids etc.

4.2.5 Identification, characterization and expression analysis of a novel mRNA line non-coding RNA gene of crocodile

One of the 339 bp subtracted cDNA clone 'a4r' was validated to be expressing at considerable higher levels in male GAM tissues (see section 4.2.3). BLAST homology revealed it to share high homology with *SLF35F5* gene along its 39-189 bp stretch (figure 4.35) and no similarity to any database sequences for its two flanks. This novel and differentially expressed cDNA clone obtained in the SSH global screen was thus selected for detailed studies by isolation and characterization of its coding region. The effort led to identification of a novel, male specific mRNA like non-coding RNA (mRNA like ncRNA) candidate gene having possible role in TSD. The results are described below.

4.2.5.1 Tissue Specific Expression of cDNA clone A4r

Expression analysis of a4r was followed by semi-quantitative RT-PCR using primers SC35F and SC35R (only primer SC35F was designed from the 160 bp region matching to *SLF35F5* gene) and cDNA of five embryonic tissues of male and female embryos (pooled over TSP *i.e.* 21-25developmental stages) as template. The analysis revealed almost comparable expression of 339 bp a4r, between MPT and FPT, in embryonic, heart, liver and kidney. In contrast, its expression in the GAM tissues was much higher at MPT than that at FPT. Moreover, its expression in GAM tissues at MPT was highest compared to any other tissue (figure 4.36A). Subsequently, the temporal expression profiling of clone a4r by semi-quantitative RT-PCR, revealed its higher expression in the GAM tissues of the embryos kept at MPT compared to those at FPT all through TSP (figure 4.36B). A parallel real-time RT-PCR study revealed expression of a4r to be 17-23 folds higher in the GAM tissues at MPT through TSP (figure 4.36C).

4.2.5.2 RACE analysis to isolate the full-length cDNA of *SLF35F5* like a4r clone

To pull out complete transcript, 5' and 3' RACE were performed using primers SC35R, SC35R2 (nested) for 3' RACE; and SC35F2 and, SC35F3 (nested) for 5' RACE. Interestingly, no amplification was observed in the primary PCR of 5'/3' RACE and even during the secondary PCR before 18 cycles. This suggests that a4r transcript represents a relatively poorly expressed gene; as in RACE reaction for an abundantly expressing gene, amplification is generally seen after only 11-13 cycles. Furthermore, no difference was seen in the amplification profiles obtained from the RACE ready cDNA made from putative male and female GAM (figure 4.37A). RACE products after cloning and sequencing resulted in a 2552 bp long sequence having a poly A tail in the end. BLAST based homology analysis indicated that the 2.5 kb full length transcript did not share significant homology to any database sequence except original 160 bp sequence match to *SLF35F5* gene.

Primers SCINTF and SCINTR were thus designed to amplify a larger 1700 bp part of complete 2552 bp cDNA, to test the fidelity (especially, because the isolated sequence did not represent the crocodile homologue of the *SLF35F5* gene). Expression analysis of the bigger fragment was carried out using RT products of the GAM tissue RNA as template, which revealed its higher expression at the MPT (figure 4.37B).

Presence of the full length transcript was further confirmed by Northern hybridization. For the purpose, a Northern blot prepared from male and female GAM RNA (each pooled

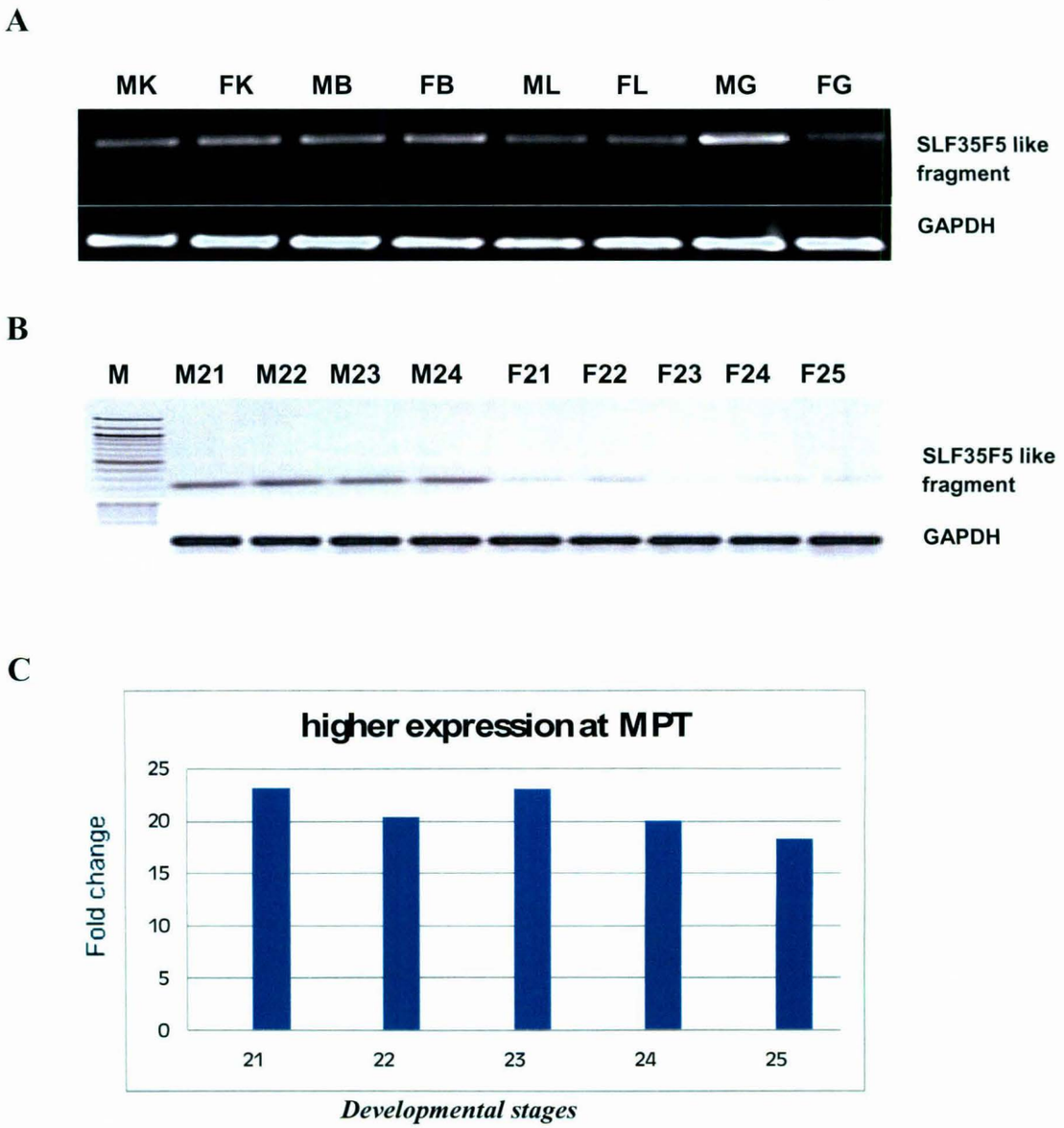


Figure 4.36: **Expression analysis of *SLF35F5* like sequence for ascertaining tissue and sex specific expression**. **A)** Semi-quantitative RT-PCR indicates male sex specific higher expression of the *SLF35F5* like transcript only in GAM tissues. Lanes MK, MB, ML, MG and FK, FB, FL, FG represent Kidney, Brain, Liver and GAM tissue from the embryos kept at MPT and FPT, respectively. *GAPDH* was used as internal control; **B)** semi-quantitative RT-PCR; and **C**) real-time RT-PCR confirms its male GAM specific expression to be 20-23 fold higher compared to its expression levels in GAM of embryos at FPT.

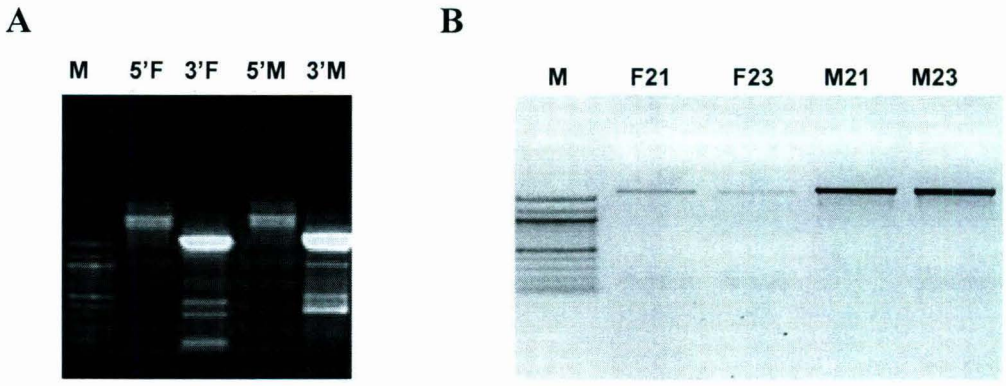


Figure 4.37: **RACE analysis and RT-PCR based conformation of the resulting fragment for differential expression. A)** 5' and 3' RACE with male and female GAM RACE ready cDNA using primers designed from *SLF35F5* like transcript. 5'F and 3'F and 5'M, 3'M represent female and male specific 5' and 3' RACE ready cDNA, respectively; and **B)** RACE extended longer sequence of the *SLF35F5* like transcript was amplified by RT-PCR to confirm its sex specific expression in male GAM. Lanes: M21-23 and F21-23 represent developmental stages 21 to 23 at MPT and FPT, respectively, M represents 100 bp DNA ladder marker.

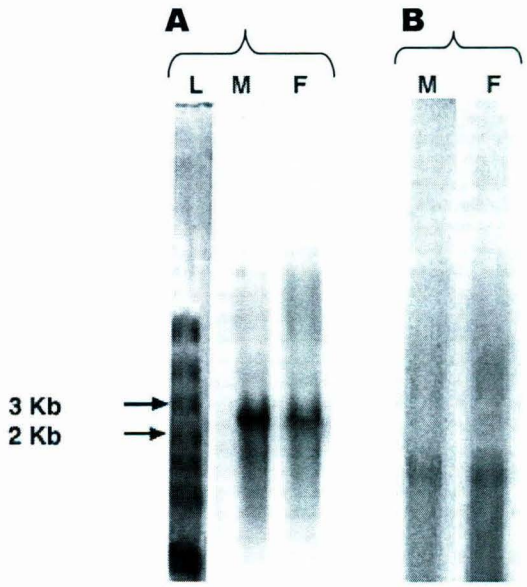


Figure 4.38: **Northern analysis of the *SLF35F5* like transcript in male and female GAM tissues. A)** RNA blot hybridized with the probe prepared from the unique region of *SLF35F5* like EST clone a4r. Note relatively higher expression in male GAM, and the size of the hybridized transcript to be approx. 2.7 kb; and **B)** blot in 'A' after stripping and rehybridization with *GAPDH* showing comparable signal in both M and F lanes. L represents RNA ladder whereas M and F represent male and female GAM RNA (pooled over 21st to 23rd stage), respectively.

from all developmental stages spanning TSP), was hybridized with 600 bp long cDNA prepared/labeled using SC35FN and SCINT primers (section 3.3.8). The hybridized blot revealed a transcript of nearly 2.7 kb, expression level of which was significantly higher in the GAM tissue from MPT in comparison to the same from FPT (figure 4.38). Thus RT-PCR, as well as, Northern analysis, conclusively demonstrated that the 2552 bp complete cDNA obtained above was not an RACE artifact, and represent a novel gene.

4.2.5.3 *SLF35F5* like mRNA is transcribed from both strands

The full length transcript obtained by RACE (see above) was neither the crocodilian homologue of *SLF35F5* nor it matched to any database sequences. Hence RACE was repeated with other primer combinations also, RACE products were sequenced every time to check the reproducibility and specificity of amplification. In addition, control RACE reactions were carried out to check the specificity of the primers (wherein primer(s) proceeding towards 5' end should work only with the 5' RACE ready cDNA, similarly another primer traversing towards 3' end should only work with the 3' RACE ready cDNA). Bur contrary to the expected, primers designed for 3' RACE were able to generate amplicons with both 5' and 3' RACE ready cDNA and the same was observed for the primers designed for the 5' RACE. End-specific RACE primers always, amplified in both directions with male and female RACE ready cDNAs (figure 4.39) even cycling parameters of RACE PCR were made stringent enough to effectively prevent amplification of false positives. End specific primers designed from the other regions of the 2552 bp fragment *i.e.* primer combinations SCINTR/SC35FN and SCINTF/SC35RN were also tested. The products amplified in all above RACE reactions were cloned and 300 clones in this effort were sequenced, interestingly, which revealed the presence of two near complimentary sense and antisense transcripts of 2552 bp and 2685 bp, respectively. Both the transcripts shared 2249 bp overlapping region and have their unique non-overlapping 3' ends, which were 313 bp and 436 bp long for the sense and antisense transcripts, respectively (schematic representation in figure 4.40). Presence of sense and antisense transcript was subsequently validated by RT-PCR, Northern analysis and real-time RT-PCR (see below).

4.2.5.3.1 Validation of sense/antisense transcript using strand Specific RT-PCR

Control RT-PCR was run to ascertain the efficacy of DNAase treatment and to rule out possibility of DNA (figure 4.41A). Subsequently, four cDNA templates were synthesized from RNA of GAM of embryos kept at MPT and FPT, separately with forward and

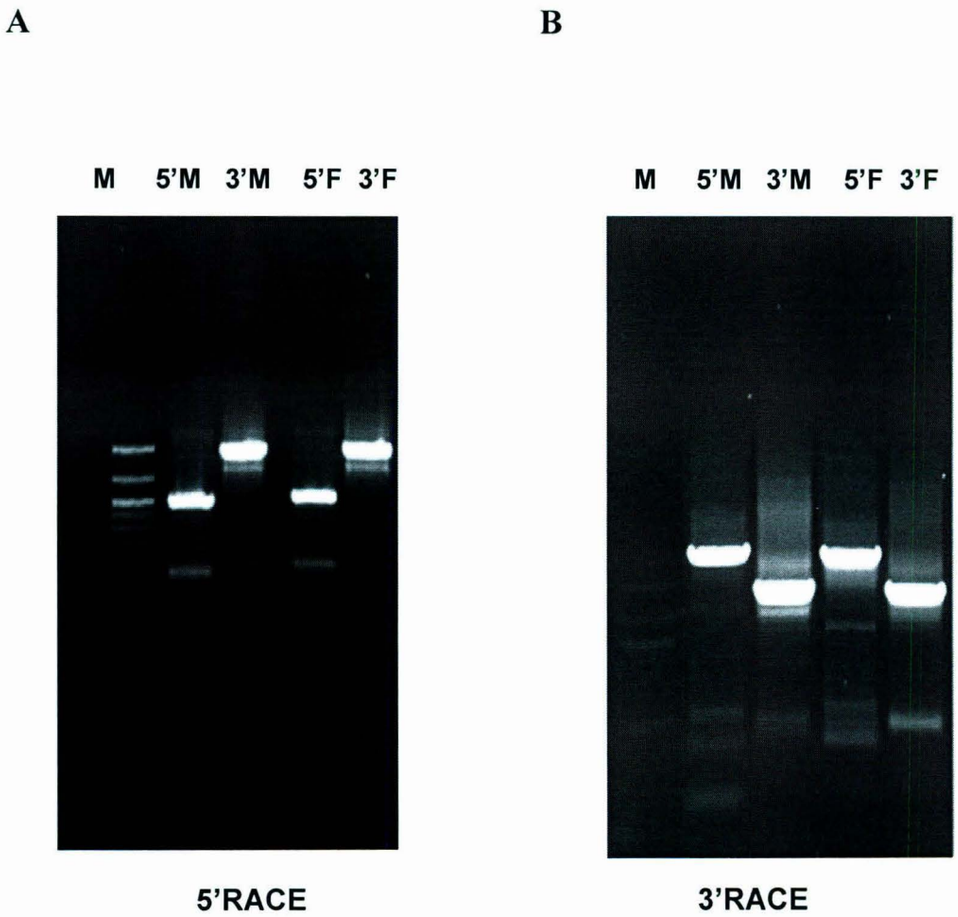


Figure 4.39: **5' and 3' RACE with male and female RACE ready cDNA to isolate sense and antisense transcripts sequences of the *SLF35F5* like transcript.** Primers used for 5' RACE were also used for 3' RACE and *vice versa* (as a 5' RACE primer for sense strand can amplify the 3' end of antisense transcript with 3' RACE-ready cDNA and *vice-versa*). **A)** 5' RACE reactions showing amplified 5' ends of sense and antisense transcripts; and **B)** 3' RACE leading to 3' end of the sense and antisense transcripts. Note: in each case amplified ends of sense/antisense transcript were of equal comparable size independent of source template from male and female GAM. 5F, 3'F and 5'M, 3'M represent female and male specific 5' and 3' RACE ready cDNAs, and M represents 100bp DNA marker lane.

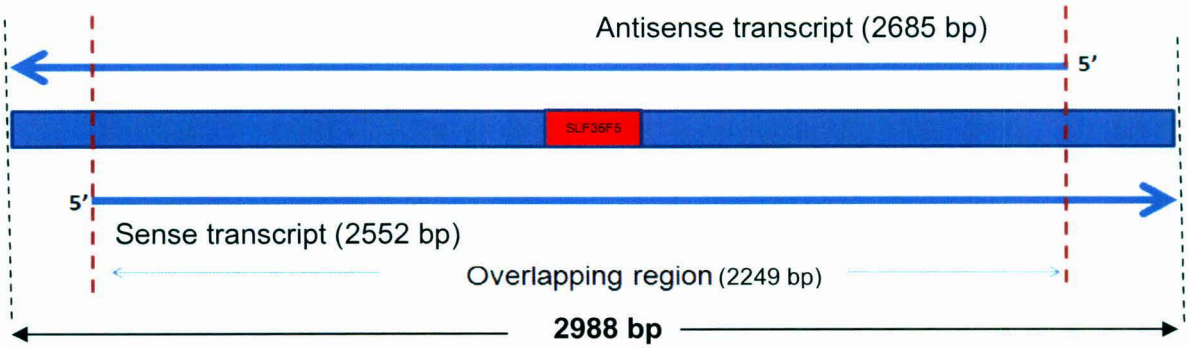


Figure 4.40: **Schematic representation of the sense and antisense transcripts** resulting from *SLF35F5* like ncRNA coding region, isolated in the study.

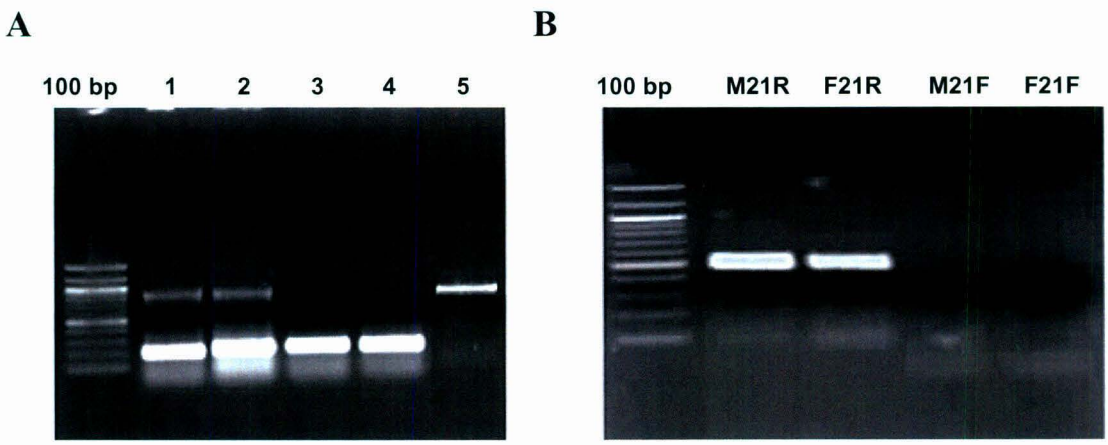


Figure 4.41: **RT-PCR analysis to ascertain purity of RNA samples (absence of contaminating genomic DNA) using *cpWt1* gene (as control) specific amplification. A)** RT-PCR with male and female GAM untreated RNA (lanes 1-2), or treated (lanes 3-4) with DNase, using primers Wt1BF1 and Wt1bigR while lane 5 represents genomic PCR with same primers. Note that lane 3, 4 (DNAse treated RT products) do not show amplification resulting from the genomic DNA; and **B)** RT-PCR to show lack of antisense transcript for *cpWt1*. Note *cpWt1* specific RT-PCR products (using WtF1/R1 primers) only in M21R and F21R, wherein the template comprised RT products of male and female GAM of 21st stage respectively, generated with Wt1R1. In contrast, no amplification was seen in lanes M21F and F21F where the template comprised RT products generated using 21st stage RNA of male and female GAM with Wt1F1 (which would have worked if antisense transcript of *cpWt1* was available).

reverse primers for the *SLF35F5* like mRNA, as well as, and *cpWt1***a** isoform used as control; primers used were SC35INTF/SC35INTR and Wt1F1/Wt1R1 (section 3.5.6.5). cDNA template made with WT1 primers, were subjected to PCR with Wt1F1 and Wt1R1 primers. Amplification was observed only in cDNA template primed with Wt1r1 (reverse primer) and none with cDNA template made using forward primer Wt1F1 (figure 4.41B), indicating absence of antisense transcript for *cpWt1* in cDNA pool. In the contrary, when cDNA templates made by SC35INTF or SC35INTR primers were subjected to PCR with SC35F/SC35R and SC35INTF/SC35INTR primers pairs separately, amplification was seen in each of the cDNA template irrespective of the forward and reverse primers of *SLF35F5* like sequence (figure 4.42). The results thus strongly suggested that *SLF35F5* like mRNA has both sense and antisense transcripts. The analysis thus further revealed that the sense transcript (relative to the mRNA orientation of the *SLF35F5* sequence) expression levels compared were considerably lower than that of antisense transcript (lanes M21F, F21F, M21R and F21R; figure 4.42).

4.2.5.3.2 Validation of sense/antisense transcripts by Northern hybridization

Primers T7SC35F and T7SF35R having the T7 core promoter region at the 5' ends were used separately for *in-vitro* transcription to generate sense and antisense strand specific probes (section 3.5.6.3). Two separate Northern blots made with approximately equal quantities of RNA from GAM tissues (of embryos kept at MPT and FPT) were individually hybridized with strand specific probes. A transcript of approximately 2.7 kb was detected after hybridization with the antisense probe with the higher signal intensity in the male GAM (figure 4.43A); it may be noted that a similar size transcript was detected by the cDNA probe (figure 4.38). In contrast the sense transcript specific probe did not reveal any signal for the same exposure period (figure 4.43B). As a control for RNA loading, blots probed with the sense specific probe was stripped and reprobed with GAPDH specific probe that revealed comparable signals in both male/female GAM lane (figure 4.43C), suggesting that nearly equal quantity of RNA was loaded in each lane and there was no degradation of the RNA loaded.

4.2.5.3.3 Real time PCR analysis of sense and antisense transcripts in GAM

Failure of detection of the sense transcript in Northern blotting could be because of low level of expression, which was indicated by strand specific RT-PCR. Primers specific to the sense and antisense transcripts designed from their non-overlapping 3' ends (*viz.*, senseR/SCINT1R for the sense transcript and antisenseF/SCINT1F for the antisense

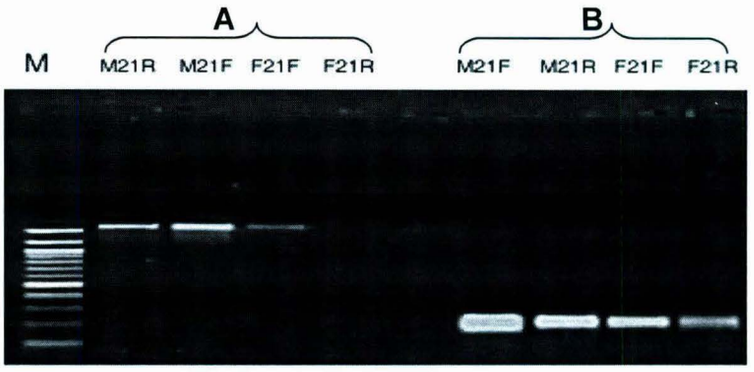


Figure 4.42: **Strand specific RT-PCR to detect transcription from both strands of *SLF35F5* like ncRNA gene**. Analysis was done using M21R, F21R and M21F, F21F as template, which represent strand specific RT products of 21st stage RNA from male and female GAM tissues using strand specific primers SCINTR1 (sense) and SCINTF (antisense), respectively. RT-PCR products obtained using primers: **A)** SCINTF/R that define 1700 bases on ncRNA; and **B)** SC35F/R which span 160 bp core of ncRNA. Note that: 1) amplification in each lane indicating transcription from both strands of ncRNA; 2) much higher expression of antisense strand (from M21F/F21F template compared to M21R/F21R); and 3) significantly more expression in lanes with male GAM cDNA.

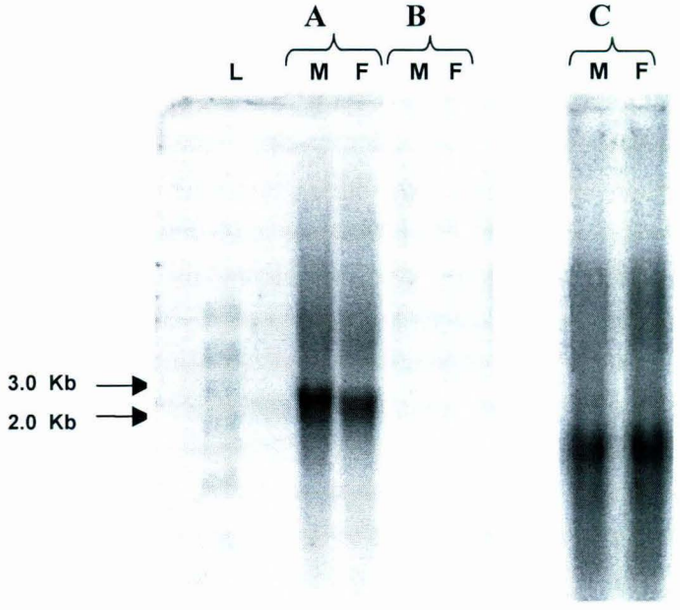


Figure 4.43: **Northern hybridization to ascertain expression of sense and antisense strands of *SLF35F5* like ncRNA gene in GAM tissues. A)** RNA blot hybridized with probe specific to antisense transcript; **B)** similar blot hybridized with probe specific to sense transcript; and **C)** blot in 'B' rehybridized (after stripping) to *GAPDH* probes (for loading control). Note: hybridization in 'A' showing antisense transcripts with abundance in male GAM; no signal in 'B' suggesting undetectable expression of sense transcript for same exposure time; and comparable signals in 'C' suggesting intact and equal amount of RNA loaded in M/F lanes. Lanes: L represents RNA ladder, while M and F represents male and female GAM RNA (pooled over 21st to 23rd stages), respectively.

transcript) were used for real-time RT-PCR that was done in triplicate and repeated twice with two sets of GAM cDNAs. The analysis robustly revealed that antisense transcript was expressed at much higher levels compared to that of sense transcript (figure 4.44). Relative quantification of expression levels exhibited, antisense transcript to be 35-40 times higher than sense transcript in GAM at MPT; in contrast antisense transcript was 3-4 folds more than sense transcript in GAM tissues at FPT (figure 4.44A). On the other hand, relative expression levels of antisense transcript between MPT and FPT varied from 17 to 23 folds in the GAM tissue through TSP (figure 4.44B), where as expression levels of sense transcripts between MPT and FPT changed only 2 to 2.5 times (figure 4.44C). The analysis thus unequivocally confirms that both strands of *SLF35F5* ncRNA locus are transcribed and the transcription from the antisense strand is much higher in GAM at MPT through TSP. Real time RT-PCR results also revealed that the sense strand is transcribed at very low levels, which made it difficult to be detected by Northern hybridization.

4.2.5.4 Determination of crocodilian homologue of genuine *SLF35F5*

Experiments were undertaken to demonstrate that there exists a genuine crocodilian homologue of *SLF35F5* like gene, which is not related to full length cDNA isolated in the present study, except for 160 bp region shared homology. For the purpose partial 1450 bp sequence of crocodilian homologue of the *SLF35F5* like gene was amplified/isolated by 3' RACE using primers designed from 160 bp long *SLF35F5* matching core of a4r cDNA clone (figure 4.35). The partial crocodilian *SLF35F5* revealed it to be highly conserved sharing up to 96% similarity to *SLF35F5* gene from the other vertebrates including mammals. Subsequently primers SCGF and SCGR were designed to analyze crocodilian *SLF35F5* expression in the GAM tissues at MPT and FPT through TSP. Real-time RT-PCR with above primers indicate that *cpSLF35F5* expressed 2-3 fold more in the female GAM tissues (FPT) during TSP (figure 4.45).

4.2.5.5 *SLF35F5* like transcript represents a non-coding RNA

BLAST analysis revealed that both the sense and antisense transcripts had no significant homology with any sequence in database (like EST, WGS and environmental), except for the 160 bp region which showed similarity to the *SLF35F5*. Moreover, both sense and the antisense transcripts did not yield any putative protein translation as a large number of stop codons (40-50) were present all along the length in each of the possible three reading frames (schematic representation in figure 4.46). The

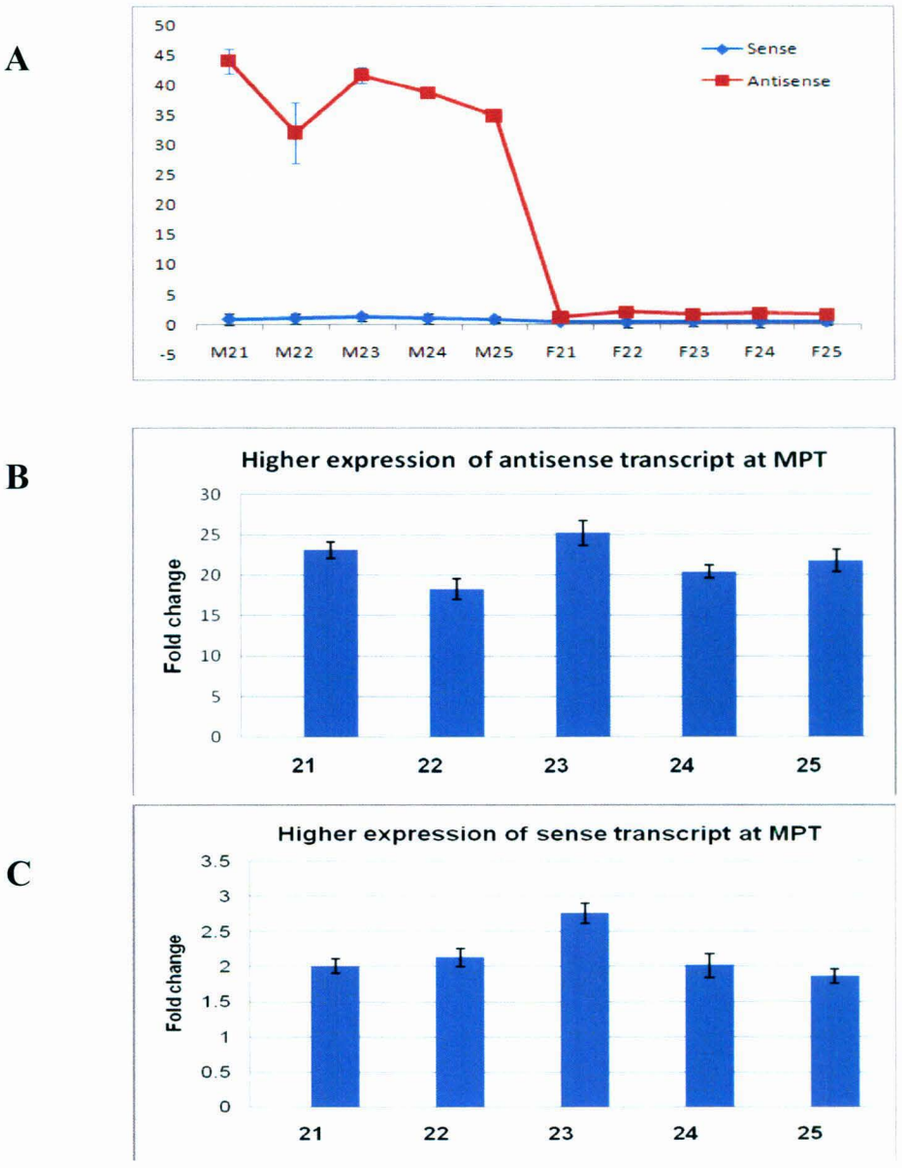


Figure 4.44: **Real-time RT-PCR analysis for expression of sense and antisense strands of the *SLF35F5* like ncRNA in the male and female GAM tissues through TSP using strand specific primer pairs**: senseR/SCINT1R (for sense) and antisenseF/SCINT1F (for antisense), respectively. The X-axis represents developmental stages and Y-axis represents fold change in expression. Analysis revealed; **A)** significant higher expression of antisense strand and almost basal expression of sense strand in male GAM tissues through TSP. Relative quantification of expression showed; and **B)** 19-25 fold higher expression of antisense transcript, and; **C)** 1.8 to 2.8 fold higher expression of sense strand, in the male GAM compared to female GAM through TSP. M21 - 25 and F21 - F25 represent male and female embryos of 21st to 25th developmental stages.

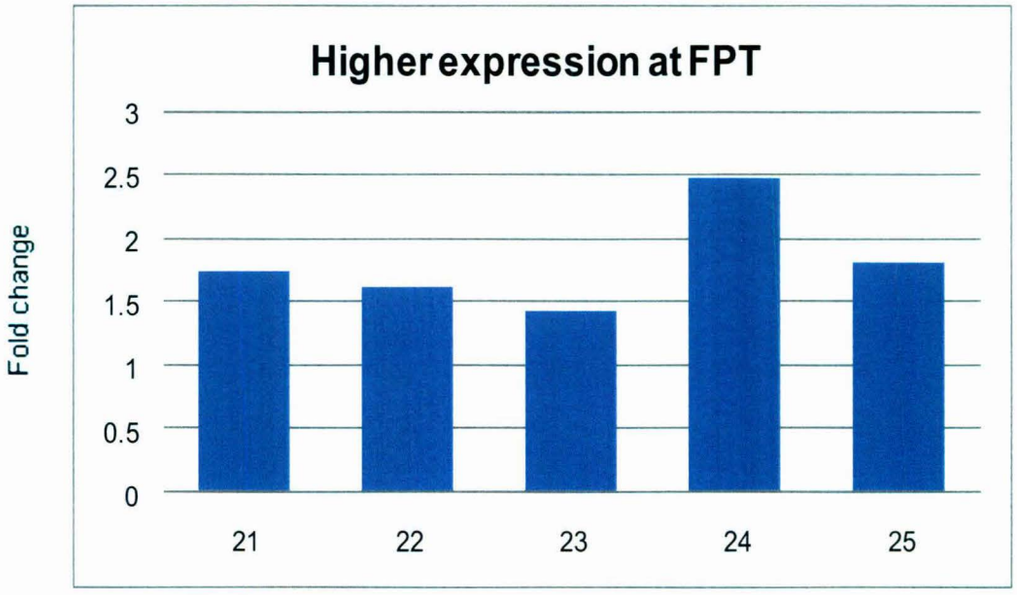


Figure 4.45: A real-time RT-PCR to analyze expression levels of genuine *SLF35F5* transcript in male and female GAM tissues during TSP with SCGF and SCGR primers. Note slightly higher expression in the female GAM tissue in comparison to male GAM.

Figure 4.46: **Predicted protein translations of the sense and antisense stands of *SLF35F5* like novel ncRNA gene of crocodile isolated in the present study.** Predicted translations in all three possible reading frames of the: **A)** sense transcript; and **B)** antisense transcript. Note a large number of stop codons in each of the six predicted translations.

biggest possible ORF (coding around 60 amino acids) was spanned the region that matched to the *SLF35F5* gene in both the sense and the antisense transcripts. All these features of two strands strongly suggested, these to be 'mRNA like non coding RNA', which are characterized by: very low sequence conservation across species, splicing, normal polyadenylated mRNA, larger exons with smaller introns or a single exon without protein coding ability owing to presence of a large number of stop codons.

4.2.5.6 Putative coding locus *of SLF35F5* like ncRNA transcript

Genomic PCR and genome walking approaches were attempted to identify the possible coding region of the *SLF35F5* like ncRNA transcript. Genomic PCR using primers designed from the ends of the sense and antisense ncRNA transcripts (FE_SEN/ RE_SEN and FE_ANSEN/ RE_ANSEN) resulted into amplicons similar in size to the respective transcripts. Sequencing of the amplicons revealed these to be identical to the sense/antisense transcript, suggesting that there are no intron(s) in the ncRNA coding locus.

In a separate study, genomic PCR done using primers SC35F and SCGR (designed from *cpSLF35F5* mRNA), one from the region that matched to the 160 bp region of *SLF35F5* like ncRNA and ~500 bp downstream region, 1200 bp region was amplified that on sequencing revealed an intronic region, which had no similarity to the *SLF35F5* like ncRNA. Similarly, more than one kb of flanking regions of ncRNA locus obtained using genome walking, did not match with the coding region *cpSLF35F5* locus.

4.2.5.7 Localization of *SLF35F5* like ncRNA in the GAM tissues of Putative male and female embryos

Localization of the novel ncRNA was monitored in the genital ridge and mesonephros regions. A fluorescent labeled, cDNA probe, designed from the non-homologus region of the non-coding RNA (to avoid picking any signals from the *SLF35F5* transcripts), was taken instead of strand specific mRNA probes (keeping in consideration the expression data that revealed bulk of ncRNA expression only from the antisense strand).

Localization signals of the ncRNA were revealed to be significantly high in the developing gonads of the male embryos (from MPT), compared to the signals from the developing gonads of the female embryos (from FPT) throughout the TSP (figure 4.47). Further, the ncRNA localization signals was seen to be coming from all cells of the developing gonads during early developmental (21st and 22nd) stages; but appear to

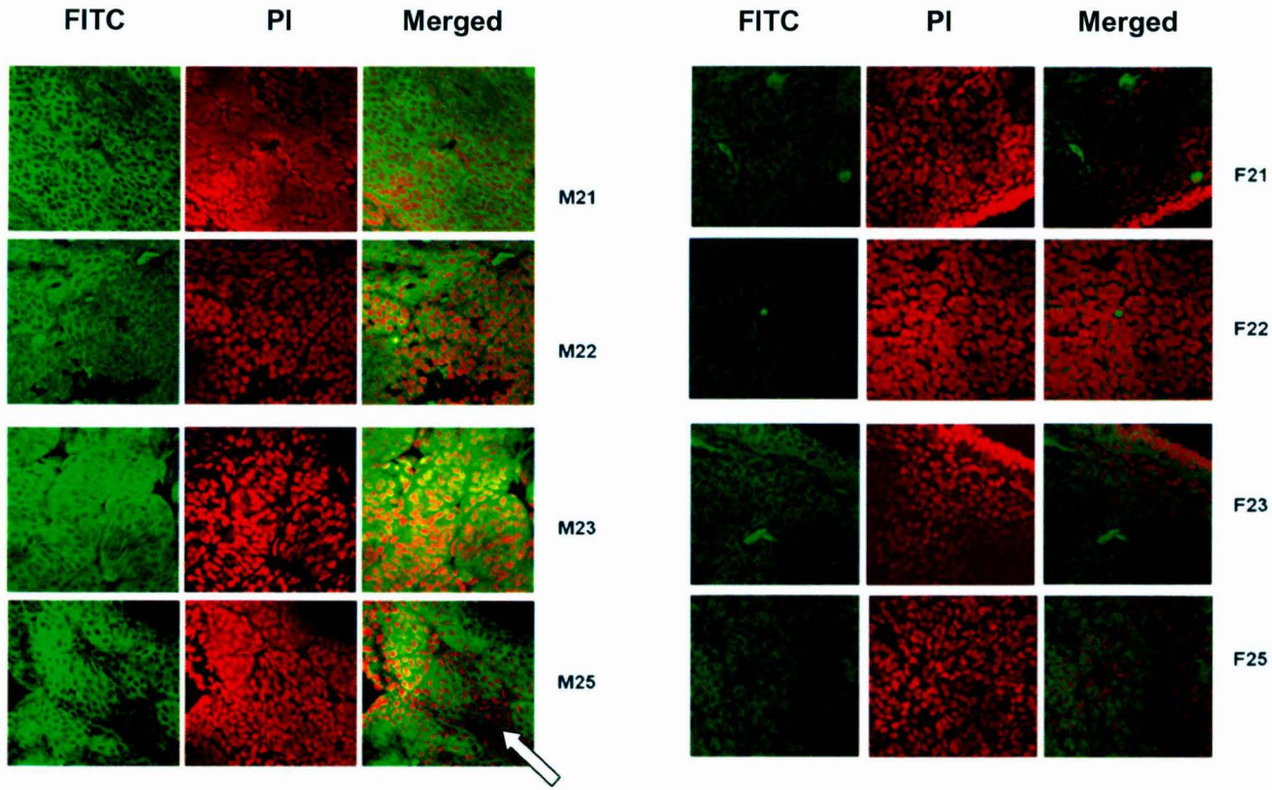


Figure 4.47: ***In-situ* hybridization showing localization of *SLF35F5* like *ncRNA* in the developing gonads of the embryos kept at MPT and FPT through TSP using FITC labelled probe** (amplified with primers SC35RN and SCINTR). Note higher expression in the male gonads, with specific localization in the testicular cords from 23rd stage onwards, correlating with thickening of cords and development of interstitium (by 25th stage, represented by arrow). In comparison, ncRNA localization remains low in developing female gonads and localization does not indicate development of female gonads. M21 to 25 and F21 to F25 represent 21st to 25th developmental stages of male and female embryos, respectively.

follow a sex-/cell type- specific pattern afterwards *i.e.* 22nd stage onwards till 25th stage that defined TSP. Most interestingly the analysis revealed an apparent correlation between the ncRNA localization and important cellular events occurring during the male gonad development, especially from 22nd stage onwards till 25th stage. ncRNA localization was observed to be cytoplasmic, increasing with the developmental stages only in male gonads, faithfully tracking the appearance/development of testicular cords (in 23rd stage), and not any other cell type like interstitium that develops between the testicular cords and traceable by PI staining in 24-25th stage. In contrast, no such cellular event/patterning was revealed by ncRNA localization in the developing gonads of the embryos at FPT, wherein differences (i.e. development of cortical region) were otherwise traceable by PI staining during later stages (25th stage).

Localization studies suggest that ncRNA may have an important role during early determination/differentiation of male gonads that involves migration/differentiation of pre-Sertoli cells, Sertoli cells and germ cells in the sex/testicular cords (all apparently have ncRNA localization).

Compared to the developing gonadal tissues, ncRNA localization did not show any significant difference(s), either in signal strength or in localization pattern, in the mesonephric region of male and female GAM through TSP (figure 4.48). In general, ncRNA signals were spread over all the cells of the mesonephric tubules, albeit at low levels compared to developing gonad tissue; and there existed no difference in its signals in comparable male and female sections of mesonephros.

4.2.5.8 Secondary structure analysis

Secondary structure analysis for the *SLF35F5* like non coding RNA was done using MFOLD v2.3 to have an idea about the RNA folding at MPT and FPT; and to identify conserved miRNA precursors, if any. The secondary structure analysis of the antisense RNA (which was significantly upregulated at MPT), did not reveal any difference in RNA folding at 30^{0}C or 32.5^{0}C, but predicted several stem loop structures (figure 4.49, 4.50). Further, the predicted stem-loop structures did not show any homology with the known stem loop miRNA precursors from other species. The same was found true for the sense RNA at MPT and FPT. The analysis thus, indicated that the ncRNA probably do not code for any known/conserved miRNA precursor, albeit, the possibility cannot be ruled out till experimental validations by Northern hybridization or RT-PCR are not attempted (which could not be done in the present study).

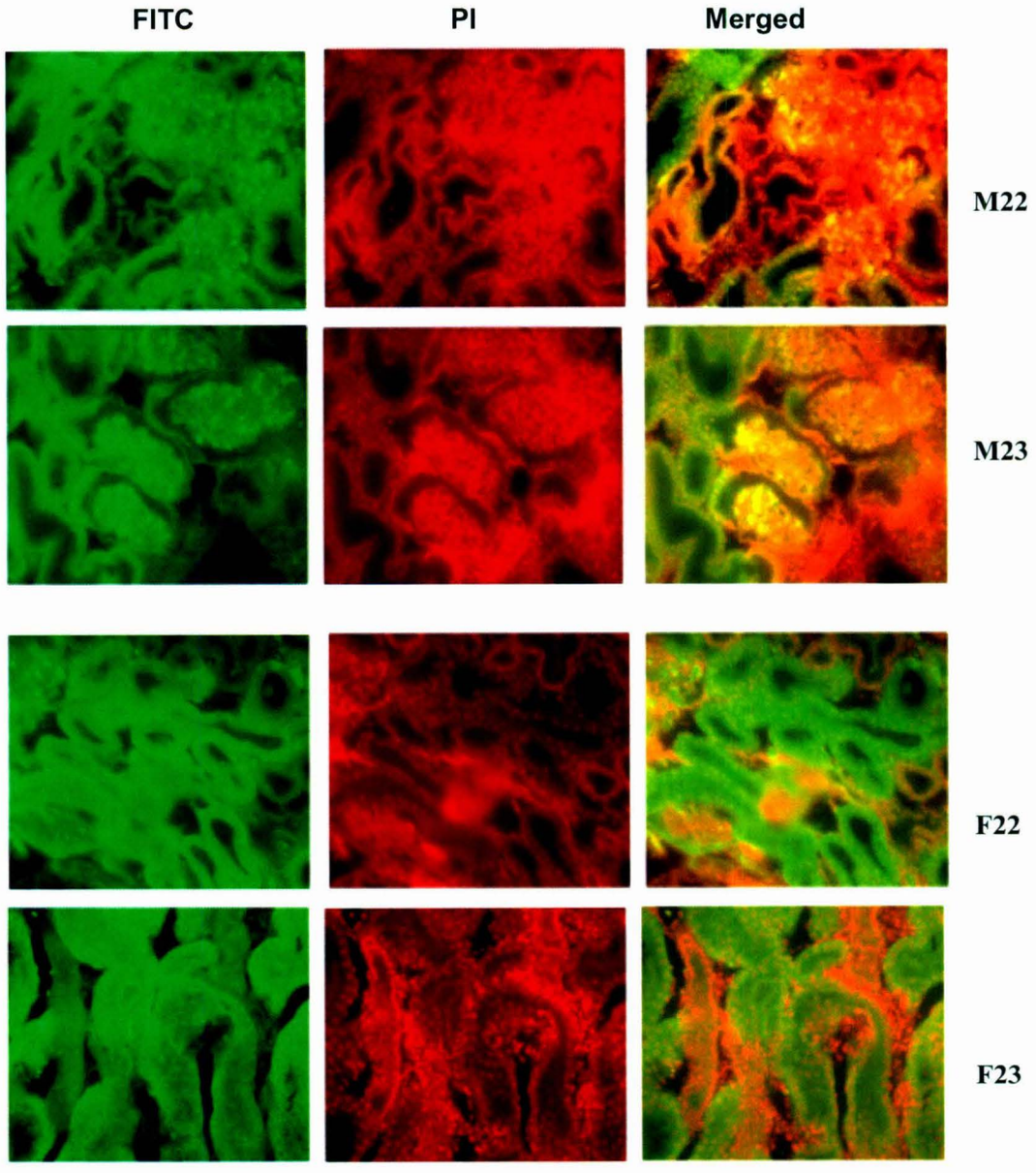


Figure 4.48: ***In-situ*** **hybridization showing localization of** ***SLF35F5*** **like** ***ncRNA*** **in the mesonephros of the embryos kept at MPT and FPT through TSP** using FITC labelled probe (amplified with primers SC35RN and SCINTR). Note: no sex/stage specific difference in localization/expression. M22, M23 and F22, F23 represent 22nd and 23rd developmental stages of male and female embryos, respectively.

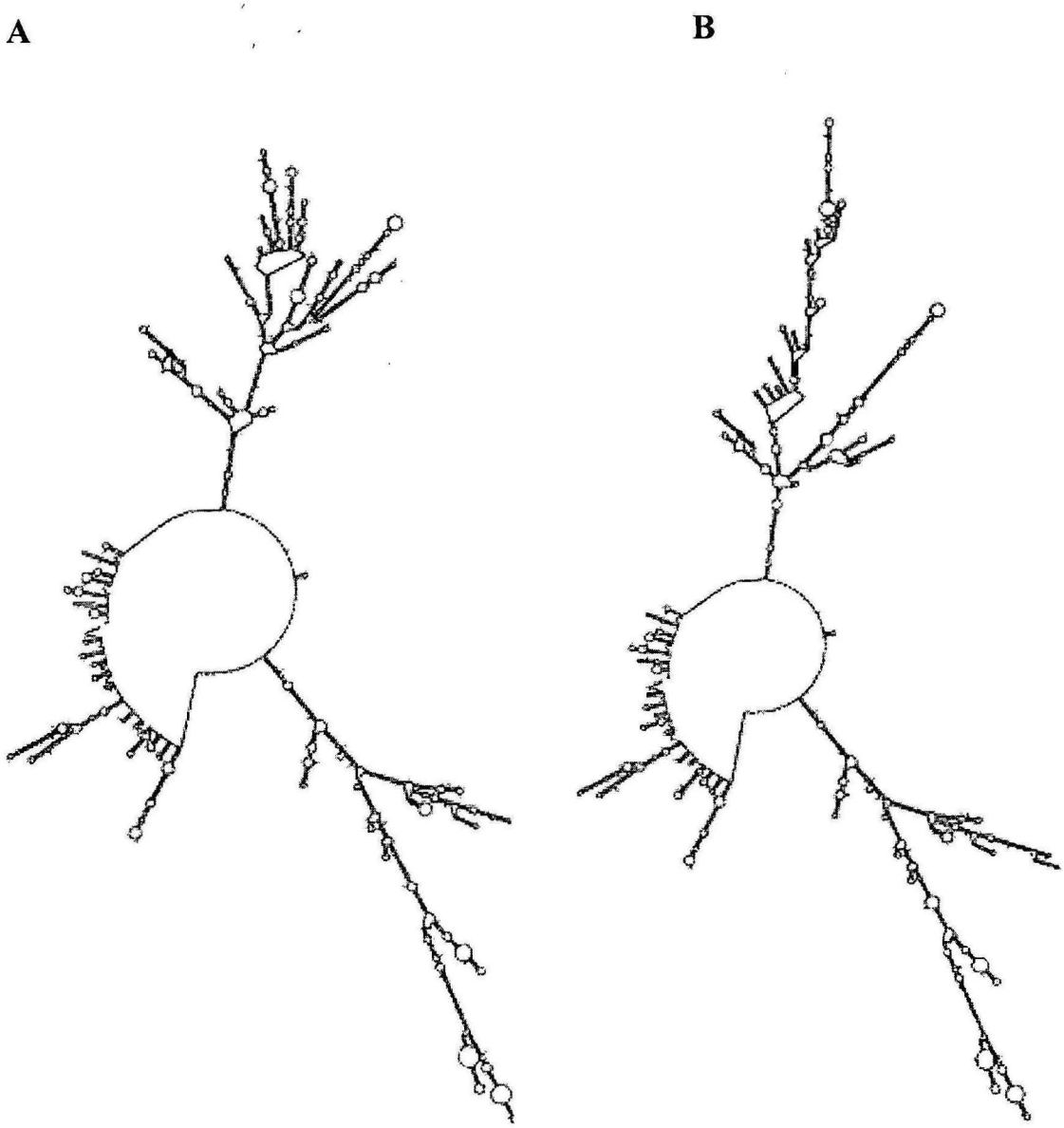


Figure 4.49: **Predicted secondary structures of the two stands of *SLF35F5* like novel *ncRNA* gene of crocodile isolated in the present study at 30^0C.** A) antisense; and B) sense strand. Note the presence of many stem loops, but overall almost comparable structures for both the strands.

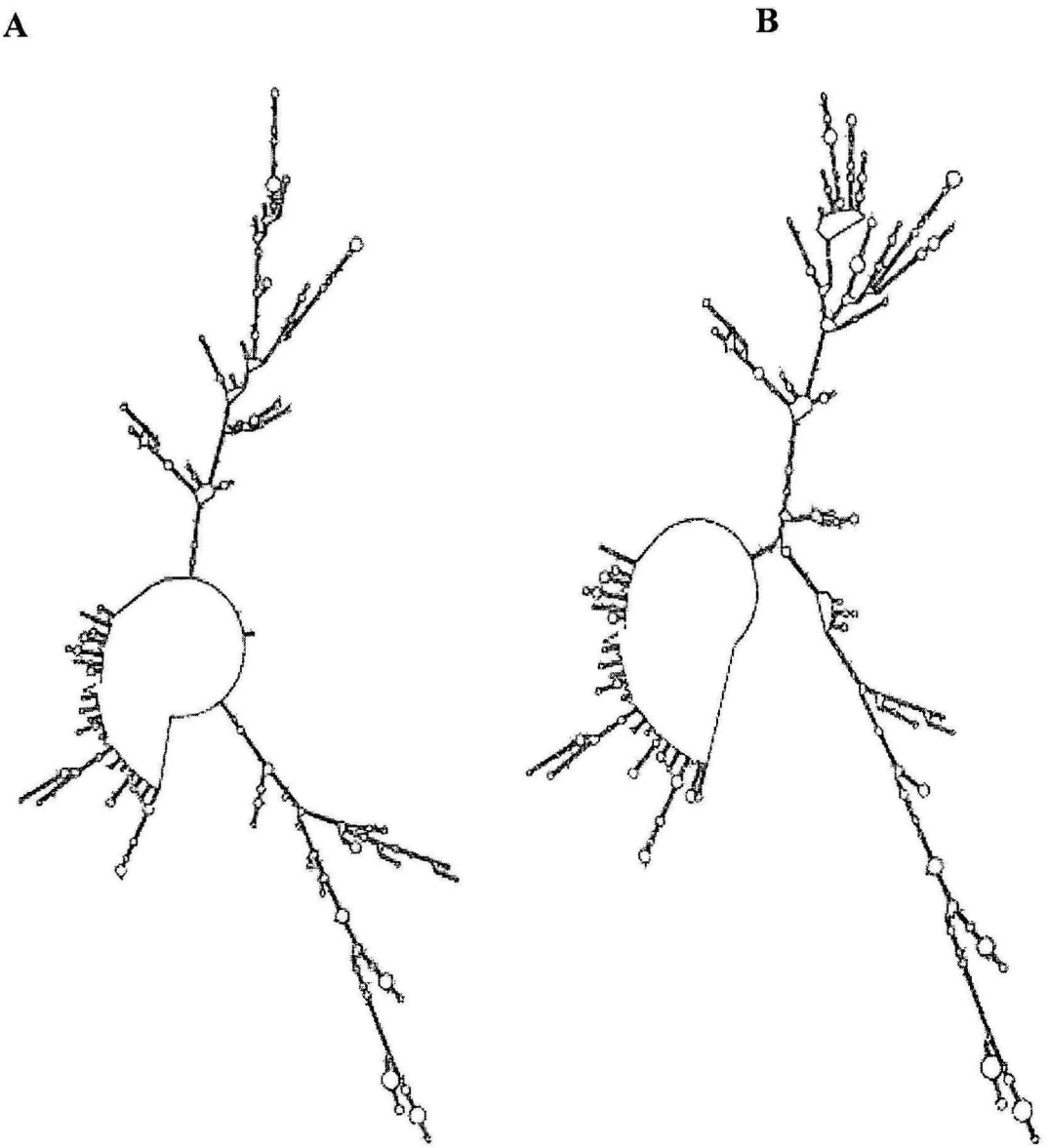


Figure 4.50: **Predicted secondary structures of the two stands of *SLF35F5* like novel *ncRNA* gene of crocodile isolated in the present study at 32.5^{0}C.** A) antisense; and B) sense strand. Note the presence of many stem loops, but overall almost comparable structures for both the strands.

Chapter 5

Discussion

Sex determination is a phenomenon of profound significance in continuation and evolution of biological life. During evolution, nature has manifested an amazing variety of sex determination mechanisms encompassing, genotypic sex determination (GSD) wherein the sex is determined by specific genes and/or chromosomes generally at the time of fertilization, to environmental sex determination (ESD) where sex is seemingly determined by specific environmental cues during embryonic development post-fertilization. Temperature dependent sex determination (TSD) is one of the most prevalent type of ESD, where it is the incubation temperature of embryos that seemingly decides the sex during a certain period of embryonic development generally referred as temperature sensitive window or period (TSW or TSP). The phenomenon of TSD is largely documented in reptilians including all the extant crocodilians, many turtle species and few species of lizards, amphibians and birds.

A large number of studies have been carried out to explore the molecular/genetic mechanism(s) underlying the sex determination, but these mainly concerned GSD in mammalian system. However, in recent years an increasing number of studies have also been addressing TSD, which still remains an unsolved, challenging problem of developmental biology. Most of such studies, generally dealing with the hormonal regulation during sex determination, have revealed hormones to play an early and important role in the TSD and even in few GSD exhibiting bird and amphibian species (section 2.5.6), which is one major difference from GSD in mammals wherein hormones do not play an early role in the sex determination. In addition, these recent studies have also shown a broad conservation of many of the major sex determining genes that have been originally defined in the mammals and other GSD vertebrates, to have sex specific or stage specific expression even in TSD species albeit with a significant difference in the expression pattern indicating differences in their function (Yao & Capel 2005; Place & Lance V.A. 2004; Valenzuela & Lance 2004; Agrawal 2006). Despite newer insights, the overall scope of most of such studies on TSD, have remained limited to identification/expression profiling of few conserved sex determining genes, while the major quest of identification of possible genetic/molecular triggers(s) that senses/translate temperature stimuli, remains completely unanswered. Thus, TSD and its underlying molecular mechanism(s) remains largely unsolved, and the present study was an step to this end using Indian mugger (*Crocodylus. palustris*) as a model system.

Success of any effort to identify gene(s) underlying expression regulation is expected to depend critically on the exact characterization of cell, tissue, and stage encompassing

the event of interest. Based on field studies (Lang *et al.*, 1989), the TSD in *C. palustris* is shown to occur at some time point between developmental stages 21 – 25 (spanning 22 - 30 days of development for male and female embryos, respectively), making it an herculean task to search for the novel sex related genes. However, a recent study in the lab (Agrawal 2006) has comprehensively shown that TSD in Indian mugger occurs between stages 21 to 23. This new insight about the TSP is expected to help ease up the search for differentially expressed genes by setting more meaningful global screens. Furthermore, there is a general lack of biological marker(s) to correctly sex the male/female embryos of *C. palustris*, which brings in additional challenge in undertaking comparative studies.

The present study was undertaken with two major aims: 1) to identify/isolate/characterize crocodile homologues of three important sex determining genes (*Aromatase, Dmrt1* and *Wt1*), which have been implicated to play important roles in determination/differentiation of sex in mammalian as well as, some other GSD and even TSD vertebrates; and 2) to search for novel differentially expressed genetic factors in the developing GAM tissue of putative female and male embryos during TSP using SSH global screens. It was anticipated that characterization of above genes would help in ascertaining their: possible role(s) in TSD, evolutionary conservation with GSD counterparts, and potential as good biological marker(s) for sexing of developing embryos and/or to delineate the determinate/indeterminate phases of TSD thus narrowing down the long TSP. On the other hand, search for differentially expressed gene(s) was expected to help in identification of novel candidate(s) that may be the master regulator(s) of TSD and/or have an important role in temperature sensing and deciding the fate of the developing bipotential gonads and their subsequent differentiation. These results are discussed in detail in the following sections.

5.1 Crocodile homologues of a few important sex-determining genes

In this study, attempts were made to identify/isolate homologues of three important sex-determining genes from the TSD exhibiting crocodilian species, *Crocodylus palustris.* These included genes namely, *Wt1*, *Dmrt1*, and *Aromatase*, which have been originally identified and implicated to play important roles in sex-determination/differentiation in GSD mammals.. In general, the strategy employed to identify/isolate the gene homologues comprised PCR based approaches i.e., initial RT-PCR using degenerate

heterologous primers designed from the apparently most conserved/shared domains of known sequences of most closely related vertebrates; followed by isolation of complete transcripts/genomic sequences by RACE, genomic-PCR and genomewalking approaches. The present study amply demonstrated the overall success of the strategy, and also revealed that; 1) primers designed from the closely related species work better for initial identification/RT-PCR; 2) use of nested primers for end specific RACE gives a higher efficiency with less chances of false positives; 3) more importantly, if carefully done, RACE can help identify important unexpected features of the target gene, such as, identification of transcriptional diversity/splice variants and/or both strand transcription (as seen in the present study, discussed below); and 4) a systematic combined use of genomic PCR and Genomewalking can help to explore the coding locus with great success even in a poorly characterized species with limited available genomic resources.

Accordingly, in the present study, all the three targeted genes i.e., *cpWt1*, *cpDmrt1*, *cpAromatase*, could be successfully isolated (transcribed homologues, as well as near complete coding regions alongwith partial upstream regions for two genes). Most interestingly, the study revealed multiple transcribed isoforms for each of the three genes, which were hitherto not reported in any of the TSD species. Subsequently, transcribed isoforms of each of the genes were characterized for their sex-/stage-/tissue-specific expression using multiple approaches that involved: RT-PCR, Northern hybridization, Real-Time PCR and RNA or immunolocalization studies. The detailed characterization, thus, revealed many novel insights strongly suggesting the important roles for some of the specific isoforms of the above genes (such as *cpWt1***b**, *cpDmrt1***e**, *cpAromatse***b**, see following sections) in sex-determination/differentiation.

Similar success and overall utility of the strategy have been demonstrated in case of *cpSox9*, another important sex-related gene studied earlier in our lab (Agrawal 2006). In that study also, a similar experimental approach has led to the first documentation of extensive alternative splicing resulting in the transcriptional diversity, and the presence of sex-specific isoforms in case of crocodile (*cpSox9*), as well as, mammalian model mouse (*Sox9*).

5.1.1 *cpWt1* (*Wt1* homologue of *Crocodylus palustris*)

5.1.1.1 *cpWT1* isoforms and coding locus

Wt1 is very important for embryonic development of the urogenital organs in mammals. It undergoes extensive alternative splicing and different *Wt1* isoforms play separate roles

in the development of gonads, kidney and adrenal (Hastie 2001; Hammes *et al.,* 2001). In the present study, an approximately 5 kb sequence (figure 4.5A) of *cpWt1* has been obtained which is comprised of 3.7 kb long coding region (spanning at least four exons and three introns), and an 1121 bp region upstream to the *cpWt1* transcriptional start site. To our understanding, this is the first study that provides insights about the *Wt1* gene structure in any reptilian and/or TSD species. Sequence comparisons show that *cpWT1* is comparable to *Wt1* gene described in other vertebrates, e.g., *Wt1* is coded by 10 exons in human (exon five being specific to only mammals) and 6 exons in case of chicken (Kent *et al.,* 1995). The analysis also revealed a broad conservation between the important Wt1 domain and the four zincfinger domains throughout the vertebrates, though the same was not true for the position and/or sequence of the intronic regions. Accordingly, the *cpWT1*introns (two complete and one partially sequenced) neither share any similarity with the database sequences, nor their positions in the *cpWt1* coding locus are conserved with that of the Gallus and mammals. The intronic sequence and positions of splice sites (except two) are also not conserved among chicken and human.

Furthermore, the present study revealed two *cpWt1* isoforms, *cpWt1***a** (1560 bp) and *cpWt1***b** (1088 bp), coded by four and one exon, respectively. Sequence comparison with the coding region revealed these to be the result of alternative splicing. This is similar to the case of *Wt1* in the mammals, which is known to have at least 24 splice variants arising by various types of alternative splicing and RNA editing (Laity *et al.,* 2000); and the transcriptional diversity of Wt1 in chicken which remains unexplored except for the presence of +KTS and –KTS isoforms (Lalli *et al.,* 2003). However, comparative status in the reptilian/TSD species is not known as no study (to the best of our knowledge) has attempted to explore the transcriptional diversity or coding region of their *Wt1*.

5.1.1.1.1 5`-upstream region of *cpWT1* coding locus

The novel features of *cpWt1* i.e., multiple isoforms and sex-specific expression suggesting an important role for the isoform *cpWt1***b** in the male sex determination (discussed in the following sections), led us to explore the upstream region to have some insight(s) about regulation of its expression. Here, it is noteworthy that *Wt1* promoter or upstream sequence has not been reported for any reptilian/TSD species till to date. In this study, 1121 bp region upstream to transcription start site of *cpWt1* was identified by genomewalking. Interestingly, this upstream region has important hallmarks of a promoter region such as, CCAAT box and TATA like sequences, as well as (rather more significantly), several binding sites for Heat shock transcription factor 1 (HSF1) and its

homologues. HSF1 acts as a primary agent which receives environmental stimulus like heat shock and then activates HSPs (heat shock proteins) to cope up with the environment. Hence HSF1 is essential for the heat shock response and is also required for several developmental processes (Akerfelt *et al.,* 2007). Developmentally dictated expression of heat shock factors is also established (Pirkkala *et al.,* 2001). Therefore, presence of at least 6 HSF1 binding sites in the 5`-upstream region of *cpWt1* suggests it is probably an early target of the machinery that underlies TSD and senses temperature; an interesting aspect that needs validation by further studies.

5.1.1.1.2 Diversity in *Wt1* sequences and putative function during vertebrate evolution

The two *cpWt1* isoforms share 625 bp at their 5' end, but have completely different 3' tails. Homology analysis reveals the bigger isoform *cpWt1***a** to be more conserved than the shorter isoform *cpWt1***b** to their vertebrate counterparts; interestingly, the two isoforms differ significantly in their similarity even to the reptilian *Wt1* which is ~ 61% for the bigger isoform (similarity includes most of its 3' UTR also) and only 9% region of the smaller isoform (as it has a unique 519 bp region at the 5' end and 480 bp region at 3' end). Furthermore, the homology analysis that revealed a part of the unique 519 bp 5'-end of the *cpWt1* isoforms to share an intronic region in *Gallus* but an exon in human, suggests the interesting possibility of exon shuffling in *Wt1* during vertebrate evolution. On the other hand, part of the considerable diversity (revealed in the homology analysis) at the 5' ends of the *Wt1* transcripts of non-mammalian vertebrates (*Trachemys*, *Alligator*, *Gallus*), appears to be more due to the incomplete status of the available sequence information.

Compared to the gene sequences, relatively less but considerable heterogeneity was evident even between the putative Wt1 proteins of reptiles, *Gallus* and mammals. Comparison and Wt1 proteins of *Trachemys scripta* and *Gallus* indicates absence of 25 aa in *Trachemys* Wt1 domain (contributed by exon 4 of *Gallus*), and the two lacking a proline-glutamine rich region, as well as, 17 aa (encoded by mammal specific exon 5) compared to that of mammalian Wt1. The proline-glutamine rich region in the mammalian Wt1 serves as a transactivation domain for many proteins (Wang *et al.,* 2001). Similarly, +KTS/–KTS isoforms of mammalian Wt1 are present in *Gallus* but not reported in any other non-mammalian vertebrate. Overall protein sequence comparisons revealed Wt1 proteins of reptilian and birds to be more similar to each other, and of

amphibian and fish as a separate close group (Bollig *et al.,* 2006); in turn this may suggests functional diversity/role(s) of *Wt1* during vertebrate evolution.

The coding regions for isoforms *cpWt1***a** and *cpWt1***b** span 885 and 483 bp predicted to code putative proteins of 294 and 160 aa, respectively. The first 108 aa of both these putative proteins are unique and do not match with corresponding region of any other WT1 protein, which may suggest different unique function for *cpWt1*. The bigger cpWt1**a** protein has the characteristic conserved signatures of Wt1 proteins i.e., a Wt1 domain (from 108-158 aa) and zinc finger domain (from 175 to 294 aa), and thus expected to have the similar conserved DNA binding (transcription factor) activity of Wt1 protein. In comparison, the putative protein translation of smaller isoform (cpWt1**b**) has a truncated Wt1 domain (108-143 aa), 17 unique aa at the C-terminal end and no DNA binding zinc finger domain.

Till now, no *Wt1* isoform has been described without DNA binding zinc finger domains, not even in human where 24 alternatively spliced isoforms of Wt1 are known. Absence of the zinc finger domain suggests that the smaller isoform (cpWt1**b**) may have a transactivation property and may not act as a transcription factor. Interestingly, expression, localization and preliminary functional studies in primary gonadal cell lines (discussed in the following sections) indicate the smaller novel *cpWt1***b** isoform to be an important candidate gene having an important role in the male gonadal development. A changed functional role due to a modification of zinc finger domain is well documented for mammalian +KTS Wt1 isoform; wherein three additional amino acids (K,T and S) between third and fourth zinc fingers makes it a less efficient DNA binding protein, and more a protein involved in mRNA splicing (Davies *et al.,* 1998). It is noteworthy that this +KTS isoform has also specifically been implicated for male sex determination (Hammes *et al.,* 2001).

5.1.1.2 Expression profiling of *cpWt1* isoforms

Both the *cpWt1* isoforms are expressed in all the embryonic tissues *viz.,* Heart, Liver, Kidney, Brain and GAM. However, it was only the shorter *cpWt1***b** isoform that expressed at significantly higher levels in the GAM tissues at MPT (approximately 40 folds higher compared to that at FPT) since the very beginning of the TSP, and not in any other tissue. This observation strongly indicates that *cpWt1***b** has an important role in the sex determination/differentiation. Although in the fish species, *Cynops pyrrhogater,* differential expression of *Wt1* isoform during spermatogenesis has been reported earlier

(Nakayama *et al.,* 1998), this is the first study which has revealed a male specific expression of a *cpWt1* isoform since the early bipotential stages through TSP. Moreover, its regulated expression since the beginning of the TSP also suggests that it may be an early player or target of the machinery that modulates the expression of genes in the bipotential gonads under the influence of incubation temperature resulting in TSD.

Localization studies in the developing GAM during TSP, carried out to further explore the sex specific expression/putative role of the novel isoform *cpWt1***b**, also strongly suggest an important role for *cpWt1* in the male gonadal development. It was amply evidenced by a close correspondence between the *cpWt1***b** localization (indicative of its expression) and the formation and development of testicular cords that represent the early and important cellular events occurring in the developing male gonads at MPT. In contrast, no such sex-/stage-/cell type specific localization was observed for the *cpWt1***a** isoform that expressed poorly and evenly at basal level all through the developing/differentiating male and female gonads. This relatively ubiquitous expression pattern of *cpWt1***a** indicates that it may have an important role in the developing urogenital tissues of *Crocodylus palustris*, similar to that of the –KTS isoform of *Wt1* that has been found responsible for general cell survival during early development of urogenital tissues in other vertebrates (Hammes *et al.,* 2001).

Contrary to the gonadal tissue, no cell-/stage-/sex-specificity was evident in the expression/localization of both the *cpWt1* isoforms in the developing mesonephros at MPT and FPT.

Furthermore, it is noteworthy that the expression of *cpWt1***b** observed in the real-time RT-PCR was significantly higher despite the fact that 80-85% of the template GAM cDNA was contributed by the mesonephros wherein the expression was comparable (as evidenced from the localization studies) over stages and putative sex of the embryo. This would, thus suggest that the observed higher expression of the *cpWt1***b** is contributed by its higher expression in the developing gonads, and infact is many fold more than the one ascertained using the cDNA from the total GAM complex, again suggesting an important role for *cpWt1***b** in the male sex determination.

5.1.1.3 Novel *cpWt1*b isoform expression tracks early events during male gonad development

A close scrutiny of the localization pattern of *cpWt1***b** through TSP at MPT, reveals that during initial stages (21-22) it appears (expressed) in all the cells scattered throughout

the gonads but from 23rd stage onward, *cpWt1***b** expressing cells start organizing themselves in tubule like structures in the gonads, which by the 25th stages develops/differentiate into well defined testicular cords and interstitium between the cords. However, it was interesting to observe that the cells inside the cords were not expressing *cpWt1***b** (figure 4.10, figure 4.11 and figure 4.12). This pattern of *cpWt1***b** localization closely corresponds to the localization/expression of *Sox9* in the mammalian male gonad development (figure 5.1), wherein *Sox9* expression starts throughout the developing gonads by 11.5 dpc, appears to organize in the pre-Sertoli and Sertoli cells marking the beginning of testicular cord formation by 12.5 dpc, and marks the completion of testicular cords by 14.5 dpc (Wilhelm *et al.,* 2007).

Studies in few TSD species have shown that cellular events are largely conserved in early gonad development (Smith & Joss 1993; Smith & Joss 1994b; Morrish & Sinclair 2002; Sinclair *et al.,* 2002). In TSD species like alligator and turtles, the cells remain scattered without any visible signs of morphological differences in the developing gonads during early stages of TSP (comparable to stages 21-22 in the *C. palustris*) (Smith C.A. & Joss M.P. 1994; Wibbels *et al.,* 1991; Schmahl *et al.,* 2003; Yao & Capel 2005). During this period, cells from coelomic and germinal epithelium migrate in to the thickened genital ridge and start developing in to early sex cords which cannot be distinguished because of their very small size (Wibbels *et al.,* 1993; Schmahl *et al.,* 2003; Smith & Joss 1993). Recently, in a gonad explants culture study it is shown that the migrating cells do not come from the mesonephros (Moreno-Mendoza *et al.,* 2001). Studies in turtle and alligator has also demonstrated that pre-Sertoli cells, epithelial cells and the germ cells remain intermixed at MPT and FPT, and that the number of pre-Sertoli cells (which has even be speculated to decide the sex of the developing embryos) is much higher in the gonad from MPT than at FPT during the initial stages of TSP (Smith & Joss 1994b; Wibbels *et al.,* 1991; Schmahl *et al.,* 2003; Yao *et al.,* 2004; Yao & Capel 2005). Subsequently, with progress of development through TSP, differentiation of male gonad at MPT is marked with appearance and gradual maturation of sex cords /testicular cords, with Sertoli cell and germ cells in the lumen (Smith & Joss 1994b; Wibbels *et al.,* 1991; Shoemaker *et al.,* 2007; Schmahl *et al.,* 2003; Yao *et al.,* 2004). Interestingly, the spatio-temporal pattern of *cpWt1***b** localization (discussed above) closely parallels the above cellular events, thus suggesting that gonad development in *C. palustris* is conserved similar to other TSD species.

It has been unequivocally demonstrated in vertebrates (both TSD and GSD) that testicular cords are made up of Sertoli cells and peritubular myoid cells (Wilhelm *et al.,*

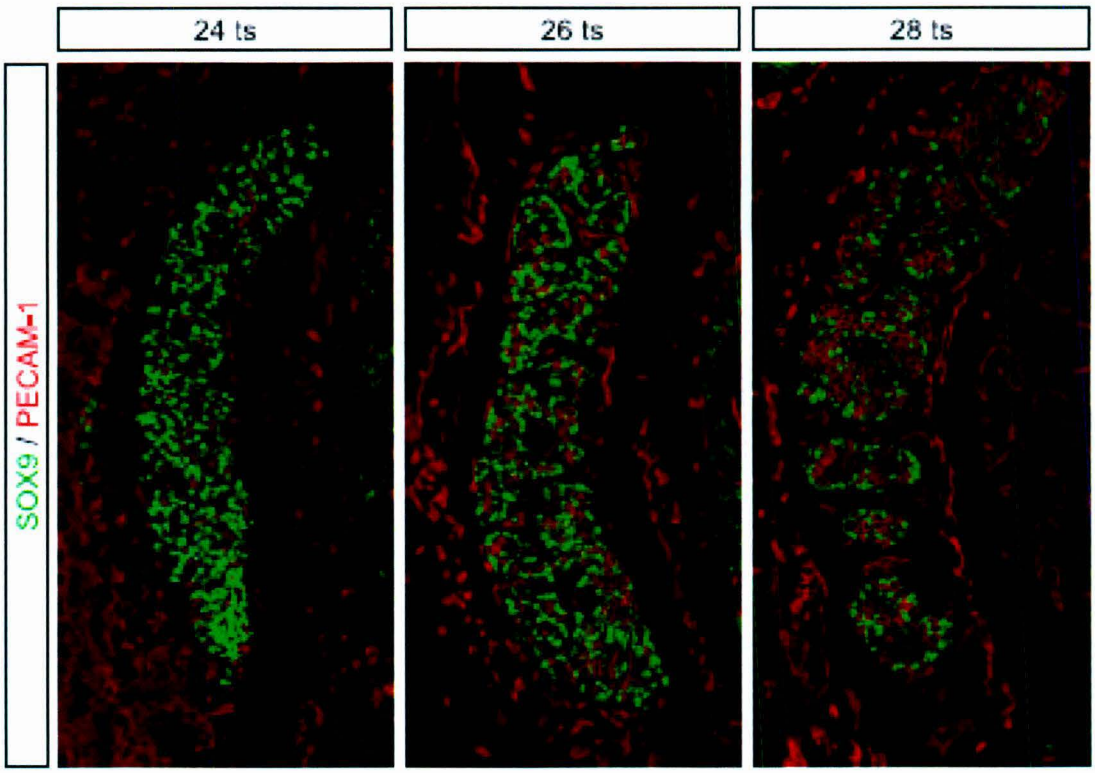


Figure 5.1: **SOX9 expression in mouse tracks differentiation of pre-Sertoli cells into Sertoli cells and testicular cords formation in mouse.** Nonpolarized, dispersed somatic cells visualized by SOX9 immunofluorescence (green) represent pre-Sertoli cells at the 24-tail somite (ts) stage. A few hours later, by 28 ts, these cells become polarized, forming epithelial aggregates that assemble into testis cords; at this stage they are referred to as Sertoli cells. PECAM-1 counterstaining (red) marks PGCs and endothelial cells (adapted from Wilhelm *et al.*, 2007).

2007; Yao & Capel 2005; Yao *et al.,* 2004; Schmahl *et al.,* 2003; Smith & Joss 1994b; Sekido & Lovell-Badge 2007). Hence, it is reasonable to conclude that *cpWt1***b** is probably expressed in the pre-Sertoli cells, Sertoli cells and peritubular myoid cells, which is an important finding of this study. Further, the results of present study suggest that the *cpWt1***b** is probably required for the development/maintenance/proliferation of Sertoli cells in higher numbers (the earliest events in the male sex determination in vertebrates preceding the development of sex cords). This inference gets further support from the fact that *cpWt1***b** induces expression of many of the genes that are indicated to be Sertoli cells specific, e.g., *Sox9* and *Amh.*

The localization studies also suggest different functions for the two *cpWt1* isoforms, as *cpWt1***a** was seen in all the cells of developing gonads and germ cells while *cpWt1***b** was revealed only in cells forming testicular cords and not the other cells/germ cells. Recently, absence of *Wt1* expression in the PGCs and Germ cells of developing male gonads has recently been shown in turtle (Schmahl *et al.,* 2003). The study has also revealed that primordial germ cells appear near the coelomic surface, which is in contrast to mouse wherein PGCs are located in the interior of the gonad (Wilhelm *et al.,* 2007). The study further proposes that primitive sex cords in turtle are formed by invagination of coelomic epithelium that encircle the PGCs (Schmahl *et al.,* 2003) rather than the PGCs migration inside the sex cords as seen in mammals. The present study suggests that the chain of events in *C. palustris* (as indicated by the localization of *cpWt1***b**) is similar to that of turtle wherein the emerging testicular chords seem to harbour cells from the very beginning. On the other hand, the *cpWt1***a** localization pattern remains comparable to that seen in mammalian cells where *Wt1* is present even in the germ cells (Natoli *et al.,* 2004).

5.1.1.3.1 Preliminary functional validation of *cpWt1*b as a putative master regulator of male development

The present study provided the first ever evidence of a novel *cpWt1***b** isoform that is expressed at higher levels in male GAM significantly tracking the major cellular events occurring during early male gonad development. These findings suggested that *cpWt1***b** may be an important candidate master regulator of male sex in TSD, an aspect which prompted preliminary functional studies wherein efforts were made to explore its role in regulation of expression of other male promoting genes and those associated with the Sertoli cells development using transfection assays.

Transfection of EGFP (enhanced green fluorescent protein) tagged *cpWt1* isoforms in the GAM-derived crocodile primary cell lines, revealed contrasting localization of the two isoforms. The novel shorter isoform *cpWt1***b** was localized mainly in nucleus and certain cytoplasmic compartments, whereas near reverse pattern was evident for the longer *cpWt1***a** isoform (strong cytoplasmic and a faint nuclear localization). These data again suggest that the two isoforms probably have separate functions, which remain to be deciphered in future studies. Similarly, it is expected that ascertaining the cytoplasmic compartments where cpWt1**b** was found localized would provide better insights about its function; but the same remains to be answered in future as in the present study this could not be done due to unavailability of characterized markers for the various cellular compartments of crocodile.

It was important to observe that the primary cell lines used for transfection studies, did not show any sign of imprinting (of expression pattern) expected due to the sex- and/or incubation temperature of the source GAM tissues (section 4.1.1.7.1); and thus represented a test system more akin to the undifferentiated gonads. Importantly, in these untransfected control cell lines, expression of *cpSox9* and the *SLF35F5* like *ncRNA* were not detectable even by the highly sensitive real-time RT-PCR. Compared to the control cells, *cpWt1***a** expression level went up >2.5 fold after transfection, but without any detectable effect on the expression of the male favouring genes checked in the analysis. Contrary to this, a similar upregulation of *cpWt1***b** after transfection however, induced the expression of all the tested male favouring genes; and most importantly, it also revived the expression of the *cpSox9* and *SLF35F5* like *ncRNA*. Many of the tested genes express at significantly higher levels in the GAM tissue since beginning of the MPT (Agrawal 2006, and present study), and their induction by *cpWT1***b** thus, suggest that *cpWT1* is an important gene very high in the hierarchy of the cascade of sex determining genes underlying the TSD.

The above functional assay when considered together with other features of *cpWt1***b**, *viz.*, 1) presence of many HSF1 binding sites in the upstream region; 2) presence of transactivation domains for HSF1 in its 108 amino acids long unique region at the N-terminal end; 3) very high expression from the beginning of TSP in GAM from MPT; and 4) localization pattern similar to that of a tracking marker of the cellular events of male gonad development comparable to that of *Sox9* in mammals, strongly suggest it to be the probable master regulator and/or component of the earliest genetic machinery that senses temperature stimulus for male sex determination in TSD.

5.1.2 *cpDmrt1* (*Dmrt1* homologue of *Crocodylus palustris*)

Dmrt1 is a conserved gene suggested to be important in the sex determination and early gonad differentiation in vertebrates. Across the evolutionary strata, it is known to play varying role i.e. from being a master regulator of male sex in teleost Medaka (Matsuda et al., 2002) to an important sex regulator on the Z-chromosome of chicken (Shan et al., 2000) and to a gene involved in embryonic testis differentiation in mammals (Raymond et al., 2000). Recent studies have shown significant transcriptional diversity for this important gene during gonadal differentiation in many GSD species (Huang et al., 2005; Zhao et al., 2007; Guo et al., 2005; Cheng et al., 2006; Lu et al., 2007). Preliminary studies done in trachemys, turtle and alligator show upregulation of *Dmrt1* during male gonadogenesis, suggesting its possible role in male sex-development even in TSD species (Kettlewell et al., 2000; Sreenivasulu et al., 2002; Shoemaker et al., 2007). However in general, these studies lack organizational and functional details of the gene, which are important to ascertain its functional role and evolutionary significance in TSD.

In the present study, for the first time a complete homologue of the *Dmrt1* gene along with its multiple transcribed isoforms, has been isolated from Indian mugger that is a TSD species. Detailed characterization of the *cpDmrt1*coding locus and its transcribed isoforms resulting from the alternative splicing suggest *cpDmrt1* to be an early and important candidate gene for male development even in TSD.

5.1.2.1 *cpDmrt1* coding locus and its transcribed isoforms

The *cpDmrt1* (*Dmrt1* homologue of *C. palustris*) has been identified and characterized in the present study. The 3.8 kb long *cpDmrt1* locus is revealed to be comprised of three exons and two introns; is observed to undergo extensive splicing leading to eight expressed isoforms. The exon-intron structure was inferred by comparison of genomic locus and isoforms obtained in the study.

A comparison of the *Dmrt1* gene sequences of GSD exhibiting vertebrate species revealed significant variations in the overall organization of the locus. In mammals and chicken, *Dmrt1* is coded by five exons and the positions of introns are conserved (Shan *et al.,* 2000; Raymond *et al.,* 2000). However, the gene homologue *DMY* in Medaka has six exons (Matsuda *et al.,* 2002) and similar homologue DMW of *Xenopus* has four exons, with no conservation across their intronic regions in sequence and/or positions (Yoshimoto *et al.,* 2008). Unfortunately, similar comparisons are presently not feasible for *Dmrt1* organization in TSD vertebrates, as these remain to be described. Therefore,

the *cpDmrt1* isolated/analysed in the present study from Indian mugger assumes importance as it represents the first case of a *Dmrt1* from a TSD species; it would be interesting to study the extent of organizational and functional conservation among other similar TSD species.

It was noteworthy that the sequencing of the large number of cloned products (form multiple RACE/genome walking PCR efforts with several primer combinations) did not reveal any other related sequence that could have suggested the possibility of another *Dmrt1* gene. This observation assumed more significance, as the RACE and genome walking, the two PCR-based approaches used in the study to isolate the *cpDmrt1* locus, are expected to be many-fold more sensitive than Southern/FISH methods to pick up any related sequences in the explored template (i.e. genomic DNA in this case). Therefore, the present study strongly suggests that there is only one *cpDmrt1* locus and all the eight transcribed isoforms identified in the study must have been generated from the same coding locus possibly by alternative splicing (see section below). *Dmrt1* has also been shown to be a single copy gene in most other GSD vertebrates except for the recent reports of Medaka and *Xenopus* (Nanda *et al.,* 2002; Osawa *et al.,* 2005) wherein additional copies of *Dmrt1* have been reported to be present on the sex chromosome resulting in different transcripts (from these copies) having different sex specific functions (Yoshimoto *et al.,* 2008; Matsuda *et al.,* 2002).

5.1.2.2 *cpDmrt1* transcribed isoforms

A number of recent studies have revealed extensive transcriptional diversity of *Dmrt1* during gonadogenesis in many GSD species like, human, mouse, chicken, zebrafish, rice eel (Huang *et al.,* 2005a; Guo *et al.,* 2005; Lu *et al.,* 2007; Zhao *et al.,* 2007). These studies also suggest that differential expression of the transcribed isoforms have significance in male sexual development, albeit exact mechanism is yet to be discovered. Compared to this, no such diversity has been reported for *Dmrt1* of any of the reptilian or TSD species. In this context, findings of the present study assume great significance. In this study, a total of eight isoforms of *cpDmrt1* have been identified in the developing GAM tissues of Indian mugger. These isoforms that are named as *cpDmrt1***a1** (longest isoform), *cpDmrt1***a2**, *cpDmrt1***b1**, *cpDmrt1***b2**, *cpDmrt1***c**, *cpDmrt1***d**, *cpDmrt1***e** and *cpDmrt1***f** (shortest isoform), ranging in size from 482 to 2060 bp, putatively encode six predicted proteins, all having the characteristic DM domain unique to *Dmrt1*. BLAST analysis revealed *cpDmrt1*a2 as the most conserved isoform, whereas all other isoforms differed significantly and had unique regions compared to the

published vertebrate *Dmrt1* sequences. The major heterogeneity in the isoforms (also predicted proteins) is seen only in their C-termini and 3'-UTRs (resulting from alternate splice sites/change in the reading frames; see next section) which in general, are unique and do not match to any similar sequences reported for other vertebrates. The study, thus, provides the first evidence that transcriptional diversity of *Dmrt1* is equally conserved even in TSD like GSD; these important findings also provide support for *cpDmrt1* to be an important candidate/master regulator of male development even in TSD.

5.1.2.3 Alternative splicing underlying the observed transcriptional diversity of *cpDmrt1*

Comparison of the eight *cpDmrt1* isoforms and *cpDmrt1* coding region indicated beyond doubt that all the isoforms were essentially generated from the same one locus by alternate splicing, but involving multiple mechanisms of alternative splicing, *i.e.*, use of alternative 5' and 3' splice sites, exonization of intronic sequences, exon skipping and alternative polyadenylation sites. Also, splicing was mainly observed towards the 3' ends or more specifically only after the conserved DM domain, and in many cases it was accompanied with a change in reading frame (for isoforms *cpDmrt1***a1**, *cpDmrt1***b1**, *cpDmrt1***d** and *cpDmrt1***e**) leading to unique regions that appeared to be possibly reptilian-specific. Extensive alternative splicing leading to *Dmrt1* transcriptional diversity has earlier been demonstrated in many of the GSD species, which include both mammalian and non-mammalian vertebrates, like, human, mouse, chicken, rice eel and zebrafish species (Zhao *et al.,* 2007; Cheng *et al.,* 2006; Guo *et al.,* 2005; Lu *et al.,* 2007; Cheng *et al.,* 2006); interestingly, in all these vertebrates also splicing has been shown to occur mostly in the region after the DM domain and in few cases with a reading frame shift. *Dmrt1* isoforms described in the mouse (Lu *et al.,* 2007) arise by the alternative 5' and 3' splice sites, exon skipping and alternative polyadenylation sites. Inclusion of a novel region has also been reported in at least one isoform of mouse *Dmrt1*, which is not found in the *Dmrt1* sequence of any other vertebrate. The putative protein translations of the mouse isoforms also showed shift in reading frames and interestingly one isoforms lacked DM domain. Alternative splicing in chicken produces several isoforms of *Dmrt1* (Zhao *et al.,* 2007) which have novel regions arisen by the exon skipping and/or exonization of the intronic sequence. Predicted proteins again were found to have many novel regions because of change in reading frame by splicing. All these examples amply suggest that alternative splicing, in general, underlies the observed transcriptional diversity of *Dmrt1* in vertebrates producing several novel

isoforms through varied patterns of alternative splicing which possibly also contributes to the diversity in the proteome.

The longer 3'UTRs of isoforms *cpDmrt1***a1** and *cpDmrt1***b1** were found to be unique, which do not show any evolutionary conservation to any other vertebrate *Dmrt1* isoform, albeit *Dmrt1* in other reptilian species remain to be explored. However, it was interesting to note that 3'UTR region from 1806 bp to the end of *cpDmrt1***a1** and 1528 bp to the end of *cpDmrt1***b1** has approximately 85% sequence conservation with corresponding 3' UTR region of *Gallus* and mammalian *Dmrt1*. Such a conservance in the 3' UTR across taxa, especially for a gene, which undergoes extensive alternative splicing and is indicated to perform different functions, is intriguing warranting further studies.

5.1.2.4 Expression profiling of cpDmrt1 isoforms indicates it to be a possible early regulator of male gonad development

Expression analysis of the *cpDmrt1* in the developing gonads at MPT and FPT during TSP, revealed significant differences in the pattern and levels of expression of individual isoforms. Despite differences, In general, the expression of *cpDmrt1* (and its individual isoforms) was invariably found to be significantly higher throughout TSP only in GAM of embryos reared at MPT, suggesting it is important for the determination/differentiation of male gonads. Significantly, *cpDmrt1***e** was revealed to be the most important isoform contributing approximately three fourth of the overall *cpDmrt1* expression since the beginning of the TSP at MPT. In comparison, the most conserved isoforms *cpDmrt1***a1/a2** were expressed at much lower levels although more abundant than the remaining isoforms *(***a2, b2, c, d and f)** which were barely detectable. These observations are in contrast to other vertebrates (like human, mouse, chicken and zebrafish) wherein the most abundant isoforms are generally found to be the ones that are most conserved (Guo *et al.,* 2005; Lu *et al.,* 2007; Zhao *et al.,* 2007).

Although, similar data is not available from any other reptilian or TSD species, we strongly believe that the above difference is significant, and presence of an apparently novel (less conserved) isoform highly expressing in male GAM may have a specific important role in the temperature dependent sex-determination. This notion gets further support from the fact that the *cpDmrt1***e** isoform codes for the longest putative cpDmrt1 protein that carries a unique 15 aa long region having a novel PDZ binding domain (common domain for protein-protein interaction) as well as, recognition site for PIKK (phosphoinositide 3-kinase related kinase) directed phosphorylation. Proteins containing

PDZ domains function as scaffolds for formation/spatial distribution of large signalling complexes and play an important role in cell polarity and growth (Jemth & Gianni 2007), while PIKK members coordinate the cellular response to DNA damage (Hiom 2005). It is possible that these novel features of the unique 15 aa region of *cpDmrt1***e** underlies the machinery that regulates its dramatic early upregulation in response to the temperature to decide the sex of developing embryo- a proposition which may be worth testing.

Furthermore, as become evident from the forgoing, identification of the novel isoform *cpDmrt1***e** is one of the most important finding in the present study; and this became possible only due to comprehensive expression studies, especially efforts to isolate complete transcripts by RACE, that led to the identification of multiple isoforms/transcriptional diversity. An similar important example was the identification of sex-specific isoforms in medaka, which was later identified as a master regulator for the male sex determination (Guan *et al.,* 2000)Therefore, we suggest that any future study directed towards the identification and function of *Dmrt1* in a new unexplored species, must start with conscious efforts to explore transcriptional diversity to have more valuable information about its role during the gonadogenesis.

5.1.2.5 Multiple isoforms of *cpDmrt1*: possible significance

Transcriptional diversity is expected to be for multiple roles that are suggested for *Dmrt1* during sex-determination and differentiation. Recent reports suggest that *Dmrt1* may have a role in cell fate determination, postnatal differentiation of Sertoli cells and also in primordial germ cells maintenance and their meiotic progression (Kim *et al.,* 2007b). Such multiple roles for *Dmrt1* have also been suggested in the case of a lizard *Calotes versicolor* (Sreenivasulu *et al.,* 2002) that exhibits hormonal/temperature sex-determination. It is now well documented that DM domain genes in GSD vertebrate species (human, mouse, chicken, rice-eel and zebrafish) undergo splicing like their invertebrate counterparts (*dsx* and *mab3*) wherein male/female specific isoforms control different aspects of male and female development (Hodgkin 2002). Therefore, it s expected that transcriptional diversity leading to multiple isoforms contributes to the multiple roles of *Dmrt1* either by binding to different upstream and downstream targets or by differential regulation of specific isoforms, as has been seen in case of *Wt1* in mammals (Hammes *et al.,* 2001). Here, it may be noted that many of the *cpDmrt1* isoforms (*cpDmrt1***b1**, *cpDmrt1***b2**, *cpDmrt1***c** and *cpDmrt1***d**) identified in the present study are indicated to have unique regions that appear to be possibly reptilian-specific (as these are not seen in the isoforms and/or genomic DNA of human, *Gallus* and

Xenopus), and may be responsible for the specific function(s) that may be needed in developmental decision like TSD (as seems to be the case with isoform *cpDmrt1***e** discussed above). Thus, it may be speculated that diversity in isoforms may be for generation of different interacting partners for cofactors or binding proteins among themselves or for making heterodimers as seen in case of doublesex (Bayrer *et al.*, 2005) although such possibilities are yet to be demonstrated for *Dmrt1* isoforms in vertebrates and in specific context of TSD.

5.1.3 *cpAromatase* (*Aromatase* homologue of *Crocodylus palustris*)

Aromatase refers to a complex that mediates synthesis of estrogens from C19 steroids that includes androgens. Aromatase complex comprises of ubiquitous flavoprotein, NADPH Cytochrome P450 reductase and a unique aromatase P450 (coded by CYP19). As described in section 2.5.5 and 2.5.6, unlike mammals, Aromatase plays an early and important role in TSD in feminization of the bipotential gonad. In this study, for the first time a successful attempt has been made to identify/characterize the homologue of Aromatase from Indian mugger that is a model TSD species. Briefly, in this study, two expressed isoforms of *cpAromatase* have been identified that show sex-/stage-/tissue-specificity in their expression, and also possible potential as markers of female gonad development. In addition, a part of the upstream region has been isolated that appears to be specific to reptilian species.

5.1.3.1 *cpAromatase* isoforms

Efforts to isolate *cpAromatase* led to the identification of two isoforms of 2133 and 1977 bp in size that we named as *cpAromatase***a** *and cpAromatase***b**, respectively. A 156 bp region towards the 5`-end of the bigger isoform was absent in the smaller isoform. To our understanding, this is first observation of multiple isoforms of aromatase in a TSD species. Sequence comparisons revealed >95% similarity of the *cpAromatase* nucleotide sequence to that of alligator aromatase, and interestingly the similarity spanned even the 5' and 3' UTR regions. However, similarity with *Trachemys scripta* and *Gallus* was only 86% and 85%, respectively, with no observed conservation of 5' and 3' UTR regions. *In silico* translation of the *cpAromatase* isoforms indicated two putative proteins of 504 and 452 aa, 90% of which corresponded to the p450 domain that is generally conserved during vertebrate evolution. Further, the putative proteins of the *cpAromatase* isoforms were found to be more similar to the aromatase proteins of other vertebrates compared to corresponding similarities seen at the nucleotide level; this is as expected as a large

proportion of the DNA variation of functional gene domains is masked during translation due to the codon degeneracy. Accordingly, cpAromatase shared ~ 98% similarity with Aromatase of alligator, ~88% of other reptile, 86% of *Gallus* and a conservance of only ~75% with mammalian Aromatase protein.

5.1.3.2 Upstream region of *cpAromatase*: features suggest it to be unique to reptiles

In the present study a 1405 bp region upstream to transcription start site of *cpAromatase* was identified and characterized by *in silico* analysis. Comparative analysis of this 1405 bp region with other aromatase upstream sequences of GSD species like, Zebrafinch, *Gallus, Courtnix* and mammals revealed similarity only for the first ~300 bp (i.e., uptp – 300 upstream region), beyond which the *cpAromatase* upstream region was generally unique compared to the data base sequences. On the other hand, similar upstream sequences are not available from other TSD and/or reptilian species for comparison, and thus comparative status therein remains a matter of speculation.

In silico prediction of protein binding sites in the *cpAromatase* upstream region revealed presence of TATA box at -29 position and many other conserved regulatory elements and also binding sites for the *SF1* and *COUP*. Interestingly, *Sf1* binding site is seen to be conserved in the Aromatase promoter of the mammals and in GSD vertebrates in general; it is also indicated to be necessary for the functional promoter activity of *Aromatase* gene, and is shown to regulate many other steroidogenic enzymes (Young & McPhaul 1998). Similarly COUP binding site of the Aromatase is also evolutionarily conserved (Tong & Chung 2003; Chen *et al.,* 2005). COUP also has been speculated to have a role in the female gonad development. . The regulation of *Sf1* and *Aromatase* is antagonistic in the GAM tissues of the TSD species *i.e. Sf1* expression is higher at masculinising temperature and *Aromatase* expression is higher at feminizing temperatures (Ramsey *et al.,* 2007). Similarly COUP expression was considered necessary for the differentiation of Leydig cells from the progenitors and for maintaining proper balance of female sexual hormones, fecundity, regularity of estrous cycles and onset of puberty (Takamoto *et al.,* 2005; Qin *et al.,* 2008). The above-suggested roles of COUP in the steroidogenic pathways make it an important candidate for following *Aromatase* regulation. However, presently nothing is known about the genetic/molecular network underlying the temperature regulated expression of Aromatase in GAM tissue, although a number of recent studies have shown *Aromatase* differential expression in TSD species (Desvages & Pieau 1992; Elf 2003; Murdock & Wibbels 2003a; Pieau *et al.,*

1998). In this context, the insights from the partially characterized upstream region of *cpAromatase* assume importance as these suggest an important role for *Sf1* and *COUP* in aromatase regulation. It would be interesting to explore the similar upstream regions of *Aromatase* coding locus from few other TSD species.

5.1.3.3 *cpAromatase* expression is sex-/stage-/tissue- specific

Expression analysis revealed sex-/stage-/tissue-specific expression of *cpAromatase*. Of all the five tissues tested, expression was detectable only in GAM and Brain tissues in sex-/stage-specific manner. While, RACE and RT-PCR studies conclusively revealed that both the isoforms (*cpAromatase***a** and *cpAromatase***b**) are expressed only in the GAM tissue of the putative female embryos kept at FPT through TSP. In comparison, expression of the smaller isoform (*cpAromatase***b**) was detected in the brain tissues of the embryos both at MPT and FPT. Apart from brain and GAM *cpAromatase* expression could not be detected in any other tissue during TSP. Furthermore, in both the tissues, *cpAromatase* expression was seen only from 23rd stage onward increasing with the progression of development. However, at the 23rd stage, Aromatase expression levels were comparatively low and could be seen only after 36 cycles of PCR amplification. During later stages Aromatase expression increases significantly, which probably indicates onset of female gonad development.

However, a review of similar studies in TSD species reveals a general lack of consensus about the expression pattern of Aromatase in the developing gonads, both with respect to the stage at which the expression starts and the sex-specificity of expression. Indeed, studies in few turtle species, *viz., Emys orbicularis* (Desvages & Pieau 1992), *Dermochelys coriacea* (Desvages *et al.,* 1993a) and *Malaclemys terrapin* (Jeyasuria & Place 1998a), a female-specific differential Aromatase activity/gene expression has been reported in the developing gonads at MPT and FPT during the TSP; which led to the belief that *Aromatase* is an important upstream element in the sexual differentiation cascade under TSD. In contrast, no such differential "expression" during the TSP was seen in the developing gonads of some other TSD species like, *Trachemys scripta* (White & Thomas 1992; Willingham *et al.,* 2000a; Murdock & Wibbels 2003b); *Lepidochelys olivacea* (Salame-Mendez *et al.,* 1998) *Alligator mississipiensis* (Gabriel *et al.,* 2001) and *Crocodylus porosus* (Smith & Joss 1994a). Furthermore, no difference was reported in the expression pattern of *Aromatase* in two closely related taxa of turtle one having TSD and other having GSD (Valenzuela *et al.,* 2006). Similar to the above studies reporting lack of sex-specificity, other studies indicated wide variation in the time

of onset of Aromatase expression. While in some of the turtle species *Aromatase* expression was observed to start only during later stages of TSP albeit in the putative female gonads at FPT (Desvages & Pieau 1992; Desvages *et al.,* 1993a), it was seen from the very beginning of TSP in the *Trachemys scripta* (Willingham *et al.,* 2000b). Based on our earlier experience in the the lab (Agrawal 2006) and present study, we strongly believe that most of the above contradictions are probably the fall out of a relatively poor experimental strategy and the apparent differences may not be real. Some of the factors that probably led to these apparent failures are: 1) limited analysis targeting only the most conserved domains; 2) no effort or strategy to explore the transcriptional diversity (that is emerging as an integral feature of the major regulatory eukaryotic genes including the ones inlvoved in sex-determination); 3) a deficient expression analysis using sub-optimised conditions of RT-PCR/RACE etc.; and 4) a general lack of interest to isolate/study the complete gene/transcripts. This conclusion gets some support from the above studies itself, wherein no sex-specificity was observed in many studies, but none such study revealed a reverse sex-specificity i.e. male-gonad specific expression.

As discussed above, a number of studies have shown apparent differences in the temporal and spatial pattern of Aromatase expression and activity across taxa. However, the results obtained in the present study for the *C. palustris* are much in accordance with the other TSD species where Aromatase expression becomes higher during the later stages of TSP only at FPT. The observed pattern of *Aromatase* expression in all such studies (including the present one with *cpAromatase*), strongly suggest that it probably is not involved in the initial decision-making process; however it may be an important key player involved in the regulation of steroidogenic pathway that becomes operational probably immediatly downstream of the temperature-sensing machinery for female gonad development. It is also noteworthy that studies in TSD species have shown *Aromatase* mRNA in brain (as seen here for *cpAromatase***b**), accompanied with high oestrogen synthesis in the brain at FPT (Jeyasuria & Place 1998b). Such studies considered together with ones involving explants cultures suggest that it is the gonadal endogenous oestrogens (and not the extragonadal oestrogens contained in the egg yolk or synthesized in brain) that is involved in TSD (Pieau & Dorizzi 2004; Elf 2003).

Expression of isoforms *cpAromatase***a** and *cpAromatase***b** in GAM contributed to higher expression levels in GAM compared to brain. Estrogen is essential in females for the development of reproductive organs, especially in TSD species where hormone play a very important role in the early sex determination/differentiation process (Pieau & Dorizzi

2004). Presence of multiple isoforms and tissue specific expression of isoforms has been reported in several vertebrate species. i.e. in pigs there are separate embryo, placenta and ovary specific isoforms (Corbin *et al.,* 1995). In humans where Aromatase expression is recorded in at least ten tissues, Aromatase expression in each tissue is regulated by tissue specific promoter and first non-coding exon (Kamat *et al.,* 2002). However in certain conditions like breast cancer (prototype estrogen-dependent malignancy) Aromatase expression is misregulated and takes advantage of four promoters (Bulun *et al.,* 2003). Similarly expression of both isoforms in the female GAM tissues may be to meet the higher demand of the estrogen. Moreover the isoform *cpAromatase***b** expressing in brain lacks myristoylation and phosphorylation sites (present in the bigger isoform), which are required for trafficking of Aromatase protein. This probably indicates lesser need of Aromatase activity in brain during TSP or a different kind function/regulation of *cpAromatase* in brain.

5.1.3.4 Aromatase localization in GAM is conserved

Immunolocalization of the cpAromatase didn't reveal any expression in the GAM tissues at MPT, but the antibody detected its expression since stage 23rd in the entire cortical regions of the developing female gonads 23rd stage onwards without any cell-type specificity. A similar localization pattern of Aromatase in the entire cortical-medullary region of female gonads increasing during later stages of TSP has recently been described in the TSD species *Trachemys scripta* (Ramsey *et al.,* 2007).

Expression of a gene associated with male/female gonad determination/differentiation indicates the fate of the morphologically indistinguishable developing gonad. *cpAromatase* expression from 23rd stage indicates commitment of the gonads to the ovarian fate which also implies sex is determined by 23rd stage. Studies in several TSD species have revealed high oestrogens level in the separated gonads in the beginning of TSP at FPT (Dorizzi *et al.,* 1991; Desvages *et al.,* 1993b; Nicolas *et al.,* 1979). Rise in the oestrogen levels closely and directly corresponds to the upregulation of the Aromatase, an observation that has been irrefutably confirmed by inhibition-experiments involving treatments with antioestrogens and Aromatase inhibitors prior to and /or during TSP (Elf *et al.,* 2002; Elf 2003). More over, detailed studies done in alligator and turtle have confirmed that gonad start developing by the 23rd stage at MPT and FPT and become cytologically distinct by 25-26th developmental stages (Smith C.A. & Joss 1993). At FPT the cortical region becomes thick and sex cords disappear (Schmahl *et al.,* 2003; Yao & Capel 2005), otherwise events like development of follicles in female gonad

development in TSD occurs by the time of hatching (Wibbels *et al.,* 1991; Wibbels *et al.,* 1993; Smith & Joss 1993). This narrows down the TSP to only about 12-15 days (stage 21 and 22 and part of 23rd), which makes it easy to look for the differentially expressed genes during sex determination. A previous study in the lab (Agrawal 2006) has also demonstrated that sex is decided by 23rd stage using *Sox9* as a marker. However, absence of Aromatase expression at MPT and expression at FPT makes it a very important molecular marker for the sexing and stage validation of the developing embryo. Taking lead from this study and the result from the previous study (Agrawal 2006) the differentially expressed genes in the GAM were only screened during 21st and 22nd stage (see the following section).

5.2 Search for novel differentially expressed genes in the GAM tissues of *C. palustris* during TSP

In recent years, a large number of studies have amply demonstrated the importance of differential screens leading to the identification of important candidate genes in several biological contexts; to mention a few *e.g.,* genes involved in the cell cycle control (Okamoto & Beach 1994), *p53* network (Okamoto *et al.,* 2002) or induction of cancer (Okamura *et al.,* 2001). *Cyclin G1* as *p53* downstream target was originally discovered using differential display (Okamoto & Beach 1994) and later found to be a very important regulator of *p53* and *mdm2* network, involving dephosphorylation of *mdm2* by recruiting *PP2A* (protein phosphatase 2A) leading to *p53* inhibition (Okamoto *et al.,* 2002). Similarly, several other genes, *p53DINP1*, *P53R2*, *p53AIP1,* all having important roles in mediating the p53 dependent apoptosis, were also identified in various differential screens (Okamura *et al.,* 2001). These and similar examples in other cellular contexts strongly emphasize that global transcriptional screens provide very effective means to search for differentially expressed novel gene candidate(s) especially in completely unexplored contexts. Nevertheless, no such attempt for the identification of differentially expressed genes that may be involved in TSD has been reported in the published literature till to date, warranting studies to this end.

Therefore, as mentioned in the beginning, one of the main aims of the present study was to undertake global transcriptional screens to search for novel genetic factors/candidate(s) that are differentially expressed in the developing GAM tissue of putative female and male embryos; and may have an important role to play as the

master regulator(s) of TSD and/or in temperature sensing and deciding the fate of the developing bipotential gonads and their subsequent differentiation. Accordingly, in this study, an attempt was made to identify/isolate genes differentially expressed in GAM during the early TSP stages (*i.e.,* 21st and 22nd) using SSH PCR strategy. Briefly, the analysis resulted in a number of clones which were partially characterized by sequencing/sequence comparisons, followed by experimental validation of a few clones for their differentially expression. The study, thus, revealed a number of novel candidates, of which one promising clone was selected and pursued for detailed characterization leading to the identification of a novel ncRNA having a putative important role in male gonad development/TSD.

5.2.1 SSH: an amenable screen to search for differentially expressed gene(s)

In our lab few years back, a preliminary attempt was made to search for novel differentially expressed candidate genes (having possible role in TSD) by comparing the transcriptomes of GAM tissues of embryos kept at MPT and FPT, using the approach of DD-PCR (differential-display PCR) (Liang & Pardee 1992). The study, although amply demonstrated the potential of the DD-PCR based differential screen to reveal novel gene candidates, the overall success and efficiency of the approach was found to be limited mainly because of very high proportion of the false positives (Agrawal 2006). Therefore, in the present study, the attempt was made using the more recent approach of SSH hybridization (cDNA subtraction using suppression PCR) that has been demonstrated to be relatively efficient in reducing the number of the false positives (Diatchenko *et al.,* 1999). The SSH screen resulted in a number of candidate genes, few of which were found to be promising when validated for their differential expression in the GAM of male and female embryos through TSP (discussed in the following sections).

Microarray technology provides another powerful approach to study/identify the global changes in the transcriptomes, but can only be employed in systems where transcriptome is known and high-resolution cDNA/oligo arrays are available. Therefore, microarray approach cannot be used for model systems like *C. palustris,* and for that matter, any of the reptilian/TSD species, where requisite genomic resources (transcriptome/cDNA chips) are presently not available. Moreover, recent studies suggest that despite larger representation of genes, a cDNA/oligo array does not comprehensively cover the transcriptome of the target cell/tissue, and thus may leave a significant number of genes out of the scope of the investigation. Interestingly, in a case study comparing the relative efficiencies of SSH approach (selected for the present

study) to microarray analysis, 25 to 30% of putative differentials identified in the SSH screens were found missing in the Affymetrix gene chip with >10% of them being completely novel transcripts. The study, thus, suggested that SSH approach may provide relatively better potential for novel gene discovery (Cao *et al.,* 2004). Furthermore, a number of recent studies have shown that non-coding RNA are important players of the transcription regulation controlling very important cellular processes (Zaratiegui *et al.,* 2007), which are generally left out of the scope of microarray analysis. Contrary to the later, such novel RNA targets are expected to be amenable to an approach like SSH, as also demonstrated in the present study (see following sections).

5.2.2 Differentially expressed gene candidate(s) identified using SSH approach

The expression data from the *cpAromatase* in the present study and *cpSox9* in an earlier study (Agrawal 2006) have strongly suggested that the sex-determination in Indian mugger occurs during the early TSP stages prior to the 23rd developmental stage. Hence, GAM transcriptomes of only the early TSP stages (21+22 developmental stages) were compared by SSH approach to search for the novel candidate genes differentially expressed in the male GAM (used as the tester transcriptome) at MTP compared to that of female GAM (used as the driver transcriptome) at FTP.

SSH subtraction was done in two rounds, starting with the double stranded RT-PCR amplified cDNA pool as source starting materials in place of RNA pools that were available limiting amounts. Furthermore, the two rounds of subtractions were performed as per kits instructions but at relatively relaxed conditions of hybridization so that the transcripts that may have even small differences in their expression levels at MPT and FPT could also be detected; albeit the conditions also meant possibilities of a larger numbers of false positives. The subtraction, thus achieved, resulted in a small library of ~700 clones, which on sequencing followed by BLAST based homology search indicated a total of 167 unigenes. Although, these unigenes represented the functional diversity expected of an developing/differentiating tissue, interestingly a reasonably large proportion (around 22%) of these were revealed to be putative novel genes having no significant homology in the database. Similarly, an almost equal proportion represented important genes like transcription factors, regulatory elements etc. and some with regulatory class of proteins. On the other hand, some of the selected fragments (that were present in higher proportion of the total 700 sequenced clones) showed homology with metabolic or mitochondrial genes, which is rather expected as the overall

metabolism /developmental growth generally occurs at higher rate in the putative male embryos at MTP.

The validation studies (see the following section) suggested the presence of a large proportion of selected genes to be rather false positives. Although, this outcome was at variance to our initial expectations, but not quite unexpected, as: 1) the hybridisation conditions used in the two rounds of subtractions were slightly relaxed; and, 2) the source of the compared transcriptome, *i.e.* GAM represented a highly complex tissue comprising mainly of very large mesonephros (and other tissue such as adrenal) and only a small representation from genital ridge –the real target; making the identification/selection of novel differentially expressed transcripts therein a daunting challenge. One other very important attribute of the expressed genes that can crucially constrain the efficient selection of novel differentially expressed gene(s) in any global screen like SSH, is them being the context-specific isoforms resulting by the alternative splicing (rather than be an altogether different gene). Testis is well known for the high occurrence of alternatively spliced isoforms (Huang *et al.,* 2005b). Presence of sex-specific isoform in male GAM along with other isoforms expressed in both sexes (male/female GAM) have been documented in the Indian mugger for *cpSox9* (Agrawal 2006), and even in the present study, for *cpWT1*. Presence of such sex-specific alternatively spliced isoforms of an otherwise important gene, may make them difficult to be detected by differential global screening approaches. In addition, use of genital ridge alone as a source tissue for transcriptome screening may provide a better option for comparative gene expression studies related to TSD.

5.2.3 Validation of subtracted clones/sequences identified in the SSH screen for their differential expression in GAM tissues

The limitations of resources, especially, the availability of male/female GAM RNA in sufficient quantities, made it impossible to undertake validation studies on all the above-referred 167 subtracted/selected unigenes. Preliminary validation studies, that could be done only for 23 subtracted clones (generally represented genes like transcription factors, regulatory elements), revealed at least four of them to be differentially expressed at higher levels in the male GAM. Reverse Northern approach, whereby, several genes can be analyzed for their differential expression in a single experiment without spending a huge amount of the target RNA samples, was found to be of lesser utility than the semi-quantitative RT-PCR approach, due to its poor sensitivity to reflect the subtle differences in the expression levels of the tested genes. The approach could clearly

reveal only one gene/clone 'a4r' showing differential male GAM specific expression, but failed to identify the other three clones/genes (a16, d4, 2n11), which could only be validated by the RT-PCR that provided a more sensitive approach for validation studies. Compared to the experimental validation, all the 167 indicated genes could be tested computationally to predict their tissues specificity, if any, by comparisons with the tissue specific EST databases available in the public domain. *In silico* analysis revealed several interesting candidates including one (from four of the above 23) that was earlier experimentally validated to be specific to male GAM. It was interesting to observe that almost 50% of the 167 selected unigene sequences (obtained in the present study) showed similarity to the genes having brain and testis specific expression. This observation is significant as it probably suggests the possibility of: 1) a broad evolutionary conservance of genes that may have been functionally selected to play a role in sex-determination/ gonadal differentiation; and also 2) the importance of brain transcriptome in the process of sex-determination/differentiation; the aspects that may be interesting to explore further in the context of vertebrate sex-determination. The analysis also indicated few instances of possible domain sharing among the subtracted/selected cDNA clones of crocodile (see section 4.2.4). An interesting example was the selected sequence (b5f) which matched to *NR2F2* in the BLASTn, however it matched with the testis-specific gene *PKNOX2* (PBX/knotted-1 homeobox-2) along the same coordinates in the above computational analysis. The above (experimental/*in silico*) validation studies thus, revealed a number of potential new candidates, which may be worth pursuing in future studies for their potential role in male gonad development/TSD.

5.2.4 Isolation/characterization of a novel *mRNA-like-ncRNA* candidate gene having role in male gonad development (originally identified in the SSH screen as a differentially expressed clone a4r)

In validation studies discussed above, the EST-clone a4r was revealed as the most promising candidate that was found to be expressed at significantly higher levels in the GAM of embryos kept at MPT since the beginning of TSP. It was, thus, selected for in depth studies comprising, isolation of its full-length expressed transcript(s) followed by expression profiling through TSP in GAM and other embryonic tissues. These attempts led to the identification of a very interesting, completely novel *mRNA-like-ncRNA* candidate gene that is indicated to have a major role in the determination/development of male gonad in the Indian mugger. To our understanding, this is the first ever finding of a

novel regulatory RNA gene in the context of sex-determination in vertebrates irrespective of the mechanism being TSD or GSD.

5.2.4.1 A novel *mRNA-like-ncRNA* gene having an anti-sense counterpart

The 339 bp long clone A4r, identified in the SSH screen and later validated to be male GAM specific in its expression, has a 160 bp core matching to *SLF35F5* gene, a solute-carrier family gene which is conserved across vertebrates but whose function remains undefined. Validation studies further revealed a sex-specific expression of A4r only in GAM and not in any other embryonic tissue at MPT or FPT.

RACE reactions using 5`-/3`-end specific primers, which were designed considering the polarity of the 160 bp *SLF35F5* core, resulted in the isolation of a 2552 bp long full-length transcript of the a4r. This sequence did not show similarity to any database sequence except for the short 160 bp region that defined the core of the original a4r clone matching to *SLF35F5*; and attempts to amplify the whole sequence indicated it to be a transcript having feeble expression. However, semi-quantitative RT-PCR and real-time RT-PCR used to monitor its expression by amplifying a shorter region indicated otherwise. The baffling observation provided the first indication that probably both the strands are being transcribed; which turned out to be the case, as amplifications were seen in control RACE reactions that were set to check the fidelity of the end-specific RACE primers with both 5' and 3' RACE ready cDNA templates. Subsequently, the analysis involving extensive rounds of RACE, each time cloning/sequencing of RACE products, resulted in the identification of another transcript of 2695 bp (which was anti-sense to the first transcript); and a complete region of 2988 bp that defined the two near complementary transcripts sharing an overlap of 2249 bp.

In silico analysis of the two transcripts revealed these to carry all the hallmarks of non-coding RNA, *i.e.,* no evolutionary conservation and more importantly presence of ~50 stop codons with no putative protein translation bigger than 80 amino acids in each of the three possible reading frames. A number of recent reports have described non-coding RNAs that play important roles in various cellular processes, *e.g., Xist* (for X-chromosome inactivation), *Air* (for controlling genes like *Igfr2*, *Slc22a2* and *Slc22a3* genes), *Rox* in drosophila (for X-chromosome hyperactivation), *H19* (Imprinting control for *IGF2*), regulation of bi-thorax complex by intergenic transcription, (for review see (Prasanth & Spector 2007). In the beginning, it was thought that such novel ncRNAs represent a rare class of regulatory messages, but surprisingly recent studies

encompassing detailed analysis of eukaryotic transcriptome suggest otherwise. The recent computational studies have now revealed that eukaryotic genome is pervasively transcribed, which is quite contrary to the earlier conventional view that only 1-2% of genome is transcribed and codes for the protein coding RNA (Cheng *et al.,* 2005; Kapranov *et al.,* 2005; Birney *et al.,* 2007; Katayama *et al.,* 2005). Furthermore, such studies show that a lion's share of the transcriptome (65-70% of the vertebrate and invertebrate transcriptome) comprises ncRNA and most of it is made up of *mRNA-like-ncRNA* (Frith *et al.,* 2005; Carninci *et al.,* 2005).

As mentioned above, attempt to amplify the complete sense–strand transcript indicated it to be expressed at very lower level, as it could only be seen after 18 cycles of secondary PCR. Lesser expression of ncRNA like transcripts can sometime be due to transcriptional noise, which arises because of the genomic DNA contamination or unprocessed pre-mRNA (and internal priming of the pre-mRNA) and/or the unexplained section of transcriptome (Okazaki *et al.,* 2002). In the present study, such a possibility was ruled out by: amplifying a bigger region followed by sequencing based confirmation of the amplified expression product to be genuine transcript. RT-PCR and Northern blotting experiments also proved that there were two genuine transcripts (one expressing at much lower levels) with regulated expression at MPT and FPT only in GAM tissues.

The approach used in the present study to explore the antisense transcription has successfully been used in the several studies (Kapranov *et al.,* 2005). Infact, transcription from both strands represent one another property of the ncRNA (Ravasi *et al.,* 2006). Antisense transcription has been reported in case of several ncRNA that have a known function, *e.g., XIST* is one such known ncRNA that has an antisense transcript called *TSIX* which auto-regulates *Xist* and decides the X-chromosome to be inactivated (for review see (Thorvaldsen *et al.,* 2007). Similarly, *Air* is another ncRNA that is transcribed from the antisense strand of the gene it regulates (Sleutels *et al.,* 2002). Furthermore, most of the imprinted loci in the case of human, have been detected to have an antisense transcript or a ncRNA attached to it which brings about the silencing of one of the copies (Andersen & Panning 2003). The reports describing the presence of ncRNA (*mRNA-likencRNA*), antisense transcripts and their role had been sparse till recent past, till a massive cDNA sequencing effort (FANTOM3 project; (Maeda *et al.,* 2006) revealed about 180,000 genes of which only about 16,000 represented new protein coding transcripts and the remaining big proportion of the transcriptome was found to be mainly occupied by the ncRNA. Another astounding observation from this project was the revelation that ~70% of the mapped transcriptional units overlap to some

extent with transcript(s) which seems to represent the opposite strand. The sense and antisense transcripts were overlapping in the head to head manner or divergent manner, this kind of overlap among the cis sense/antisense pairs is most common in the mammalian transcriptome (Katayama *et al.,* 2005), however transcriptome of other species were not analyzed.

To our understanding, this is the first ever finding of a novel *mRNA-like-ncRNA* which has developing male gonad specific expression; however, regulated expression of *mRNA-like-ncRNA* has been reported in other biological context in other species. In *Drosophila*, a number of *mRNA-like-ncRNA* were predicted from cDNA library, and expression analysis of 35 candidates revealed that 27 of them had tissue-specific expression (Inagaki.S. *et al.,* 2005). Similarly, a large-scale study in mouse showed that ~70% of the 1602 suggested ncRNA (*mRNA-like-ncRNA*) were found to have tissue specific expression (Ravasi *et al.,* 2006). In both the studies, expression of *mRNA-like-ncRNA* was found during the crucial phases of the development of the tissues, suggesting these belong to the regulated component of the transcriptome.

5.2.4.2 Only antisense strand of the novel *mRNA-like-ncRNA* gene is transcribed at higher levels at MPT

First indication about the higher expression of antisense transcript of the *SLF35F5-like-ncRNA* at MPT came from the strand specific RT-PCR done with two different sets of primers. This was further confirmed by both Northern hybridization, as well as, real-time PCR wherein its expression was found to be similar/slightly more than that of GAPDH at MPT. In contrast, the sense strand specific transcript could not be detected both at MPT and FPT by Northern hybridization indicating very low levels of its expression. Subsequently, real-time PCR showed it to be expressed at very low levels that were about $1/30^{th}$ of the levels at which the GAPDH was expressed in GAM both at MPT/FPT. Thus, the analysis amply demonstrated that only the transcription of antisense strand of the novel ncRNA gene was regulated by the temperature of incubation, and thus it may represents an important candidate master regulator involved in TSD.

There are only few studies wherein the functional significance/regulatory role(s) of the ncRNA have been understood and described (few like *Air, Xist, H19* etc.), although many recent studies have shown their widespread and abundant presence in eukaryotic transcriptomes (see earlier section). Nonetheless, it has become clear from these studies that these novel ncRNAs regulate gene expression almost at every level *i.e.,*

from transcription, splicing, stability and maturation, translation to gene function. On the otherhand, antisense transcripts were initially thought to make sense-antisense pairs which were implicated in the modulation of the gene expression by dsRNA mediated gene silencing or methylation mediated silencing (Andersen & Panning 2003; Shamovsky & Nudler 2006; Pauler *et al.,* 2007). However, a separate type of regulation by sense-antisense transcripts (that we believe is the case with sense-antisense transcription of the *SLF35F5-like-ncRNA*) is seen for *IME4* in Yeast. *IME4* (initiator of meiosis) is one of the genes required for the entry of the dividing diploid cell to the meiosis mode, which appeared to be controlled by the switching of the transcription of *IME4* from sense to antisense strand or *vice-versa*. It is shown that the diploid state is marked by the sense transcription from *IME4* locus, while its antisense transcription marks the haploid state. Further, upregulation of the antisense *IME4* confers haploid like properties in the diploid cells and reverse is true for upregulation of the sense strand that could confer diploid like properties in the haploid cells. Thus, *IME4* represents an important example wherein antisense transcription decides the cell fate (Hongay *et al.,* 2006).

5.2.4.3 Increase in antisense transcription of the novel *mRNA-like-ncRNA* gene occurs only in the developing gonads not in the adjacent mesonephros

Localization studies carried out in the GAM tissues using a antisense transcript specific cDNA probe, clearly indicates that higher expression of novel *mRNA-like-ncRNA* occurs only in the developing gonads, and not in the adjacent mesonephros that comprises ~85% of the total GAM. These data, thus, suggest that the higher expression of the ncRNA recorded by RT-PCR/real-time PCR using GAM RNA as template source, is infact represents the increased antisense expression in the developing gonads that specify only 10-15% of the GAM.

Further, the ncRNA localization was observed to follow an interesting pattern through TSP in the GAM tissue suggesting, it plays a very important role in the development/differentiation of male gonad. During the early 21st and 22nd stages, when the gonads are bipotential, the *SLF35F5* like ncRNA localization was recorded in all the cells (without any cell type specificity) of the developing gonads at MPT, and with much lower levels at FPT. From 23rd stage onward, ncRNA localization in the developing gonads reveals appearance of sex cords that were otherwise not visible even by PI staining. This localization of ncRNA was comparable to that of *cpWt***b** (also analysed in this study; discussed in earlier sections) that is revealed to track the gonadal cells in

loosely organized sex cords by 23rd stage; but differed from *cpWt***b** at later stages of TSP as (unlike *cpWt***b)** it also localized inside the lumen of the developing sex-cords. Cellular events during male gonad differentiation encompassing emergence of the sex cords and their development in to testicular cords at MPT are well worked out in TSD species like turtle and alligator (Smith & Joss 1994b; Wibbels *et al.,* 1991). From these studies it becomes clear that proliferation of scattered pre-Sertoli cells occurs during early TSP stages, and subsequently developing Sertoli cells start organizing themselves in the testicular cords with germ cells inside their lumen by the middle of TSP (Yao *et al.,* 2004; Moreno-Mendoza *et al.,* 2001). Thus, the ncRNA localization data considered along with these studies, imply that *mRNA-like-ncRNA* expresses in both the cell types *i.e.,* Sertoli cells and Germ cells.

Overall, the expression and localization studies amply demonstrate that antisense strand of the newly identified mRNA-*like-ncRNA* expresses at significantly high levels from the very beginning of TSP mainly in the cells of bipotential gonads, and later in the differentiating male gonadal cells. These observations strongly suggest that antisense transcription of *mRNA-like-ncRNA* gene is very important for the initial sex determination/differentiation events. However, the machinery that regulates the transcription of the antisense strand under influence of temperature remains to be addressed in future studies.

5.2.4.4 *SLF35F5 mRNA-like-ncRNA* does not originate from the intronic transcription

Some of the ncRNA may have an intronic origin from a locus that codes for a protein coding mRNA (Ravasi *et al.,* 2006). The *mRNA-like-ncRNA* identified in the study has a 160 bp long core that matched with the *SLF35F5* coding region. Therefore, we tried to investigate its origin using PCR-based approaches, all of which ruled out its transcription from the *SLF35F5* locus, and also indicated that the identified *mRNA-like-ncRNA* is transcribed from another locus. In absence of the information about the *C. palustris* genome, although it is not possible for us to speculate about the distance between the two loci (the novel identified *mRNA-like-ncRNA* and the native *cpSLF35F5* locus), the data generated for the purpose could at least rule out the possibilities of the identified ncRNA to be of the intronic or internal priming origin. Furthermore, in the present study, the native homologue of *SLF35F5 i.e., cpSLF35F5* could be partially isolated, and the same was found to be upregulated only in GAM at FPT, also indicating that the two loci are differentially regulated.

The above studies also suggest that the 2988 bp long *mRNA-like-ncRNA* gene identified in this study is intron-less, a feature that resembles another general property of the ncRNAs which are shown to have shorter or no introns (Ravasi *et al.,* 2006). Presences of short conserved stretches (like the short 160 bp long *SLF35F5* matching core of the identified *mRNA-like-ncRNA*) have also been observed for other ncRNA, like, *Xist, Air* etc., and importantly, all these have been ruled out to be the transcription product from the intronic region of the genes to which their short stretches were matching. Such shorter stretches may possibly arise because of duplications (of coding or non-coding region of DNA), as are commonly found in the human genome (Birney *et al.,* 2007).

5.2.4.5 *mRNA*-like-ncRNA transcripts do not harbour stem-loops precursors for miRNA

Recent studies have revealed another class of small novel RNA molecules *i.e,* miRNA, as important regulators of gene expression in diverse biological contexts (Flynt & Lai 2008). Also, in some bacteria, it has been shown that mRNA can sense temperature because of changes in its secondary structure at the 5' end, which could lead to their translation (Chowdhury *et al.,* 2006). Most of miRNA discovered so far has been found to be conserved across vertebrates (Lee *et al.,* 2007), which can be identified by conservative *in silico* approaches searching for stem-loops as miRNA precursors with limited success (Chaudhuri & Chatterjee 2007). Therefore, in this study, attempts were also made to analyze the secondary structure of sense and antisense transcripts of the *mRNA-like-ncRNA* to ascertain if they have any conserved stem-loop structure(s) that may code for a potential miRNA. The *in-silico* predictions to this end did not indicate any conserved stem-loops indicative of the potential miRNA(s) in the sense and antisense transcripts; moreover, no noticeable difference was seen in the overall folding pattern of the two transcripts both at MPT and FPT. The analysis thus largely discounts the possibility of miRNA involvement in ncRNA expression/function. Nonetheless, there may exist some non-conserved miRNA which can be identified from the stem loops based on criteria like, secondary structure, free energy and size etc. Bentwich and coworkers discovered at least 50 novel microRNAs which were not conserved though possessed all the features of the miRNA stem loops (Bentwich *et al.,* 2005). Similar stem loops are also observed in the secondary structures of the two strand specific transcripts of *mRNA-like-ncRNA,* however further experiments are required to accept or reject the possibility of these to be miRNA, if any.

5.3 Sex is determined before 23rd stage in *C. palustris*

Notwithstanding, male sex specific expression of *cpWt1, cpDmrt1* and *mRNA-like-ncRNA* from the very beginning of the TSP (21st stage), many of the subtracted ESTs, (like subtractive library clones 2n11 and d4) were indicated to have male GAM specific expression only after 22nd stage. This expression pattern is similar to the expression of a male specific isoform of *cpSox9* (Agrawal 2006), and also to the female specific *cpAromatase***b** isoform (present study) in the GAM from 23rd stage onwards; which signifies the onset of male/female gonad development, respectively. Taken together, the above expression patterns strongly suggest that sex in Indian mugger is decided somewhere before 23rd stage, as also concluded earlier (Agrawal 2006). This time point of sex determination further corroborates well with the cytological events (appearance of testicular cords at MPT by 23rd stage, present study) of the developing gonad, as well as, egg shifting field experiments that also suggest that sex is decided by 23rd stage (Lang *et al.,* 1989). In light of the above, it is suggested that TSP in Indian mugger span 12-15 days (instead of 22-30 days) corresponding to 21-23 developmental stages at MPT and FPT, respectively.

5.4 Perspective

As becomes clear from the forgoing, both the identified aims of the present study *i.e,* characterization of crocodile homologues of three of the known sex-determination related genes, as well as, search for few new putative candidate gene(s) that may have a role in TSD, were undertaken with rigor and (to our belief) were completed successfully. In this effort, the study not only led to the enumeration of most comprehensive accounts of *cpWt1*, *cpDmrt1*, *cpAromatase* gene homologues from the TSD species Indian mugger, but also was successful in the identification/characterization of a completely novel candidate gene indicated to be an important player in the TSD. The detailed studies also provided a number of novel and interesting insights related to the expression patterns of the tested genes through TSP, and many more potential novel candidates differentially expressed in gonads at MPT and FPT. Overall, to our understanding, the study has been successful in creating a knowledge base, as well as, new genomic resources, which are expected to be very useful for unravelling the TSD in future.

Identification and characterization of *cpWt1* (coding locus and transcriptional diversity) has provided many new attributes (novel isoform, sex-specific expression etc.) so far undocumented for *Wt1* in a TSD species. The findings, thus, provide an altogether new perspective of the *cpWt1* complexity and its possible role in sex determination/differentiation events. Specifically, it highlights the need for extending studies encompassing much more critical appraisal of *cpWt1* function, especially because the present study in a TSD species for the first time provides evidence that it can induce expression of many male favouring genes. The novel isoform ***cpWt1*b** indicated to have this capability, may not act as transcription factor as it lacks the DNA binding region. All the results for *cpWt1*, from homology analysis to sex specific expression of novel isoform, indicate a different and early role for *cpWt1* in TSD. It may be interesting to explore the significance of 108 unique amino acids at the N-terminal, which may act as unique trans-activation/interacting domain in a reptilian specific manner in context of TSD. The trans-activation property of this region probably can be addressed by making deletion constructs and testing the same in the primary cell lines of crocodile (now available in CCMB) to see how the change spanning a specific part of the 108 aa region may effect the expression of other sex determining genes. Another challenge for the future studies may be addressing function of male sex specific ***cpWt1*b**, with the added complexity that it may not directly act as transcription factor. Studies involving ablation of the expression to identify the downstream targets, search of interacting partners/genes regulating its expression to identify upstream effectors, may help in understanding the cell type specific transcription of novel sex-specific isoform ***cpWt1*b** and its role in the cell fate determination.

Non-availability of an amenable model system (that can be genetically manipulated) for functional studies is one of the biggest constrain for studying TSD. However, establishment of primary cell lines can meet this long standing demand to a certain extent, especially, if these can survive for long time as observed for the crocodilian primary cell lines established at CCMB. Such cell lines provide the possibility of undertaking global differential screens to identify the differentially expressed gene(s) that may be modulated/influenced by the over expression and/or deletion (knock out) of ***cpWt1*b**. The results are expected to reveal the downstream targets of ***cpWt1*b**, Here, it may be emphasized that the cDNA subtraction based global screens (as used in present study) provide quite efficient screens for identification of differentials (in an unbiased way), even better than gene discovery using genechip technology (Cao *et al.,* 2004). Similarly, studies like immuno-precipitation followed by Mass spectroscopy can be done for identification of the interacting genes.

Other significant insight from the present study is the prediction of binding sites for a number of important regulatory elements like, HSF1 and AR/PR/ER in the 5`-upstream region of *cpWt1*, which can be directly tested after standardization/production of antibodies to detect crocodilian HSF1 and AR (androgen receptor). Chromatin-immunoprecipitation (ChIP) can be done to examine whether these proteins bind to the *cpWt1* promoter or not. Importantly the ChIPs can be done in both using cell lines and in tissues, and experiments can be designed to understand the binding *vis-a-vis* transcription of *cpWt1***b**. It will be really very interesting to examine, whether HSf1 has any role in the *cpWt1***b** regulation.

In mammals, a great deal of information is available about *Wt1* function, and one of its isoform *Wt1+KTS* is shown to have an early and important role in the male sex determination; therefore, it would be interesting to examine some of these functions even in TSD. It is now well documented that *WT1* expression starts early within the undifferentiated gonad in mammals, and it's knock-outs (*Wt1–/–*) in mice undergo gonadal apoptosis along with atrophy of adrenal glands (Hatano *et al.,* 1996; Moore *et al.,* 1999). This phenotype is also seen in homozygous *Sf1*-knockout animals, which also lack gonads and adrenals owing to massive apoptosis in the adreno-genital primordium (Luo *et al.,* 1994). Moreover, recently it is demonstrated that *WT1*: can activate the *Sf1* promoter both *in vitro* and *in vivo*; is responsible for transcriptional activation of *Sry* (Wilhelm & Englert 2002) and gonads in mice lacking *WT1 (+KTS)* lack *Sox9*, as well as *AMH* expression (Hammes *et al.,* 2001). Other potential targets during gonad formation include *Dax1* and *MIS,* that seem to be activated in *WT1(–KTS)* variants (Kim *et al.,* 1999; Hammes *et al.,* 2001). Considering all these reports, the observations made in the present study assumes even more significance, which show *cpSox9* and *cpAmh* and *cpSf1* as possible targets of *cpWt1***b**. However, since *cpWt1***b** is probably not a transcription factor (as it lacks zinc finger domains), it is expected that it probably causes transcriptional activation of the *cpSf1, cpSox9* and *cpAmh* by acting as a trans-activator, an aspect that need to be explored by looking for its interacting partners.

The present study also suggested that the second longer isoform *cpWt1***a** may be an important factor needed for the cell survival, which probably can be tested by knockdown experiments. In case of mouse also, *WT1* is speculated to activate a *Sf1*-independent pathway for adrenogenital survival that probably include the *WT1* activated antiapoptotic factor *Bcl-2* (Mayo *et al.,* 1999). Overall, the study reveals several interesting features of *cpWt1* many of which may be unique to reptiles, but it also leaves many excitements to follow.

Studies on *cpDmrt1* for the first time revealed extensive alternative splicing and characterization of *cpDmrt1* locus in a TSD species; however the most important finding was the identification of the most abundant isoform *cpDmrt1***e**, which carried a unique 45 bases near its 3' end. We expect this unique region to be central to the higher expression of this novel isoform. The putative translation of this unique region defines a 15 aa domain that is predicted to harbour PIKK directed phosphorylation domains. Therefore, in our opinion studies to identify the interacting proteins to this novel domain may hold the key for *cpDmrt1* function in TSD.

Dmrt1 is a conserved gene suggested to be important in the sex determination and early gonad differentiation in vertebrates. Across the evolutionary strata, it is known to play varying role i.e. from being a master regulator of male sex in teleost Medaka (Matsuda *et al.,* 2002) to an important sex regulator on the Z-chromosome of chicken (Shan *et al.,* 2000) and to a gene involved in embryonic testis differentiation in mammals (Raymond *et al.,* 1998). Recent studies have shown significant transcriptional diversity for this important gene during gonadal differentiation in many GSD species (Huang *et al.,* 2005a; Guo *et al.,* 2005; Cheng *et al.,* 2006; Lu *et al.,* 2007; Zhao *et al.,* 2007). Based on these studies it is speculated that probably different isoforms are central to the varied function of *Dmrt1* (cell autonomous and non-cell autonomous roles) observed in context of male gonad development (Kim *et al.,* 2007a). Hence, it may be pertinent to undertake detailed localization/expression studies to characterize individual isoforms during gonadogenesis in TSD.

The other important outcome of the present study is the identification of a number of novel genes that show differential expression in the GAM tissue of male/female embryos. Some of these could be validated in preliminary studies suggesting them to be putative candidates having role in TSD; and it would be very interesting to characterize each of these putative new candidates individually in detail like, *cpDmrt1* and *cpWt1* in this study, to understand their role in TSD, if any. The another important discovery of this effort was the identification of the novel candidate gene, *SLF35F5 mRNA-like-ncRNA*; detailed characterization of which revealed it to be an very important player akin to a master candidate gene involved in the temperature dependent sex-determination. This important novel gene (and its transcribed sense/anti-sense strands) need to be explored in detail for ascertaining their exact biological significance by undertaking functional studies involving over-expression studies, subcellular localization etc., which can be initially addressed in the primary cell lines.

In nutshell, the present study has revealed many small cogs of the machinery that governs TSD, and it would be interesting to see if (and how) these are related (i.e. as upstream, downstream or interacting partners) to each other. It is expected that the leads obtained in this study would give the desired impetus to more exhaustive studies on TSD and thus may help solve this long standing enigma of the developmental biology.

5.5 A plausible model to explain underlying mechanism of TSD

Based on some of the new significant leads from the present study (considered in light of the published knowledge about sex-determination mechanisms), a working model indicating the major components of underlying molecular mechanism of TSD (figure 5.2), is proposed; which if validated, would be a singular contribution of this study. Some of the important findings (from this study) and the published knowledge that are considered to develop this model are as follow:

1. *Identification of novel isoforms of cpWt1 and cpDmrt1 that are expressed from the very beginning of TSP at significantly higher levels in the putative male GAMs;*

2. *A novel isoform cpWt1***b** can lead to transcriptional activation of many male sex-related genes;

3. *A close correspondence between the localization of the above novel isoforms with cellular events occurring during early indeterminate TSP stages and later during the development/differentiation of male gonad from mid to late TSP;*

4. *Identification of a novel mRNA-like-ncRNA showing sex-specific expression in GAM tissues from the very beginning of TSP at MTP;*

5. *Identification of regulatory elements/binding sites for factors like HSF-1 in the 5`-upstream regions of cpWt1;*

6. *Alternate splicing central to the transcriptional diversity of important regulatory genes (this study; and recent literature (Huang et al., 2005b; Okazaki et al., 2002; Brett et al., 2002; Ben-Dov et al., 2008), including ones involved in GSD seen in drosophila to humans (Lalli et al., 2003; Forch & Valcarcel 2003)*

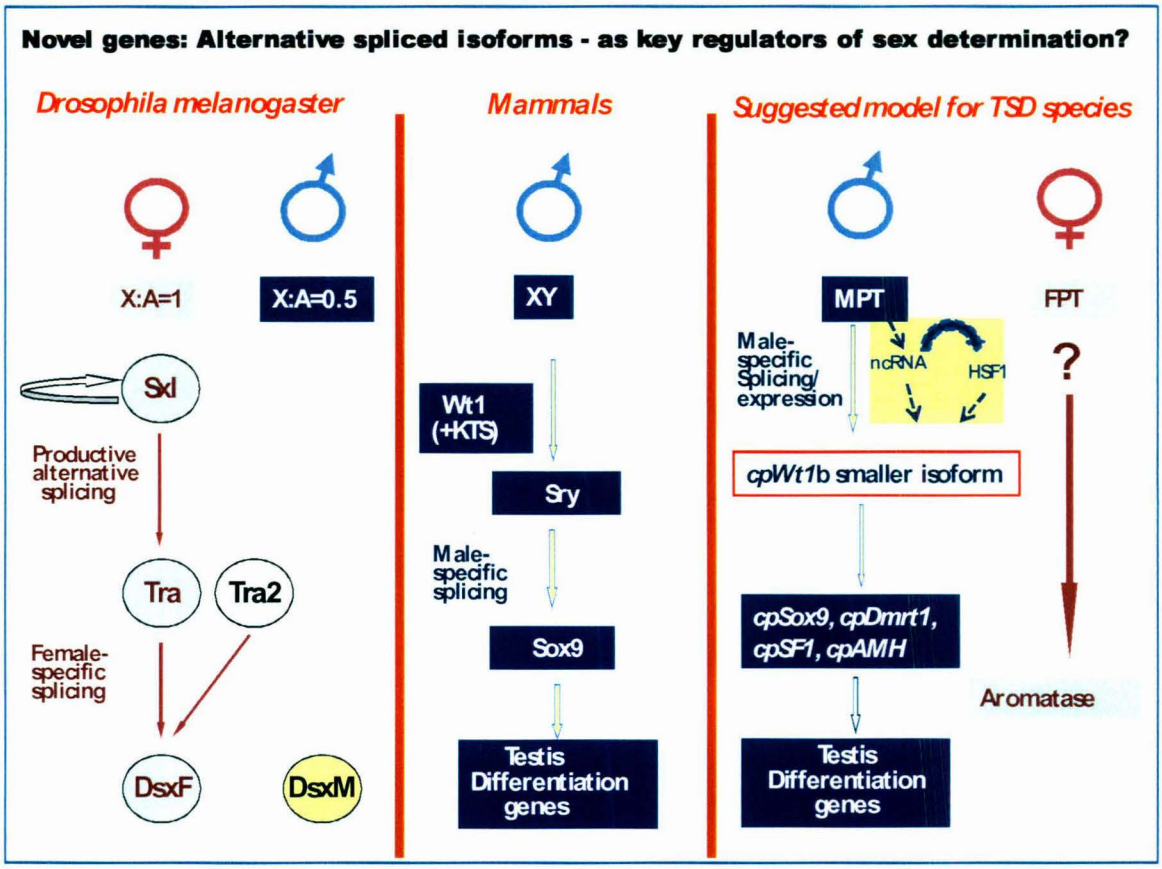


Figure 5.2: Proposed molecular mechanism underlying TSD based on the findings of the present study *vis-a-vis* other sex determining mechanisms. In this model proposed here, female sex as the default sex, and generation of sex-specific novel transcripts (by alternate splicing) are considered integral to the sex-determination process as has been suggested by many published and also our own studies in CCMB. Furthermore, considering the conserved role of heat shock factor-1 and recently discovered role of an *mRNA-like-ncRNA* in heat shock response (Shamovsky *et al.*, 2006), we speculate that:

a novel regulatory RNA like the ncRNA (identified/suggested to play an important role in TSD in the present study) interacts with HSF1 factor(s) (yet to be identified), and together constitute the temperature sensing machinery underlying TSD; furthermore, the complex directs the generation of sex specific isoforms and/or regulates the sex-specific expression of certain isoform of a nodal gene, thus affecting the sex-determination (generally towards the male sex). Our study suggests this nodal gene (downstream of HSF1-ncRNA) to be the cpWt1, which is found to harbour many HSF1 binding sites in its upstream region; moreover, one of its novel isoform cpWt1b expresses at significantly higher levels from the very beginning of TSP in the GAM tissues; most importantly, this isoform is also suggested to be upstream of many early male favouring genes and can induce their expression.

7. *The recent elegant demonstration that perception of heat shock signal evoking subsequent response of the cellular machinery can be mediated by mRNA-like-ncRNA and HSF1 interactions (Shamovsky et al., 2006).*

8. *Important role for WT1, Dmrt1, Sox9 genes/isoforms to varying extent in GSD during vertebrate evolution (Wagner et al., 2003; Hastie 2001; Ferguson-Smith 2007; Guo et al., 2002; Kanai et al., 2005; Koopman 2005; Yao & Capel 2005)*

9. *SRY* (the accepted master regulator) and other HMG-box containing SOX6/SOX9 genes play a direct role in pre-mRNA splicing (Ohe *et al.*, 2002).

Collating the above, we propose that: a novel regulatory RNA like the ncRNA *(identified/suggested to play an important role in TSD in the present study) might interact with HSF1 factor(s) (yet to be identified), and together constitute the temperature sensing machinery underlying TSD; furthermore, the complex directs the generation of sex specific isoforms and/or regulates the sex-specific expression of certain isoform of a nodal gene, thus affecting the sex-determination (generally towards the male sex). Based on the present work, we suggests this nodal gene (downstream of HSF1-ncRNA) to be the cpWt1, which is found to harbour many HSF1 binding sites in its upstream region; moreover, one of its novel isoform* cpWt1*b expresses at significantly higher levels from the very beginning of TSP in the GAM tissues; most importantly, this isoform is also suggested to be upstream of many early male favouring genes and can induce their expression. Furthermore, we would also like to suggest the novel isoform cpDmrt1***e*** *(showing high expression from early TSP in GAM at MTP) as another possible candidate for this nodal gene role.*

5.6 Conclusions

The present study successfully describes the transcriptional diversity, expression patterning, important novel isoforms and coding regions of the crocodilian homologues of three important sex-related genes *Wt1, Dmrt1* and *Aromatase* for the first time in TSD species. In addition, it has revealed a number of novel candidate genes by global comparisons of the GAM transcriptome at the MPT and FPT, which appear to be

differentially regulated under the influence of temperature and thus, may have a role in TSD. Most importantly, the search for novel candidates has led to the identification/characterization of a completely new *mRNA-like-ncRNA* that appears to play a very crucial role in TSD; to our understanding this is the first ever discovery of a novel ncRNA in the context of sex determination of any kind.

Study describes *cpWt1* coding region and two isoforms of *cpWt1* arising from alternative splicing. The *cpWt1* showed considerable sequence diversity in the 5' region, homology analysis of the region showed possibility of exon shuffling during vertebrate evolution. Among the two isoforms, *cpWt1***b** lacked a part of conserved WT1 domain and all four zinc finger domains, present in the *cpWt1***a**. Isoform *cpWt1***b** has a novel domain (probably reptilian specific) and showed sex specific expression in the male GAM (which makes it the only known *Wt1* isoform having early sex specific expression). The localization of *cpWt1***b** tracked the events during male sex determination/differentiation, and showed specific localization to the testicular cords (most likely in the pre-Sertoli cells, Sertoli cells and PM cells). The transfection studies in the primary cell lines show that *cpWt1***b** expression can induce expression of male promoting genes, and importantly genes associated with Sertoli cells development and function, which strengthens its pivotal role in male gonad determination in TSD.

Compared to *cpWt1*, more extensive transcriptional diversity was revealed in case of *cpDmrt1* in the GAM tissue; subsequent characterization of ~ 3.7 kb coding region of *cpDmrt1* explained the origin of eight isoforms by various mechanisms of alternative splicing. The isoforms showed varying degree of sequence similarity with the other vertebrate *Dmrt1*. Expression profiling showed an overall male specific higher expression of *cpDmrt1* isoforms, and revealed its novel isoform, *cpDmrt1***e** to be the most abundant from the beginning of TSP. Interestingly, *cpDmrt1***e** was completely similar to the most conserved isoforms (*cpDmrt1***a2)** but carried a unique 45 bp region near its 3' end, which specifies 15 aa domain. The later appears to be reptilian specific and also harbours regulatory elements that probably underlie its significantly high expression (contributing >70% to overall *cpDmrt1* levels) during TSP.

Study also revealed two transcribed isoforms for *cpAromatase* showing sex-/stage-specific expression of the isoform *cpAromatase***a**, and only stage-specific expression for the novel isoform *cpAromatase***b** during TSP. Study also suggested a tissue-specific regulated expression of the two isoforms *i.e.* both isoforms in the female GAM and only one in brain (*cpAromatase***b)**. Partial upstream region of *cpAromatase* was also obtained

that carried regulatory signatures of *Sf1* and *COUP*, which are transcriptional factors. *cpAromatase* localization in the developing GAM was conserved with other TSD species, and suggests its potential as a good tracking marker for female gonadogenesis.

Another very important finding of the study was the identification of novel differentially expressed genes in the male GAM, most important of which was the discovery of an novel *mRNA-like-ncRNA* (*SLF35F5* like ncRNA). This novel candidate gene is revealed to be transcribed from both strands (more from the anti-sense strand). Interestingly, expression of its anti-sense strand was found to be regulated during TSP, as it showed overwhelmingly high expression only in the developing gonads/GAM at MPT. Moreover, the ncRNA localization in the GAM during TSP tracks the development of testicular cords right from 23rd stage and its expression remained restricted to the testicular cords only. Importantly, its localization reveals the development of the male gonads from indeterminate stage till development of testicular cords and interstitium (the later does not express this ncRNA). The present study tracking expression of candidate and novel sex related genes, and study of developing gonads at cellular levels, narrows down the TSP from 25-30 days to 10-12 days.

It is expected that the leads obtained in this study would give the desired impetus to more exhaustive studies on TSD and thus may help solve this long standing enigma of the developmental biology. Further, based on some of the new significant leads from the present study (and recent published knowledge about eukaryotic transcriptome and sex-determination mechanisms), we have proposed a plausible underlying molecular mechanism of TSD, which if validated, would be a singular contribution of this study.

Bibliography

Ackerman, R. A. (1981a). Growth and gas exchange of embryonic sea turtle (*chelonia, caretaa). Coepia,* 757-765.

Ackerman, R. A. (1981b). oxygen consumption by sea turtle (*chelonia, Caretta*) eggs during development. *Physiol.Zool.,* 54, 316-324.

Agoulnik, I. U., Krause, W. C., Bingman, W. E., III, Rahman, H. T., Amrikachi, M., Ayala, G. E. & Weigel, N. L. (2003). Repressors of androgen and progesterone receptor action. *J Biol Chem.,* 278, 31136-31148.

Agrawal, R. Expression analysis of Sox9 and search for sex specific genes in developing gonads of Crocodilus palustris: a species with temperature dependent sex determiantion. 2006. Ref Type: Thesis/Dissertation

Akerfelt, M., Trouillet, D., Mezger, V. & Sistonen, L. (2007). Heat shock factors at a crossroad between stress and development. *Ann.N.Y.Acad.Sci.,* 1113, 15-27.

Albrecht, K. H. & Eicher, E. M. (2001). Evidence that Sry is expressed in pre-Sertoli cells and Sertoli and granulosa cells have a common precursor. *Dev.Biol,* 240, 92-107.

Andersen, A. A. & Panning, B. (2003). Epigenetic gene regulation by noncoding RNAs. *Curr.Opin.Cell Biol.,* 15, 281-289.

Anderson, R., Garcia-Castro, M., Heasman, J. & Wylie, C. (1998). Early stages in male germ cell differentiation in the mouse. Review article. *APMIS,* 106, 127-132.

Arango, N. A., Lovell-Badge, R. & Behringer, R. R. (1999). Targeted mutagenesis of the endogenous mouse Mis gene promoter: in vivo definition of genetic pathways of vertebrate sexual development. *Cell,* 99, 409-419.

Bardoni, B., Zanaria, E., Guioli, S., Floridia, G., Worley, K. C., Tonini, G., Ferrante, E., Chiumello, G., McCabe, E. R., Fraccaro, M. & . (1994). A dosage sensitive locus at chromosome Xp21 is involved in male to female sex reversal. *Nat Genet,* 7, 497-501.

Bayrer, J. R., Zhang, W. & Weiss, M. A. (2005). Dimerization of doublesex is mediated by a cryptic ubiquitin-associated domain fold: implication for sex-specific gene regulation. *J.Biol.Chem.,* 280, 32989-32996.

Behringer, R. R., Finegold, M. J. & Cate, R. L. (1994). Mullerian-inhibiting substance function during mammalian sexual development. *Cell,* 79, 415-425.

Ben-Dov, C., Hartmann, B., Lundgren, J. & Valcarcel, J. (2008). Genome-wide analysis of alternative pre-mRNA splicing. *J Biol Chem.,* 283, 1229-1233.

Bentwich, I., Avniel, A., Karov, Y., Aharonov, R., Gilad, S., Barad, O., Barzilai, A., Einat, P., Einav, U., Meiri, E., Sharon, E., Spector, Y. & Bentwich, Z. (2005). Identification of hundreds of conserved and nonconserved human microRNAs. *Nat Genet,* 37, 766-770.

Bergeron, J. M., Gahr, M., Horan, K., Wibbels, T. & Crews, D. (1998). Cloning and in situ hybridization analysis of estrogen receptor in the developing gonad of the red-eared slider turtle, a species with temperature-dependent sex determination. *Dev.Growth Differ.,* 40, 243-254.

Berta, P., Hawkins, J. R., Sinclair, A. H., Taylor, A., Griffiths, B. L., Goodfellow, P. N. & Fellous, M. (1990). Genetic evidence equating SRY and the testis-determining factor. *Nature,* 348, 448-450.

Birk, O. S., Casiano, D. E., Wassif, C. A., Cogliati, T., Zhao, L., Zhao, Y., Grinberg, A., Huang, S., Kreidberg, J. A., Parker, K. L., Porter, F. D. & Westphal, H. (2000). The LIM homeobox gene Lhx9 is essential for mouse gonad formation. *Nature,* 403, 909-913.

Birney, E., Stamatoyannopoulos, J. A., Dutta, A., Guigo, R., Gingeras, T. R., Margulies, E. H., Weng, Z., Snyder, M., Dermitzakis, E. T., Thurman, R. E., Kuehn, M. S., Taylor, C. M., Neph, S., Koch, C. M., Asthana, S., Malhotra, A., Adzhubei, I., Greenbaum, J. A., Andrews, R. M., Flicek, P., Boyle, P. J., Cao, H., Carter, N. P., Clelland, G. K., Davis, S., Day, N., Dhami, P., Dillon, S.

C., Dorschner, M. O., Fiegler, H., Giresi, P. G., Goldy, J., Hawrylycz, M., Haydock, A., Humbert, R., James, K. D., Johnson, B. E., Johnson, E. M., Frum, T. T., Rosenzweig, E. R., Karnani, N., Lee, K., Lefebvre, G. C., Navas, P. A., Neri, F., Parker, S. C., Sabo, P. J., Sandstrom, R., Shafer, A., Vetrie, D., Weaver, M., Wilcox, S., Yu, M., Collins, F. S., Dekker, J., Lieb, J. D., Tullius, T. D., Crawford, G. E., Sunyaev, S., Noble, W. S., Dunham, I., Denoeud, F., Reymond, A., Kapranov, P., Rozowsky, J., Zheng, D., Castelo, R., Frankish, A., Harrow, J., Ghosh, S., Sandelin, A., Hofacker, I. L., Baertsch, R., Keefe, D., Dike, S., Cheng, J., Hirsch, H. A., Sekinger, E. A., Lagarde, J., Abril, J. F., Shahab, A., Flamm, C., Fried, C., Hackermuller, J., Hertel, J., Lindemeyer, M., Missal, K., Tanzer, A., Washietl, S., Korbel, J., Emanuelsson, O., Pedersen, J. S., Holroyd, N., Taylor, R., Swarbreck, D., Matthews, N., Dickson, M. C., Thomas, D. J., Weirauch, M. T., Gilbert, J., Drenkow, J., Bell, I., Zhao, X., Srinivasan, K. G., Sung, W. K., Ooi, H. S., Chiu, K. P., Foissac, S., Alioto, T., Brent, M., Pachter, L., Tress, M. L., Valencia, A., Choo, S. W., Choo, C. Y., Ucla, C., Manzano, C., Wyss, C., Cheung, E., Clark, T. G., Brown, J. B., Ganesh, M., Patel, S., Tammana, H., Chrast, J., Henrichsen, C. N., Kai, C., Kawai, J., Nagalakshmi, U., Wu, J., Lian, Z., Lian, J., Newburger, P., Zhang, X., Bickel, P., Mattick, J. S., Carninci, P., Hayashizaki, Y., Weissman, S., Hubbard, T., Myers, R. M., Rogers, J., Stadler, P. F., Lowe, T. M., Wei, C. L., Ruan, Y., Struhl, K., Gerstein, M., Antonarakis, S. E., Fu, Y., Green, E. D., Karaoz, U., Siepel, A., Taylor, J., Liefer, L. A., Wetterstrand, K. A., Good, P. J., Feingold, E. A., Guyer, M. S., Cooper, G. M., Asimenos, G., Dewey, C. N., Hou, M., Nikolaev, S., Montoya-Burgos, J. I., Loytynoja, A., Whelan, S., Pardi, F., Massingham, T., Huang, H., Zhang, N. R., Holmes, I., Mullikin, J. C., Ureta-Vidal, A., Paten, B., Seringhaus, M., Church, D., Rosenbloom, K., Kent, W. J., Stone, E. A., Batzoglou, S., Goldman, N., Hardison, R. C., Haussler, D., Miller, W., Sidow, A., Trinklein, N. D., Zhang, Z. D., Barrera, L., Stuart, R., King, D. C., Ameur, A., Enroth, S., Bieda, M. C., Kim, J., Bhinge, A. A., Jiang, N., Liu, J., Yao, F., Vega, V. B., Lee, C. W., Ng, P., Shahab, A., Yang, A., Moqtaderi, Z., Zhu, Z., Xu, X., Squazzo, S., Oberley, M. J., Inman, D., Singer, M. A., Richmond, T. A., Munn, K. J., Rada-Iglesias, A., Wallerman, O., Komorowski, J., Fowler, J. C., Couttet, P., Bruce, A. W., Dovey, O. M., Ellis, P. D., Langford, C. F., Nix, D. A., Euskirchen, G., Hartman, S., Urban, A. E., Kraus, P., Van, C. S., Heintzman, N., Kim, T. H., Wang, K., Qu, C., Hon, G., Luna, R., Glass, C. K., Rosenfeld, M. G., Aldred, S. F., Cooper, S. J., Halees, A., Lin, J. M., Shulha, H. P., Zhang, X., Xu, M., Haidar, J. N., Yu, Y., Ruan, Y., Iyer, V. R., Green, R. D., Wadelius, C., Farnham, P. J., Ren, B., Harte, R. A., Hinrichs, A. S., Trumbower, H. & Clawson, H. (2007). Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature,* 447, 799-816.

Bishop, C. E., Whitworth, D. J., Qin, Y., Agoulnik, A. I., Agoulnik, I. U., Harrison, W. R., Behringer, R. R. & Overbeek, P. A. (2000). A transgenic insertion upstream of sox9 is associated with dominant XX sex reversal in the mouse. *Nat Genet,* 26, 490-494.

Bollig, F., Mehringer, R., Perner, B., Hartung, C., Schafer, M., Schartl, M., Volff, J. N., Winkler, C. & Englert, C. (2006). Identification and comparative expression analysis of a second wt1 gene in zebrafish. *Dev.Dyn.,* 235, 554-561.

Bouma, G. J., Albrecht, K. H., Washburn, L. L., Recknagel, A. K., Churchill, G. A. & Eicher, E. M. (2005). Gonadal sex reversal in mutant Dax1 XY mice: a failure to upregulate Sox9 in pre-Sertoli cells. *Development,* 132, 3045-3054.

Bowles, J., Knight, D., Smith, C., Wilhelm, D., Richman, J., Mamiya, S., Yashiro, K., Chawengsaksophak, K., Wilson, M. J., Rossant, J., Hamada, H. & Koopman, P. (2006). Retinoid signaling determines germ cell fate in mice. *Science,* 312, 596-600.

Brennan, J., Tilmann, C. & Capel, B. (2003). Pdgfr-alpha mediates testis cord organization and fetal Leydig cell development in the XY gonad. *Genes Dev.,* 17, 800-810.

Brett, D., Pospisil, H., Valcarcel, J., Reich, J. & Bork, P. (2002). Alternative splicing and genome complexity. *Nat Genet,* 30, 29-30.

Buehr, M., Gu, S. & McLaren, A. (1993). Mesonephric contribution to testis differentiation in the fetal mouse. *Development,* 117, 273-281.

Bull, J. J. (1980). Sex determiantion in reptiles. *Q.Rev.Biol.,* 55, 2-21.

Bull, J. J. (1983). *Evolution of sex sex determining mechanisms*. Menlo Park, CA: Benjamin/Cummings.

Bulun, S. E., Sebastian, S., Takayama, K., Suzuki, T., Sasano, H. & Shozu, M. (2003). The human CYP19 (aromatase P450) gene: update on physiologic roles and genomic organization of promoters. *J Steroid Biochem.Mol Biol,* 86, 219-224.

Burns, R. K. (1961). in sex and internal secretions. In Williams & Wilkins (Eds) (pp. 76-156). Baltimore.

Cao, W., Epstein, C., Liu, H., DeLoughery, C., Ge, N., Lin, J., Diao, R., Cao, H., Long, F., Zhang, X., Chen, Y., Wright, P. S., Busch, S., Wenck, M., Wong, K., Saltzman, A. G., Tang, Z., Liu, L. & Zilberstein, A. (2004). Comparing gene discovery from Affymetrix GeneChip microarrays and Clontech PCR-select cDNA subtraction: a case study. *BMC.Genomics,* 5, 26.

Capel, B., Albrecht, K. H., Washburn, L. L. & Eicher, E. M. (1999). Migration of mesonephric cells into the mammalian gonad depends on Sry. *Mech.Dev.,* 84, 127-131.

Carninci, P., Kasukawa, T., Katayama, S., Gough, J., Frith, M. C., Maeda, N., Oyama, R., Ravasi, T., Lenhard, B., Wells, C., Kodzius, R., Shimokawa, K., Bajic, V. B., Brenner, S. E., Batalov, S., Forrest, A. R., Zavolan, M., Davis, M. J., Wilming, L. G., Aidinis, V., Allen, J. E., mbesi-Impiombato, A., Apweiler, R., Aturaliya, R. N., Bailey, T. L., Bansal, M., Baxter, L., Beisel, K. W., Bersano, T., Bono, H., Chalk, A. M., Chiu, K. P., Choudhary, V., Christoffels, A., Clutterbuck, D. R., Crowe, M. L., Dalla, E., Dalrymple, B. P., de, B. B., Della, G. G., di, B. D., Down, T., Engstrom, P., Fagiolini, M., Faulkner, G., Fletcher, C. F., Fukushima, T., Furuno, M., Futaki, S., Gariboldi, M., Georgii-Hemming, P., Gingeras, T. R., Gojobori, T., Green, R. E., Gustincich, S., Harbers, M., Hayashi, Y., Hensch, T. K., Hirokawa, N., Hill, D., Huminiecki, L., Iacono, M., Ikeo, K., Iwama, A., Ishikawa, T., Jakt, M., Kanapin, A., Katoh, M., Kawasawa, Y., Kelso, J., Kitamura, H., Kitano, H., Kollias, G., Krishnan, S. P., Kruger, A., Kummerfeld, S. K., Kurochkin, I. V., Lareau, L. F., Lazarevic, D., Lipovich, L., Liu, J., Liuni, S., McWilliam, S., Madan, B. M., Madera, M., Marchionni, L., Matsuda, H., Matsuzawa, S., Miki, H., Mignone, F., Miyake, S., Morris, K., Mottagui-Tabar, S., Mulder, N., Nakano, N., Nakauchi, H., Ng, P., Nilsson, R., Nishiguchi, S., Nishikawa, S., Nori, F., Ohara, O., Okazaki, Y., Orlando, V., Pang, K. C., Pavan, W. J., Pavesi, G., Pesole, G., Petrovsky, N., Piazza, S., Reed, J., Reid, J. F., Ring, B. Z., Ringwald, M., Rost, B., Ruan, Y., Salzberg, S. L., Sandelin, A., Schneider, C., Schonbach, C., Sekiguchi, K., Semple, C. A., Seno, S., Sessa, L., Sheng, Y., Shibata, Y., Shimada, H., Shimada, K., Silva, D., Sinclair, B., Sperling, S., Stupka, E., Sugiura, K., Sultana, R., Takenaka, Y., Taki, K., Tammoja, K., Tan, S. L., Tang, S., Taylor, M. S., Tegner, J., Teichmann, S. A., Ueda, H. R., van, N. E., Verardo, R., Wei, C. L., Yagi, K., Yamanishi, H., Zabarovsky, E., Zhu, S., Zimmer, A., Hide, W., Bult, C., Grimmond, S. M., Teasdale, R. D., Liu, E. T., Brusic, V., Quackenbush, J., Wahlestedt, C., Mattick, J. S., Hume, D. A., Kai, C., Sasaki, D., Tomaru, Y., Fukuda, S., Kanamori-Katayama, M., Suzuki, M., Aoki, J., Arakawa, T., Iida, J., Imamura, K., Itoh, M., Kato, T., Kawaji, H., Kawagashira, N., Kawashima, T., Kojima, M., Kondo, S., Konno, H., Nakano, K., Ninomiya, N., Nishio, T., Okada, M., Plessy, C., Shibata, K., Shiraki, T., Suzuki, S., Tagami, M., Waki, K., Watahiki, A., Okamura-Oho, Y., Suzuki, H., Kawai, J. & Hayashizaki, Y. (2005). The transcriptional landscape of the mammalian genome. *Science,* 309, 1559-1563.

Carr, J. L. & Bickham, J. W. (1981). Sex chromosomes of the Asian black pond turtle, Siebenrockiella crassicollis (Testudines: Emydidae). *Cytogenet.Cell Genet,* 31, 178-183.

Charnier, M. (1966). Action de la temperature sur la sex-ratio chez l' embryon ' Agama agama (Agamidae: Lacertilien). *C.R.Seances Soc.biol.l' Ouest Africain,* 160, 620-622.

Charnov, E. L. & Bull, J. (1977). When is sex environmentally determined? *Nature,* 266, 828-830.

Chaudhuri, K. & Chatterjee, R. (2007). MicroRNA detection and target prediction: integration of computational and experimental approaches. *DNA Cell Biol.,* 26, 321-337.

Chen, S., Ye, J., Kijima, I., Kinoshita, Y. & Zhou, D. (2005). Positive and negative transcriptional regulation of aromatase expression in human breast cancer tissue. *J Steroid Biochem.Mol Biol,* 95, 17-23.

Cheng, H. H., Ying, M., Tian, Y. H., Guo, Y., McElreavey, K. & Zhou, R. J. (2006). Transcriptional diversity of DMRT1 (dsx- and mab3-related transcription factor 1) in human testis. *Cell Res.*, 16, 389-393.

Cheng, J., Kapranov, P., Drenkow, J., Dike, S., Brubaker, S., Patel, S., Long, J., Stern, D., Tammana, H., Helt, G., Sementchenko, V., Piccolboni, A., Bekiranov, S., Bailey, D. K., Ganesh, M., Ghosh, S., Bell, I., Gerhard, D. S. & Gingeras, T. R. (2005). Transcriptional maps of 10 human chromosomes at 5-nucleotide resolution. *Science,* 308, 1149-1154.

Chowdhury, S., Maris, C., Allain, F. H. & Narberhaus, F. (2006). Molecular basis for temperature sensing by an RNA thermometer. *EMBO J,* 25, 2487-2497.

Ciofi, C. & Swingland, I. R. Enviornmental sex determination in reptiles. Behaviour Science 51, 251-265. 1997.
Ref Type: Journal (Full)

Clinton, M. (1998). Sex determination and gonadal development: a bird's eye view. *J Exp.Zool.*, 281, 457-465.

Cocquet, J., Pailhoux, E., Jaubert, F., Servel, N., Xia, X., Pannetier, M., De, B. E., Messiaen, L., Cotinot, C., Fellous, M. & Veitia, R. A. (2002). Evolution and expression of FOXL2. *J Med.Genet,* 39, 916-921.

Colvin, J. S., Feldman, B., Nadeau, J. H., Goldfarb, M. & Ornitz, D. M. (1999). Genomic organization and embryonic expression of the mouse fibroblast growth factor 9 gene. *Dev.Dyn.*, 216, 72-88.

Colvin, J. S., Green, R. P., Schmahl, J., Capel, B. & Ornitz, D. M. (2001). Male-to-female sex reversal in mice lacking fibroblast growth factor 9. *Cell,* 104, 875-889.

Conover, D. O. (2004). Temperature dependent sex determiantion in fishes. In N. Valenzuela & Lance V.A. (Eds) *Temperature dependent sex determination in vertebrates* (pp. 11-20). Washington: Smithsonian Books.

Conover, D. O. & Heins, S. W. (1987). Environmental and genetic components of of sex ratio in *Menidia menidia. Coepia,* 732-743.

Corbin, C. J., Khalil, M. W. & Conley, A. J. (1995). Functional ovarian and placental isoforms of porcine aromatase. *Mol Cell Endocrinol.,* 113, 29-37.

Crews, D., Bergeron, J. M., Bull, J. J., Flores, D., Tousignant, A., Skipper, J. K. & Wibbels, T. (1994). Temperature-dependent sex determination in reptiles: proximate mechanisms, ultimate outcomes, and practical applications. *Dev.Genet,* 15, 297-312.

Crews, D., Fleming, A., Willingham, E., Baldwin, R. & Skipper, J. K. (2001). Role of steroidogenic factor 1 and aromatase in temperature-dependent sex determination in the red-eared slider turtle. *J.Exp.Zool.,* 290, 597-606.

D'Cotta, H., Fostier, A., Guiguen, Y., Govoroun, M. & Baroiller, J. F. (2001). Aromatase plays a key role during normal and temperature-induced sex differentiation of tilapia Oreochromis niloticus. *Mol Reprod.Dev.,* 59, 265-276.

Daneau, I., Pilon, N., Boyer, A., Behdjani, R., Overbeek, P. A., Viger, R., Lussier, J. & Silversides, D. W. (2002). The porcine SRY promoter is transactivated within a male genital ridge environment. *Genesis.,* 33, 170-180.

Davies, R. C., Calvio, C., Bratt, E., Larsson, S. H., Lamond, A. I. & Hastie, N. D. (1998). WT1 interacts with the splicing factor U2AF65 in an isoform-dependent manner and can be incorporated into spliceosomes. *Genes Dev.,* 12, 3217-3225.

de Santa, B. P., Bonneaud, N., Boizet, B., Desclozeaux, M., Moniot, B., Sudbeck, P., Scherer, G., Poulat, F. & Berta, P. (1998). Direct interaction of SRY-related protein SOX9 and steroidogenic factor 1 regulates transcription of the human anti-Mullerian hormone gene. *Mol.Cell Biol.,* 18, 6653-6665.

Deeming, D. C. & Ferguson, M. W. (1988). Environmental regulation of sex determination in reptiles. *Philos.Trans.R.Soc.Lond B Biol.Sci.*, 322, 19-39.

del-Stenzel, M., Anderson, R., Heasman, J. & Wylie, C. (1998). The origin and migration of primordial germ cells in the mouse. *Semin.Cell Dev.Biol,* 9, 393-400.

Demas, S., Duronslet, M., Wachtel, S. S., Caillout, C. & Nakamura, D. (1990). Sex specific DNA in reptiles with temperature depensent sex determiantion. *Journal of Experimental Zoology,* 253, 319-324.

Desvages, G., Girondot, M. & Pieau, C. (1993a). Sensitive stages for the effects of temperature on gonadal aromatase activity in embryos of the marine turtle Dermochelys coriacea. *Gen.Comp Endocrinol.*, 92, 54-61.

Desvages, G., Girondot, M. & Pieau, C. (1993b). Sensitive stages for the effects of temperature on gonadal aromatase activity in embryos of the marine turtle Dermochelys coriacea. *Gen Comp Endocrinol.*, 92, 54-61.

Desvages, G. & Pieau, C. (1991). Steroid metabolism in gonads of turtle embryos as a function of the incubation temperature of eggs. *Journal of Steroid Biochemistry and Molecular Biology,* 39, 203-213.

Desvages, G. & Pieau, C. (1992). Time required for temperature induced chages in gonadal aromatase activity and related gonadal structure in turtle embryos. *Differentiation,* 52, 13-18.

Diatchenko, L., Lukyanov, S., Lau, Y. F. & Siebert, P. D. (1999). Suppression subtractive hybridization: a versatile method for identifying differentially expressed genes. *Methods Enzymol.*, 303, 349-380.

Dorizzi, M., Mignote, T.-M., Guichard, A., Desvages, G. & Pieau, C. (1991). Involvement of estrogens in sexual differentiation of gonads as a function of temperature in reptiles. *Differentiation,* 47, 9-17.

Dupont, S., Dennefeld, C., Krust, A., Chambon, P. & Mark, M. (2003). Expression of Sox9 in granulosa cells lacking the estrogen receptors, ERalpha and ERbeta. *Dev.Dyn.*, 226, 103-106.

Elbrecht, A. & Smith, R. G. (1992). Aromatase enzyme activity and sex determination in chickens. *Science,* 255, 467-470.

Elf, P. K. (2003). Yolk steroid hormones and sex determination in reptiles with TSD. *Gen.Comp Endocrinol.*, 132, 349-355.

Elf, P. K., Lang, J. W. & Fivizzani, A. J. (2002). Dynamics of yolk steroid hormones during development in a reptile with temperature-dependent sex determination. *Gen.Comp Endocrinol.*, 127, 34-39.

Engel, W., Klemme, B. & Ebrecht, A. (1981). Serological evidence for H-Y antigen in XO-female mice. *Hum.Genet,* 57, 68-70.

Engel, W. & Schmid, M. (1981). H-Y antigen as a tool for the determination of the heterogametic sex in Amphibia. *Cytogenet.Cell Genet,* 30, 130-136.

Ewert, M. A., Etchberger, C. R. & Nelson, C. E. (1990). An apparent co-occurence of genetic and environmental sex determiantion in a turtle. *American zoologist,* 30, 56A.

Ferguson, M. W. (1987). Post-laying stages of embryonic development for crocodilians. In G. J. W. Webb, S. C. Manolis & P. J. Whitehead (Eds) *Wildlife management: crocodiles and alligators* (pp. 427-444). Chipping Norton; Surrey Beatty.

Ferguson, M. W. Method of hatching avian eggs. [WO 94/13132]. 1994a.
Ref Type: Patent

Ferguson, M. W. Temperature dependent sex determination and growth in reptiles and manipulation of poultry sex byincubation temperature. Proceedings of the 9th European Poultry conference in Glassgow , 380-382. 1994b.
Ref Type: Conference Proceeding

Ferguson, M. W. & Joanen, T. (1982). Temperature of egg incubation determines sex in Alligator mississippiensis. *Nature,* 296, 850-853.

Ferguson-Smith, M. (2007). The evolution of sex chromosomes and sex determination in vertebrates and the key role of DMRT1. *Sex Dev.,* 1, 2-11.

Fleming, A., Wibbels, T., Skipper, J. K. & Crews, D. (1999). Developmental expression of steroidogenic factor 1 in a turtle with temperature-dependent sex determination. *Gen.Comp Endocrinol.,* 116, 336-346.

Flynt, A. S. & Lai, E. C. (2008). Biological principles of microRNA-mediated regulation: shared themes amid diversity. *Nat Rev Genet,* 9, 831-842.

Forch, P. & Valcarcel, J. (2003). Splicing regulation in Drosophila sex determination. *Prog.Mol Subcell.Biol,* 31, 127-151.

Foster, J. W., Dominguez-Steglich, M. A., Guioli, S., Kowk, G., Weller, P. A., Stevanovic, M., Weissenbach, J., Mansour, S., Young, I. D., Goodfellow, P. N. & . (1994). Campomelic dysplasia and autosomal sex reversal caused by mutations in an SRY-related gene. *Nature,* 372, 525-530.

Francis, R. C. & Barlow, G. W. (1993). Social control of primary sex differentiation in the Midas cichlid. *Proc.Natl.Acad.Sci.U.S.A,* 90, 10673-10675.

Frith, M. C., Pheasant, M. & Mattick, J. S. (2005). The amazing complexity of the human transcriptome. *Eur.J Hum.Genet,* 13, 894-897.

Gabriel, W. N., Blumberg, B., Sutton, S., Place, A. R. & Lance, V. A. (2001). Alligator aromatase cDNA sequence and its expression in embryos at male and female incubation temperatures. *J Exp.Zool.,* 290, 439-448.

Ganesh, S., Choudhary, B. & Raman, R. (1999). Temporal difference between testis and ovary determinations with possible involvement of testosterone and aromatase in gonadal differentiation in TSD lacking lizard, Calotes versicolor. *J Exp.Zool.,* 283, 600-607.

Ganesh, S., Mohanty, J. & Raman, R. (1997). Male-biased distribution of the human Y chromosomal genes SRY and ZFY in the lizard Calotes versicolor, which lacks sex chromosomes and temperature-dependent sex determination. *Chromosome.Res.,* 5, 413-419.

Girondot, M., Zaborski, P., Servan, J. & Pieau, C. (1994). Genetic contribution to sex determiantion in turtles wih environmental sex determiantion. *Genet.Res.,* 63, 117-127.

Goto, R., Kabaya, T., Adachi, S. & Yamauchi, K. (2000). Effects of temperature on sex determiantion in Marbled sole *Limanda yokohamae*. *Fish.Sci.,* 66, 400-402.

Grutzner, F., Rens, W., Tsend-Ayush, E., El-Mogharbel, N., O'Brien, P. C., Jones, R. C., Ferguson-Smith, M. A. & Marshall Graves, J. A. (2004). In the platypus a meiotic chain of ten sex chromosomes shares genes with the bird Z and mammal X chromosomes. *Nature,* 432, 913-917.

Guan, G., Kobayashi, T. & Nagahama, Y. (2000). Sexually dimorphic expression of two types of DM (Doublesex/Mab-3)-domain genes in a teleost fish, the Tilapia (Oreochromis niloticus). *Biochem.Biophys.Res.Commun.,* 272, 662-666.

Guo, J. K., Hammes, A., Chaboissier, M. C., Vidal, V., Xing, Y., Wong, F. & Schedl, A. (2002). Early gonadal development: exploring Wt1 and Sox9 function. *Novartis.Found.Symp.,* 244, 23-31.

Guo, Y., Cheng, H., Huang, X., Gao, S., Yu, H. & Zhou, R. (2005). Gene structure, multiple alternative splicing, and expression in gonads of zebrafish Dmrt1. *Biochem.Biophys.Res.Commun.,* 330, 950-957.

Gutzke, W. H. & Chymiy, D. B. (1988). Sensitive periods during embryogeny for hormonally induced sex determination in turtles. *Gen.Comp Endocrinol.,* 71, 265-267.

Gutzke, W. H. & Paukstis, G. L. (1983). Influence of the hydric environment on sexual differentiation of turtles. *J.Exp.Zool.,* 226, 467-469.

Hammes, A., Guo, J. K., Lutsch, G., Leheste, J. R., Landrock, D., Ziegler, U., Gubler, M. C. & Schedl, A. (2001). Two splice variants of the Wilms' tumor 1 gene have distinct functions during sex determination and nephron formation. *Cell,* 106, 319-329.

Hanley, N. A., Hagan, D. M., Clement-Jones, M., Ball, S. G., Strachan, T., Salas-Cortes, L., McElreavey, K., Lindsay, S., Robson, S., Bullen, P., Ostrer, H. & Wilson, D. I. (2000). SRY, SOX9, and DAX1 expression patterns during human sex determination and gonadal development. *Mech.Dev.,* 91, 403-407.

Harley, V. R., Clarkson, M. J. & Argentaro, A. (2003). The molecular action and regulation of the testis-determining factors, SRY (sex-determining region on the Y chromosome) and SOX9 [SRY-related high-mobility group (HMG) box 9]. *Endocr.Rev,* 24, 466-487.

Harrington, R. W. Jr. (1967). Environmentally controlled induction of primary bmale gonochorists from eggs of self fertilizing hermaphroditic fish *Rivulus marmoratus*. *Poey.Biol.Bull.,* 132, 174-199.

Hastie, N. D. (2001). Life, sex, and WT1 isoforms--three amino acids can make all the difference. *Cell,* 106, 391-394.

Hatano, O., Takakusu, A., Nomura, M. & Morohashi, K. (1996). Identical origin of adrenal cortex and gonad revealed by expression profiles of Ad4BP/SF-1. *Genes Cells,* 1, 663-671.

Hilscher, B., Hilscher, W., Bulthoff-Ohnolz, B., Kramer, U., Birke, A., Pelzer, H. & Gauss, G. (1974). Kinetics of gametogenesis. I. Comparative histological and autoradiographic studies of oocytes and transitional prospermatogonia during oogenesis and prespermatogenesis. *Cell Tissue Res.,* 154, 443-470.

Hiom, K. (2005). DNA repair: how to PIKK a partner. *Current Biology,* 15, R473-R475.

Hodgkin, J. (2002). The remarkable ubiquity of DM domain factors as regulators of sexual phenotype: ancestry or aptitude? *Genes Dev.,* 16, 2322-2326.

Holmgren, K. & Mosegaard, H. (1996). Implication of individual growth status on the future sex of the European eel. *Journal of Fish Biology,* 49, 910-925.

Hongay, C. F., Grisafi, P. L., Galitski, T. & Fink, G. R. (2006). Antisense transcription controls cell fate in Saccharomyces cerevisiae. *Cell,* 127, 735-745.

Huang, B., Wang, S., Ning, Y., Lamb, A. N. & Bartley, J. (1999). Autosomal XX sex reversal caused by duplication of SOX9. *Am.J Med.Genet,* 87, 349-353.

Huang, X., Guo, Y., Shui, Y., Gao, S., Yu, H., Cheng, H. & Zhou, R. (2005a). Multiple alternative splicing and differential expression of dmrt1 during gonad transformation of the rice field eel. *Biol.Reprod.,* 73, 1017-1024.

Huang, X., Li, J., Lu, L., Xu, M., Xiao, J., Yin, L., Zhu, H., Zhou, Z. & Sha, J. (2005b). Novel development-related alternative splices in human testis identified by cDNA microarrays. *J Androl,* 26, 189-196.

Ikeda, Y., Swain, A., Weber, T. J., Hentges, K. E., Zanaria, E., Lalli, E., Tamai, K. T., Sassone-Corsi, P., Lovell-Badge, R., Camerino, G. & Parker, K. L. (1996). Steroidogenic factor 1 and Dax-1 colocalize in multiple cell lineages: potential links in endocrine development. *Mol Endocrinol.,* 10, 1261-1272.

Inagaki.S., Numata, K., Kondo, T., Tomita, M., Yasuda, K., Kanai, A. & Kageyama, Y. (2005). Identification and expression analysis of putative mRNA-like non-coding RNA in Drosophila. *Genes to Cells,* 10, 1163-1173.

Iyer, A. K., Zhang, Y. H. & McCabe, E. R. (2006). Dosage-sensitive sex reversal adrenal hypoplasia congenita critical region on the X chromosome, gene 1 (DAX1) (NR0B1) and small heterodimer partner (SHP) (NR0B2) form homodimers individually, as well as DAX1-SHP heterodimers. *Mol Endocrinol,* 20, 2326-2342.

Jaccarini, V., Agius, L., Schembre, P. J. & Rizzo, M. (1983). Sex determiantion and larval sexual interaction in *Bonellia virdis* Ronaldo (Echiura: Bonelliidae). *Exp.Mar.Biol.Ecol.,* 66, 25-40.

Janzen, F. J. & Paukstis, G. L. (1991). Environmental sex determination in reptiles: ecology, evolution, and experimental design. *Q.Rev.Biol.,* 66, 149-179.

Janzen, F. J. & Phillips, P. C. (2006). Exploring the evolution of environmental sex determination, especially in reptiles. *J Evol.Biol,* 19, 1775-1784.

Jemth, P. & Gianni, S. (2007). PDZ doamins:folding and binding. *Biochemistry,* 46, 8701-8708.

Jeyasuria, P. & Place, A. R. (1997). Temperature-dependent aromatase expression in developing diamondback terrapin (Malaclemys terrapin) embryos. *J Steroid Biochem.Mol Biol,* 61, 415-425.

Jeyasuria, P. & Place, A. R. (1998a). Embryonic brain-gonadal axis in temperature-dependent sex determination of reptiles: a role for P450 aromatase (CYP19). *J.Exp.Zool.,* 281, 428-449.

Jeyasuria, P. & Place, A. R. (1998b). Embryonic brain-gonadal axis in temperature-dependent sex determination of reptiles: a role for P450 aromatase (CYP19). *J Exp.Zool.,* 281, 428-449.

Jones, K. W. & Singh, L. (1981). Conserved repeated DNA sequences in vertebrate sex chromosomes. *Hum.Genet,* 58, 46-53.

Jones, K. W. & Singh, L. (1985). Snake and evolution of sex chromosomes. *Trends Genet,* 1, 55-61.

Jost, A. (1972). A new look at mechanisms controlling sex sex differentiation in mammals. *John Hokins Medical Journal,* 130, 38-53.

Just, W., Rau, W., Vogel, W., Akhverdian, M., Fredga, K., Graves, J. A. & Lyapunova, E. (1995). Absence of Sry in species of the vole Ellobius. *Nat Genet,* 11, 117-118.

Kamat, A., Hinshelwood, M. M., Murry, B. A. & Mendelson, C. R. (2002). Mechanisms in tissue-specific regulation of estrogen biosynthesis in humans. *Trends Endocrinol.Metab,* 13, 122-128.

Kanai, Y., Hiramatsu, R., Matoba, S. & Kidokoro, T. (2005). From SRY to SOX9: mammalian testis differentiation. *J.Biochem.(Tokyo),* 138, 13-19.

Kapranov, P., Drenkow, J., Cheng, J., Long, J., Helt, G., Dike, S. & Gingeras, T. R. (2005). Examples of the complex architecture of the human transcriptome revealed by RACE and high-density tiling arrays. *Genome Res.,* 15, 987-997.

Karl, J. & Capel, B. (1998). Sertoli cells of the mouse testis originate from the coelomic epithelium. *Dev.Biol,* 203, 323-333.

Katayama, S., Tomaru, Y., Kasukawa, T., Waki, K., Nakanishi, M., Nakamura, M., Nishida, H., Yap, C. C., Suzuki, M., Kawai, J., Suzuki, H., Carninci, P., Hayashizaki, Y., Wells, C., Frith, M., Ravasi, T., Pang, K. C., Hallinan, J., Mattick, J., Hume, D. A., Lipovich, L., Batalov, S., Engstrom, P. G., Mizuno, Y., Faghihi, M. A., Sandelin, A., Chalk, A. M., Mottagui-Tabar, S., Liang, Z., Lenhard, B. & Wahlestedt, C. (2005). Antisense transcription in the mammalian transcriptome. *Science,* 309, 1564-1566.

Katoh-Fukui, Y., Owaki, A., Toyama, Y., Kusaka, M., Shinohara, Y., Maekawa, M., Toshimori, K. & Morohashi, K. (2005). Mouse Polycomb M33 is required for splenic vascular and adrenal gland formation through regulating Ad4BP/SF1 expression. *Blood,* 106, 1612-1620.

Katoh-Fukui, Y., Tsuchiya, R., Shiroishi, T., Nakahara, Y., Hashimoto, N., Noguchi, K. & Higashinakagawa, T. (1998). Male-to-female sex reversal in M33 mutant mice. *Nature,* 393, 688-692.

Kaufmann, E., Paul, H., Friedle, H., Metz, A., Scheucher, M., Clement, J. H. & Knochel, W. (1996). Antagonistic actions of activin A and BMP-2/4 control dorsal lip-specific activation of the early response gene XFD-1' in Xenopus laevis embryos. *EMBO J,* 15, 6739-6749.

Kent, J., Coriat, A. M., Sharpe, P. T., Hastie, N. D. & van, H., V (1995). The evolution of WT1 sequence and expression pattern in the vertebrates. *Oncogene,* 11, 1781-1792.

Kent, J., Wheatley, S. C., Andrews, J. E., Sinclair, A. H. & Koopman, P. (1996). A male-specific role for SOX9 in vertebrate sex determination. *Development,* 122, 2813-2822.

Kettlewell, J. R., Raymond, C. S. & Zarkower, D. (2000). Temperature-dependent expression of turtle Dmrt1 prior to sexual differentiation. *Genesis.*, 26, 174-178.

Kim, J., Prawitt, D., Bardeesy, N., Torban, E., Vicaner, C., Goodyer, P., Zabel, B. & Pelletier, J. (1999). The Wilms' tumor suppressor gene (wt1) product regulates Dax-1 gene expression during gonadal differentiation. *Mol.Cell Biol.*, 19, 2289-2299.

Kim, S., Bardwell, V. J. & Zarkower, D. (2007a). Cell type-autonomous and non-autonomous requirements for Dmrt1 in postnatal testis differentiation. *Dev.Biol,* 307, 314-327.

Kim, S., Namekawa, S. H., Niswander, L. M., Ward, J. O., Lee, J. T., Bardwell, V. J. & Zarkower, D. (2007b). A mammal-specific Doublesex homolog associates with male sex chromatin and is required for male meiosis. *PLoS Genet.*, 3, e62.

Kim, Y., Bingham, N., Sekido, R., Parker, K. L., Lovell-Badge, R. & Capel, B. (2007c). Fibroblast growth factor receptor 2 regulates proliferation and Sertoli differentiation during male sex determination. *Proc.Natl.Acad.Sci.U.S.A,* 104, 16558-16563.

Kim, Y. & Capel, B. (2006). Balancing the bipotential gonad between alternative organ fates: a new perspective on an old problem. *Dev.Dyn.*, 235, 2292-2300.

Kitamura, K., Yanazawa, M., Sugiyama, N., Miura, H., Iizuka-Kogo, A., Kusaka, M., Omichi, K., Suzuki, R., Kato-Fukui, Y., Kamiirisa, K., Matsuo, M., Kamijo, S., Kasahara, M., Yoshioka, H., Ogata, T., Fukuda, T., Kondo, I., Kato, M., Dobyns, W. B., Yokoyama, M. & Morohashi, K. (2002). Mutation of ARX causes abnormal development of forebrain and testes in mice and X-linked lissencephaly with abnormal genitalia in humans. *Nat Genet,* 32, 359-369.

Koopman, P. (2005). Sex determination: a tale of two Sox genes. *Trends Genet,* 21, 367-370.

Koopman, P., Bullejos, M., Loffler, K. & Bowles, J. (2002). Expression-based strategies for discovery of genes involved in testis and ovary development. *Novartis.Found.Symp.*, 244, 240-249.

Koopman, P., Gubbay, J., Vivian, N., Goodfellow, P. & Lovell-Badge, R. (1991). Male development of chromosomally female mice transgenic for Sry. *Nature,* 351, 117-121.

Koopman, P., Munsterberg, A., Capel, B., Vivian, N. & Lovell-Badge, R. (1990). Expression of a candidate sex-determining gene during mouse testis differentiation. *Nature,* 348, 450-452.

Kreidberg, J. A., Sariola, H., Loring, J. M., Maeda, M., Pelletier, J., Housman, D. & Jaenisch, R. (1993). WT-1 is required for early kidney development. *Cell,* 74, 679-691.

Kurokawa, H., Aoki, Y., Nakamura, S., Ebe, Y., Kobayashi, D. & Tanaka, M. (2006). Time-lapse analysis reveals different modes of primordial germ cell migration in the medaka Oryzias latipes. *Dev.Growth Differ.*, 48, 209-221.

Ladomery, M. R., Slight, J., Mc, G. S. & Hastie, N. D. (1999). Presence of WT1, the Wilm's tumor suppressor gene product, in nuclear poly(A)(+) ribonucleoprotein. *J Biol Chem.*, 274, 36520-36526.

Laity, J. H., Chung, J., Dyson, H. J. & Wright, P. E. (2000a). Alternative splicing of Wilms' tumor suppressor protein modulates DNA binding activity through isoform-specific DNA-induced conformational changes. *Biochemistry,* 39, 5341-5348.

Laity, J. H., Dyson, H. J. & Wright, P. E. (2000b). Molecular basis for modulation of biological function by alternate splicing of the Wilms' tumor suppressor protein. *Proc.Natl.Acad.Sci.U.S.A,* 97, 11932-11935.

Lalli, E., Ohe, K., Latorre, E., Bianchi, M. E. & Sassone-Corsi, P. (2003). Sexy splicing: regulatory interplays governing sex determination from Drosophila to mammals. *J.Cell Sci.*, 116, 441-445.

Lance, V. A. (1997). reptilian sex determiantion: An update. *American zoologist,* 37, 504-513.

Lance, V. A. (2003). Alligator physiology and life history: the importance of temperature. *Exp.Gerontol.*, 38, 801-805.

Lang, J. W. & Anderson, H. V. (1994). Temperature dependent sex determiantion in crocodialians. *Journal of Experimental Zoology,* 270, 28-44.

Lang, J. W., H.Andrews & R.Whittaker (1989). Sex determination and sex ratios in Crocodylus palustris. *American zoologist,* 29, 935-952.

Larsson, S. H., Charlieu, J. P., Miyagawa, K., Engelkamp, D., Rassoulzadegan, M., Ross, A., Cuzin, F., van, H., V & Hastie, N. D. (1995). Subnuclear localization of WT1 in splicing or transcription factor domains is regulated by alternative splicing. *Cell,* 81, 391-401.

Lawson, K. A. & Hage, W. J. (1994). Clonal analysis of the origin of primordial germ cells in the mouse. *Ciba Found.Symp.,* 182, 68-84.

Lee, C. T., Risom, T. & Strauss, W. M. (2007). Evolutionary conservation of microRNA regulatory circuits: an examination of microRNA gene complexity and conserved microRNA-target interactions through metazoan phylogeny. *DNA Cell Biol.,* 26, 209-218.

Liang, P. & Pardee, A. B. (1992). Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. *Science,* 257, 967-971.

Lovell-Badge, R., Canning, C. & Sekido, R. (2002). Sex-determining genes in mice: building pathways. *Novartis.Found.Symp.,* 244, 4-18.

Lu, H., Huang, X., Zhang, L., Guo, Y., Cheng, H. & Zhou, R. (2007). Multiple alternative splicing of mouse Dmrt1 during gonadal differentiation. *Biochem.Biophys.Res.Commun.,* 352, 630-634.

Luo, X., Ikeda, Y. & Parker, K. L. (1994). A cell-specific nuclear receptor is essential for adrenal and gonadal development and sexual differentiation. *Cell,* 77, 481-490.

Maeda, N., Kasukawa, T., Oyama, R., Gough, J., Frith, M., Engstrom, P. G., Lenhard, B., Aturaliya, R. N., Batalov, S., Beisel, K. W., Bult, C. J., Fletcher, C. F., Forrest, A. R., Furuno, M., Hill, D., Itoh, M., Kanamori-Katayama, M., Katayama, S., Katoh, M., Kawashima, T., Quackenbush, J., Ravasi, T., Ring, B. Z., Shibata, K., Sugiura, K., Takenaka, Y., Teasdale, R. D., Wells, C. A., Zhu, Y., Kai, C., Kawai, J., Hume, D. A., Carninci, P. & Hayashizaki, Y. (2006). Transcript annotation in FANTOM3: mouse gene catalog based on physical cDNAs. *PLoS Genet,* 2, e62.

Matsuda, M., Nagahama, Y., Shinomiya, A., Sato, T., Matsuda, C., Kobayashi, T., Morrey, C. E., Shibata, N., Asakawa, S., Shimizu, N., Hori, H., Hamaguchi, S. & Sakaizumi, M. (2002). DMY is a Y-specific DM-domain gene required for male development in the medaka fish. *Nature,* 417, 559-563.

Mayo, M. W., Wang, C. Y., Drouin, S. S., Madrid, L. V., Marshall, A. F., Reed, J. C., Weissman, B. E. & Baldwin, A. S. (1999). WT1 modulates apoptosis by transcriptionally upregulating the bcl-2 proto-oncogene. *EMBO J,* 18, 3990-4003.

McLaren, A. (1991). Development of the mammalian gonad: the fate of the supporting cell lineage. *Bioessays,* 13, 151-156.

McLaren, A. (2000). Germ and somatic cell lineages in the developing gonad. *Mol Cell Endocrinol.,* 163, 3-9.

McLaren, A. & Southee, D. (1997). Entry of mouse embryonic germ cells into meiosis. *Dev.Biol,* 187, 107-113.

Meeks, J. J., Crawford, S. E., Russell, T. A., Morohashi, K., Weiss, J. & Jameson, J. L. (2003a). Dax1 regulates testis cord organization during gonadal differentiation. *Development,* 130, 1029-1036.

Meeks, J. J., Weiss, J. & Jameson, J. L. (2003b). Dax1 is required for testis determination. *Nat Genet,* 34, 32-33.

Menke, D. B. & Page, D. C. (2002). Sexually dimorphic gene expression in the developing mouse gonad. *Gene Expr.Patterns.,* 2, 359-367.

Merchant, H. (1975). Rat gonadal and ovarioan organogenesis with and without germ cells. An ultrastructural study. *Dev.Biol,* 44, 1-21.

Merchant-Larios, H. & Moreno-Mendoza, N. (2001). Onset of sex differentiation: dialog between genes and cells. *Arch Med.Res.,* 32, 553-558.

Merchant-Larios, H. & Villalpando, I. (1990). Effect of temperature on gonadal sex differentiation in the sea turtle, *Lepidochelys olivacea,* an organ culture study. *Journal of Experimental Zoology,* 254, 327-331.

Miller, D., Summers, J. & Silber, S. (2004). Environmental versus genetic sex determination: a possible factor in dinosaur extinction? *Fertil.Steril.,* 81, 954-964.

Miyamoto, N., Yoshida, M., Kuratani, S., Matsuo, I. & Aizawa, S. (1997). Defects of urogenital development in mice lacking Emx2. *Development,* 124, 1653-1664.

Moore, A. W., McInnes, L., Kreidberg, J., Hastie, N. D. & Schedl, A. (1999). YAC complementation shows a requirement for Wt1 in the development of epicardium, adrenal gland and throughout nephrogenesis. *Development,* 126, 1845-1857.

Morais da, S. S., Hacker, A., Harley, V., Goodfellow, P., Swain, A. & Lovell-Badge, R. (1996). Sox9 expression during gonadal development implies a conserved role for the gene in testis differentiation in mammals and birds. *Nat.Genet.,* 14, 62-68.

Moreno-Mendoza, N., Harley, V. R. & Merchant-Larios, H. (1999). Differential expression of SOX9 in gonads of the sea turtle Lepidochelys olivacea at male- or female-promoting temperatures. *J.Exp.Zool.,* 284, 705-710.

Moreno-Mendoza, N., Harley, V. R. & Merchant-Larios, H. (2001). Temperature regulates SOX9 expression in cultured gonads of Lepidochelys olivacea, a species with temperature sex determination. *Dev.Biol.,* 229, 319-326.

Morrish, B. C. & Sinclair, A. H. (2002). Vertebrate sex determination: many means to an end. *Reproduction.,* 124, 447-457.

Mrosovsky, N. & Pieau, C. (1991). Transitional change of temperature, pivotal temperature and thermosensitive stages for sex determiantion in reptiles. *Chelon.COns.Biol.,* 1, 238-240.

Mundlos, S., Pelletier, J., Darveau, A., Bachmann, M., Winterpacht, A. & Zabel, B. (1993). Nuclear localization of the protein encoded by the Wilms' tumor gene WT1 in embryonic and adult tissues. *Development,* 119, 1329-1341.

Murdock, C. & Wibbels, T. (2003b). Cloning and expression of aromatase in a turtle with temperature-dependent sex determination. *Gen Comp Endocrinol.,* 130, 109-119.

Murdock, C. & Wibbels, T. (2003a). Cloning and expression of aromatase in a turtle with temperature-dependent sex determination. *Gen.Comp Endocrinol.,* 130, 109-119.

Murdock, C. & Wibbels, T. (2003c). Expression of Dmrt1 in a turtle with temperature-dependent sex determination. *Cytogenet.Genome Res.,* 101, 302-308.

Murdock, C. & Wibbels, T. (2006). Dmrt1 expression in response to estrogen treatment in a reptile with temperature-dependent sex determination. *J.Exp.Zoolog.B Mol Dev.Evol.,* 306, 134-139.

Nachtigal, M. W., Hirokawa, Y., Enyeart-VanHouten, D. L., Flanagan, J. N., Hammer, G. D. & Ingraham, H. A. (1998). Wilms' tumor 1 and Dax-1 modulate the orphan nuclear receptor SF-1 in sex-specific gene expression. *Cell,* 93, 445-454.

Nakamura, D., Wachtel, S. S., Lance, V. & Becak, W. (1987). On the evolution of sex determiantion. *Proceedings of the Royal Society of London B Biological Sciences,* 232, 159-180.

Nakamura, Y., Yamamoto, Y., Usui, F., Mushika, T., Ono, T., Setioko, A. R., Takeda, K., Nirasawa, K., Kagami, H. & Tagami, T. (2007). Migration and proliferation of primordial germ cells in the early chicken embryo. *Poult.Sci.,* 86, 2182-2193.

Nakayama, Y., Yamamoto, T., Matsuda, Y. & Abe, S. I. (1998). Cloning of cDNA for newt WT1 and the differential expression during spermatogenesis of the Japanese newt, Cynops pyrrhogaster. *Dev.Growth Differ.*, 40, 599-608.

Nanda, I., Kondo, M., Hornung, U., Asakawa, S., Winkler, C., Shimizu, A., Shan, Z., Haaf, T., Shimizu, N., Shima, A., Schmid, M. & Schartl, M. (2002). A duplicated copy of DMRT1 in the sex-determining region of the Y chromosome of the medaka, Oryzias latipes. *Proc.Natl.Acad.Sci.U.S.A,* 99, 11778-11783.

Nanda, I., Schartl, M., Feichtinger, W., Epplen, J. T. & Schmid, M. (1992). Early stages of sex chromosome differentiation in fish as analysed by simple repetitive DNA sequences. *Chromosoma,* 101, 301-310.

Natoli, T. A., Alberta, J. A., Bortvin, A., Taglienti, M. E., Menke, D. B., Loring, J., Jaenisch, R., Page, D. C., Housman, D. E. & Kreidberg, J. A. (2004). Wt1 functions in the development of germ cells in addition to somatic cell lineages of the testis. *Dev.Biol,* 268, 429-440.

Nef, S., Schaad, O., Stallings, N. R., Cederroth, C. R., Pitetti, J. L., Schaer, G., Malki, S., Dubois-Dauphin, M., Boizet-Bonhoure, B., Descombes, P., Parker, K. L. & Vassalli, J. D. (2005). Gene expression during sex determination reveals a robust female genetic program at the onset of ovarian development. *Dev.Biol,* 287, 361-377.

Nef, S., Verma-Kurvari, S., Merenmies, J., Vassalli, J. D., Efstratiadis, A., Accili, D. & Parada, L. F. (2003). Testis determination requires insulin receptor family function in mice. *Nature,* 426, 291-295.

Nicolas, J. C., Boussioux, A. M., Decomps, B. & Crastes de Paulet, A. (1979). Enzymatic determiantion of estradiol and estrone in plasma and urine. *Clinica Chemica Acta,* 92, 1-9.

Nishikimi, H., Kansaku, N., Saito, N., Usami, M., Ohno, Y. & Shimada, K. (2000). Sex differentiation and mRNA expression of P450c17, P450arom and AMH in gonads of the chicken. *Mol Reprod Dev.,* 55, 20-30.

Oh, H. J., Li, Y. & Lau, Y. F. (2005). Sry associates with the heterochromatin protein 1 complex by interacting with a KRAB domain protein. *Biol Reprod,* 72, 407-415.

Ohe, K., Lalli, E. & Sassone-Corsi, P. (2002). A direct role of SRY and SOX proteins in pre-mRNA splicing. *Proc.Natl.Acad.Sci.U.S.A,* 99, 1146-1151.

Okamoto, K. & Beach, D. (1994). Cyclin G is a transcriptional target of the p53 tumor suppressor protein. *EMBO J,* 13, 4816-4822.

Okamoto, K., Li, H., Jensen, M. R., Zhang, T., Taya, Y., Thorgeirsson, S. S. & Prives, C. (2002). Cyclin G recruits PP2A to dephosphorylate Mdm2. *Mol Cell,* 9, 761-771.

Okamura, S., Arakawa, H., Tanaka, T., Nakanishi, H., Ng, C. C., Taya, Y., Monden, M. & Nakamura, Y. (2001). p53DINP1, a p53-inducible gene, regulates p53-dependent apoptosis. *Mol Cell,* 8, 85-94.

Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D. P., Bult, C., Hume, D. A., Quackenbush, J., Schriml, L. M., Kanapin, A., Matsuda, H., Batalov, S., Beisel, K. W., Blake, J. A., Bradt, D., Brusic, V., Chothia, C., Corbani, L. E., Cousins, S., Dalla, E., Dragani, T. A., Fletcher, C. F., Forrest, A., Frazer, K. S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I. J., Jarvis, E. D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R. M., King, B. L., Konagaya, A., Kurochkin, I. V., Lee, Y., Lenhard, B., Lyons, P. A., Maglott, D. R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W. J., Pertea, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J. U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J. C., Reed, D. J., Reid, J., Ring, B. Z., Ringwald, M., Sandelin, A., Schneider, C., Semple, C. A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M. S., Teasdale, R. D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L. G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N.,

Hirozane-Kishikawa, T., Konno, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E. S., Rogers, J., Birney, E. & Hayashizaki, Y. (2002). Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs. *Nature,* 420, 563-573.

Oreal, E., Mazaud, S., Picard, J. Y., Magre, S. & Carre-Eusebe, D. (2002). Different patterns of anti-Mullerian hormone expression, as related to DMRT1, SF-1, WT1, GATA-4, Wnt-4, and Lhx9 expression, in the chick differentiating gonads. *Dev.Dyn.,* 225, 221-232.

Oreal, E., Pieau, C., Mattei, M. G., Josso, N., Picard, J. Y., Carre-Eusebe, D. & Magre, S. (1998). Early expression of AMH in chicken embryonic gonads precedes testicular SOX9 expression. *Dev.Dyn.,* 212, 522-532.

Ornitz, D. M. & Itoh, N. (2001). Fibroblast growth factors. *Genome Biol,* 2, REVIEWS3005.

Osawa, N., Oshima, Y. & Nakamura, M. (2005). Molecular cloning of Dmrt1 and its expression in the gonad of Xenopus. *Zoolog.Sci.,* 22, 681-687.

Ottolenghi, C., Fellous, M., Barbieri, M. & McElreavey, K. (2002). Novel paralogy relations among human chromosomes support a link between the phylogeny of doublesex-related genes and the evolution of sex determination. *Genomics,* 79, 333-343.

Ottolenghi, C., Omari, S., Garcia-Ortiz, J. E., Uda, M., Crisponi, L., Forabosco, A., Pilia, G. & Schlessinger, D. (2005). Foxl2 is required for commitment to ovary differentiation. *Hum.Mol Genet,* 14, 2053-2062.

Ottolenghi, C., Uda, M., Crisponi, L., Omari, S., Cao, A., Forabosco, A. & Schlessinger, D. (2007). Determination and stability of sex. *Bioessays,* 29, 15-25.

Palmer, S. J. & Burgoyne, P. S. (1991a). In situ analysis of fetal, prepuberal and adult XX----XY chimaeric mouse testes: Sertoli cells are predominantly, but not exclusively, XY. *Development,* 112, 265-268.

Palmer, S. J. & Burgoyne, P. S. (1991b). XY follicle cells in the ovaries of XO/XY and XO/XY/XYY mosaic mice. *Development,* 111, 1017-1019.

Pannetier, M., Mandon-Pepin, B., Copelli, S. & Fellous, M. (2004). Molecular aspects of female and male gonadal development in mammals. *Pediatr.Endocrinol.Rev,* 1, 274-287.

Patek, C. E., Kerr, J. B., Gosden, R. G., Jones, K. W., Hardy, K., Muggleton-Harris, A. L., Handyside, A. H., Whittingham, D. G. & Hooper, M. L. (1991). Sex chimaerism, fertility and sex determination in the mouse. *Development,* 113, 311-325.

Pauler, F. M., Koerner, M. V. & Barlow, D. P. (2007). Silencing by imprinted noncoding RNAs: is transcription the answer? *Trends Genet,* 23, 284-292.

Pieau, C. Temperature variation and sex determination in reptiles. Bioessays 18[19], 26. 1996. Ref Type: Journal (Full)

Pieau, C. & Dorizzi, M. (1981). Determination of temperature sensitive stages for sexual differentiation of the gonads in embryos of the turtle, *Emys orbicularis. J.Morphol.,* 170, 373-383.

Pieau, C. & Dorizzi, M. (2004). Oestrogens and temperature-dependent sex determination in reptiles: all is in the gonads. *J.Endocrinol.,* 181, 367-377.

Pieau, C., Dorizzi, M. & Richard-Mercier, N. (1999). Temperature-dependent sex determination and gonadal differentiation in reptiles. *Cell Mol Life Sci.,* 55, 887-900.

Pieau, C., Dorizzi, M., Richard-Mercier, N. & Desvages, G. (1998). Sexual differentiation of gonads as a function of temperature in the turtle Emys orbicularis: endocrine function, intersexuality and growth. *J.Exp.Zool.,* 281, 400-408.

Pieau, C., Girondot, M., Desvages, G., Dorizzi, M., Rechard-Mercier, N. & Zaborski, P. (1994). Environmental control of gonadal differentiation. In R. V. Short & E. Balaban (Eds) *The difference between the sexes* (pp. 433-448). Cambridge: Cambridge University Press.

Piferrer, F., Blazquez, M., Navarro, L. & Gonzalez, A. (2005). Genetic, endocrine, and environmental components of sex determination and differentiation in the European sea bass (Dicentrarchus labrax L.). *Gen Comp Endocrinol.*, 142, 102-110.

Pirkkala, L., Nykanen, P. & Sistonen, L. (2001). Roles of the heat shock transcription factors in regulation of the heat shock response and beyond. *FASEB J,* 15, 1118-1131.

Place, A. R. & Lance V.A. (2004). The Temperature dependent Sex Determination Drama: Same Cast, Different Stars. In N. Valenzuela & Lance V.A. (Eds) *Temperature Dependent Sex Determiantion in vertebrates* (pp. 99-110). Washington: Smithsonian books.

Prasanth, K. V. & Spector, D. L. (2007). Eukaryotic regulatory RNAs: an answer to the 'genome complexity' conundrum. *Genes Dev.*, 21, 11-42.

Qin, J., Tsai, M. J. & Tsai, S. Y. (2008). Essential roles of COUP-TFII in Leydig cell differentiation and male fertility. *PLoS ONE.*, 3, e3285.

Ramsey, M., Shoemaker, C. & Crews, D. (2007). Gonadal expression of Sf1 and aromatase during sex determination in the red-eared slider turtle (Trachemys scripta), a reptile with temperature-dependent sex determination. *Differentiation.*

Ravasi, T., Suzuki, H., Pang, K. C., Katayama, S., Furuno, M., Okunishi, R., Fukuda, S., Ru, K., Frith, M. C., Gongora, M. M., Grimmond, S. M., Hume, D. A., Hayashizaki, Y. & Mattick, J. S. (2006). Experimental validation of the regulated expression of large numbers of non-coding RNAs from the mouse genome. *Genome Res.*, 16, 11-19.

Raymond, C. S., Kettlewell, J. R., Hirsch, B., Bardwell, V. J. & Zarkower, D. (1999). Expression of Dmrt1 in the genital ridge of mouse and chicken embryos suggests a role in vertebrate sexual development. *Dev.Biol.*, 215, 208-220.

Raymond, C. S., Murphy, M. W., O'Sullivan, M. G., Bardwell, V. J. & Zarkower, D. (2000). Dmrt1, a gene related to worm and fly sexual regulators, is required for mammalian testis differentiation. *Genes Dev.*, 14, 2587-2595.

Raymond, C. S., Shamu, C. E., Shen, M. M., Seifert, K. J., Hirsch, B., Hodgkin, J. & Zarkower, D. (1998). Evidence for evolutionary conservation of sex-determining genes. *Nature,* 391, 691-695.

Romer, U. & Beisenhertz, W. Environmental determination of sex in Apistogramma (Cichlidae) and two other fresh water fishes (Teleostei). Journal of Fish Biology 48, 714-725. 1996. Ref Type: Journal (Full)

Ross, A. J., Tilman, C., Yao, H., MacLaughlin, D. & Capel, B. (2003). AMH induces mesonephric cell migration in XX gonads. *Mol.Cell Endocrinol.*, 211, 1-7.

Salame-Mendez, A., Herrera-Munoz, J., Moreno-Mendoza, N. & Merchant-Larios, H. (1998). Response of diencephalon but not the gonad to female-promoting temperature with elevated estradiol levels in the sea turtle Lepidochelys olivacea. *J Exp.Zool.*, 280, 304-313.

Sarre, S. D., Georges, A. & Quinn, A. (2004). The ends of a continuum: genetic and temperature-dependent sex determination in reptiles. *Bioessays,* 26, 639-645.

Schartl, M. (2004). A comparative view on sex determination in medaka. *Mech.Dev.*, 121, 639-645.

Schmahl, J., Eicher, E. M., Washburn, L. L. & Capel, B. (2000). Sry induces cell proliferation in the mouse gonad. *Development,* 127, 65-73.

Schmahl, J., Kim, Y., Colvin, J. S., Ornitz, D. M. & Capel, B. (2004). Fgf9 induces proliferation and nuclear localization of FGFR2 in Sertoli precursors during male sex determination. *Development,* 131, 3627-3636.

Schmahl, J., Yao, H. H., Pierucci-Alves, F. & Capel, B. (2003). Colocalization of WT1 and cell proliferation reveals conserved mechanisms in temperature-dependent sex determination. *Genesis.*, 35, 193-201.

Schmidt, D., Ovitt, C. E., Anlag, K., Fehsenfeld, S., Gredsted, L., Treier, A. C. & Treier, M. (2004). The murine winged-helix transcription factor Foxl2 is required for granulosa cell differentiation and ovary maintenance. *Development,* 131, 933-942.

Sekido, R., Bar, I., Narvaez, V., Penny, G. & Lovell-Badge, R. (2004). SOX9 is up-regulated by the transient expression of SRY specifically in Sertoli cell precursors. *Dev.Biol,* 274, 271-279.

Sekido, R. & Lovell-Badge, R. (2007). Mechanisms of gonadal morphogenesis are not conserved between chick and mouse. *Dev.Biol.,* 302, 132-142.

Sekido, R. & Lovell-Badge, R. (2008). Sex determination involves synergistic action of SRY and SF1 on a specific Sox9 enhancer. *Nature,* 453, 930-934.

Shamovsky, I., Ivannikov, M., Kandel, E. S., Gershon, D. & Nudler, E. (2006). RNA-mediated response to heat shock in mammalian cells. *Nature,* 440, 556-560.

Shamovsky, I. & Nudler, E. (2006). Gene control by large noncoding RNAs. *Sci.STKE.,* 2006, e40.

Shan, Z., Nanda, I., Wang, Y., Schmid, M., Vortkamp, A. & Haaf, T. (2000). Sex-specific expression of an evolutionarily conserved male regulatory gene, DMRT1, in birds. *Cytogenet.Cell Genet,* 89, 252-257.

Shibata, K., Takase, M. & Nakamura, M. (2002). The Dmrt1 expression in sex-reversed gonads of amphibians. *Gen.Comp Endocrinol.,* 127, 232-241.

Shine, R. (1999). Why is sex determined by nest temperature in many reptiles. *Trends Ecol Evol.,* 14, 186-189.

Shinoda, K., Lei, H., Yoshii, H., Nomura, M., Nagano, M., Shiba, H., Sasaki, H., Osawa, Y., Ninomiya, Y., Niwa, O. & . (1995). Developmental defects of the ventromedial hypothalamic nucleus and pituitary gonadotroph in the Ftz-F1 disrupted mice. *Dev.Dyn.,* 204, 22-29.

Shoemaker, C., Ramsey, M., Queen, J. & Crews, D. (2007a). Expression of Sox9, Mis, and Dmrt1 in the gonad of a species with temperature-dependent sex determination. *Dev.Dyn.,* 236, 1055-1063.

Shoemaker, C. M., Queen, J. & Crews, D. (2007b). Response of Candidate Sex-Determining Genes to Changes in Temperature Reveals Their Involvement in the Molecular Network Underlying Temperature-Dependent Sex Determination. *Mol Endocrinol.,* 21, 2750-2763.

Sinclair, A., Smith, C., Western, P. & McClive, P. (2002). A comparative analysis of vertebrate sex determination. *Novartis.Found.Symp.,* 244, 102-111.

Sinclair, A. H., Berta, P., Palmer, M. S., Hawkins, J. R., Griffiths, B. L., Smith, M. J., Foster, J. W., Frischauf, A. M., Lovell-Badge, R. & Goodfellow, P. N. (1990). A gene from the human sex-determining region encodes a protein with homology to a conserved DNA-binding motif. *Nature,* 346, 240-244.

Singh, L. & Jones, K. W. (1982). Sex reversal in the mouse (Mus musculus) is caused by a recurrent nonreciprocal crossover involving the x and an aberrant y chromosome. *Cell,* 28, 205-216.

Sleutels, F., Zwart, R. & Barlow, D. P. (2002). The non-coding Air RNA is required for silencing autosomal imprinted genes. *Nature,* 415, 810-813.

Smith C.A. & Joss M.P. (1994). **Sertoli Cell Differentiation and Gonadogenesis in *Alligator mississippiensis*.** *Journal of Experimental Zoology,* 270, 57-70.

Smith C.A. & Joss, J. M. P. (1993). Gonadal sex differentiation in *Alligator mississipiensis,* a species with temperature dependent sex determiantion. *Cell Tissue Research,* 273, 149-162.

Smith, C. A., Elf, P., Lang, J. W. & Joss, J. M. P. (1995). Aromatase enzyme activity during gonadal sex differentiation in Alligator embryos. *Differentiation,* 58, 281-290.

Smith, C. A. & Joss, J. M. (1994a). Steroidogenic enzyme activity and ovarian differentiation in the saltwater crocodile, Crocodylus porosus. *Gen.Comp Endocrinol.,* 93, 232-245.

Smith, C. A. & Joss, J. M. P. (1993). Gonadal sex differentiation in *Alligator mississipiensis,* a species with temperature dependent sex determiantion. *Cell Tissue Research,* 273, 149-162.

Smith, C. A. & Joss, M. P. (1994b). **Sertoli Cell Differentiation and Gonadogenesis in *Alligator mississipiensis*.** *Journal of Experimental Zoology,* 270, 57-70.

Smith, C. A., Katz, M. & Sinclair, A. H. (2003). DMRT1 is upregulated in the gonads during female-to-male sex reversal in ZW chicken embryos. *Biol.Reprod.,* 68, 560-570.

Smith, C. A. & Sinclair, A. H. (2004). Sex determination: insights from the chicken. *Bioessays,* 26, 120-132.

Solari, A. J. (1994). *Sex chromosomes and sex determination in vertebrates.* Boca Raton: CRC press.

Spotila, L. D. & Hall, S. E. (1998). Expression of a new RNA-splice isoform of WT1 in developing kidney-gonadal complexes of the turtle, Trachemys scripta. *Comp Biochem.Physiol B Biochem.Mol.Biol.,* 119, 761-767.

Spotila, L. D., Spotila, J. R. & Hall, S. E. (1998). Sequence and expression analysis of WT1 and Sox9 in the red-eared slider turtle, Trachemys scripta. *J.Exp.Zool.,* 281, 417-427.

Sreenivasulu, K., Ganesh, S. & Raman, R. (2002). Evolutionarily conserved, DMRT1, encodes alternatively spliced transcripts and shows dimorphic expression during gonadal differentiation in the lizard, Calotes versicolor. *Mech.Dev.,* 119 Suppl 1, S55-S64.

Suzuki, T., Kasahara, M., Yoshioka, H., Morohashi, K. & Umesono, K. (2003). LXXLL-related motifs in Dax-1 have target specificity for the orphan nuclear receptors Ad4BP/SF-1 and LRH-1. *Mol Cell Biol,* 23, 238-249.

Swain, A. (2002). Vertebrate sex determination: a new player in the field. *Curr.Biol.,* 12, R602-R603.

Swain, A. & Lovell-Badge, R. (1999). Mammalian sex determination: a molecular drama. *Genes Dev.,* 13, 755-767.

Swain, A., Narvaez, V., Burgoyne, P., Camerino, G. & Lovell-Badge, R. (1998). Dax1 antagonizes Sry action in mammalian sex determination. *Nature,* 391, 761-767.

Swain, A., Zanaria, E., Hacker, A., Lovell-Badge, R. & Camerino, G. (1996). Mouse Dax1 expression is consistent with a role in sex determination as well as in adrenal and hypothalamus function. *Nat.Genet.,* 12, 404-409.

Takamoto, N., Kurihara, I., Lee, K., Demayo, F. J., Tsai, M. J. & Tsai, S. Y. (2005). Haploinsufficiency of chicken ovalbumin upstream promoter transcription factor II in female reproduction. *Mol Endocrinol,* 19, 2299-2308.

Taketo, T., Thau, R. B., Adeyemo, O. & Koide, S. S. (1984). Influence of adenosine 3':5'-cyclic monophosphate analogues on testicular organization of fetal mouse gonads in vitro. *Biol Reprod,* 30, 189-198.

Tevosian, S. G., Albrecht, K. H., Crispino, J. D., Fujiwara, Y., Eicher, E. M. & Orkin, S. H. (2002). Gonadal differentiation, sex determination and normal Sry expression in mice require direct interaction between transcription partners GATA4 and FOG2. *Development,* 129, 4627-4634.

Thevenet, L., Albrecht, K. H., Malki, S., Berta, P., Boizet-Bonhoure, B. & Poulat, F. (2005). NHERF2/SIP-1 interacts with mouse SRY via a different mechanism than human SRY. *J Biol Chem.,* 280, 38625-38630.

Thorvaldsen, J. L., Verona, J. I. & Bartolomei, M. S. (2007). X-tra! X-tra! News from the Mouse X Chromosome. *Developmental Biology,* 298, 344-353.

Tong, S. K. & Chung, B. C. (2003). Analysis of zebrafish cyp19 promoters. *J Steroid Biochem.Mol Biol,* 86, 381-386.

Torres Maldonado, L. C., Landa, P. A., Moreno, M. N., Marmolejo, V. A., Meza, M. A. & Merchant, L. H. (2002). Expression profiles of Dax1, Dmrt1, and Sox9 during temperature sex determination in gonads of the sea turtle Lepidochelys olivacea. *Gen.Comp Endocrinol.,* 129, 20-26.

Uda, M., Ottolenghi, C., Crisponi, L., Garcia, J. E., Deiana, M., Kimber, W., Forabosco, A., Cao, A., Schlessinger, D. & Pilia, G. (2004). Foxl2 disruption causes mouse ovarian failure by pervasive blockage of follicle development. *Hum.Mol Genet,* 13, 1171-1181.

Ukeshima, A., Kudo, M. & Fujimoto, T. (1987). Relationship between genital ridge formation and settlement site of primordial germ cells in chick embryos. *Anat Rec.,* 219, 311-314.

Vainio, S., Heikkila, M., Kispert, A., Chin, N. & McMahon, A. P. (1999). Female development in mammals is regulated by Wnt-4 signalling. *Nature,* 397, 405-409.

Valenzuela, N. & Lance, V. A. (2004). *Temperature dependent sex determination in vertebrates.* Washington D.C.: Smithsonian Books.

Valenzuela, N., LeClere, Shikano & T. (2006). comaparative gene expression of steroidogenic factor1 in *Chrysemys picta* and *Apalone mutica* turtles with temperature dependent sex determiantion and genotypic sex determiantion. *Evolution and Development,* 8, 424-432.

Vidal, V. P., Chaboissier, M. C., de Rooij, D. G. & Schedl, A. (2001). Sox9 induces testis development in XX transgenic mice. *Nat Genet,* 28, 216-217.

Viets, B. E., Ewert, M. A., Talent, L. G. & Nelson, C. E. (1994). Sex-determining mechanisms in squamate reptiles. *Journal of Experimental Zoology,* 270, 45-56.

Wagner, K. D., Wagner, N. & Schedl, A. (2003). The complex life of WT1. *J.Cell Sci.,* 116, 1653-1658.

Wagner, K. D., Wagner, N., Vidal, V. P., Schley, G., Wilhelm, D., Schedl, A., Englert, C. & Scholz, H. (2002). The Wilms' tumor gene Wt1 is required for normal development of the retina. *EMBO J,* 21, 1398-1405.

Wagner, T., Wirth, J., Meyer, J., Zabel, B., Held, M., Zimmer, J., Pasantes, J., Bricarelli, F. D., Keutel, J., Hustert, E. & . (1994). Autosomal sex reversal and campomelic dysplasia are caused by mutations in and around the SRY-related gene SOX9. *Cell,* 79, 1111-1120.

Wang, W., Lee, S. B., Palmer, R., Ellisen, L. W. & Haber, D. A. (2001). A functional interaction with CBP contributes to transcriptional activation by the Wilms tumor suppressor WT1. *J Biol Chem.,* 276, 16810-16816.

Warne, G. L. & Kanumakala, S. (2002). Molecular endocrinology of sex differentiation. *Semin.Reprod Med.,* 20, 169-180.

Warner, D. A. & Shine, R. (2008). The adaptive significance of temperature-dependent sex determination in a reptile. *Nature,* 451, 566-568.

Webb, G. J. W., Beal, A. M., Manolis, S. C. & DempseyK.E. (1987). The effect of incubation temperature on sex determiantion and embryonic development rate in *Crocodylus jonnstoni* and *Crocodylus porosus.* In G. J. W. Webb, S. C. Manolis & P. J. Whitehead (Eds) *Wildlife Management* (pp. 507-531). Sidney: Surrey Beatty and sons.

Wellins, D. J. (1987). Use of an H_Y antigen assay for sex determination in sea turtles. *Coepia,* 46-53.

Wennstrom, K. L. & Crews, D. (1995). Making males from females: the effects of aromatase inhibitors on a parthenogenetic species of whiptail lizard. *Gen Comp Endocrinol.,* 99, 316-322.

Western, P. S., Harry, J. L., Graves, J. A. & Sinclair, A. H. (1999a). Temperature-dependent sex determination in the American alligator: AMH precedes SOX9 expression. *Dev.Dyn.,* 216, 411-419.

Western, P. S., Harry, J. L., Graves, J. A. & Sinclair, A. H. (1999b). Temperature-dependent sex determination: upregulation of SOX9 expression after commitment to male development. *Dev.Dyn.,* 214, 171-177.

Western, P. S., Harry, J. L., Marshall Graves, J. A. & Sinclair, A. H. (2000). Temperature-dependent sex determination in the American alligator: expression of SF1, WT1 and DAX1 during gonadogenesis. *Gene,* 241, 223-232.

Western, P. S. & Sinclair, A. H. (2001). Sex, genes, and heat: triggers of diversity. *J.Exp.Zool.,* 290, 624-631.

White, R. B. & Thomas, P. (1992a). Adrenal-kidney and gonadal steroidogenesis during sexual differentiation of a reptile with temperature-dependent sex determination. *Gen.Comp Endocrinol.,* 88, 10-19.

White, R. B. & Thomas, P. (1992b). Adrenal-kidney and gonadal steroidogenesis during sexula differentiation in reptiles differentiation in reptile with teperature-dependent sex determiantion. *General and Comparative Endocrinology,* 88, 10-19.

Wibbels, T., Bull, J. J. & Crews, D. (1991). Chronology and morphology of temperature-dependent sex determination. *J.Exp.Zool.,* 260, 371-381.

Wibbels, T., Gideon, P., Bull, J. J. & Crews, D. (1993). Estrogen- and temperature-induced medullary cord regression during gonadal differentiation in a turtle. *Differentiation,* 53, 149-154.

Wilhelm, D. (2007). R-spondin1--discovery of the long-missing, mammalian female-determining gene? *Bioessays,* 29, 314-318.

Wilhelm, D. & Englert, C. (2002). The Wilms tumor suppressor WT1 regulates early gonad development by activation of Sf1. *Genes Dev.,* 16, 1839-1851.

Wilhelm, D., Martinson, F., Bradford, S., Wilson, M. J., Combes, A. N., Beverdam, A., Bowles, J., Mizusaki, H. & Koopman, P. (2005). Sertoli cell differentiation is induced both cell-autonomously and through prostaglandin signaling during mammalian sex determination. *Dev.Biol,* 287, 111-124.

Wilhelm, D., Palmer, S. & Koopman, P. (2007). Sex determination and gonadal development in mammals. *Physiol Rev,* 87, 1-28.

Willingham, E., Baldwin, R., Skipper, J. K. & Crews, D. (2000b). Aromatase activity during embryogenesis in the brain and adrenal-kidney-gonad of the red-eared slider turtle, a species with temperature-dependent sex determination. *Gen.Comp Endocrinol.,* 119, 202-207.

Willingham, E., Baldwin, R., Skipper, J. K. & Crews, D. (2000a). Aromatase activity during embryogenesis in the brain and adrenal-kidney-gonad of the red-eared slider turtle, a species with temperature-dependent sex determination. *Gen Comp Endocrinol.,* 119, 202-207.

Yao, H. H. & Capel, B. (2005). Temperature, genes, and sex: a comparative view of sex determination in Trachemys scripta and Mus musculus. *J.Biochem.(Tokyo),* 138, 5-12.

Yao, H. H., DiNapoli, L. & Capel, B. (2004). Cellular mechanisms of sex determination in the red-eared slider turtle, Trachemys scripta. *Mech.Dev.,* 121, 1393-1401.

Yao, H. H., Whoriskey, W. & Capel, B. (2002). Desert Hedgehog/Patched 1 signaling specifies fetal Leydig cell fate in testis organogenesis. *Genes Dev.,* 16, 1433-1440.

Yntema, C. L. Effects of incubation temperatures on sexual differentiation in the turtle, *Chelydra serpentina.* J Morph. 150[453], 462. 1976.
Ref Type: Journal (Full)

Yoshimoto, S., Okada, E., Umemoto, H., Tamura, K., Uno, Y., Nishida-Umehara, C., Matsuda, Y., Takamatsu, N., Shiba, T. & Ito, M. (2008). A W-linked DM-domain gene, DM-W, participates in primary ovary development in Xenopus laevis. *Proc.Natl.Acad.Sci.U.S.A,* 105, 2469-2474.

Young, M. & McPhaul, M. J. (1998). A steroidogenic factor-1-binding site and cyclic adenosine 3',5'-monophosphate response element-like elements are required for the activity of the rat aromatase promoter in rat Leydig tumor cell lines. *Endocrinology,* 139, 5082-5093.

Zaborski, P. (1986). Temperature and estrogen dependent changes of sex phenotype and HY antigen expression in gonads of a newt. *progress in developmental biology* (pp. 163-169). NewYork: Alan R. Liss, Inc.

Zaratiegui, M., Irvine, D. V. & Martienssen, R. A. (2007). Noncoding RNAs and gene silencing. *Cell,* 128, 763-776.

Zhang, H., Thomsen, J. S., Johansson, L., Gustafsson, J. A. & Treuter, E. (2000). DAX-1 functions as an LXXLL-containing corepressor for activated estrogen receptors. *J Biol Chem.,* 275, 39855-39859.

Zhang, H., Zhang, R. & Liang, P. (1996). Differential screening of gene expression difference enriched by differential display. *Nucleic Acids Res.,* 24, 2454-2455.

Zhao, Y., Lu, H., Yu, H., Cheng, H. & Zhou, R. (2007). Multiple alternative splicing in gonads of chicken DMRT1. *Dev.Genes Evol.,* 217, 119-126.