Molecular mechanism of sex determination in silkmoths: A Lepidopteran model

Thesis submitted to

Manipal University

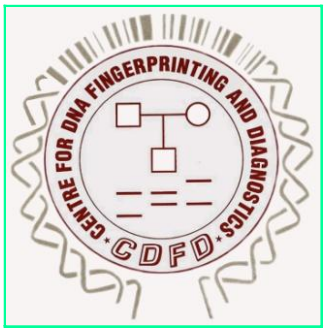
for the degree of

Doctor of Philosophy

By

Jayendra Nath Shukla

(Regd.No.050100016)



Centre of Excellence for Genetics and Genomics of Silkmoths

Laboratory of Molecular Genetics

Centre for DNA Fingerprinting and Diagnostics

Hyderabad-500 001

August 2009

DECLARATION

The research work embodied in this thesis entitled **"*Molecular mechanism of sex determination in silkmoths: A Lepidopteran model*"** has been carried out by me at the Centre of Excellence for Genetics and Genomics of Silkmoths, Laboratory of Molecular Genetics, Centre for DNA Fingerprinting and Diagnostics (CDFD), Hyderabad, under the guidance of Dr. J. Nagaraju. I hereby declare that this work is original and has not been submitted in part or full for any degree or diploma of any other University or Institution.

Jayendra Nath Shukla

Regd.No.050100016

Centre of Excellence for Genetics and Genomics of Silkmoths
Laboratory of Molecular Genetics
Centre for DNA Fingerprinting and Diagnostics
Hyderabad-500 001

Content

Acknowledgements

Every effort in this world comes to its fruitful culmination not because of sincere work of one but only due to the combined support and endeavour of many in whose reality one lives his life. It is my immense pleasure to utilize this space of my work to acknowledge all those who directly or indirectly brought light to this mass of work.

First and foremost I thank my guide Dr. J. Nagaraju for allowing me to work in his laboratory, for his full fledged cooperation, endless patience, affection and constructive suggestions. He not only taught me to work independently but also trained in the ethics of science. His critical comments always worked as a booster dose for me to work in a better fashion. I would like to pay my sincere regards to Savitri mam, for her caring nature and supportive talks.

I thank Dr. Hasnain for considering me eligible and providing me the platform to conduct my research in CDFD. I thank Dr. J Gowrishankar for his intellectual ideas and suggestions during Genetic club seminars, briefings and other scientific meetings. I pay my regards to Dr. Shekhar Mande for his guidelines to complete this highest academic degree. I thank Dr. Sanjeev khosla Dr. Gayatri ramakrishna, Dr. Murli bhashyam and Dr. Shubhodeep chatterji for their scientific suggestions, whenever needed.

My sincere thanks to my seniors- Dr. Sriramana, Dr. Mrinal, Dr. Satish, Dr.VLN reddy and EVS for their kind help and valuable guidance during this study. It is hard to define cooperation, love and support provided by all my batch mates (JRF 2003), Geetha, Mrk, Gokul, Arun, Santosh, Madhav, Pankaj, Uma, Nasreena, Bibhusita, Rakesh, Smanala and Noor. I really enjoyed having them as my batchmate and friends. I can never forget the time spend together by us during our coursework, gatherings and stage performances from time to time.

When it comes to my stay in Hyderabad, I can never forget my golden bachelor days spend with Gokul, Vaibhav, Anoop and Sudhish. I am very thankful to them for their fun-filled arguments and care.

I thank my lab mates, Arun, Jyoti, Asha, Suresh, Chandrapal, Deepa, Varsha, Muthulakhsmi, Satyavati, Shobhan, Nagamanju, Aadarsh, Annapurna, Kali, Swarna, Vijay, Archana, Srividhya, Murli, Venkateshwary, Jv and Gopinath for creating the convenient work atmosphere in the lab. I thank my group (Vijay, Kali, Swarna, Asha, Santosh) with whom I enjoyed daily lunch in CDFD canteen, the talks during which were quite rejuvenating after an exhaustive session of lab work. I thank Vijay and Archana for solving all my doubts regarding bioinformatics, and giving my thesis a better shape. Words fail to express my feeling for Santosh Jadhav whose sincere help, cooperation and cosy affection made my work much easier.

My special thanks to Ramesh Babu, Chandrashekhar and Vijay bhaskar for their daily assistance in the lab. I thank Mr. Sanjeev Rao, Mr. Ravinder, Mr. Santosh, Mr. Siddiqui, Ms. Prasanthi, Ms. Punitha and all other administrative staff for their help throughout. I am thankful to the security staff, drivers, maintenance staff and dispatch section people for their assistance. I also thank the research committee at Manipal University for registering me with them for my PhD degree.

I thank my seniors and juniors of Pantnagar (Brajesh Sir and Pratibha mam, Arvind Sir, Amit Sir, Kt Sir, Pankaj and Anoop) who were always there to help and support me with their valuable suggestions, both scientific as well as personal. I really cherished the moments spend with them. I appreciate the friendship of Ira and Ramji with whom I have spend memorable moments.

I am lucky enough to get the homely atmosphere away from my home provided by my didi (Ranjana) and jijaji (Mr. G. N Shukla). The love, affection and their caring nature supported me a lot during my stay in Hyderabad. I really enjoyed watching Chinu's witty activities and sincere talks with Mili.

I like to acknowledge my love, my wife (Nidhi) who came in my life as a sign of good luck, filling the empty space of my life and providing me all the happiness. I thank her for being with me during my happy and tough times with equal firmness. I expect the same support, love and care from her throughout life.

My parents, the most important people of my life provided me unlimited love and support throughout this study. Their constant encouragement and moral boosting talks made me to walk with patience even in most difficult situations. I thank all my family members and my ancestors for their blessings which always motivated me to choose the right path with endurance and never let me down.

My elder sister (Dr. Prachi) and jijaji (Dr. Pankaj) helped me to solve my problems and always supported me with their sincere suggestions. I pay my regards to them for being with me. My talks with my niece Shreya always proved to be refreshing for me whenever I was down. The love and affection by my sisters (Richa and Jyoti) always helped me to move further and trace the path without any hiccups.
The support from my in laws was a dose of positive thinking and determination to achieve my goal. Especially the innocent talks with Nikky were quite revitalizing. Thanks to Ranjana didi and Piyush bhaiya for their loving and caring attitude which always kept me optimistic.
My heart expresses sincere thanks to all those whose names have been unable to displayed here but without whom either directly or indirectly this work could not have seen the light of success.
Finally I thank the almighty to fill the spaces of my life with loving and caring people without whom it was hard to accomplish my goal.

Jayendra

List of Abbreviations

1. ESD- Environmental sex determination
2. TSD- Temperature-dependent sex determination
3. GSD- Genotypic sex determination
4. CSD- Complementary Sex Determination
5. TSP- Temperature sensitive period
6. GFP- Green fluorescent pigment
7. dsRNA- Double stranded RNA
8. dNTP- Deoxyribonucleotide triphosphate.
9. Dpc- Days post coitum
10. RE- Repeat elements
11. PRE- Proline rich element
12. SDS- Sodium dodecyl sulfate
13. RNAi- RNA interference
14. GO- Gene ontology
15. IBS- Insect buffer saline
16. DTT- Dithiothreitol
17. DNA- Deoxyribonucleic acid
18. cDNA- Complementary DNA
19. PCR- Polymerase chain reaction
20. RT-PCR- Reverse transcription-PCR
21. DOP-PCR- Degenerate oligonucleotide primed PCR
22. RAPD- Randomly amplified polymorphic DNA
23. RACE- Rapid Amplification of cDNA Ends
24. PESS- Putative exonic splicing silencer
25. PESE- Putative exonic splicing enhancer
26. TDF- Testis determining factor
27. EST- Expressed sequence tag
28. SIT- Sterile insect technology
29. ORF- Open reading frame
30. TDY- Testis determining region on Y
31. PAR- Pseudo autosomal region
32. HMG- High mobility group

33. DEPC- Diethylpyrocarbonate
34. GRA- Gel retardation assay
35. µl- Microliter
36. pm- Picomoles
37. U- Unit

Synopsis

Self perpetuation by sexual reproduction is the rule in the animal kingdom. Sexual reproduction is characterized by two basic features: sex determination and sexual differentiation. Sex determination, the primary decision made very early during embryonic development results in the development of anatomically and physiologically distinct male and female organisms and ensures sexual reproduction. In this process the undifferentiated "bipotential" embryonic gonads are committed to either testes or ovaries, guided by genetic or environmental cues. The genes involved in this developmental pathway are '*sex determination genes*'. Sexual differentiation, the process of formation of sexually dimorphic structures that define the male and the female adults, are governed by the '*sex-cytodifferentiation genes*' which are in turn regulated by the *sex determination* genes. The sex determining genes vary remarkably between species and a variety of mechanisms are known to operate to produce two distinct sexes. A very well characterised gene *doublesex* (*dsx*), the most downstream gene of the '*Drosophila* sex determination cascade', is found to be conserved in many different insect species, it is also conserved in worms and higher organisms where it is known as *Mab-3* (in *C. elegans*) and *Dmrt1* (in vertebrates), respectively. Other genes of the *Drosophila* sex determination cascade are *Sex lethal* (*Sxl), transformer* (*tra), fruitless (fru)* and *intersex* (*ix*). In the order Lepidoptera (moths and butterflies) where a female heterogamety exists, the sex determination studies are still in their infancy. The only lepidopteran where sex determination studies have been initiated recently is the silkworm, *Bombyx mori* because of its economic importance, ease of handling, availability of a large number of genetic mutants, linkage maps, whole genome sequencing and diverse amount of genetic stocks. Efforts to study the sex determination pathway genes in *B. mori* have lead to the characterization of *dsx* homologue i.e., *Bmdsx*. Other related wild silkmoth species are completely unexplored and no information on sex determination mechanisms is available. In order to understand the mechanism of sex determination operative in silkmoths, two species i.e., *Antheraea assama* and *Antheraea mylitta* were selected with the objective of characterizing their *dsx* homologues.

Chapter I

Diverse variety of sex determination mechanisms exists in nature. Molecular similarities between sexual regulatory genes, in different phyla, have begun to indicate that sex determination might be ancient but highly diverged, and seem to support evolutionary models in which sex determination pathways are formed by sequential addition of upstream regulators. In Chapter I, the existence of various sex determination systems studied in model organisms, outlining the main emphasis on the molecular mechanisms of sex determination in vertebrate and invertebrate systems including, *Drsosphila*, *C. elegans*, *Apis mellifera* and human are described. The sex determination pathways have been discussed in comparative terms, owing to conservation and diversity of the pathway genes. Finally, recent advances in the lepidopteran sex determination studies which mostly include the studies on the domesticated silkmoth *Bombyx mori* are discussed.

Chapter II

In Chapter II, the isolation and structural characterization of *dsx* homologues from two economically important wild silkmoth species, *Antheraea assama* and *Antheraea mylitta* are discussed. *A. assama,* confined to Assam state of India, is known for its golden coloured, glossy and fine textured silk threads, known as Muga silk. The females are heterogametic (ZO) and males are homogametic (ZZ) sex. Absence of W chromosome in this species makes it a very interesting system to study the mechanism of sex determination. *A. mylitta*, widely distributed in the tropical states of India such as Bihar, Jharkhand, Madhya Pradesh, Orissa, Andhra Pradesh and Karnataka, produces silk known as tasar silk. In order to obtain the *dsx* homologues, *Aadsx* and *Amdsx* from *A. assama* and *A. mylitta*, respectively, degenerate primers were designed for the most conserved regions (DM domain and OD2 domain) on the basis of sequence alignment of Dsx protein from different insects. RT-PCR of cDNA from these species using degenerate primers led to the identification of *dsx* homologues in these silkmoths. 5’ and 3’ RACE was done to obtain the full length transcripts of *Aadsx* and *Amdsx*. Sex-specific splice forms were identified and further confirmed by RT-PCRs using specific primers followed by sequencing. Interestingly, the pre-mRNAs of *Aadsx* and *Amdsx* produce

multiple sex specific variants, the conceptual translation of which suggested the presence of two female specific and one male specific Dsx proteins in these species.

Chapter III

Sex-specific Dsx proteins are known to be involved in the regulation of downstream sex differentiation genes. The mechanism of this regulation has been well studied, both in the case of *Drosophila* and *Bombyx*. In this chapter a detailed functional analysis of *Aadsx* gene is presented. Unlike the other known insect species including *B. mori* where only one male-specific and one female-specific Dsx proteins are produced, the wild silkmoths (*A. assama* and *A. mylitta*) produce two female and one male-specific proteins. To analyse the functional significance of the female specific proteins, knockdown experiments were conducted: in one experiment all the six female and one male transcripts of *Aadsx* together were targeted and in the second experiment combinations of different female transcripts on the basis of their ORFs were knocked down. The dsRNA mediated knockdown of *Aadsx* was highly effective as suggested by the abolishment of the target transcripts. Further the effect of knock down was assayed by analyzing the expression profile of its downstream target genes (*vitellogenin* and *hexamerin*), and by observing morphological defects, if any, in the gonads. Interestingly, our results suggest that both the female specific Dsx proteins are required for the expression of downstream target genes essential for sexual differentiation. *Aadsx* knockdown was quite effective even in the next generation as evident by the significant reduction in fecundity and hatchability.

Chapter IV

In this chapter identification and characterization of the new splice form of *B. mori dsx* (*Bmdsx*) gene is described. According to the previous reports, the primary transcript of *Bmdsx* gene splices to generate a male specific and a female specific transcript, which ultimately generate male- and female-specific BmDsx proteins. Search for the 15bp sequence unique to a group of female specific *Aadsx* and *Amdsx* transcripts in the EST database of *B. mori*, lead to the identification of an EST (BP121180), having this sequence. Alignment of the EST sequence with the previously reported *Bmdsx* transcripts confirmed the presence of a

third splice form of *Bmdsx*, having an additional stretch of 15bp sequence. This is a novel splice form we are reporting for the first time. Protein formed by virtual translation of this newly identified splice form confirmed it to be female specific. Expression analysis confirmed the presence of two female splice forms of *Bmdsx*.

Results obtained from this study point towards the diversity in the *dsx* gene within the insect community. Contrary to the existence of one female and one male-specific Dsx proteins in all the reported insect species, results from this study confirm the presence of two female and one male-specific Dsx proteins in silkmoths. The functional analysis of female-specific Dsx proteins in *A. assama* confirms their role in sexual differentiation.

Chapter - I

General Introduction

1.1 Introduction

Silk and the silk trade have contributed to an early form of globalization for nearly 2000 years during the 'Silk Road Era', enriching the human endeavour through art and culture (Kurin, 2002). Silkworms belong to the order Lepidoptera, which includes more than 160,000 species (http://www.nhm.ac.uk/entomology/lepindex), of which Bombycoid moths include silkmoths of economic importance. Bombycoid moths secrete diverse varieties of silk fibers. On the basis of their food plants, the different silkmoths have been grouped into two major types i.e., the mulberry and the non-mulberry silkmoths. The mulberry silk is produced by the domesticated silkworm, *Bombyx mori*, which belongs to family Bombycidae and the non-mulberry silks are produced by different species of the family Saturniidae. *B. mori* is intimately connected with humans from the time of the legendry fall and unravelling of a cocoon in the tea cup of Chinese princess, Xi Lingshi (Kurin, 2002). Saturniid moths, estimated to be around 1,300 to 1,500 species with worldwide distribution, are among the largest and most spectacular of the order Lepidoptera which includes giant silkmoths, royal moths and emperor moths. The major species of wild silkmoths include *Antheraea assama* (Indian golden silkmoth), *Antheraea mylitta* (Indian tropical tasar silkmoth), *Antheraea roylei* (Indian oak tasar silkmoth), *Antheraea yamamai* (Japanese oak silkmoth), *Antheraea pernyi* (Chinese oak silkmoth) and *Samia Cynthia ricini* (Indian castor silkmoth) (Grimaldi and Engel, 2005).

The study of silkworm biology and its genetics is as old as the history of silkmoths. Beginning in the nineteenth century, the silkworm became a model for scientific discovery in microbiology, physiology and genetics at a period when enormous paradigm changes took place in our understanding of biology (Willis et al., 1995). The domesticated silkmoth, *B. mori* became an important species for research in insect physiology and developmental biology (Goldsmith and Kafatos, 1984) owing to its large body size, ease of rearing, short life span, and relatively simple body plan. As a result, *B. mori* has continuously attracted researchers and emerged as a model organism, second only to *Drosophila melanogaster* (Goldsmith et al., 2005; Tazima, 1964.). The wealth of knowledge of genetics and genomics of *B. mori*, includes 400 mutations distributed over 28 linkage groups that correspond to 28 haploid sets of chromosomes, more than 3000 silkworm strains that represent diverse geographic origins, high resolution molecular linkage maps, EST databases including more

than 2.5 lakh ESTs, draft genome sequence, germ line transformation system and RNAi technologies (Goldsmith et al., 2005; Nagaraju and Goldsmith, 2002).

The order Lepidoptera also includes many pests, among which the family Noctuidae represents some of the most devastating pests of agriculture, particularly the *Helicoverpa armigera*, *Heliothis zea*, *Heliothis virescens* (Heliothines) and many others including *Mamestra*, *Spodoptera* and *Trichoplusia*. Although lepidopteran pests of agriculture, forestry and food storage are found in diverse families (*Noctuidae, Pyralidae, Arctiidae, Sphingidae, Tortricidae, Lymantridae, and Yponomeutidae*) they all belong to the same suborder, Ditrysia, and are not very divergent phylogenetically. Thus the genomic and genetic information generated for the model species of lepidopterans i.e., *B. mori* and related species can be applied to the lepidopteran pest species in order to design genetic strategies to control them.

The silkworm life cycle comprises four morphologically distinct stages: egg, larva, pupa and moth. The silkworms have been broadly classified into diapausing and non-diapausing strains, based on their hibernation property. The diapausing strains hibernate during winter and are referred to as uni (one) or bivoltine (two) depending on the number of life cycles completed per year. Non-diapausing strains, also referred to as polyvoltines, do not hibernate and go through multiple generations in a year. The larval stage comprised of five instars, is the only feeding stage. *B. mori* larvae are monophagous and feed only on mulberry (*Morus alba*). Wild silkworms, on the other hand, are polyphagous in nature and feed on multiple host plant species. The Indian tasar silkmoth, *Antheraea mylitta* feeds on the leaves of *Terminalia arjuna*, *Terminalia tomentosa* and *Shorea robusta*. It undergoes hibernation at pupal stage. The Indian golden silkmoth, *Antheraea assama*, a continuous breeder without any intervening hibernation, feeds on *Litchia polyantha* and *Machilus bombycina*. The life cycles of *A. assama*, *A. mylitta* and *B. mori* are shown in **Fig. 1.1a, 1.1b** and **1.1c**, respectively.

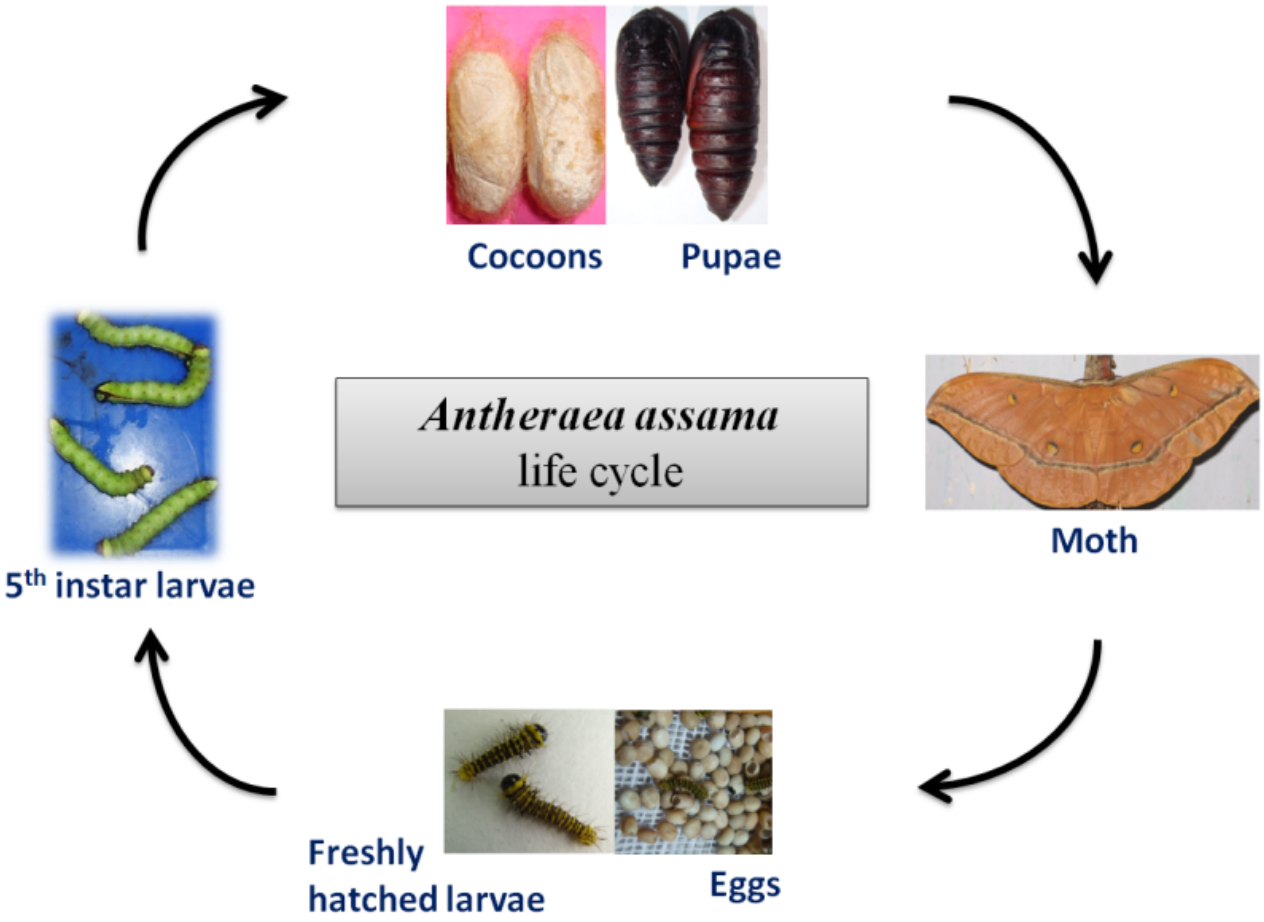


Fig. 1.1a Life cycle of Indian muga silkworm, *Antheraea assama*. Each fertilized female moth lays approximately 180-200 eggs which hatch in 9-10 days. Larvae pass through four moults to reach to the fully grown fifth instar stage in 25-26 days. The matured fifth instar larva secretes silk fiber of about 500-800 meters in order to encapsulate itself within the cocoon. The larva metamorphoses into pupa which rests within the cocoon for 15-30 days after which the adult moths emerge by dissolving one end of cocoon. Moths are non-feeding and mate to produce eggs and live up to one week.

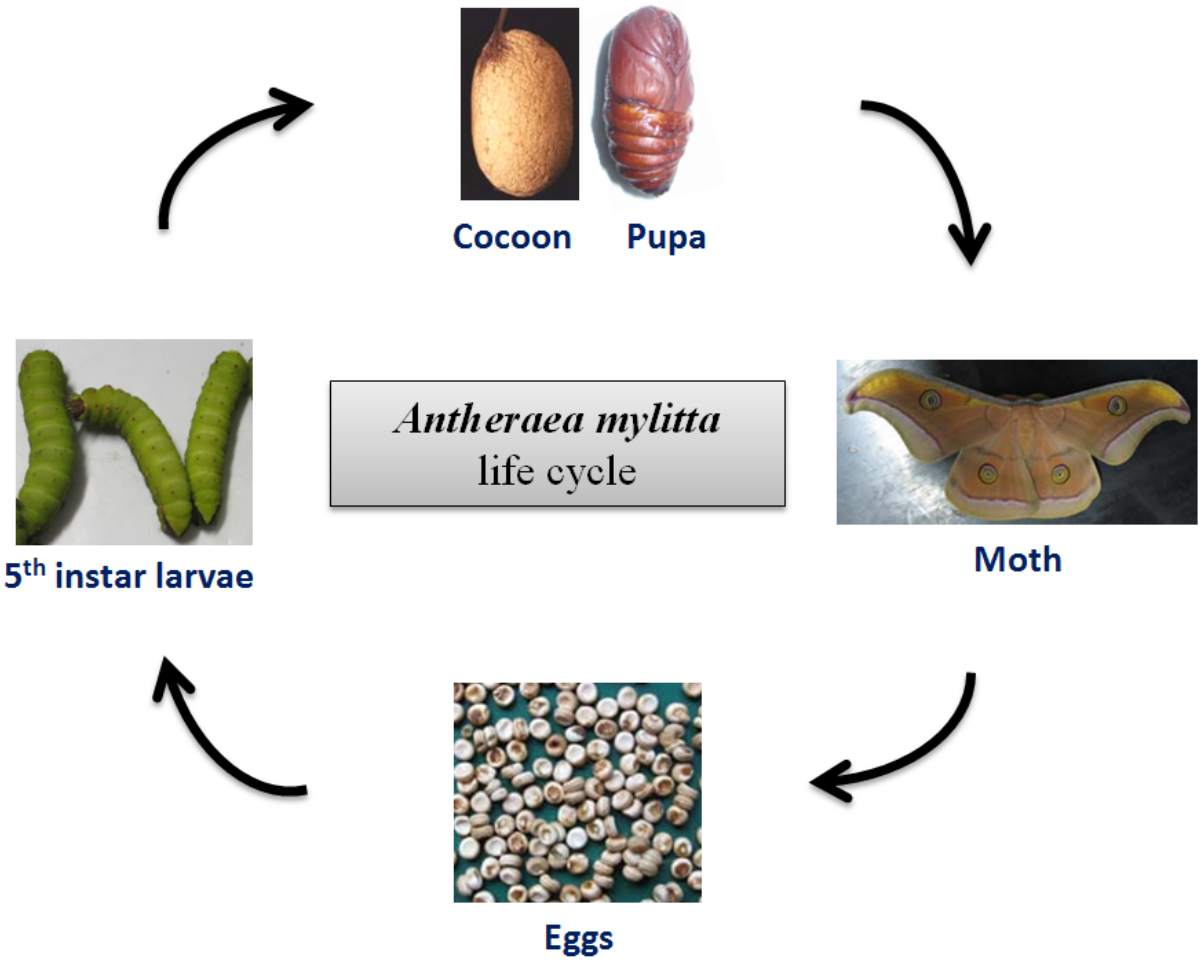


Fig. 1.1b Life cycle of Indian tasar silkworm, *Antheraea mylitta*. Each fertilized female moth lays approximately 180-200 eggs which hatch in 9-10 days. Larvae pass through four moults to reach to the fully grown fifth instar stage in 25-30 days. The matured fifth instar larva secretes silk fiber of about 800-1000 meters in order to encapsulate itself within the cocoon. The larva metamorphoses into pupa which rests within the cocoon for 15-30 days after which the adult moths emerge by dissolving one end of cocoon. Moths are non-feeding and mate to produce eggs and survive up to a week.

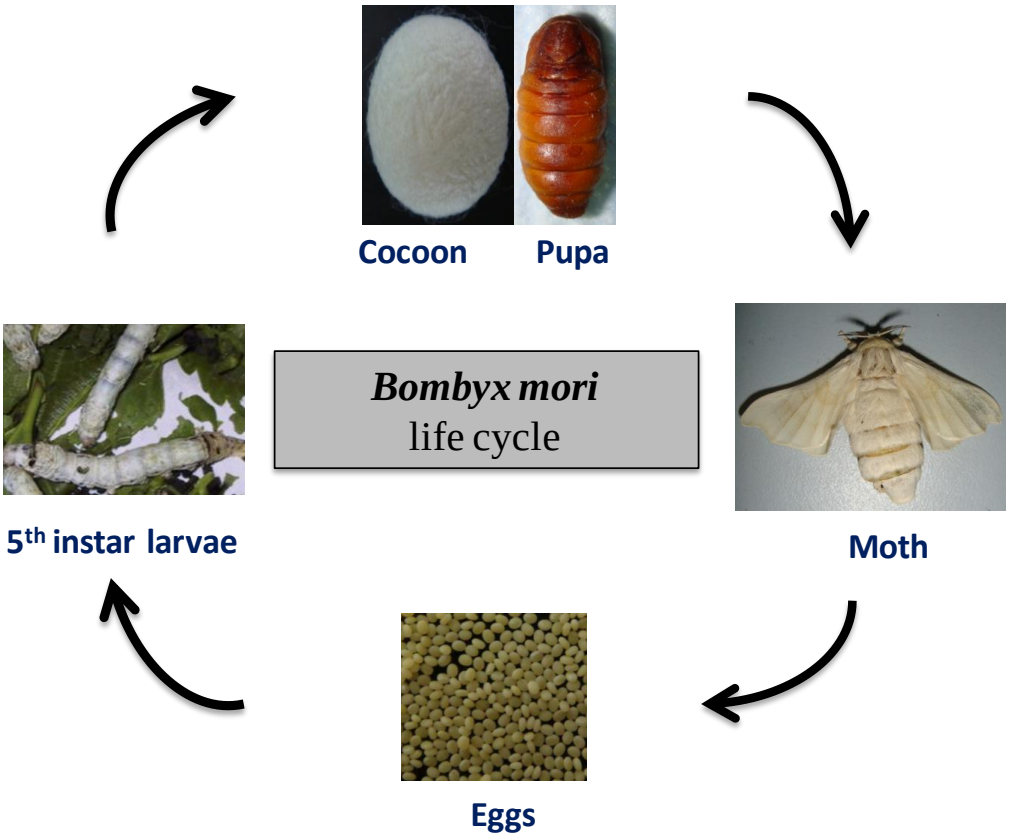


Fig. 1.1c Life cycle of mulberry silkworm, *Bombyx mori*. Each fertilized female moth lays approximately 300 eggs which hatch in 9-10 days. Larvae pass through four moults to reach to the fully grown fifth instar stage in 17-23 days. The matured fifth instar larva secretes silk fiber of about 400-700 meters in order to encapsulate itself within the cocoon. The larva metamorphoses into pupa which rests within the cocoon for 10-12 days after which the adult moths emerge by dissolving one end of cocoon. Moths are non-feeding and mate to produce eggs and survive for a week.

Comparative studies that focus on behavioural and developmental features at the molecular level in different organisms are always illuminating because such studies provide evolutionary insights into the molecular players involved in the developmental process. In this respect sex determination process seems to be exceptionally suitable for comparative studies, given the great variety of sex determination mechanisms that exist in the animal kingdom (Sanchez, 2008). Apart from this, in recent years molecular understanding of sex determination mechanism has been successfully employed in sterile insect technology (SIT) programs, for the control and eradication of insect pests and insect vectors of human diseases (Ruiz et al., 2005). The SIT program can also be applied to insect pests belonging to the order Lepidoptera, once the sex determination mechanism and the genes involved in the pathway are known for the insects belonging to this order. Besides, sex determination, from the evolutionary point of view, is an example of extremes and much can be gained from the study of opposite extremes.

1.2 Sex Determination Mechanisms

Sex determination is a fundamental biological process of profound significance in development, genome evolution and maintenance of proper sex ratio in the population (West et al., 2002). It is among the most important primary decisions made during the embryonic phase of life, ensuring the development of anatomically and physiologically distinct male and female organisms thus allowing sexual reproduction to prevail. The animal kingdom possesses myriad mechanisms via which sex of the individual is decided early in development (Bull, 1983; Zarkower, 2001). In many organisms, the sex of an offspring is irreversibly determined by its sex chromosomes, or rather, a set of genes on these 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 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 to as "temperature-dependent sex determination" (TSD). Studies have shown that genes involved in primary sex determination evolve rather rapidly (de Bono and Hodgkin, 1996; Lundrigan and Tucker, 1997; Tucker and Lundrigan, 1993) and the same set of sex determining genes of one species may not have similar role in the process of sex determination in related species (Graves, 1995; Jimenez et al., 1996). Various sex determination mechanisms exist in different groups of organisms which include haplodiploidy (males derived from haploid eggs, females from diploid eggs), paternal genome loss (sex determined by loss of paternal chromosomes after fertilization), male heterogamety (males with heteromorphic XY sex chromosomes and female with homomorphic XX), female heterogamety (ZW females and ZZ males), polygenic sex determination, environmental sex determination, and a variety of other mechanisms (Bull, 1983; Uller et al., 2007). Study of sex determination on different model organisms has shown conservation in the basic mechanism of sex determination process which follows any one of the following three mechanisms:

1.2.1 Environmental Sex Determination (ESD)

Here, the sex of the individual is influenced by the environmental factors ranging from the temperature at which the embryo develops, to the social behaviour of the individual

after birth. This system operates in reptiles (Ewert and Nelson, 1991), some amphibians (Dournon et al., 1990) and fishes (Baroiller and D'Cotta, 2001).

1.2.2 Chromosomal Sex Determination (CSD)

In this case the sex chromosome constitution of the individual decides its sex. The CSD can be subdivided into various categories, the identical pair of sex chromosomes may be present in both males (birds) and females (mammals) and their major sex-determining gene may be, either known (e.g., mammalian SRY) or still suspected (Bull, 1987).

It is generally accepted that ESD is the ancestral state and the GSD is evolved as a derived condition. It is also recognized that GSD is evolutionarily highly labile, having evolved multiple times independently across diverse taxa (Sarre et al., 2004).

a) XY system:

XX Females and XY Males: Here homogametic sex is female (only one type of gametes i.e., 100% X) and the heterogametic sex is male (makes two types of gametes i.e., 50% X gametes and 50% Y gametes). Found in mammals (e.g., Humans and mice) and in some insects (e.g., *Drosophila*).

b) XO system:

XX Female and XO Male: In this system one of the sex chromosomes is absent in individuals of a particular sex (e.g., grasshoppers and crickets). Generally females are euploids whereas males are aneuploids for the sex chromosome constitution. All the female gametes carry X chromosome whereas, only 50% of male gametes carry X chromosome, the remaining 50% male gametes do not carry the sex chromosome.

c) ZW system:

ZW Females and ZZ Males: Here the female is a heterogametic sex which makes two types of gametes i.e., 50% gametes having Z chromosome and the other 50% having W chromosome. Males are homogametic sex making only one type of gamete (all having Z chromosome). This system is found in birds and some insects, (e.g, moths and butterflies).

1.2.3 Haplodiploidy Sex Determination

There are a few animal species, notably the hymenopteran insects (ants, wasps, bees) which exhibit haplodiploidy i.e., unfertilized eggs develop into haploid (at all genetic loci) males and fertilized eggs develop into diploid (at all genetic loci) females. This is an extreme form of genotypic sex determination in which all chromosomes behave effectively as sex chromosomes. Individuals homozygous or hemizygous at a sufficient number of sex determining loci, become males and individuals heterozygous at a sufficient number of these loci become females.

1.3 Sex determination in vertebrates

Sex determining mechanism varies considerably across vertebrates. During early embryonic stages gonads appear as bipotential primordia. The undifferentiated "bipotential" embryonic gonads are committed either to become testes or ovaries, depending on genetic or environmental cues (Sekido and Lovell-Badge, 2007; Wilhelm, 2007; Yao and Capel, 2005). This process occurs both in GSD and TSD species. Once committed, the bipotential gonad differentiates into, either testis or ovary following complex cellular events programmed by an intricate regulatory network of genetic factors.

1.3.1 Environmental/Temperature dependent sex determination

In this case, the external environment during early embryonic development irreversibly determines the sex of an individual. The existence of ESD is clearly evident when sex ratio is altered significantly by the environmental factors, in which the developing embryos are incubated. A wide range of environmental factors has been shown to influence the sex determination process. For instance, factors such as osmotic stress, level of oxygen and carbon dioxide may affect sex determination as in the case of turtles (Ackerman, 1981a; Ackerman, 1981b; Gutzke and Paukstis, 1983). Effect of pH and social conditions on sex determination in some fish species have also been reported (Francis and Barlow, 1993; Holmgren and Mosegaard, 1996; Romer and Beisenhertz, 1996). Among various environmental factors, influence of temperature on sex determination (TSD) has been observed to be common, especially in a number of reptilian species. In addition, a few species of birds, fishes and amphibians have also been reported to have TSD (Valenzuela and Lance, 2004). 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 groups showed that TSD is widespread in reptiles mostly in all crocodilians and tuataras, and in many turtle species (Valenzuela and Lance 2004), and it occurs in natural as well as under artificial conditions. However, a vast majority of reptilians are yet to be investigated for the type of sex determination they undergo (Janzen and Paukstis, 1991).

In TSD, the prevailing temperature during the early embryonic development acts as a trigger for the sex determination. The specific period of development, when temperature trigger operates, is called temperature sensitive period (TSP). Temperature acts as trigger in at least three patterns, that are characterized by the sex ratios produced as a 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 most of the turtle species it is the other way round: 100% males at low temperature and 100 % females at high temperature (M→F). In some 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) (Bull, 1983; Lang et al., 1989). 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. These, as a consensus, have been classified as TSD-Ia, TDS-Ib and TSD-II. TSD-Ia is called 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 and Pieau, 1991). Accordingly, TSD-Ia/TSD-Ib has one pivotal temperature, while TSD-II 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).

Most of the members of three reptilian groups (crocodilians, lizards, turtles), that are known to exhibit TSD, follow TSD-ll pattern with the exception of sea turtles and 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-ll), there is only a very narrow temperature range at which only males or both sexes are produced. Transition zone can be defined as intermediate pivotal temperature at which 50 males: 50 females can be obtained. This transitional temperature zone 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 within species. For example, in a study by Janzen and Paukstis, 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). During the TSP, aromatase a cyt450 enzyme, responsible for the conversion of androgens into estrogens, appears in many organisms. In reptiles where steroidogenesis begins very early, prior even to the thermosensitive period, aromatase activity remains universally low (Belaid et al., 2001; Dorizzi et al., 1996; Dournon et al., 1990; Pieau et al., 1998). Aromatase activity increases with the onset of the TSP, the temperature for which varies for different species. For example, in marine and freshwater turtles, higher temperatures cause an exponential increase of aromatase activity, whereas at lower temperatures aromatase activity remains low. The undifferentiated gonads differentiate into testis or ovary guided by the different levels of aromatase activity (**Fig. 1.2**). Any change in the temperature after the onset of TSP (once the fate of the gonad is established) is not known to have any effect.

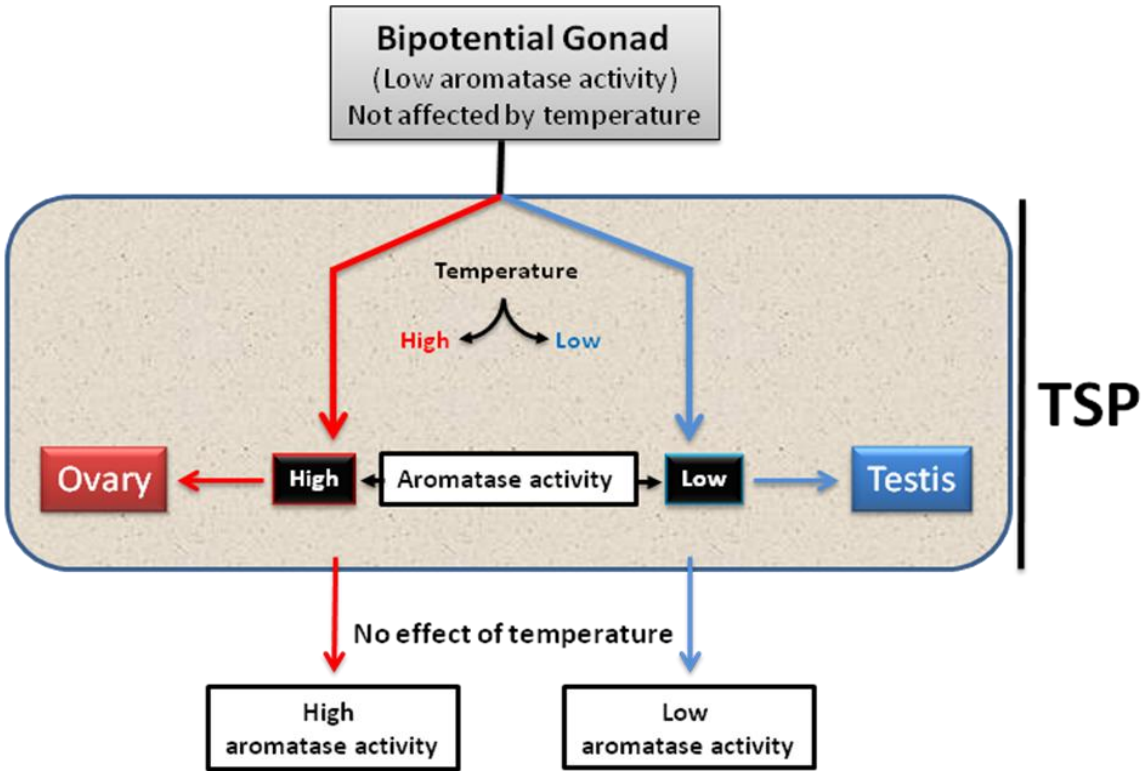


Fig. 1.2 Aromatase activity levels during the thermosensitive period (TSP) are regulated by the temperature of the environment and control gonadal differentiation. Changes in the environmental temperature before and after TSP do not seem to affect sex of the individual.

Interestingly, a number of genes originally described as part of the GSD system of sexual development in men and other mammals have also been detected in reptiles following TSD. For instance, in the sea turtle *Lepidochelys olivacea*, several genes related to mammalian sex determination are expressed, including *Dax1* (*dosage-sensitive sex reversal 1*), *Dmrt1* (*doublesex- and mab-3 related transcription factor 1*) and *Sox9* (*Sry related HMG box 9*) (Ferguson-Smith, 2007). In particular, *Dax1*, a known regulator of gonadal development in mice and other mammals, considered to be an "anti-testis" gene is not differentially expressed in response to temperature variation during the TSP. Therefore, its role in reptile sex determination is unclear. *Dax1* is also expressed in crocodilians such as *Alligator mississippiensis*, with temperature-dependent sex determination, but whether this gene could indeed be a target for androgen or estrogen-related actions following the TSP remains unknown. As far as *Dmrt1* is concerned, the gene was initially related to sex determination in *D. melanogaster*, due to the presence of a domain comparable to the downstream sex determination gene *doublesex*. Subsequent investigations have revealed the expression of *Dmrt1* homologue in several other species as well, including birds, fish and

reptiles. In alligators, such as *A. mississippiensis*, *Dmrt1* is expressed exclusively in the male gonads. Moreover, its expression appears to precede that of *Sox9*, another testis-specific gene conserved from reptiles to mammals. The latter is originally expressed in the bipotential gonad of reptile embryos, but following the TSP, it remains active only in males, suggesting it to be a candidate gene for sex steroid-induced regulation. In alligators, *Sox9* is also related to increased Anti-Mullerian Hormone (*Amh*) levels, but, contrary to mammals, *Amh* induction precedes that of *Sox9* (Western et al., 1999a; Western et al., 1999b). In case of lizards, attempts made to examine sexual dimorphism in the brain shows distinct differences in estrogen receptor expression and progesterone concentrations in specific areas of the central nervous system. These findings imply that aromatase regulation is only the first step in a sequence of several more complex sex-specific/dimorphic genetic phenomena that still remain to be examined. Finally, it has recently been suggested that aromatase may also be regulated by secondary parameters, other than temperature. This has been described for instance, in the case of Prostaglandin E2, which appears to be associated with increased aromatase activity (Morrish and Sinclair, 2002; Sinclair et al., 2002; Valleley et al., 2001).

1.3.2 Genetic sex determination

In the species following "genetic sex determination" (GSD), 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; Grutzner et al., 2004; Solari, 1994). GSD is seen in all the 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 mammals, sexual fate is determined genetically by the presence of the *Sry* (Sex determining Region of Y-chromosome) gene on the Y chromosome. In contrast to males, female heterogamety is found in birds and snakes with ZZ males and ZW females. To date, it is not clear whether the sex determining mechanism in ZW systems is due to a dominant ovary-determining gene residing on the W chromosome or the double dose of Z-linked genes which may trigger 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 et al., 1981; Engel and Schmid,

1981). However, it gave conflicting results, 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 and 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 this DNA as probe, sex reversal phenomenon was shown to be due to translocation of Y DNA onto the X chromosome in mouse (Jones and Singh, 1981; Jones and Singh, 1985; Singh and Jones, 1982). The sex-specificity of Bkm-related satellite DNA was also shown in a few species of sea turtle and fishes (Demas et al., 1990; Nanda et al., 1992). Notwithstanding these positive results, an unequivocal role of Bkm sequences in GSD is yet to be demonstrated. Studies on XXY/XO and abnormal individuals with sex chromosomes anomalies established that, sex in human is not decided by the X chromosomes and/or autosomes ratio, and that Y chromosome contains a dominant locus for sex determination. Later it became obvious that development of testis is associated with the presence of a Y linked locus called 'testis determining factor' (TDF) in human or 'testis determining region on Y' (TDY) in mouse. This was adjacent to the pseudo autosomal region (PAR), 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 in 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 in several 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 its role in sex determination. *Sry* expression is detected in the genital ridge at 10.5 days post coitum (dpc); it peaks at 11.5 dpc and by 12.5 dpc its expression

declines (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, many evidences support Sry as a key factor essential 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, who showed testicular dysgenesis leading to sex reversal in 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 downregulated in female (Kent et al., 1996; Morais da Silva et al., 1996). *Sox9* upregulation has been demonstrated by transient expression of *Sry* in Sertoli cells precursors (Saccone et al., 1998). The observations that *Sox9* is a target for *Sry* was recently demonstrated in an experiment where Sry binding on the *Sox9* promoter was shown to be essential for *Sox9* expression (Sekido and Lovell-Badge, 2008).

To date, a number of genes have been implicated in the process of genetic sex determination in mammals (**Fig. 1.3**). Potential role of some of these genes in gonadal development is now better understood by the genetic manipulation studies using mouse as a model. As indicated earlier, in GSD, sex of the embryo is decided at the time of fertilization however manifestation of sex determination occurs later in embryonic development. Before sex determination of the developing gonad a bipotential gonad forms which shows no morphological difference between male and female. Several of the genes (*Wt1, Sf1, Emx2,*

M33, *Igf1r* and *Lhx9*) are required for the maintenance of the bipotential primordium (Wilhelm et al., 2007).

In the female embryo, the Y chromosome is not present and, therefore, *Sry* is not expressed. The genetic cascade, regulating the differentiation of female reproductive system, is not as extensively studied as in men, but *Dax1* (and its regulatory system, including genes such as *Wnt4* and *Sf1*) is generally considered as a significant player in this process, which is how it came to acquire the rather simplistic description of the "antitestis gene". Regulation of sex steroid production is also important for the establishment of a normal female phenotype and is mediated via *Sf1* expression and aromatase enzyme complex induction. Two relatively recently described genes with a potential role in sex determination and differentiation are *Dmrt1* and *Stra8* (stimulated by retinoic acid gene 8) (Raymond et al., 1999). The first has been already discussed in previous paragraph as a conserved sex-related gene, bearing a DM domain originally studied in nematodes. In humans, XY sex reversal in cases of 9p chromosome deletions have been attributed to impaired action of *Dmrt1* or its homologue, *Dmrt2*. Still, the exact involvement of these genes in the sex determination circuit has not been clarified. *Stra8*, on the other hand, is exclusively expressed in female germ cells and its presence signals their gradual sexual differentiation, in an anterior to posterior direction. However, it has not yet been established whether the gene's product directly induces sex determination towards the female pathway, or rather acts as a simple marker of this phenomenon, without active participation in the process *per se*.

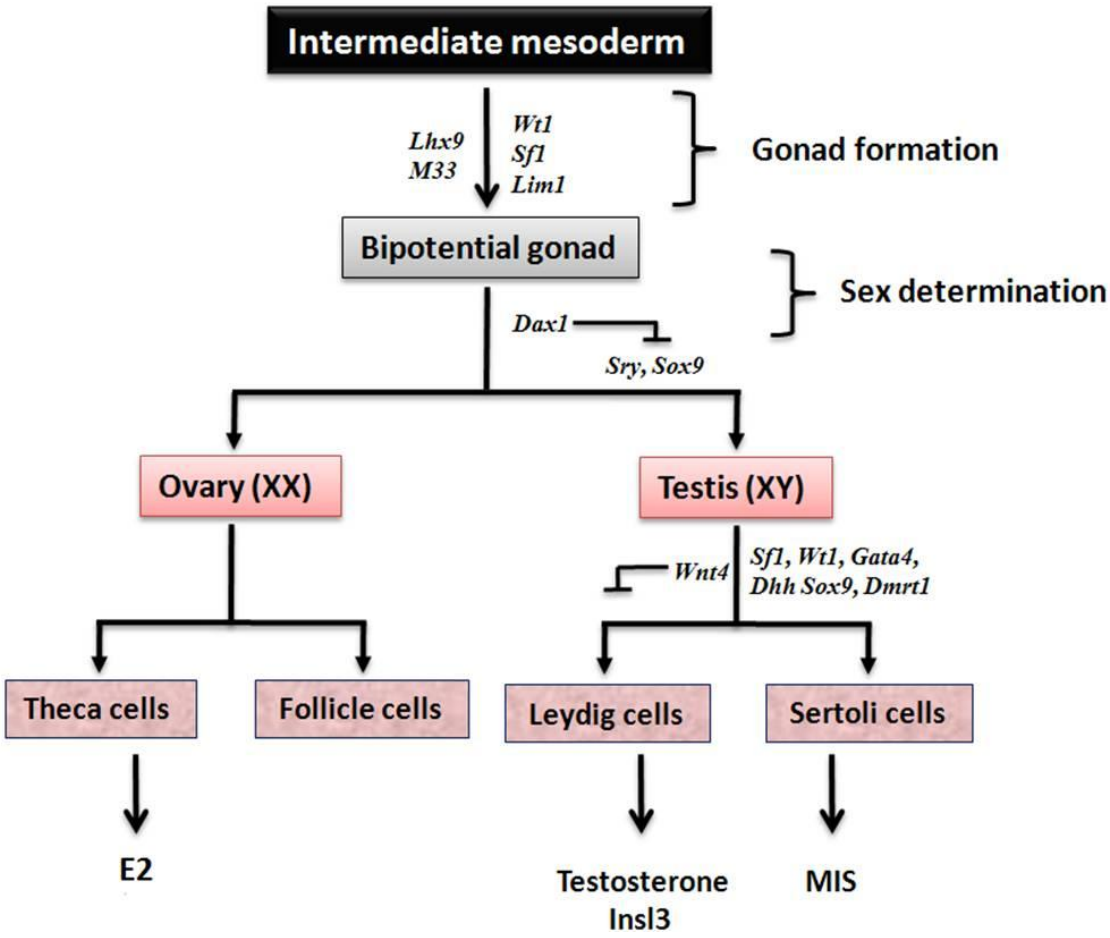


Fig. 1.3 Intermediate mesoderm gives rise to undifferentiated bipotential gonad, by the activity of the *Wilms tumour 1* (*Wt1*), *steroidogenic factor 1* (*Sf1*), LIM homeobox protein 1 (*Lim1*, also called *Lhx1*), LIM homeobox protein 9 (*Lhx9*), polycomb group M33 genes and possibly some other unknown genes. The sex determining region gene on the Y chromosome (*Sry*) and SRY-box containing gene 9 (*Sox9*) in XY embryos lead to the development of testis formation from the undifferentiated gonads. Lack of *Sry* in XX embryos results in ovary development. No ovary-determining gene comparable to *Sry* has been found, but *DAX1* (nuclear receptor subfamily 0, group B, member 1) seems to act as an 'antitestis' gene, and *Wnt4* (wingless-related MMTV integration site 4) suppresses Leydig-cell differentiation in the ovary, among other functions. The period till the formation of differentiated gonads, is the period of sex determination, after this, gonad differentiation starts. For different cell type differentiation of testis, several genes including *Sf1, Wt1, Gata4* (GATA-binding protein 4), *Dhh* (desert hedgehog), *Sox9* and *Dmrt1* (*dsx* and *mab-3*-related transcription factor 1), are required. Testis secretes the hormones, Mullerian inhibiting substance (MIS, also called anti-Müllerian hormone or AMH), from the Sertoli cells, testosterone and Insl3 (insulin-like 3) from the Leydig cells. The MIS and testosterone suppress female differentiation and support male differentiation, respectively, elsewhere in the body.

1.4 Sex determination in invertebrates

1.4.1 Sex determination in *Drosophila*

Unlike mammals, sex determination in flies is accomplished on a cell-by-cell basis. Each cell independently accesses its chromosome complement and establishes its sexual identity. Within the insect kingdom the sex determination pathway has been well elucidated in *Drosophila melanogaster*, a dipteran insect (Cline and Meyer, 1996; MacDougall et al., 1995; Sanchez, 2008; Schutt and Nothiger, 2000). Here the sex of the individual is decided by a well characterized pathway known as "sex determination pathway" (**Fig. 1.4**). The two sexes in *Drosophila* are XX females and XY males, where the Y chromosome is passive in the process of sex determination and is actively involved in the process of spermatogenesis, later in the differentiation of the male germ line (Hackstein, 1987; Lifschytz, 1987; Zhang and Stankiewicz, 1998). The X/A ratio, the ratio of number of X chromosomes to the number of autosomal sets, is the primary signal governing the process of sex determination (Bridges and Anderson, 1925); females are 2X:2A and males are X:2A. Thus regardless of the ploidy, the individuals with the X/A ratio of 1.0 are females and the individuals with the ratio of 0.5 are males. The individuals with the X/A ratio more than one, less than 0.5 or intermediate values between 0.5 and 1 are superfemales, supermales and intersexes, respectively. The signal generated by the X:A ratio, in the early developmental process, is communicated to the sex determination pathway through a master regulatory switch gene *Sex-lethal* (*Sxl*) (Cline, 1984; Penalva and Sanchez, 2003). The *Sxl* gene gets activated in females by the delicate balance between the dose-sensitive X chromosome numerator elements, which includes genes such as *sis-a*, *sis-b*, *runt* and less so *sis-c*, and the autosomal denominator genes such as *dead-pan* (*dpn*) in conjunction with the maternally derived products of the *daughterless* (*da*) gene and some other genes like *emc*, *groucho, her* and *snf* (MacDougall et al., 1995). The feminizing effect of the numerator elements is titrated against the masculinising denominators, with the maternally derived products of the rest of the genes acting as reference points (Manolakou et al., 2006). The *Sxl* gene is having two promoters, the early promoter (*pe*) and the late promoter or the maintenance promoter (*pm*) (Salz et al., 1989).

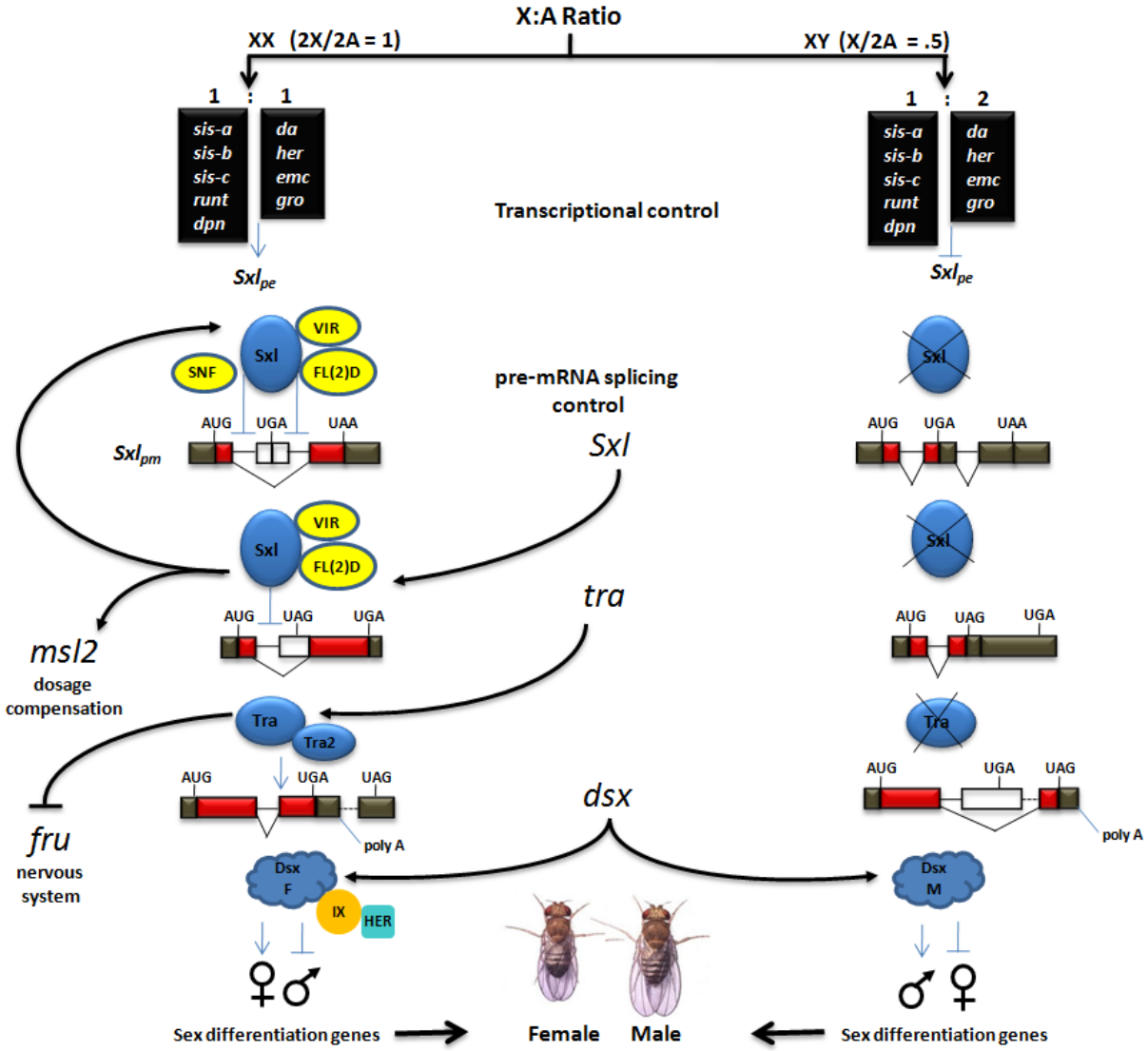


Fig. 1.4 Sex determination cascade of *D. melanogaster*. The X:A signal initiates the cascade by regulating the expression of the *Sxl* gene, the activation of which occurs only in embryos with the chromosomal constitution 2X:2A, so that the Sxl protein is produced only in females. Sxl regulates the splicing of its own pre-mRNA through a positive-feedback loop. The products of the *fl(2)d*, *vir* and *snf* genes are also necessary for this splicing regulation. The Sxl protein functions in females to control the activity of two different pathway genes: i) Sxl protein regulates *transformer* (*tra*) pre-mRNA splicing so as to generate a female-specific mRNA encoding the Tra protein, involved in somatic sexual differentiation pathway and ii) Sxl prevents dosage compensation in females by blocking the translation of *msl-2* mRNA. Tra forms a heterodimer with the Transformer-2 (Tra-2) protein that modulates the splicing of two genes: *doublesex* (*dsx*) and *fruitless* (*fru*). The female-specific *fru* mRNAs are not translated to sex-specific products. In the case of *dsx*, both the male- and female-specific *dsx* mRNAs encode zinc finger proteins that have identical DNA-binding domains but different carboxy termini. The *intersex* (*ix*) and *hermaphrodite* (*her*) genes are required for Dsx activity in females. In X(Y):2A embryos, no Sxl protein is produced, as a consequence, *tra* pre-mRNA follows a default splicing pattern and no functional product is generated, whereas *msl-2* mRNA is translated and dosage compensation ensues. In the absence of Tra, *dsx* and *fru* pre-mRNAs are spliced in the default male mode, producing male-specific *dsx* and *fru* mRNAs. FruM and DsxM control the expression of target genes necessary for male sex differentiation and behaviour. The *fru* branch of the sex determination hierarchy is responsible for nearly all CNS-dependent aspects of sexual differentiation.

At the blastoderm stage, the X:A ratio of 1 in females activates the *Sxl* gene through its early promoter i.e., *Sxl-pe*, (Keyes et al., 1992; Torres and Sanchez, 1991) whereas in case of males, *Sxl-pe* remains inactive. This leads to the production of an early pulse of Sxl protein only in females and once the activity of *Sxl* is established, the X/A signal is no longer needed (Bachiller and Sanchez, 1991; Sanchez and Nothiger, 1983). The *Sxl* transcript from its late promoter is produced in both male and female, which contains an in-frame stop codon. The early pulse of Sxl protein, in case of females, orchestrates the splicing of the *Sxl-pm* transcript to produce the functional Sxl protein (Bell et al., 1988; Bopp et al., 1991); males being deficient in the early pulse of Sxl protein, default splicing of the *Sxl* transcript, from late promoter, and production of the truncated non-functional protein occurs. Through an auto regulatory feedback loop *Sxl* manages to remain in an active state throughout the development only in female, through its sex-specific splicing (Bell et al., 1991; Cline, 1984; Keyes et al., 1992; Penalva and Sanchez, 2003). Therefore, the expression and regulation of *Sxl* from its late promoter *Sxl-pm*, throughout adult life after the blastoderm stage, occurs by the sex-specific splicing of its transcript, such that functional Sxl proteins are produced only in females. The activity of other genes, such as *sans fille* (*snf*) (Albrecht and Salz, 1993; Flickinger and Salz, 1994; Salz and Flickinger, 1996; Samuels et al., 1998), *femalelethal-2-d* (*fl(2)d*) (Granadino et al., 1990; Granadino et al., 1996; Ortega et al., 2003; Penalva et al., 2000), and *virilizer* (*vir*) (Hilfiker et al., 1995; Niessen et al., 2001) are also required for the female-specific splicing of late *Sxl* pre-mRNA, in addition to the Sxl protein. These genes, however, play no role in the splicing pattern of *Sxl* transcripts produced from the early promoter, *Sxl-pe* (Horabin and Schedl, 1996). The functional Sxl protein in females acts as a splicing regulator for its immediate downstream target gene *transformer* (*tra*). The *tra* transcript is produced in both males and females, but its pre-mRNA follows alternative splicing pathways. The *tra* pre-mRNA has two alternative 3' splice sites in the second intron; one is female-specific (distal 3'splice site) and the other non sex-specific (proximal 3'splice site). In females due to the intervention of the functional Sxl protein, the splicing of *tra* transcript uses the distal 3' splice site leading to the removal of the stretch containing the termination codon in the mature *tra* transcript and synthesis of full length Tra protein. In case of males, since there is no Sxl protein, the *tra* splicing is default, leading to the selection of proximal 3' splice site and inclusion of the stretch having in frame stop codon which ultimately leads to the formation of truncated non-functional protein (Belote et al., 1989; Boggs et al., 1987; Valcarcel et al., 1993). The genes *fl(2)d* (Granadino et al., 1996) and *vir* (Hilfiker et al., 1995) are also required for female-specific splicing of the *tra* pre-mRNA.

Genetic analyses have ruled out a direct role of *snf* in *tra* pre-mRNA splicing (Cline et al., 1999). Tra protein along with the protein product of the constitutive gene *tra-2* controls the sex-specific splicing of pre-mRNA from the gene *doublesex* (*dsx*). The *dsx* gene of *Drosophila* is the most downstream gene of the hierarchy of sex determination pathway. It is a binary switch gene which controls somatic sexual differentiation by producing sex-specific alternatively-spliced mRNAs which produce functional proteins, both in males and females, with common amino termini but sex-specific carboxyl termini (Baker et al., 1989; Burtis and Baker, 1989; Hoshijima et al., 1991; Slee and Bownes, 1990; Steinmann-Zwicky et al., 1990). Sex-specific Dsx proteins act on the sex differentiating genes to attribute the sex-specific characters. Another two important genes, *intersex (ix)* and *hermaphrodite (her)* are also required for the proper sex determination in *Drosophila,* but they are not the direct players in the sex determination pathway. The *ix* gene is transcribed in both the sexes, but is not sex-specifically spliced hence the Ix protein is formed in both the sexes. Ix protein is required for the proper functioning of DsxF to control female terminal differentiation whereas it is not required in males hence Ix interacts with DsxF but not with DsxM (Chase and Baker, 1995; Garrett-Engele et al., 2002; Waterbury et al., 1999). The gene *her* has a dual function. Maternal expression of *her* is required for the activation of the *Sxl* gene through the "early" promoter in females, during early developmental process, whereas its zygotic expression is necessary for the terminal sexual differentiation in both males and females (Li and Baker, 1998; Pultz and Baker, 1995).

The sex determination mechanism in the germ cells is different from that in somatic cells. The sex of the germ cell is induced, both by the signal from the surrounding cells and by its cell autonomous components (Cline and Meyer, 1996; Waterbury et al., 2000). XY germ cells are male, irrespective of the sex of the surrounding somatic cells, indicating that X:A ratio is operative cell autonomously. The operationalization of the X:A ratio is known to be contributed by a multitude of numerator elements on the X chromosome including *sis-a* (Cline, 1988; Erickson and Cline, 1991), *sis-b* (Cline, 1988; Parkhurst et al., 1990), *runt* (Duffy and Gergen, 1991; Torres and Sanchez, 1992), *sans fille* (Salz, 1992), *flex* (Bhattacharya et al., 1999) and denominator elements on autosomes-e.g., *deadpan* (Younger-Shepherd et al., 1992). At an X:A ratio of 1, in the presence of maternal transcription factor daughterless, *Sxl* is activated. Once activated, the Sxl protein directs the female-specific splicing of its own transcripts and becomes independent of the X:A ratio (Bell et al., 1991). The female-specific message of *Sxl* codes for an RNA binding protein which functions by

alternate splicing of the target genes. In males, the *Sxl* transcript undergoes constitutive splicing resulting in the addition of an extra exon containing an in-frame stop codon. The Sxl performs four vital functions. It regulates itself; plays a role in somatic sex determination i.e., the sex-specific differentiation in all somatic tissues and also germline sex determination, i.e., the decision of the germline to enter spermatogenesis or oogenesis. Finally, it plays an important role in dosage compensation- a process by which the levels of transcription from two X chromosomes in one sex are made equivalent to the amount expressed from a single X (Mukherjee and Beermann, 1965).

1.4.1.1 Genes homologous to the *Drosophila* sex determination cascade in other dipteran and non-dipteran insects

The mechanism of sex determination in *D. melanogaster* has been well understood and acts as a reference system for the study of sex determination in other insects. The comparative evaluation of sex determination mechanisms operative in different organisms, should throw light on the evolution of genes involved in sex determination in different insect species. Several genes homologous to the sex determination cascade genes of *D. melanogaster* have been identified and an interesting fact has emerged from these studies: genes at the bottom of the cascade are more conserved as compared to the upstream hierarchy genes.

X:A signal genes:

Among the genes that form the X:A signal, *scute* of *D. subobscura* (Botella et al., 1996) and *sisterless-a* of *D. pseudoobscura* and *D. virilis* (Erickson and Cline, 1998) have been characterized. These genes show significant amount of structural as well as functional conservation.

The *Sex lethal* gene:

Among the different *Drosophila* species, the *Sxl* gene has been characterized in *D. virilis* (Bopp et al., 1996) and *D. subobscura* (Penalva et al., 1996) and it has been found that, as in case of *D. melanogaster*, the *Sxl* transcript is sex-specifically spliced: the *Sxl* transcripts in males have an additional exon containing multiple stop translation codons, producing non functional truncated protein. However, in *D. virilis* males also produce the functional Sxl protein which differs from female Sxl protein by 25 amino acids (encoded by differentially

spliced exon) from the starting of the amino terminal region. This male-specific Sxl protein predominantly accumulates in the embryonic ectoderm suggesting its involvement in the development of the central nervous system (Bopp et al., 1996).

Other non-drosophild dipterans, in which *Sxl* has been characterized include *Chrysomya rufifacies* (blowfly) (Muller-Holtkamp, 1995), *Megaselia scalaris* (the phorid fly) (Sievert et al., 2000; Sievert et al., 1997), *Musca domestica* (the housefly) (Meise et al., 1998), *Ceratitis capitata* (Medfly) (Saccone et al., 1998) and *Bactrocera oleae* (olive fly) (Lagos et al., 2005) (all of which belong to the suborder Brachycera), and *Sciara ocellaris* (fungus gnat fly) (Ruiz et al., 2003), *Sciara coprophila, Rynchosciara americana* and *Trichosia pubescens* (Serna et al., 2004), which belong to the suborder Nematocera. The only lepidopteran where *Sxl* has been characterized is the silkmoth, *Bombyx mori* (Niimi et al., 2006), but its transcript is not sex-specifically spliced excluding its probable involvement in the sex determination pathway. All the published results suggest that the *Sxl* is recruited as the master regulatory gene in the sex determination cascade only in the drosophilids.

Outside drosophilids, primary function of *Sxl* seems to modulate the gene activity through inhibition of mRNA translation in both the sexes (Saccone et al., 1998). This preposition is based on the following observations: first, *Sxl* controls dosage compensation in *Drosophila* through the translational regulation of *msl-2* mRNA (Kelley et al., 1997), second, Sxl protein gets accumulated at many transcriptional active sites in the polytene chromosomes of females (Samuels et al., 1994) and third, ectopic expression of *C. capitata* (Saccone et al., 1998) and *M. domestica* (Meise et al., 1998) Sxl protein in *D. melanogaster* is lethal in both sexes, presumably by interfering with certain cellular functions since *Drosophila*, *Ceratitis*, and *Musca* Sxl proteins have conserved RNA-binding domains. In *D. melanogaster,* the activity of *Sxl* gene is required only in females since males lacking this function are viable and fertile (Cline, 1978). However, the alternative, non-mutually exclusive possibility of Sxl being a general splicing factor cannot be ruled out since both functions are exerted through its two RNA-binding domains (Sanchez, 2008). Results obtained from the study of *Sxl* in drosophilids and nondrosophilids suggest that initially it was involved in the regulation of the gene at the splicing and/or translational level in a non sex-specific manner. During the course of evolution, *Sxl* gene was recruited in drosophilids for controlling the process of sex determination and dosage compensation via specific splicing or translation inhibition, respectively.

The *transformer* gene:

Apart from *D. melanogaster*, the *tra* gene has been characterized in many other drosophilids including *D. simulans, D. mauritiana, D. sechellia*, *D. erecta* (Kulathinal et al., 2003; O'Neil and Belote, 1992), *D. hydei* and *D. virilis* (O'Neil and Belote, 1992). The *tra* gene from these organisms show considerable divergence compared to the *tra* of *D. melanogaster*. Despite the divergence, the heterologous gene can rescue *tra* mutation in *D. melanogaster* (O'Neil and Belote, 1992). The *tra* gene has been characterized in many of the non-drosophilids including the tephritids *C. capitata* (Pane et al., 2002), *B. oleae* (Lagos et al., 2007) and *Anastrepha obliqua* (Ruiz et al., 2007b), and the calliphorid *Lucilia cuprina* (Concha and Scott, 2009). The putative Tra protein of tephritids contains an extra tract of amino acids in the N-terminal domain compared to that of *Drosophila* because of which they are longer than the drosophilids Tra protein. In tephritids also, the *tra* gene expresses constitutively in both the sexes but its primary transcript, as in drosophilids, shows sex-specific alternative splicing.

The regulation of sex-specific splicing of *tra* pre-mRNA is different in drosophilids and non-drosophilids. In contrast to drosophilids where the Sxl protein regulates the splicing of *tra* pre-mRNA, the sex-specific splicing of the *tra* transcript in calliphorid (*L. cuprina*) and tephritids is autoregulated by a positive feedback loop which produces the functional Tra protein only in females. In this case, the *tra* contains a male-specific exon having multiple in-frame stop codons, due to which, a truncated non-functional Tra protein is produced in males. The male-specific exon and its surrounding introns harbour multiple putative Tra/Tra-2 binding sites to which Tra/Tra-2 protein complex binds, preventing the male-specific exons from being incorporated into mature *tra* mRNA in females, ensuring the continuous production of functional Tra protein only in females. In *Ceratitis*, the *tra* plays a key regulatory role and acts as the memory device for sex determination via its autoregulatory function (Pane et al., 2002). The dsRNA mediated knock-down of *tra* in *Ceratitis* (Pane et al., 2005) and *Bactrocera* (Lagos et al., 2007) resulted in the complete transformation of females into fertile males whereas dsRNA mediated knockdown of *tra* in *L. cuprina* lead to the development of male genitalia in XX adults (Concha and Scott, 2009).

Ectopic expression of *Ceratitis* Tra protein in *D. melanogaster* transformed the males into females, though this transformation appears to be incomplete (Pane et al., 2005). The Tra

proteins in the tephritids *Ceratitis*, *Bactrocera* and *Anastrepha* show a dual splicing role. On one hand Tra behaves as a splicing activator of *dsx* pre-mRNA, the binding of which to the female-specific exon promotes the inclusion of this exon into the mature mRNA. On the other hand, Tra acts as a splicing inhibitor of its own pre-mRNA, the binding of which to the male-specific exons prevents the inclusion of these exons into the mature mRNA. It has been proposed that the Tra-2-ISS binding sites, which have been found in the splicing regulatory region of the *tra* pre-mRNA, but not in *dsx* pre-mRNA of the tephritids, provide the distinguishing marker for the dual splicing function of the Tra/Tra-2 complex in tephritids (Ruiz et al., 2007b).

The *transformer-2* gene:

The *tra-2* has been characterized in *D. virilis* (Chandler et al., 1997), which encodes a set of protein isoforms analogous to those of *D. melanogaster*, and can rescue *tra-2* mutations in the latter species. Outside the drosophilids, *tra-2* has been characterized in *M. domestica* (Burghardt et al., 2005). The *tra-2* of *Musca* is involved in sex determination is evident by the fact that dsRNA mediated knock-down of *Musca tra-2* results in the complete transformation of genotypically female embryos into fertile adult males. As in case of *Drosophila*, *Musca tra-2* is transcribed in both the sexes and its function is required for the female-specific splicing of *Musca dsx* pre-mRNA. Genetic data suggest that *tra-2* participates in the autocatalytic activity of gene *F* (Burghardt et al., 2005), the key sex-determining gene in the housefly (Dubendorfer et al., 2002). *tra-2* has also been characterized in the domesticated silkmoth, *B. mori* where its pre-mRNA produces multiple mRNAs through alternative splicing encoding six distinct BmTra-2 proteins as in the case of *tra-2* in *Drosophila*, but they do not seem to be required for the sex-specific splicing of the *Bmdsx* pre-mRNA (Niu et al., 2005).

The *doublesex* gene:

Outside drosophilids, the homolog of *dsx* gene has been found in many other dipterans, including *M. scalaris* (Kuhn et al., 2000; Sievert et al., 1997), *M. domestica* (Hediger et al., 2004), and *Anopheles gambiae* (Scali et al., 2005), in the tephritids *B. tryoni* (Shearman and Frommer, 1998), *B. oleae* (Lagos et al., 2005), *C. capitata* (Saccone *et al.*, cited in Pane *et al.*, 2002) and *A. obliqua* (Ruiz et al., 2007a; Ruiz et al., 2005), in the lepidopteran *B. mori* (Ohbayashi et al., 2001; Suzuki et al., 2001) and in the hymenopteran

Apis mellifera (Cristino et al., 2006) and *Nasonia vitripennis* (Oliveira et al., 2009). The *dsx* gene structure and the molecular organisation of the *dsx* ORF varies among these insects, but in all the cases, as in *Drosophila, dsx* produces sex-specific RNAs that encode putative sex-specific Dsx proteins, sharing the N-terminal region but differing in their C-terminal regions. Sex-specific Dsx proteins are produced as a result of sex-specific splicing of the *dsx* pre-mRNA. Like in *Drosophila,* in all other dipteran insects the female-specific exon of *dsx* contains putative Tra/Tra-2 binding sites, suggesting the requirement of Tra protein for female-specific splicing of *dsx* pre-mRNA, whereas male-specific splicing represents the default mode. The sex-specific alternative splicing of the pre-mRNA of *Bombyx dsx* (*Bmdsx*) is very different from that of the dipteran *dsx*: *Bmdsx* does not have a weak female-specific splice site, and does not carry the putative Tra/Tra-2 binding motif related sequence, thus the processing of *Bmdsx* pre-mRNA would need splicing repressor(s) rather than splicing activator(s) (Suzuki et al., 2001).

The *fruitless* gene:

The gene *fru* of *D. simulans, D. yakuba, D. pseudoobscura, D. virilis* and *D. suzukii* (Billeter et al., 2002) and that of *A. gambiae* and *Tribolium castaneum* (Gailey et al., 2006) has been characterized. In all the cases, the *fru* shows a conserved molecular structure and the male-specific FruM protein arise by conserved mechanisms of sex-specifically activated and alternative exon splicing, which is controlled by *tra*. Moreover, the ectopic expression of *A. gambiae* Fru^{MC} in *D. melanogaster* was sufficient to rescue the muscle of Lawrence (MOL)-less phenotype–fru^{3} mutant males and induce the formation of a MOL-like muscle in females (Gailey et al., 2006). The MOL in *D. melanogaster* is under the control of *fru* gene (reviewed in Billeter *et al.*, 2002).

The *intersex* gene:

The gene *ix* has been characterized in the dipterans *D. virilis* and *M. scalaris*, and in the lepidopteran, *B. mori* (Siegal and Baker, 2005) and *Maruca vitrata* (Cavaliere et al., 2009). The Ix protein from these organisms shows a conserved organization: the amino-terminal region is rich in glutamine, proline, glycine and serine residues which is a feature of known transcription regulators. The carboxy-terminal region contains a high proportion of polar amino acids and two conserved phenylalanine residues, which appear to be specific to the Ix proteins. The Ix function in *ix* mutant *D. melanogaster* flies, is completely restored by

the *D. virilis* and the *M. scalaris ix* genes, suggesting that these Ix homologs can interact with the female-specific Dsx protein (DsxF) just as *Drosophila* Ix protein does. The *B. mori ix* gene only partially rescues female sexual development of *Drosophila ix* mutant females, suggesting that co-evolution has occurred for DsxF and Ix proteins, and that sufficient time has elapsed since the separation of the lepidopteran and dipteran lineages to allow the accumulation of divergence between the DsxF and Ix proteins of these species (Siegal and Baker, 2005).

1.4.2 Sex-determination in *C. elegans*

C. elegans occurs in two natural sexes, the hermaphrodite (XX) and the male (XO), which differ extensively in anatomy, physiology and behaviour. Hermaphrodites are somatically females but can reproduce either by self-fertilization or by mating with males. If both male and the hermaphrodite sperms are present in the spermatheca, the male sperm is used preferentially (Ward and Carrel, 1979). All somatic differences between the sexes result from the differential activity of a "global" sex determination regulatory pathway which is governed by a single switch gene (Hodgkin, 1987; Villeneuve and Meyer, 1990) (**Fig. 1.5**). As in the case of *Drosophila*, here also the X:A ratio triggers the sex determination cascade (Nigon, 1951) by regulating the expression of upstream regulator gene XO-lethal (*xol-1*) (Miller et al., 1988; Rhind et al., 1995). The X:A ratio regulates the activity of *xol-1* via the combined action of a set of "numerators" or X-signal elements on the X chromosome and "denominators," or autosomal signal elements elsewhere in the genome. This signal is very sensitive which in normal case can distinguish between ratios of 0.5 (XO) and 1.0 (XX), but can reliably discriminate a ratio of 0.67 from one of 0.75 also (Hodgkin et al., 1979; Madl and Herman, 1979; Nigon, 1949). Genetic studies have indicated the presence of at least four "X-signaling elements" (Akerib and Meyer, 1994) among which Sex-1 (signal element on X) protein acts at the level of transcription (Carmi et al., 1998), and the Fox-1 (feminizing locus on X), a putative RNA binding protein acts post-transcriptionally (Hodgkin et al., 1994; Skipper et al., 1999). The X-signal elements are opposed by denominator or autosomal signal elements, among which two genes have been cloned and found to encode likely transcriptional regulators: *sea-1* encodes a T-box protein, and *sea-2* a novel protein with zinc fingers and a metalloprotease domain (Powell et al., 2005).

The X:A ratio of 0.5, as in case of males, results in a high level of *xol-1* transcripts, whereas a hermaphrodite dose of X:A ratio results in a low level of these transcripts (Minniti

et al., 1996; Rhind et al., 1995). Low level of Xol-1 in hermaphrodites' activates the three *sdc* genes (sex determination and dosage compensation defective), out of which *sdc-2* is expressed only in hermaphrodites and this gives the SDC complex sex-specificity (Dawes et al., 1999). The role of *sdc-3* in sex determination is not very clear since the null alleles of this do not have obvious sex determination defects (DeLong et al., 1993). The SDC proteins in XX animals binds to the promoter of the male-specific gene *hermaphrodite-1* (*her-1*) and downregulates its expression to one-twentieth that of XO animals (Chu et al., 2002; Dawes et al., 1999; Schauer and Wood, 1990; Trent et al., 1991). In the absence of *her-1* product, the *tra* genes (*tra-2 and tra-3*) are transcribed at high levels, which lead to the hermaphrodite development. Her-1 inhibits the function of the Tra-2A and promotes male development in a cell non-autonomous manner (Hunter and Wood, 1992; Perry et al., 1993). Tra-2A binds Fem-3 (Feminization protein) and inhibits its ability to promote male development (Mehra et al., 1999); in turn the transcriptional regulator Tra-1 promotes female development. Tra-2A is inactivated by the presence of high levels of Her-1, thus allowing the Fem proteins to inactivate Tra-1, which ultimately leads to the male developmental pathway. The Zinc finger transcription factor, Tra-1, is the master regulator controlling all somatic sexual fate decisions at the cellular level (Hodgkin, 1987; Zarkower and Hodgkin, 1992). Tra-1 directly targets *mab-3* gene (male abnormal) and represses its transcription (Yi et al., 2000) in order to prevent male differentiation. In males, *mab-3* encodes a protein with sequence and functional similarity to *Drosophila* Dsx and vertebrate Dmrt1 (Raymond et al., 1998; Wilkins, 1995).

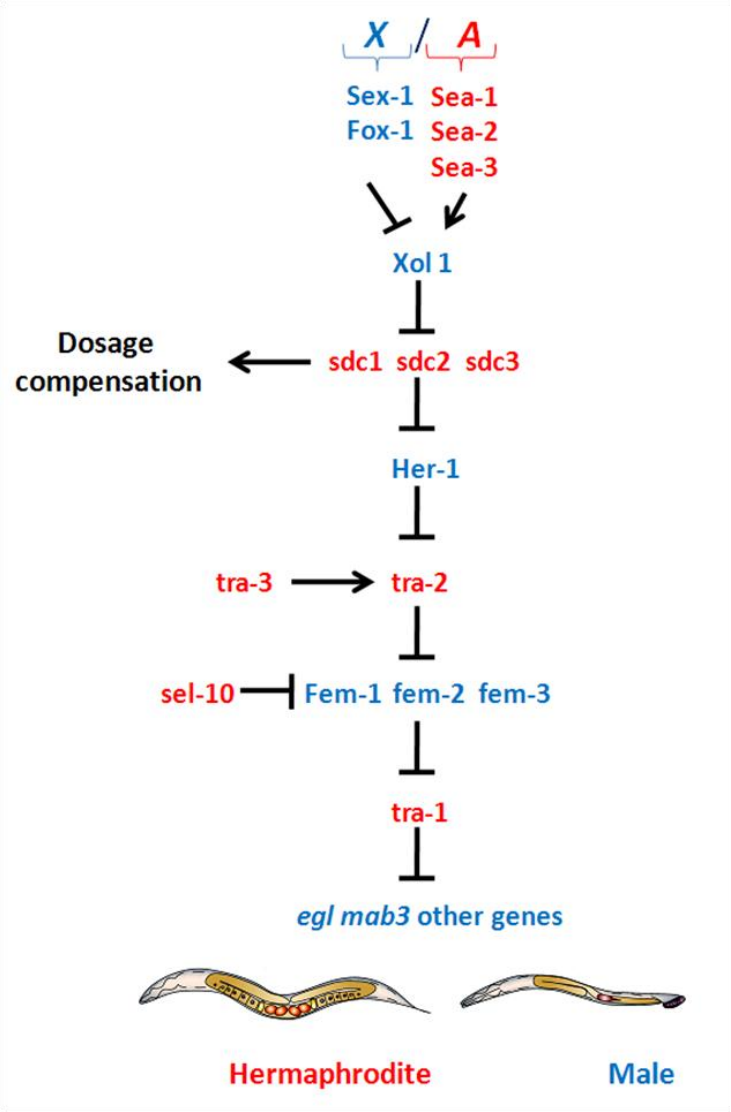


Fig. 1.5 Sex determination pathway in *C. elegans*. Masculinising factors are shown in blue and feminizing factors are in red. Arrows represent positive regulatory interactions and bars represent negative interactions. The X-signal elements, such as the Sex-1 and Fox-1 proteins, control the levels of Xol- 1 and are determining factors for sex. In hermaphrodites (XX), the *XO-lethal 1* (*xol-1*) gene is repressed permitting the Tra-1 protein to be active through a series of negative regulatory interactions. Active Sdc-1, -2 and -3 repress *her-1*. They also control sex determination and activate dosage compensation, leading to hypo-transcription of X-chromosome-linked genes. Active Tra-1 represses the expression of *egg-laying defective 1* (*egl-1*), *male abnormal 3* (*mab-3*), and other unidentified genes, allowing female somatic development to occur. (Hermaphrodites in *C. elegans* are somatically females but produce a limited amount of sperms, so they are self-fertile as well as cross-fertile). In males, *xol-1* is active, which leads to post-translational inactivation of Tra-1, thus resulting in the expression of *egl-1*, *mab-3* and other genes which, in turn, direct male differentiation.

1.4.3 Sex determination in *Apis mellifera*

In hymenopteran insects, such as wasps, sawflies, ants and bees, haploidy/diploidy mechanism of sex determination is found where both fertilized and unfertilized eggs are laid: fertilized eggs develop into diploid females and unfertilized eggs develop into haploid males (drones). In these insects, the sexual pathway followed by the zygote depends on a primary signal governed by heterozygosity at the *complementary sex determiner* (*csd*) gene which occurs in multiple allelic forms (Beye et al., 2003; Evans et al., 2004; Pienaar and Greeff, 2003; Stahlhut and Cowan, 2004; Zhou et al., 2006). Individuals heterozygous for this locus develop as females whereas hemizygous ones develop as males (**Fig. 1.6**). Diploid individuals homozygous for this locus, arising by mating of male and female that share the same allele, develop as sterile males (Cook, 1993). The *csd* gene encodes a member of the SR protein family, an arginine-serine rich (SR) protein which is structurally similar to the protein encoded by dipteran *transformer* genes. The *csd* is transcribed in both sexes. The protein coded by different *csd* alleles is specific owing to the highly variable region present in different *csd* alleles. Heterozygosity at *csd* locus leads to the production of two functional Csd protein isoforms capable of forming a complex that determines female development whereas, allelic homozygosity generates a single isoform of Csd protein that cannot generate a functional complex, thus determining male development (Beye, 2004; Beye et al., 2003) (**Fig. 1.6**). The initial observation that Csd function is required only in females and that its product is nonfunctional when derived from only one allele comes from the suggestion of three possible models. First, that different allelic Csd proteins form active heterodimers. Second, those Csd proteins derived from the same allele form homomers, with two homomer species in females and one in males. And third, the existence of different alleles is required in females for Csd to complete its function. According to a recent report, *csd* arose from ancestrally conserved progenitor gene, *feminizer* (*fem*) (Hasselmann et al., 2008). *fem* encodes an SR-type protein, harbouring an Arg/Ser-rich domain the arrangement of which is similar to the Tra of *Drosophila*. The pre-mRNA of *fem* is sex-specifically spliced to produce male and female specific splice variants: male-splice variant contains a premature stop codon resulting in the production of non-functional protein, whereas the female-splice variant encodes the functional Fem protein. Fem is required for the female sexual development as evident by the complete sex reversal (female to male) of the *fem* knockdown females. The *fem* pre-mRNA follows the male splicing pattern in *csd* knockdown females suggesting that the *fem* is downstream to *csd* and implements the switch of female sexual development

controlled by heterozygosity at *csd* locus (Hasselmann et al., 2008). Recently *dsx* homolog has been discovered and found to be sex-specifically spliced in honey bees (Cho et al., 2007; Cristino et al., 2006), proving its possible role in honey bee sex determination pathway.

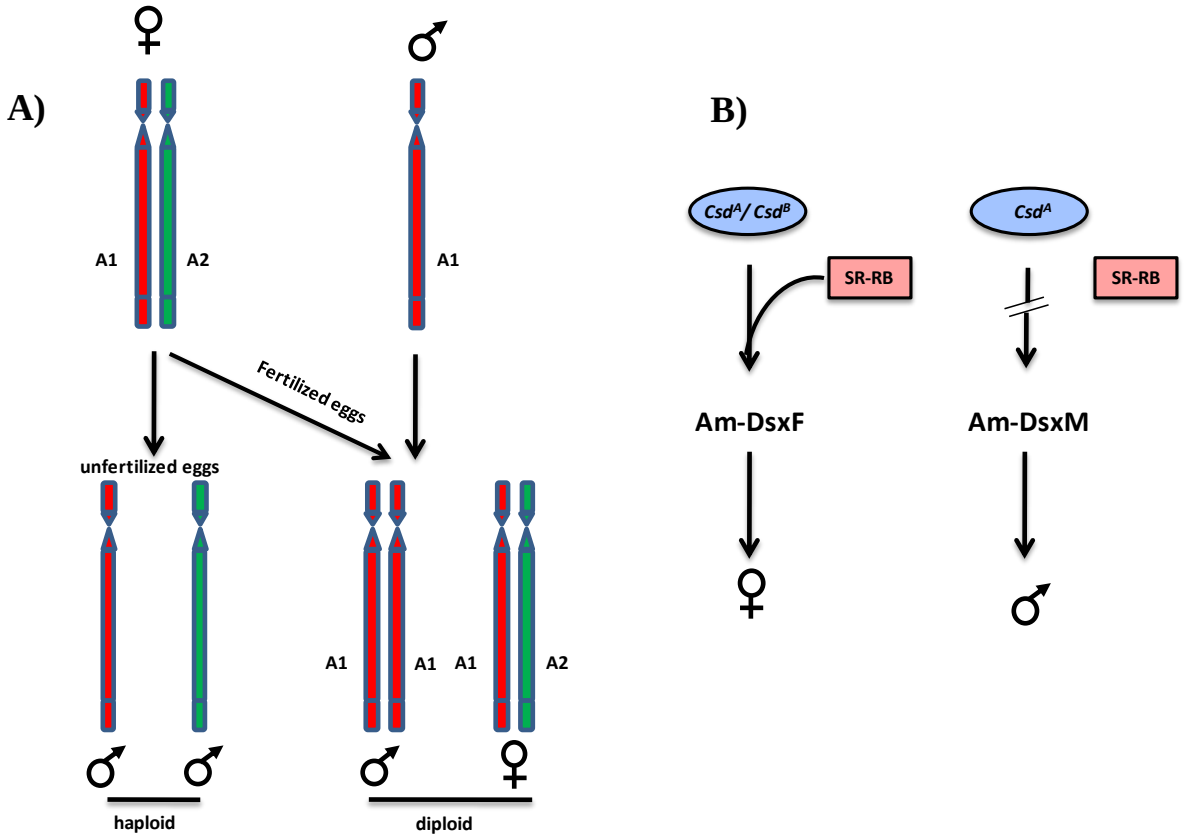


Fig. 1.6 Complementary Sex Determination (CSD) mechanism in honey bee. **A**) Fertilized eggs, homozygous at *csd* locus, develop as sterile males whereas heterozygous ones at *csd* locus develop as fertile females. Unfertilized haploid eggs, where *csd* is in hemizygous condition, develop into fertile males. **B**) Heterozygosity at *csd* locus in females produces heteromeric Csd protein which binds cooperatively to a SR-type protein containing an RNA recognition motif (SR-RB protein), that splices *Apis dsx* pre-mRNA in female mode. Female-specific *Am-dsx* transcript produces Am-DsxF protein which induces female development. Single type of Csd protein produced in males (hemi- or homozygous for *csd* locus) is non functional and default splicing of *Am-dsx* transcript leads to the production of Am-DsxM isoform, which induces male development.

1.4.4 Sex determination in *Lepidopterans*

Chromosomal sex determination is operative in lepidopterans, where females are the heterogametic and males are homogametic. Lepidopteran females are disomic (ZW) or monosomic (ZO), depending on the presence or the absence of W chromosome, whereas males are always ZZ. Lepidopterans share this property with Trichoptera (caddis flies), their closest relatives in which all the species examined to date follow the ZO/ZZ system of sex determination which might be the ancestral case (Ennis, 1976; Traut and Marec, 1996). Data from the inheritance of irradiated chromosome fragments in *B. mori* suggest that the sex is strongly controlled by the presence or absence of a specific region of the W chromosome (Fujii and Shimada, 2007); in the presence of W chromosome female develops from the embryo whereas in its absence male individual develops (Tazima, 1964). This led to the prediction long ago that the W chromosome carries an epistatic feminizing factor that determines female development (Hashimoto, 1933; Tazima, 1964). Very recently two genes (Z1 and Z20) encoding Zn finger protein motif have been identified on the W chromosome of *B. mori* that appear to act as upstream regulatory signals in the silkworm sex determination pathway (Ajimura. et al., 2006; Satish et al., 2006). A very contrasting feature is found in the Mediterranean flour moth, *Ephestia kuehniella* Zeller where a male killing factor is found on the W chromosome that might serve as a feedback control of sex rather than as a determinant in the female sex differentiation pathway (Marec et al., 2001). However, the absence of a W chromosome in primitive Lepidoptera and frequent secondary loss of the W in advanced Lepidoptera favour balance mechanisms of sex determination with Z-linked male promoting factors instead of W-linked female promoting factors (Traut and Marec, 1996).

The molecular level investigation of the W chromosome of *B. mori*, by Abe and his colleagues, led to the identification of 12W chromosome specific random amplified polymorphic DNA (RAPD) markers (Abe et al., 1998a; Abe et al., 1998b; Abe et al., 2000a; Abe et al., 1998c; Abe et al., 2005; Abe et al., 2000b). The sequences of these RAPD markers revealed that the W chromosome of *B. mori* is full of retrotransposable elements (Abe et al., 1998b). Neither a classical linkage map of the Z chromosome of *B. mori* containing 15 morphological traits dispersed over 50cM (Fujii, 1998), nor did the 18 Z-linked RAPD markers (Yasukochi, 1998) and the 16 SSR/ISSR/RAPD marker-based genetic map (Nagaraja et al., 2005) revealed any sex-determining locus. The 13 genes identified in a

contiguous 320 kb walk of the Z-chromosome are also not of significance in terms of sex determination (Koike et al., 2003).

1.4.4.1 Genes homologous to the *Drosophila* sex determination cascade in *B. mori* *sex* determination

The existence of a female-heterogametic and a cell-autonomous epistatic system of sex determination in the silkworm, *B. mori*, rather than a balance system as found in *D. melanogaster*, further adds to the amazing diversity of sex determination mechanism in insects. As compared to many well studied dipteran insects, the sex determination studies are still at a very incipient stage in lepidopteran insects, where only *B. mori* has been investigated to some extent (Traut et al., 2007). The orthologue of *Drosophila Sxl* gene is conserved in *B. mor*i but the function is not, its pre-mRNA is not sex-specifically spliced suggesting that it is probably not involved in the sex determination pathway of *B. mori* (Niimi et al., 2006; Traut et al., 2006). The orthologue of *Drosophila tra-2* gene has been identified in *B. mori*, however, the orthologue of *tra* has not been found (Niu et al., 2005). This is presumably due to its rapid sequence evolution which had previously been seen even among different *Drosophila* species (McAllister and McVean, 2000). Even though there exists the dissimilarity in the *Drosophila* and *B. mori* upstream genes involved in the sex determination, a homologue of the *dsx* gene that is the downstream member of the sex-determining cascade in *D. melanogaster* is well conserved in *B. mori*. *B. mori doublesex* (*Bmdsx*) is present on autosomes in both sexes and, like *Drosophila dsx*, is alternatively spliced to yield male and female-specific mRNAs. *Bmdsx* encodes male-specific (BmDsxM) and female-specific (BmDsxF) polypeptides, the BmDsxF protein has a regulatory function in females that acts repressively in males (Suzuki et al., 2003; Suzuki et al., 2005; Suzuki et al., 2001). The sex-specific splicing of the *Bmdsx* pre-mRNA is regulated by repression of the default female-specific processing pattern (Ohbayashi et al., 2001) and thus is very different from that in *Drosophila*, where the default state is male-specific splicing, and female-specific splicing is under the control of positive splicing cofactors (Graham et al., 2003). Thus, sex-specific splicing of *Bmdsx* is controlled not by a splicing activator like Tra, instead, by some as yet unidentified splicing repressor, indicating a difference between *Drosophila* and *Bombyx* in the mechanisms of splicing regulation of their respective *dsx* genes. In *Drosophila*, the activation of transcription of the yolk protein gene (*Yp*), by the female-type Dsx protein (DsxF), requires the interaction of DsxF with a partner protein encoded by the *intersex* (*ix*)

gene (Garrett-Engele et al., 2002). Although the homolog of *ix* gene is present in the *B. mori* WGS database, the protein encoded by the gene did not interact with BmDsxF in the yeast two hybrid system (Fujii and Shimada, 2007). Thus, it is possible that BmDsxF also needs some additional factor(s) for its proper functioning in order to bring about female-specific regulation. In *Drosophila*, sex in the nervous system is determined by the *fruitless* (*fru*) gene (Gailey et al., 2006; Kimura et al., 2005). *Bombyx* genome contains *Bmfru,* which is orthologous to *Drosophila fru*, but it is not known whether or not *Bmfru*, *Bmdsx,* and/or another factor determine sex in *Bombyx* neurons. As is true for *Drosophila fru*, *Bmfru* also exists in many isoforms produced by transcription initiation at different promoters and by alternative splicing. Moreover, expression in the adult head is quantitatively different between the sexes, suggesting a sex-dependent function of *Bmfru* in the silkworm brain (Fujii and Shimada, 2007). In *Drosophila*, control of germ line sex determination is very different from that of somatic cells. Although the mechanisms of germline sex determination in *Drosophila* are not yet fully understood, it is known that *Sxl*, *bam*, *otu*, *ovo*, and *snf* regulate sex determination of the germ line in this species (Casper and Van Doren, 2006). Although orthologues of *Sxl*, *otu*, *ovo*, and *snf* can be found in *Bombyx* WGS contig sequences, a recognizable orthologue of *bam,* critically important for determination of female germ cells in *Drosophila*, is absent in *Bombyx*, suggesting that the sexual differentiation pathway in *Bombyx* germ cells is different from that in *Drosophila* (Mita et al., 2004; Xia et al., 2004).

To understand the sex determination process in silkmoths and evolution of sex determination systems in Lepidoptera, it was necessary to study the similarities and differences of the conserved sex determining genes in different species of silkmoths. To begin with, the sex determination studies were carried out in two different wild silkmoths, *A. assama* and *A. mylitta*, with the following objectives:

a) Isolation of *doublesex* homologues from two saturniid silkmoths, *A. assama and A. mylitta*, and analysis of their sex-specific transcripts

b) Functional characterization of *A. assama dsx* gene

c) Comparison of *dsx* homologue of saturniid silkmoths, with *Bombyx dsx* and known homologues of *dsx* from other insect species

Chapter - II

Isolation of *A. assama and A. mylitta doublesex* genes and analysis of their sex-specific transcripts

2.1 Introduction

Sex determination and sexual differentiation events in *Drosophila* are governed by a complex set of well characterized cascade of genes (Baker et al., 1989; Slee and Bownes, 1990; Steinmann-Zwicky et al., 1990). *Drosophila doublesex* (*dsx*) gene, present at the bottom of the sex determination cascade, is so far the best characterized candidate followed by *doublesex*-equivalents i.e., *mab3* of *C. elegans* and *Dmrt-1* of mammals. These sex determining genes share a conserved DNA binding motif, termed as DM domain (*Doublesex* /*Mab-3 DNA-binding motif*) (Erdman and Burtis, 1993; Raymond et al., 1998; Zhu et al., 2000). The DM domain is a cystine-rich DNA binding motif first recognized in Dsx protein of *Drosophila* (Erdman and Burtis, 1993; Zhu et al., 2000). Conservation of the DM domain in metazoan proteins related to sexual differentiation suggests the existence of universal mechanism of sexual dimorphism (Raymond et al., 1998). As the name suggests, the *dsx* gene functions in both the sexes; sex-specific splicing of the *dsx* pre-mRNA generates sex-specific transcripts producing sex-specific Dsx proteins; DsxM in males and DsxF in females (Baker and Wolfner, 1988; Burtis and Baker, 1989). Sex-specific Dsx proteins ultimately direct most aspects of somatic sexual differentiation in *Drosophila* (Burtis and Baker, 1989). The DsxF and DsxM have common N-terminal DNA binding domain (Erdman and Burtis, 1993), but differ at their C-terminal dimerization domain (Erdman et al., 1996), responsible for the sex-specific regulatory properties of these proteins. Some of the sexually dimorphic features controlled by Dsx protein include the genitalia and sex combs, differences in the pigmentation of fifth and sixth abdominal tergites and certain aspects of courtship behaviour (Villella and Hall, 1996), nervous system development (Taylor and Truman, 1992), and female-specific expression of yolk protein genes (Bownes, 1994).

The pre-mRNA of *Drosophila dsx* gene contains six exons; first three exons are common to both male- and female-specific *dsx* transcripts, exon 4 is female-specific whereas exons 5 and 6 are male-specific (**Fig. 2.1**) (Burtis and Baker, 1989). The female-specific exon 4 contains six copies of identical tridecamer (13-nt) repeat sequences called *dsx* repeat elements (*dsx*RE) which act as an ESE (Exonic Splicing Enhancer). The 3' splice site preceding the female-specific exon (exon 4) contains a stretch of purine nucleotides, which makes it a weak splicing acceptor that is overlooked by the spliceosomal machinery in males.

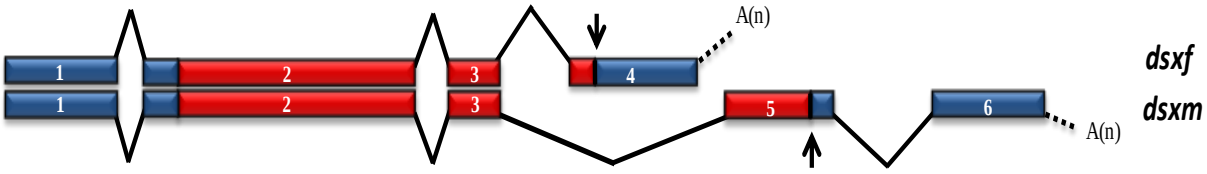


Fig. 2.1 Male and female splice forms of *Drosophila dsx*. *dsxf* and *dsxm* represent the female and male forms, respectively. Boxes are exons and lines are introns. Red coloured portion is the ORF whereas blue coloured regions are UTRs. A(n) represents poly A tail and arrows represent stop codon sites.

In females, the regulatory protein Tra promotes the cooperative binding of an SR protein, Rbp1 and an SR-like protein, Transformer 2 (Tra2) to *dsx*RE in the *dsx* pre-mRNA (**Fig. 2.2**) (Lynch and Maniatis, 1996). This binding leads to the formation of a complex containing SR proteins and interact with the U2AF small subunit through its RS domain which stimulates the use of a non-canonical weak female-specific 3'splice site, as a result the *dsx* pre-mRNA follows female mode of splicing, encoding DsxF (Hedley and Maniatis, 1991; Hoshijima et al., 1991; Ryner and Baker, 1991; Tacke and Manley, 1999; Tian and Maniatis, 1992). The efficient female-specific splicing of *dsx* is attributed to the additive contribution of each of the six 13nt elements, suggesting that each enhancer complex interacts with a single target during spliceosomal assembly (Hertel and Maniatis, 1998). Absence of Tra protein in males leads to the use of an alternative conventional splice site due to which *dsx* pre-mRNA splices in default male mode, producing DsxM protein.

The DM domain contains a non-classical Zn module (Zhu et al., 2000). DNA binding activities of the DM domain are enhanced by C-terminal dimerization domain (Cho and Wensink, 1998) and are mediated by a novel α-helical dimer containing ubiquitin-associated (UBA-like) folds (Bayrer et al., 2005). Mutations, either in the DM domain or dimerization domain of *dsx*, lead to the development of intersex phenotypes (Erdman and Burtis, 1993; Erdman et al., 1996; Narendra et al., 2002). The *dsx* gene homologues have been characterized in many other dipteran insects outside the genus *Drosophila* which include *Megaselia scalaris* (Kuhn et al., 2000), *Musca domestica* (Hediger et al., 2004) and *Anopheles gambiae* (Scali et al., 2005), in the tephritids *Bactrocera tryoni* (Shearman and Frommer, 1998), *Bactrocera oleae* (Lagos et al., 2005), *Ceratitis capitata* (Salvemini et al., personal communication) and *Anastrepha obliqua* (Ruiz et al., 2005). The pre-mRNAs of all the dipteran *dsx* studied till date, sex-specifically splice to produce sex-specific transcripts

that encode putative male- and female-specific Dsx proteins having common N-terminal but different C-terminal regions. As in the case of *Drosophila dsx*, the *dsx* pre-mRNA of other dipterans also contain Tra-Tra2 binding sites (*dsx*RE elements) in the female-specific exon, suggesting the common mode of sex-specific splicing of *dsx* in dipterans; default splicing in males whereas females recruiting Tra protein for the splicing of female-specific exon (Sanchez, 2008).

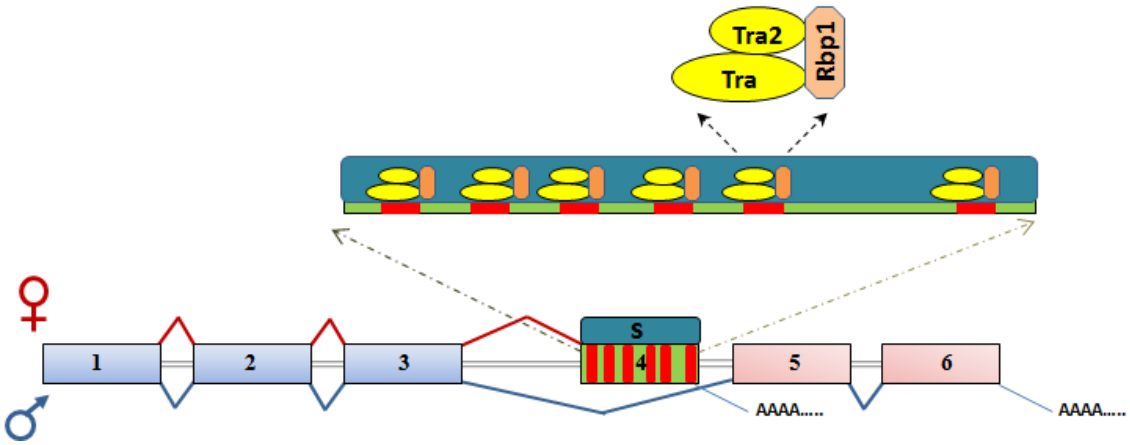


Fig. 2.2 Organization of *dsx*RE and binding of SR proteins. Boxes are exons and lines are introns. Red colour introns show the female splicing pattern whereas blue colour introns show the splicing pattern followed by males. Assembly of female-specific heterotrimeric protein complexes on ESEs regulate the splicing of *dsx* pre-mRNA in female mode. Exons 1, 2 and 3 are constitutively spliced in both sexes. The weak 3' splice site preceding exon 4 is not recognized by the splicing machinery in males, resulting in its exclusion, and splicing of exon 3 to exon 5. In females, the female-specific Tra protein promotes binding of the SR protein Rbp1, and the SR-like protein Tra2 to six copies of an ESE (indicated by red rectangles). These splicing enhancer complexes then recruit the splicing machinery to the 3' splice site preceding exon 4, leading to its inclusion in the mRNA. In females, polyadenylation (AAA…) occurs downstream of exon 4, whereas in males it occurs downstream of exon 6. 'S' designates the splicing machinery. Exons are indicated by boxes and introns are indicated by lines. Enlarged view of exon 4 is shown for clarity.

Outside dipterans, *dsx* homologues have been characterized only in *Bombyx mori* (Ohbayashi et al., 2001), and two hymenopterans *Apis mellifera* (Cristino et al., 2006) and *Nasonia vitripennis* (Oliveira et al., 2009). The pre-mRNA of *B. mori dsx* (*Bmdsx*) gene sex-specifically splices to generate male- and female-specific mRNAs which express in various tissues at larval, pupal and adult stages of the silkworm (Ohbayashi et al., 2001). *Bmdsx* gene consists of 6 exons; the female-specific *Bmdsx* transcript (*Bmdsxf*) contains all the 6 exons out of which first four exons constitute its ORF. In males the exons 3 and 4 are skipped, as a result exons 1, 2 and 5 constitute the ORF of male-specific *Bmdsx* transcript (*Bmdsxm*) (Suzuki et al., 2001) (**Fig. 2.3**). Exon 5 encodes male-specific C-terminal amino acid

sequence, while it is transcribed as a 3' UTR in females. The features of the sex-specific splicing of *Bmdsx* and *dsx* seem to be the same but the mechanism of splicing in both the cases is different. Unlike in *Drosophila*, the default form of *Bmdsx* splicing is the female form, as evident from the splicing pattern of the *Bmdsx* mini gene in the HeLa nuclear extract (Suzuki et al., 2001). *Bmdsx* neither has a weak female-specific splice site nor does its genomic DNA contain Tra/Tra-2 binding motif related sequences. Hence it is assumed that *tra* and *tra-2* genes are not required for the sex-specific splicing of *Bmdsx* pre-mRNA. Since *Bmdsx* has strong female splice site, splicing repressor(s) is needed to produce male-specific splice form of *Bmdsx* transcript (Suzuki et al., 2001). Recently it has been shown that binding of BmPSI, a *Bombyx* homolog of PSI (P-Element Somatic Inhibitor) on the 'Exonic Splicing Silencer' sequence present on exon 4 is essential for repressing female-specific splicing and skipping of exons 3 and 4 in males (Suzuki et al., 2008). PSI is a KH-domain RNA-binding protein responsible for alternative splicing of P-element transposase pre-mRNA in *Drosophila* somatic tissues (Adams et al., 1997). In male cells, the interaction of BmPSI with the cis-element (CE1) present on exon 4, splices exon 2 to exon 5 resulting in the skipping of exons 3 and 4 to produce male form of *Bmdsx*. Either because of insufficient levels of BmPSI or the presence of a *trans*-acting factor(s) that counteracts the activity of BmPSI, exons 3 and 4 are spliced in to generate female-type *Bmdsx* mRNA in female (Suzuki et al., 2008).

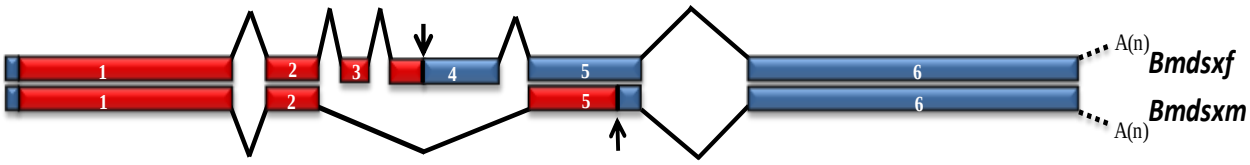


Fig. 2.3 Male and female splice forms of *Bmdsx*. Boxes are exons and lines are introns. Red coloured portion is the ORF whereas blue coloured regions are UTRs. The *Bmdsx* pre-mRNA splices to generate female-specific (*Bmdsxf*) and male-specific (*Bmdsxm*) forms. A(n) represents the poly-A tail and arrows represent the stop codon sites.

The pre-mRNA of honey bee *dsx* gene (*Am-dsx*), sex and non sex specifically splices to produce four types of transcripts (Cho et al., 2007); one male-specific (*Am-dsxm*), two female-specific (*Am-dsxf1* and *Am-dsxf2*), and one common to both male and female (*Am-dsxb*) (**Fig. 2.4**).

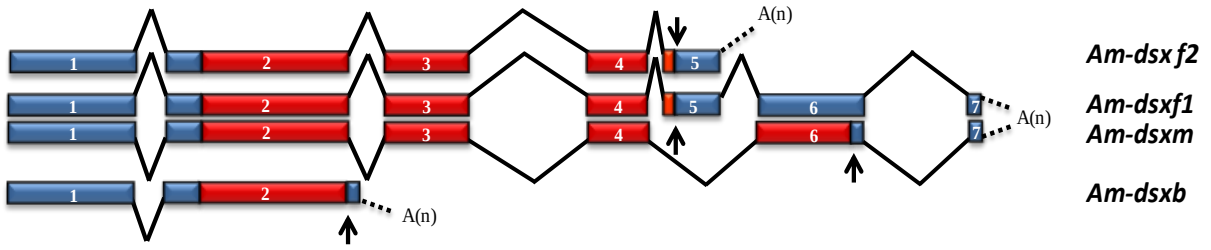


Fig. 2.4 Splice forms of *Apis dsx*. Boxes are exons and lines are introns. Red coloured portion is the ORF whereas blue coloured regions are UTRs. The *Am-dsx* pre-mRNA splices to generate 4 different splice products, two female-specific (*Am-dsxf1* and *Am-dsxf2*), one male-specific (*Am-dsxm*) and one common to both the sexes (*Am-dsxb*). A(n) represents the poly-adenylation site. Arrows represent stop codon sites.

ORFs of both the female-specific *Am-dsx* transcripts (*Am-dsxf1* and *Am-dsxf2*) are identical leading to the production of same putative protein from both the transcripts. In this way, as in other insect species where *dsx* has been studied, *A. mellifera* also produces one male- and one female-specific Dsx proteins. The splicing of *Am-dsx* is comparable to the splicing of *Drosophila dsx* and also to that of *Bombyx dsx* (*Bmdsx*) (Cho et al., 2007). Comparison of the female-specific transcript *Am-dsxf1* with the male-specific transcript (*Am-dsxm*) indicates that the splicing pattern of *Am-dsx* is similar to that of *Bmdsx*; the female-specific exon is having a strong splice site thus leading to the female default splice form and skipping of female-specific exons takes place in males because of the presence of splicing suppressor. Similarity in splicing pattern of *Am-dsx* and *Drosophila dsx* is evident from the comparison of another female-specific transcript, *Am-dsxf2* with the male-specific transcript (*Am-dsxm*); the sex-specific splice forms differ from each other by the differential poly-adenylation sites (Cho et al., 2007).

Though both fly type and moth type splicing patterns exist in the splicing of *Am-dsx* pre-mRNA, it rather resembles moth-type splicing mechanism, since as in the case of *Bmdsx*, here also no *dsx*RE or PRE is found in the female-specific exons and the 3' splice site preceding the female-specific exon is not weak (probably leading to the female splice form as the default form). Another observation of the presence of low level of female splice forms in

the male samples and absence of male splice form in the female samples in RT-PCR experiments as observed in silkworm again suggests the similarity in the sex-specific alternative splicing mechanism of *Am-dsx* and *Bmdsx* pre-mRNA (Cho et al., 2007).

Antheraea assama, considered to be the progenitor species of silkmoths, is confined to only Assam state of India (Arunkumar et al., 2008). It has low chromosome number (n=15) (**Fig. 2.5**) (Deodikar et al., 1962) as compared to other known silkmoths and follows chromosomal system of sex determination. Sexual dimorphism is observed in pupal and moth stages (**Fig. 2.6**). The female being heterogametic for the sex chromosome i.e., ZO and the male homogametic i.e., ZZ, makes it interesting for the sex determination studies. Another economically important saturniid moth is *Antheraea mylitta* chosen for the study has ZW females and ZZ males. To explore the sex determination mechanism in the two species of silkmoths that belong to the same phylogenetic lineage but possessing different chromosome composition in females, we initiated the isolation and characterization of *dsx* from these species.

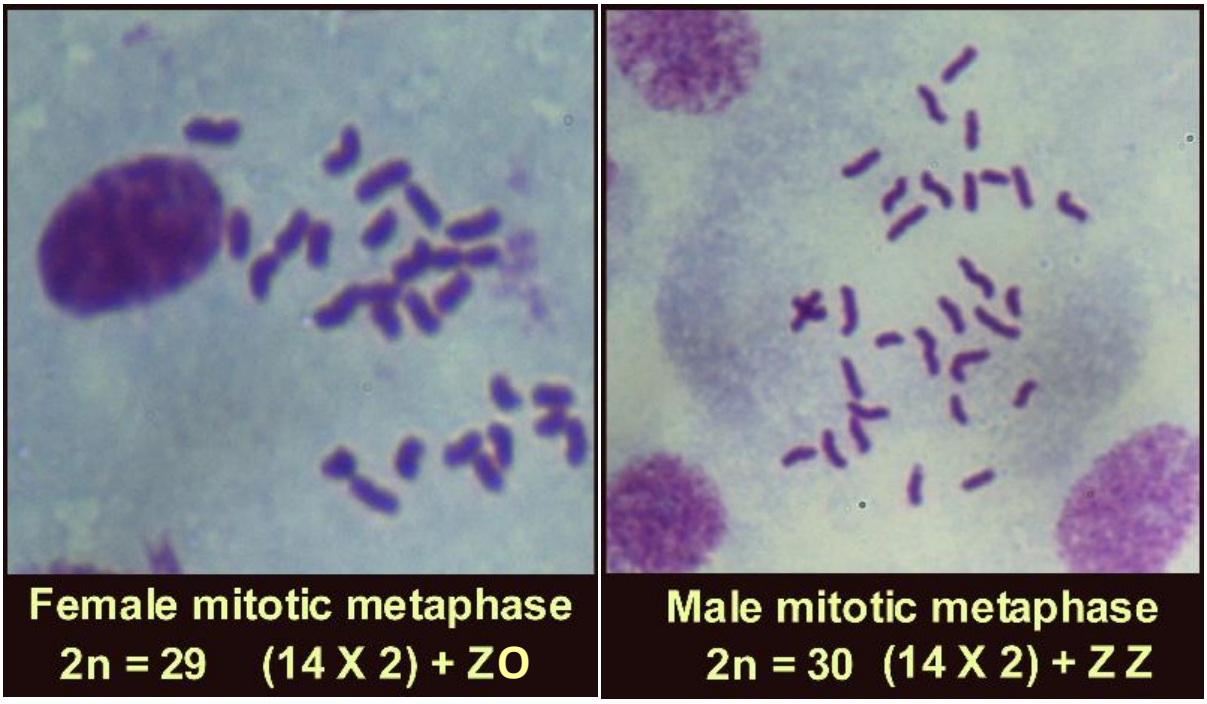


Fig. 2.5 Mitotic metaphase chromosomes of saturniid silkmoth, *A. assama*: male (2n= 30) and female (2n= 29). Figure obtained from the thesis of *Arunkumar, K.P.* Laboratory of Molecular Genetics, CDFD, India.

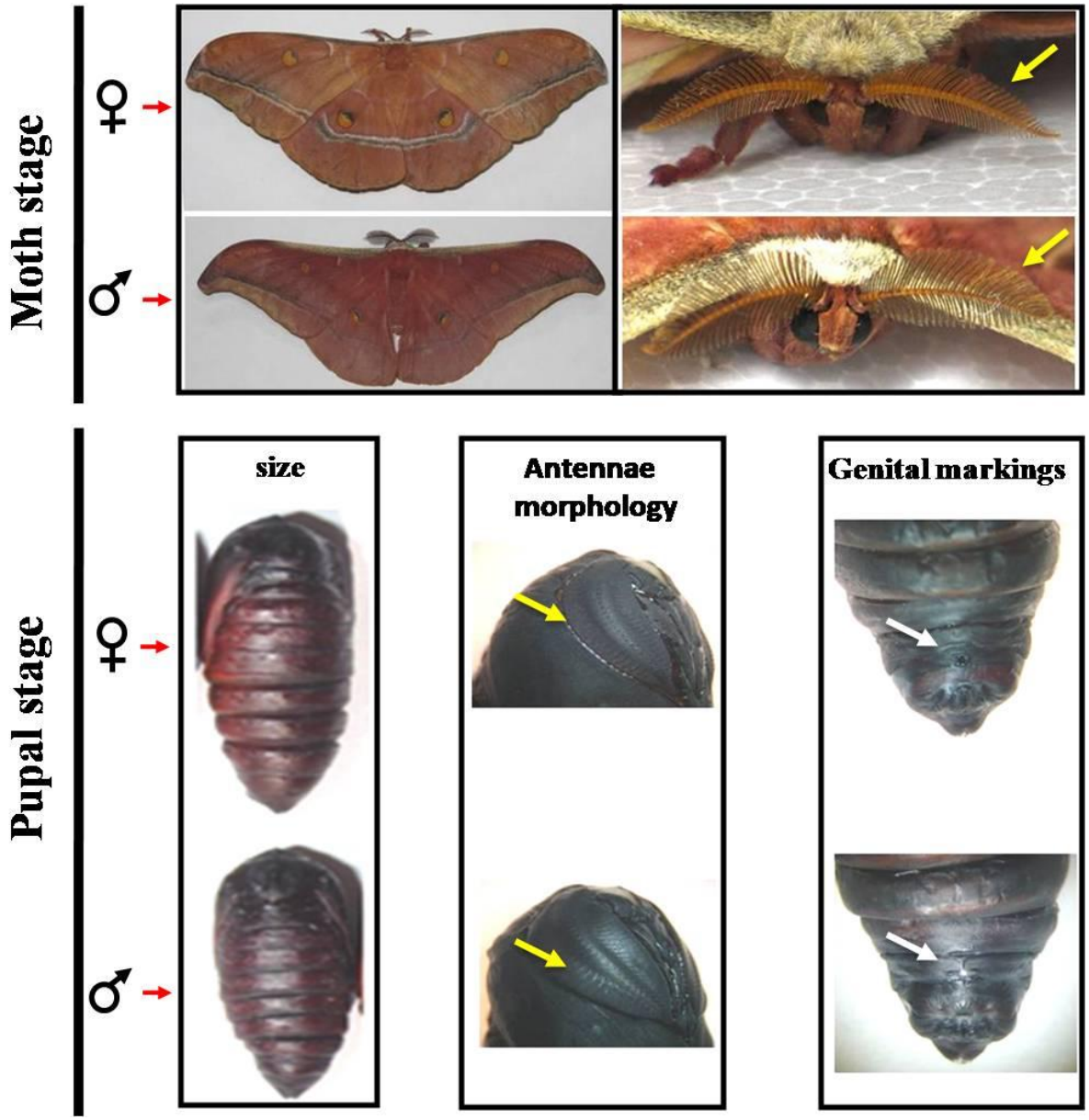


Fig. 2.6 Sexual dimorphism at moth (upper panel) and pupal (lower panel) stages. Yellow arrows show sexual dimorphism in antennae morphology; female antennae are tapered unlike males where they are broad. White arrows show sexual dimorphism in genital markings. Overall size and weight of female pupae and moths are bigger than that of males.

2.2 Materials and Methods

2.2.1 Degenerate primers

Dsx protein sequences of 7 different insect species were pulled out from the database (www.ncbi.nlm.nih.gov) and were aligned using ClustalW program (http://www.ebi.ac.uk/Tools/clustalw2/index.html). Two conserved regions, one in OD1 domain and the other in OD2 domain, were selected for designing of forward and reverse degenerate primers, respectively (**Fig. 2.7**). Nucleotide sequences, of the corresponding selected regions, were manually aligned with each other (**Fig. 2.8**).

OD1 Domain

```
Bm    PNCARCRNHRLKIELKGHKRYCKYQHCTCEKCRLTADRQRVMAKQTAIRRAQAQDEAR--
Cc    PNCARCRNHGLKITLKGHKRYCKFRYCTCEKCRLTADRQRVMALQTALRRAQAQDEQRVL
Bo    PNCARCRNHGLKITLKGHKRYCKFRYCTCEKCRLTADRQRVMALQTALRRAQAQDEQRVL
Dm    PNCARCRNHGLKITLKGHKRYCKFRYCTCEKCRLTADRQRVMALQTALRRAQAQDEQRAL
Md    PNCARCHNHGLKIKLKGHKRYCKYRFCNCEKCRLTADRQRVMALQTALRRAQQQDEARIL
Ms    PNCARCRNHSLKIALKGHKRYCKYRYCDCEKCRLTADRQKIMAAQTALRRAQAQDESRPL
Ag    PNCARCRNHGLKIGLKGHKRYCKYRTCHCEKCCLTAERQRVMALQTALRRAQTQDEQRAL
      ****** ** *** *********:: * **** ***:***:** ******** *** : :
```

OD1 Domain

```
Bm     LLEKFHYSWEMMPLVLVIMNYARSDLDEASRKIYEGKMIVDEYARKHNLNVFDGLELRNS
Cc     LLEKFRYPWETMPLMYVILKDAGADIEEASRRIEEGQHVVNEYSRQHNLNIFDGGELRST
Bo     LLEKFRYPWEMMPLMYVILKDAGADIEEASRRIEEGQHVVNEYSRQHNLNIYDGGELRST
Dm     LLEKFRYPWELMPLMYVILKDADANIEEASRRIEEGQYVVNEYSRQHNLNIYDGGELRNT
Md     LIEKFGYPWEMMPLMYVILKDAGVDIDEASKRIEEGQHVVNEYSRQHNLNIYDGCELRCA
Ms     LLEQFRFPFEMMPLMYVILKSVD-DEEEASRLISEGQYAVNEYSRQHNLNIFDGGELRSQ
Ag     LLEKLGYPWEMMPLMYVILKSADGDVQKAHQRIDEGQAVVNEYSRLHNLNMFDGVELRNT
       *:*:: :.:* ***: **:: .  : ::* : * **:  *:**:* ****::** ***
```

Fig. 2.7 ClustalW alignment of OD1 (DNA binding) and OD2 (oligomerization) domains of Dsx proteins from 7 different insect species. Bm- *Bombyx mori*, Cc- *Ceratitis capitata*, Bo- *Bactrocera oleae*, Dm- *Drosophila melanogaster*, Md- *Musca domestica*, Ms- *Megaselia scalaris*, Ag- *Anopheles gambiae*. Underlined shaded regions were selected as the conserved regions for their corresponding nucleotide alignment.

Primers were designed on the basis of consensus nucleotide sequences, obtained after aligning of nucleotide sequences corresponding to the conserved region of OD1 and OD2 domains of Dsx protein from different insects. Codon usage database (http://www.kazusa.or.jp/codon/) was used to select the preferential nucleotide (for *A. assama* and *A. mylitta*) in case of highly degenerate codons. The forward and reverse degenerate primers thus designed are: Dsx_1aF {5'-CAACTGCGCCCGGTG(CT)(AC)(AG)(AGCT) AA(CT)CA-3'} and New_DsxR {5'-CA(CT)(AT)AG(GCT)GGCATC(AG)TCTC-3'}, respectively. These primers were used to obtain the *dsx* sequences from both the wild silkmoths. To obtain the full length transcripts, 5' and 3' RACE primers were designed on the basis of specific *dsx* sequences of *A. assama* and *A. mylitta*

```
Bm  CC T  AA C  TG T  GC G    C  G C   TGC C G  C   AA T  CAC
Cc  CC G  AA T  TG C  GC T    A  G G   TGC C G  T   AA T  CAT
Bo  CC A  AA T  TG T  GC A    C  G G   TGC C G  A   AA T  CAT
Dm  CC G  AA C  TG C  GC C    C  G C   TGC C G  C   AA T  CAT
Md  CC G  AA T  TG T  GC C    C  G C   TGC C A  C   AA C  CAT
Ms  CC A  AA C  TG T  GC A    A  G G   TGC C G  G   AA T  CAT
Ag  CC G  AA C  TG C  GC C    C  G T   TGC C G  C   AA C  CAC

Consensus sequence to design forward primer

      CC(ATG) AA(CT) TG(TC) GC(ATGC) (CA)G(CGT) TGC C(AG)(ATGC) AA(TC) CA(CT)


Bm  GAG A T  G ATG CC G   C  T T    G  TG  C  T  C   GT C  AT C
Cc  GAG A C  G ATG CC G   T  T A    A  TG  T  A  T   GT G  AT A
Bo  GAG A T  G ATG CC A   T  T A    A  TG  T  A  T   GT G  AT A
Dm  GAG C T  G ATG CC A   C  T C    A  TG  T  A  T   GT G  AT A
Md  GAG A T  G ATG CC C   C  T A    A  TG  T  A  T   GT G  AT A
Ms  GAG A T  G ATG CC G   C  T A    A  TG  T  A  T   GT G  AT A
Ag  GAG A T  G ATG CC C   C  T G    A  TG  T  A  C   GT C  AT A

Consensus sequence to design reverse primer

      GAG A(TC)G ATG CC(AGC) (CT)T(TAG) (GA)TG (CT)(TA)(CT) GT(GC) AT(CA)
```

Fig. 2.8 Manual alignment of nucleotide sequences of the conserved OD1 and OD2 domains of *dsx* from 7 different insect species **(Fig. 2.7)**. Bm- *Bombyx mori*, Cc- *Ceratitis capitata*, Bo- *Bactrocera oleae*, Dm- *Drosophila melanogaster*, Md- *Musca domestica*, Ms- *Megaselia scalaris*, Ag- *Anopheles gambiae*. Consensus sequences are indicated at the bottom.

2.2.2 DNA extraction

Frozen larvae (male and female) in liquid nitrogen were used for the genomic DNA extraction, using the method described earlier (Nagaraja and Nagaraju, 1995). Briefly, the whole larvae were ground in liquid nitrogen using a pestle and mortar. Extraction buffer (100 mM Tris-HCl, pH 8.0, 50 mM NaCl, 50 mM EDTA and 1% SDS) and proteinase K (100 µg/ml) was added to the ground tissue and incubated at 37°C for 2 hours with occasional swirling. The DNA was extracted twice with phenol-chloroform-isoamyl alcohol (25:24:1) and once with chloroform. The DNA in supernatant was ethanol-precipitated, resuspended in TE (10 mM Tris-HCI, 1 mM EDTA, pH 8.0) buffer and incubated at 37°C for 1 hour after addition of RNase A (100 µg/ml). DNA was re-extracted with phenol-chloroform and ethanol-precipitated, as described earlier. The integrity of genomic DNA was verified by 0.8% agarose gel electrophoresis in TAE buffer (0.04 M Tris-Acetate, 0.001 M EDTA, pH 8.0).

2.2.3 RNA isolation and RT-PCR

Fifth instar larvae of *A. assama* were sex separated based on visualization of sex-specific glands present at their ventral surface. The sexual markings appear as a pair of milky white spots in the eighth and ninth larval abdominal segments in female, and are referred to as Ishiwata's Fore Gland and Ishiwata's Hind Gland, respectively. In males a small milky white body known as Herold's Gland appears ventrally in the centre, between eighth and ninth segments (Hiaso, 1994). Tissues were obtained after dissecting the larvae, frozen into liquid nitrogen and stored in refrigerator at -70^{o}C till further use. RNA was isolated from the stored tissues using Trizol method (Invitrogen Corporation, USA). DNAse treated total RNA was denatured at 75^{o}C for 10 minutes and immediately chilled on ice. First strand cDNA was synthesized by SuperScript III reverse transcriptase (Invitrogen, USA), according to the protocol provided by the manufacturer, using 17-mers polyT as primer. The condition for DOP-PCR to obtain *dsx* sequences from *A. assama* and *A. mylitta* was 94^{o}C for 2 minutes (initial denaturation), 32 cycles of the program 94^{o}C for 30 seconds, 50^{o}C for 30 seconds, 72^{o}C for 1 minute and final extension at 72^{o}C for 10 minutes.

The smear obtained in the primary PCR was diluted with water, in the ratio of 1:50 and one µl of this diluted product was taken as a template for the secondary PCR, using the same degenerate primers and the same PCR conditions.

2.2.4 Cloning

Multiple faint bands obtained after re-amplification using degenerate primers, were ligated by the TA Cloning® into the plasmid vector, pCR®II-TOPO® (Invitrogen, USA) exploiting the non template-dependent terminal transferase activity of *Taq* polymerase, which adds a single deoxyadenosine (A) to the 3´ ends of PCR products (Yasuda, 1996), according to the protocol provided by the manufacturer. Transformation of the ligated PCR product was done using ultra-competent cells and plated on agar plate containing ampicillin and X-gal.

2.2.5 Screening of colonies

Fifty white colonies were picked and inoculated separately in 200µl LB media containing ampicillin in sterile condition and kept at 37^{o}C for 2 hours, in shaking condition. One µl culture from each inoculated sample was taken as a template for the PCR using vector-specific, M13F and M13R primers. The PCR condition used was 94^{o}C for 2 minutes (initial denaturation), 32 cycles of the program 94^{o}C for 30 seconds, 60^{o}C for 30 seconds, 72^{o}C for 1 minute and final extension at 72^{o}C for 10 minutes. The amplicons larger than 500bp were sequenced with ABI Prism 3100 (Applied Biosystems, USA) sequencer. The vector sequences were removed from the quality sequences, using online NCBI tool VecScreen (http://www.ncbi.nlm.nih.gov/VecScreen/VecScreen.html). The sequences were analyzed by blasting them against the protein database, using BlastX program. Specific primers were designed for the sequence showing >50% homology with the Dsx protein, using Primer-Blast program (online NCBI primer designing tool) (http://www.ncbi.nlm.nih.gov/tools/ primer-blast/index.cgi? LINK_LOC=BlastHomeAd).

2.2.6 Sequencing

Sequencing of all the clones was performed by the cycle sequencing/chain termination method, on a Thermal Cycler model 9700 (Perkin-Elmer) and sequenced on an ABI Prism 3100 sequencer. The thermal conditions for cycle sequencing were, 96^{o}C for 1 minute, 25 cycles of 96^{o}C for 15 seconds and 60^{o}C for 4 minutes.

2.2.7 Obtaining full-length cDNA by 'Rapid Amplification of cDNA Ends' (RACE)

5' and 3' RACE reactions were performed using the RLM RACE kit (Ambion, USA) according to manufacturer's instructions, using one adapter-specific and the other gene-specific primer. The outer and the inner gene-specific primers for 5'RACE of *A. assama dsx* (*Aadsx*) are Aad_5'RACE1 and Aad_5'RACE2 whereas for *A. mylitta dsx* (*Amdsx*) are ML_5'RACE2 and Aad_5'RACE 2, respectively (**Table 2.1**). The outer and inner gene-specific primers used for *Aadsx* and *Amdsx* 3'RACE are Aad_3'RACE1 and Aad_3'RACE2, respectively (**Table 2.1**). Both primary and secondary RACE-PCR reactions were performed in an Eppendorf master cycler with the PCR conditions of 94^{o}C for 2 minutes (initial denaturation), 35 cycles of the program 94^{o}C for 30 seconds, 68^{o}C for 30 seconds, 72^{o}C for 2 minutes and final extension at 72^{o}C for 10 minutes. Amplicons of different sizes were gel eluted, sequenced and confirmed to be genuine 5' and 3' *dsx* sequences.

Table 2.1: Gene-specific primers used for RACE-PCR reactions

Serial No.	Primer	Sequence	Annealing temperature (^{o}C)
1	Aad 5'RACE 1	5'-ATCGATACTTGCAGTACCGTTTGTGGCC-3'	67
2	Aad 5'RACE 2	5'-AGTACCGTTTGTGGCCTTTCAGCTCGAC-3'	68
3	Aad 3'RACE 1	5'-TTAGTTGAAAATTGCCACAAACTGCTGG-3'	63
4	Aad 3'RACE 2	5'-TTGCCACAAACTGCTGGAGAAGTTCCAC-3'	67
5	ML5'RACE2	5'-TGCTTTCACTATAGGCGGCTCCGGTC-3'	68

2.2.8 Southern hybridization

A. assama genomic DNA isolated from male and female was digested separately with two different restriction enzymes, *Nde*I and *Spe*I (NEB, USA). Male and female genomic DNA (20µg) was incubated with *Nde*I or *Spe*I enzymes at a final concentration of 5U/µg, along with 1x NEBuffer 2, at 37^{o}C for 16 hours. Each individually digested genomic DNA was then loaded onto a 0.8% agarose gel in TAE buffer (0.04 M Tris-Acetate, 0.001 M EDTA, final pH 8.0) and DNA was separated by overnight electrophoresis at 40V. After capillary gel transfer under alkaline conditions (0.4 M NaOH, 1 M NaCl) to the Hybond N^{+} nylon membrane (Amersham, USA), pre-hybridization was done for 2 hours at 65°C in 10 ml

pre-hybridization solution containing 5X SSC, 5X Denhardt's solution (0.1% w/v bovine serum albumin, 0.1% w/v Ficoll, 0.1% w/v polyvinylpyrrolidone) and 0.5% SDS. 50ng of purified probe DNA was radioactively labeled with α-^{32}P-dATP using Strip EZ kit (Ambion, USA) and directly added to the pre-hybridization solution at a concentration of ~1 × 106 cpm/ml. Overnight hybridization was allowed at 65°C. After hybridization, the membrane was washed three times in 2× Standard Saline Citrate (SSC), 0.1% Sodium Dodecyl Sulfate (SDS), at 60°C for 10 minutes and then twice in 0.1× SSC, 0.1% SDS, at 60°C for 5 minutes. The radioactive membrane was then exposed to PhosphorImager (Amersham Biosciences, USA) screens. After exposure, the PhosphorImager screen was scanned and results were analyzed with Image-Quant 5.0 software (Amersham Biosciences, USA).

2.2.9 Tissue-specific expression profile

Semi quantitative RT-PCR was done to see the spatial gene expression in different tissues including fat-body, midgut, silk-gland, epidermis, gonads and head. Primers used were- "3'RACE_Aad" (forward primer) and "Aad3'm3" (reverse primer) (**Table 2.2**). Primers for *Aa*Actin were used as endogenous control in RT-PCR analysis. PCR was carried out in an Eppendorf PCR master cycler using PCR conditions of 94^{o}C for 2 minutes (initial denaturation), 40 cycles of the program 94^{o}C for 30 seconds, 60^{o}C for 30 seconds, 72^{o}C for 2 minutes and final extension at 72^{o}C for 10 minutes. All the female and male-specific amplicons were sequenced and blasted against protein dataset to confirm them to be *dsx* transcripts.

2.2.10 Alignment of different female and male-specific transcripts

Different sex-specific *Aadsx* and *Amdsx* transcripts obtained were aligned, using online software of ClustalW (http://www.ebi.ac.uk/Tools/clustalw2/index.html). Exon boundaries were marked based on the alignment results.

2.2.11 Expression profile of *Aadsx*

RT-PCR was done using combinations of common forward primer located in first exon and eight different reverse primers located in different exons, using cDNA from fat-body as template. The gene-specific primers used in different RT-PCR and PCR experiments are listed in **Table 2.2** and the control primers used are mentioned in **Table 2.3**.

Table 2.2: Gene-specific primers used in the study

Serial No.	Primer	Sequence	Annealing temperature (^oC)
1	Df1	5’-ACGGCAGACAGGCAGCG-3’	60
2	Dr1	5’CAGGAGTAGTGGAACTTCTC-3’	57
3	3’RACE_Aad (**F1**)	5’- AAGTCGCATCGTGACCGCG-3’	60.98
4	3’m1_new (**R1**)	5’-ACCCCTGACAATGTGATTACGG-3’	60.3
5	Aad 3’m3	5’-ATTTATGTCCCACACGCTTC	56
6	3’RACE_Aad	5’-GCCGTCGGTTCCGCCTTTACAGGCC-3’	72
7	Aad 3’m4 (**R2**)	5’-ATCGCATCGCTACAAGTGCG-3’	59.4
8	AadFem_7Rnew (**R3**)	5’-GTGACATTGTGGTGAGTAGTAAAGTC-3’	61.64
9	Aad_Fem_5R (**R4**)	5’-TCATTCGCAAAATAACTTCATCG-3’	55.3
10	Aad_Fem_6R (**R5**)	5’-ACACTGACACTAAACACATG-3’	53.2
11	Aad_Fem_3R (**R6**)	5’-TCCAGCATTTTCTATTAAAGTCCG-3’	57.6
12	Aad_Fem_4R (**R7**)	5’- TCGAGTTCCTCAGCTCGAGACCG-3’	66
13	Dr2 (**R8**)	5’-CCCAGGAGTAGTGGAACTTC-3’	59.34
14	Vitellogenin 1F	5’-AAAGCGGAAACCACCAGCAC-3’	59.4
15	Vitellogenin 1R	5’-CTATTGGTAATTTGGGACATC-3’	54
16	Hex 4F	5’-GGCAAGAGGGCAAGATGAGG-3’	64
17	Hex 4R	5’-GTAGCGATCCAGAGGGAAGC-3’	61
18	AmDsx_R2	5’-ATGAACGACATTGTAACAGACG-3’	56.52

Table 2.3: Control primers and universal primers used in the study

Serial No.	Primer	Sequence	Annealing temperature (^oC)
1	*A.assama* β-actin F	5’-CACTGAGGCTCCCCTGAAC-3’	60
2	*A.assama* β-actin R	5’- GGAGTGCGTATCCCTCGTA-3’	60
3	GFP F	5’-CCAAACGACTATGACGCAAATT-3’	60
4	GFP R	5’-TTGTTAAATTGGCCCACCAC-3’	60
5	M13- Forward primer	5’-GTAAAACGACGGCCAG-3’	60
6	M13- Reverse primer	5'-CAGGAAACAGCTATGAC-3'	60

2.2.12 Nuclear extracts preparation and synthesis of labelled RNA

Fatbody from female and male larvae were isolated and the nuclear extracts were separately prepared from both the samples, using the following protocol: Approximately 200mg of tissues were suspended in 500µl of nuclear lysis buffer in an eppendorf tube. The tissue was homogenised in an electric homogeniser for 2-3 seconds followed by incubation on ice for 1 hour with intermittent vortexing, after every 10 minutes. Tube was spun at maximum speed (14000 rpm) in a refrigerated centrifuge at 4^{o}C for 30 minutes. The supernatant was collected and aliquoted in 0.5ml eppendorf tubes and stored at -70^{o}C. From one of the aliquots, the concentration of nuclear extract was estimated by the Bradford reagent. The RNA oligo (UUAAUAAUAUAAGUGGUGUA) was commercially synthesized (MWG, Germany). 25µl end labelling reaction was set up on ice by mixing 1µl of RNA (12.5 pmol/µl) with 2.5µl of 10x T4 PNK buffer, .5µl of [γ-^{32}P] ATP (0.5 pmol/µl) and 1µl of T4 PNK enzyme (12 u/µl) and the volume was adjusted to 25µl by adding water. The reaction mix was incubated at 37^{o}C for 30 minutes. The reaction was stopped by incubating the mix at 70^{o}C for 5 minutes. Labelled oligos were purified by ethanol precipitation.

2.2.13 RNA mobility shift assay

Binding reactions were performed in a 30µl reaction. 10µg of nuclear extract was mixed with 1µl of labelled probe, 12 mM HEPES (pH 7.9), 4 mM Tris-Cl (pH 7.9) 50 mM KCl, 50 mM NaCl, 1 mM EDTA, 1 mM DTT, 12% glycerol and 2µg of poly-dI/dC and water to make the final reaction volume to 30µl. The reaction mix was incubated on ice for 20 minutes. For competition experiments, the indicated amount of cold competitor was incubated in the binding reaction mix, 5 minutes before the addition of the probe (labelled oligo). The sample was loaded on to a 4% native poly-acrylamide gel and run at 200V in cold room for 2 hours. The gel was pre-run at 120V for 1 hour.

2.3 Results

2.3.1 Molecular cloning of *A. assama and A. mylitta dsx* cDNAs

Degenerate oligonucleotide primed PCR (DOP-PCR) generated specific *dsx* sequence of 490bp, from both *A. assama* and *A. mylitta*. To obtain the full-length transcript including coding and UTR sequences of the *doublesex* of these wild silkmoths, RACE-PCR was performed on total RNA samples from fat body of 5^{th} instar male and female larvae. In case of *A. assama*, the 5'RACE-PCR yielded additional sequence of 225bp in the upstream, from the gene-specific primer binding site. The 3' RACE experiment yielded 6 amplicons of 1092bp, 1022bp, 959bp, 944bp, 744bp and 729bp in female and one amplicon of 690bp in male. These were further confirmed to be genuine stretches of *Aadsx* transcript by aligning the overlapping sequences and blasting (blastX) them. All the seven (six female- and one male-specific) sequences obtained revealed overlapping and non overlapping regions. Full length sequences of sex-specific *Aadsx* transcripts were confirmed after the assembly of the RACE sequences with the sequence obtained through DOP-PCR. When RT-PCR was done with the specific primers designed for the first and the last exons on the basis of sequences obtained through RACE using male and female fat-body cDNAs as template, six amplicons in female and one in male were obtained. The same result was obtained from the RT-PCR using cDNA as template from different *assama* tissues (**Fig. 2.9**). The forward and reverse primers used were "3'RACE_Aad" and "Aad3'm3", respectively. The alignment of all the six female specific transcripts with that of the male transcript is shown in **Fig. 2.10**. We observed very low level of expression of male specific transcript in female tissues **Fig. 2.10**.

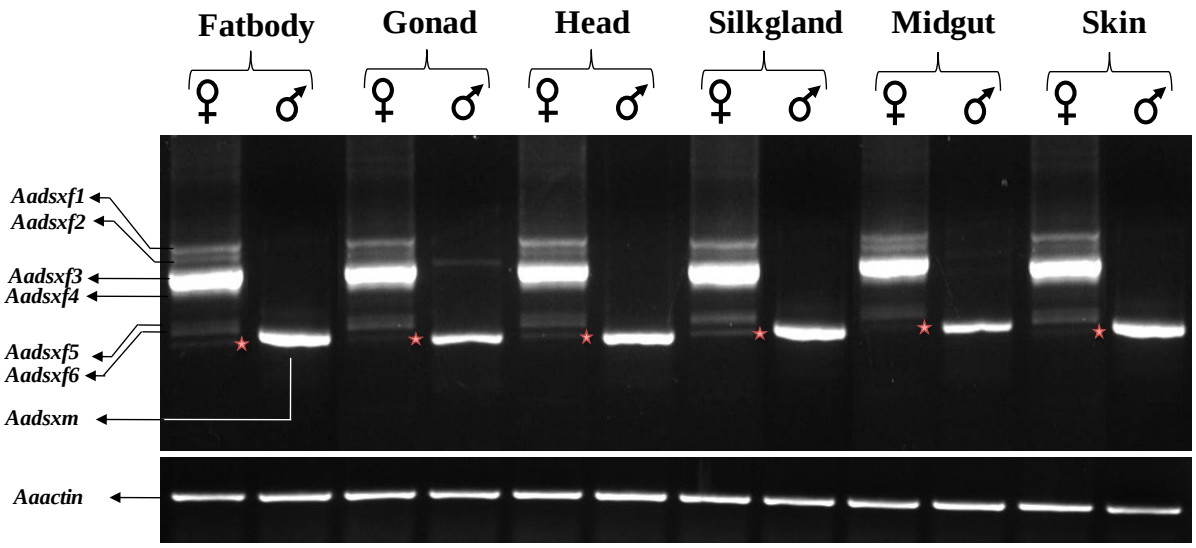


Fig. 2.9 Semi-quantitative RT-PCR using cDNA from different *A. assama* tissues as template and forward (3'RACE_Aad) and reverse (Aad3'm3) primers designed in the common region of the first and last exons of *Aadsx* (see text). Multiple bands were amplified in all the female organs whereas only single band was amplified from male organs (upper panel) which were subsequently confirmed to be *Aadsx* transcripts (see **Fig. 2.10**). The splice variants are mentioned on the left of the diagram. *A. assama* β-actin gene was amplified as a control (lower panel). Faint male specific band (indicated by star) amplified in all the female tissues.

```
Aadsxm    AAAAGTCGCATCGTGACCGCGCGCCGTGTTGTTTGTATAACGTTCGCGCGCGTCGACTTC
Aadsxf1   AAAAGTCGCATCGTGACCGCGCGCCGTGTTGTTTGTATAACGTTCGCGCGCGTCGACTTC
Aadsxf3   AAAAGTCGCATCGTGACCGCGCGCCGTGTTGTTTGTATAACGTTCGCGCGCGTCGACTTC
Aadsxf5   AAAAGTCGCATCGTGACCGCGCGCCGTGTTGTTTGTATAACGTTCGCGCGCGTCGACTTC
Aadsxf2   AAAAGTCGCATCGTGACCGCGCGCCGTGTTGTTTGTATAACGTTCGCGCGCGTCGACTTC
Aadsxf4   AAAAGTCGCATCGTGACCGCGCGCCGTGTTGTTTGTATAACGTTCGCGCGCGTCGACTTC
Aadsxf6   AAAAGTCGCATCGTGACCGCGCGCCGTGTTGTTTGTATAACGTTCGCGCGCGTCGACTTC
          ************************************************************

Aadsxm    GTCGCGATCACGGCCGAGAACGAGGGCCCAAACATGGTATCTGTGGGCGCGTGGAGGCGC
Aadsxf1   GTCGCGATCACGGCCGAGAACGAGGGCCCAAACATGGTATCTGTGGGCGCGTGGAGGCGC
Aadsxf3   GTCGCGATCACGGCCGAGAACGAGGGCCCAAACATGGTATCTGTGGGCGCGTGGAGGCGC
Aadsxf5   GTCGCGATCACGGCCGAGAACGAGGGCCCAAACATGGTATCTGTGGGCGCGTGGAGGCGC
Aadsxf2   GTCGCGATCACGGCCGAGAACGAGGGCCCAAACATGGTATCTGTGGGCGCGTGGAGGCGC
Aadsxf4   GTCGCGATCACGGCCGAGAACGAGGGCCCAAACATGGTATCTGTGGGCGCGTGGAGGCGC
Aadsxf6   GTCGCGATCACGGCCGAGAACGAGGGCCCAAACATGGTATCTGTGGGCGCGTGGAGGCGC
          ************************************************************
                                                  M  V  S  V  G  A  W  R  R

Aadsxm    CGCGCGCCCGACGACTGTGAGGAACGCTCCGACCCGGGCGCCTCCAGCTCAGCGGTGCCG
Aadsxf1   CGCGCGCCCGACGACTGTGAGGAACGCTCCGACCCGGGCGCCTCCAGCTCAGCGGTGCCG
Aadsxf3   CGCGCGCCCGACGACTGTGAGGAACGCTCCGACCCGGGCGCCTCCAGCTCAGCGGTGCCG
Aadsxf5   CGCGCGCCCGACGACTGTGAGGAACGCTCCGACCCGGGCGCCTCCAGCTCAGCGGTGCCG
Aadsxf2   CGCGCGCCCGACGACTGTGAGGAACGCTCCGACCCGGGCGCCTCCAGCTCAGCGGTGCCG
Aadsxf4   CGCGCGCCCGACGACTGTGAGGAACGCTCCGACCCGGGCGCCTCCAGCTCAGCGGTGCCG
Aadsxf6   CGCGCGCCCGACGACTGTGAGGAACGCTCCGACCCGGGCGCCTCCAGCTCAGCGGTGCCG
          ************************************************************
            R  A  P  D  D  C  E  E  R  S  D  P  G  A  S  S  S  A  V  P

Aadsxm    CGCGCGCCCCCGAACTGCGCCCGGTGCCGCAACCACAGATTGAAGGTCGAGCTGAAAGGC
Aadsxf1   CGCGCGCCCCCGAACTGCGCCCGGTGCCGCAACCACAGATTGAAGGTCGAGCTGAAAGGC
Aadsxf3   CGCGCGCCCCCGAACTGCGCCCGGTGCCGCAACCACAGATTGAAGGTCGAGCTGAAAGGC
Aadsxf5   CGCGCGCCCCCGAACTGCGCCCGGTGCCGCAACCACAGATTGAAGGTCGAGCTGAAAGGC
Aadsxf2   CGCGCGCCCCCGAACTGCGCCCGGTGCCGCAACCACAGATTGAAGGTCGAGCTGAAAGGC
Aadsxf4   CGCGCGCCCCCGAACTGCGCCCGGTGCCGCAACCACAGATTGAAGGTCGAGCTGAAAGGC
Aadsxf6   CGCGCGCCCCCGAACTGCGCCCGGTGCCGCAACCACAGATTGAAGGTCGAGCTGAAAGGC
          ************************************************************
            R  A  P  P  N  C  A  R  C  R  N  H  R  L  K  V  E  L  K  G
```

```
Aadsxm    CACAAACGGTACTGCAAGTATCGATATTGCAATTGCGAGAAATGCCGACTGACGGCAGAC
Aadsxf1   CACAAACGGTACTGCAAGTATCGATATTGCAATTGCGAGAAATGCCGACTGACGGCAGAC
Aadsxf3   CACAAACGGTACTGCAAGTATCGATATTGCAATTGCGAGAAATGCCGACTGACGGCAGAC
Aadsxf5   CACAAACGGTACTGCAAGTATCGATATTGCAATTGCGAGAAATGCCGACTGACGGCAGAC
Aadsxf2   CACAAACGGTACTGCAAGTATCGATATTGCAATTGCGAGAAATGCCGACTGACGGCAGAC
Aadsxf4   CACAAACGGTACTGCAAGTATCGATATTGCAATTGCGAGAAATGCCGACTGACGGCAGAC
Aadsxf6   CACAAACGGTACTGCAAGTATCGATATTGCAATTGCGAGAAATGCCGACTGACGGCAGAC
          ************************************************************
          H  K  R  Y  C  K  Y  R  Y  C  N  C  E  K  C  R  L  T  A  D

Aadsxm    AGGCAGCGTGTCATGGCACTGCAGACTGCTCTACGACGTGCTCAAGCGCAGGATGAGGCG
Aadsxf1   AGGCAGCGTGTCATGGCACTGCAGACTGCTCTACGACGTGCTCAAGCGCAGGATGAGGCG
Aadsxf3   AGGCAGCGTGTCATGGCACTGCAGACTGCTCTACGACGTGCTCAAGCGCAGGATGAGGCG
Aadsxf5   AGGCAGCGTGTCATGGCACTGCAGACTGCTCTACGACGTGCTCAAGCGCAGGATGAGGCG
Aadsxf2   AGGCAGCGTGTCATGGCACTGCAGACTGCTCTACGACGTGCTCAAGCGCAGGATGAGGCG
Aadsxf4   AGGCAGCGTGTCATGGCACTGCAGACTGCTCTACGACGTGCTCAAGCGCAGGATGAGGCG
Aadsxf6   AGGCAGCGTGTCATGGCACTGCAGACTGCTCTACGACGTGCTCAAGCGCAGGATGAGGCG
          ************************************************************
          R  Q  R  V  M  A  L  Q  T  A  L  R  R  A  Q  A  Q  D  E  A

Aadsxm    CGAGCGAGAGCCAAAGAGACAGGTGGTCATCTTCCAGGAGTAGAATTAGACAGACCGGAG
Aadsxf1   CGAGCGAGAGCCAAAGAGACAGGTGGTCATCTTCCAGGAGTAGAATTAGACAGACCGGAG
Aadsxf3   CGAGCGAGAGCCAAAGAGACAGGTGGTCATCTTCCAGGAGTAGAATTAGACAGACCGGAG
Aadsxf5   CGAGCGAGAGCCAAAGAGACAGGTGGTCATCTTCCAGGAGTAGAATTAGACAGACCGGAG
Aadsxf2   CGAGCGAGAGCCAAAGAGACAGGTGGTCATCTTCCAGGAGTAGAATTAGACAGACCGGAG
Aadsxf4   CGAGCGAGAGCCAAAGAGACAGGTGGTCATCTTCCAGGAGTAGAATTAGACAGACCGGAG
Aadsxf6   CGAGCGAGAGCCAAAGAGACAGGTGGTCATCTTCCAGGAGTAGAATTAGACAGACCGGAG
          ************************************************************
          R  A  R  A  K  E  T  G  G  H  L  P  G  V  E  L  D  R  P  E

Aadsxm    CCGCCTATAGTGAAAGCACCGAGGAGCCCAGTGGTGCCGCCGCCGCCGCCGCGGTCTCTC
Aadsxf1   CCGCCTATAGTGAAAGCACCGAGGAGCCCAGTGGTGCCGCCGCCGCCGCCGCGGTCTCTC
Aadsxf3   CCGCCTATAGTGAAAGCACCGAGGAGCCCAGTGGTGCCGCCGCCGCCGCCGCGGTCTCTC
Aadsxf5   CCGCCTATAGTGAAAGCACCGAGGAGCCCAGTGGTGCCGCCGCCGCCGCCGCGGTCTCTC
Aadsxf2   CCGCCTATAGTGAAAGCACCGAGGAGCCCAGTGGTGCCGCCGCCGCCGCCGCGGTCTCTC
Aadsxf4   CCGCCTATAGTGAAAGCACCGAGGAGCCCAGTGGTGCCGCCGCCGCCGCCGCGGTCTCTC
Aadsxf6   CCGCCTATAGTGAAAGCACCGAGGAGCCCAGTGGTGCCGCCGCCGCCGCCGCGGTCTCTC
          ************************************************************
          P  P  I  V  K  A  P  R  S  P  V  V  P  P  P  P  P  R  S  L

Aadsxm    GGTTCTGCGAGTTGCGACTCAGTTCCGGGTTCACCAGGAGTTTCTCCGTATGCTCCGCCA
Aadsxf1   GGTTCTGCGAGTTGCGACTCAGTTCCGGGTTCACCAGGAGTTTCTCCGTATGCTCCGCCA
Aadsxf3   GGTTCTGCGAGTTGCGACTCAGTTCCGGGTTCACCAGGAGTTTCTCCGTATGCTCCGCCA
Aadsxf5   GGTTCTGCGAGTTGCGACTCAGTTCCGGGTTCACCAGGAGTTTCTCCGTATGCTCCGCCA
Aadsxf2   GGTTCTGCGAGTTGCGACTCAGTTCCGGGTTCACCAGGAGTTTCTCCGTATGCTCCGCCA
Aadsxf4   GGTTCTGCGAGTTGCGACTCAGTTCCGGGTTCACCAGGAGTTTCTCCGTATGCTCCGCCA
Aadsxf6   GGTTCTGCGAGTTGCGACTCAGTTCCGGGTTCACCAGGAGTTTCTCCGTATGCTCCGCCA
          ************************************************************
          G  S  A  S  C  D  S  V  P  G  S  P  G  V  S  P  Y  A  P  P

Aadsxm    CCGCCGTCGGTTCCGCCTTTGCAGGCAATGCCACCACTAATGCCGCCACAACAGCCTGCT
Aadsxf1   CCGCCGTCGGTTCCGCCTTTGCAGGCAATGCCACCACTAATGCCGCCACAACAGCCTGCT
Aadsxf3   CCGCCGTCGGTTCCGCCTTTGCAGGCAATGCCACCACTAATGCCGCCACAACAGCCTGCT
Aadsxf5   CCGCCGTCGGTTCCGCCTTTGCAGGCAATGCCACCACTAATGCCGCCACAACAGCCTGCT
Aadsxf2   CCGCCGTCGGTTCCGCCTTTGCAGGCAATGCCACCACTAATGCCGCCACAACAGCCTGCT
Aadsxf4   CCGCCGTCGGTTCCGCCTTTGCAGGCAATGCCACCACTAATGCCGCCACAACAGCCTGCT
Aadsxf6   CCGCCGTCGGTTCCGCCTTTGCAGGCAATGCCACCACTAATGCCGCCACAACAGCCTGCT
          ************************************************************
          P  P  S  V  P  P  L  Q  A  M  P  P  L  M  P  P  Q  Q  P  A

Aadsxm    GTATCATTAGAGACATTAGTTGAAAATTGCCACAAACTGCTGGAGAAGTTCCACTACTCC
Aadsxf1   GTATCATTAGAGACATTAGTTGAAAATTGCCACAAACTGCTGGAGAAGTTCCACTACTCC
Aadsxf3   GTATCATTAGAGACATTAGTTGAAAATTGCCACAAACTGCTGGAGAAGTTCCACTACTCC
Aadsxf5   GTATCATTAGAGACATTAGTTGAAAATTGCCACAAACTGCTGGAGAAGTTCCACTACTCC
Aadsxf2   GTATCATTAGAGACATTAGTTGAAAATTGCCACAAACTGCTGGAGAAGTTCCACTACTCC
Aadsxf4   GTATCATTAGAGACATTAGTTGAAAATTGCCACAAACTGCTGGAGAAGTTCCACTACTCC
Aadsxf6   GTATCATTAGAGACATTAGTTGAAAATTGCCACAAACTGCTGGAGAAGTTCCACTACTCC
          ************************************************************
          V  S  L  E  T  L  V  E  N  C  H  K  L  L  E  K  F  H  Y  S
```

```
Aadsxm    TGGGAGATGATGCCCCTGGTGTTGGTCATCCTAAATTATGCAGGCAGCGACCTTGAAGAG
Aadsxf1   TGGGAGATGATGCCCCTGGTGTTGGTCATCCTAAATTATGCAGGCAGCGACCTTGAAGAG
Aadsxf3   TGGGAGATGATGCCCCTGGTGTTGGTCATCCTAAATTATGCAGGCAGCGACCTTGAAGAG
Aadsxf5   TGGGAGATGATGCCCCTGGTGTTGGTCATCCTAAATTATGCAGGCAGCGACCTTGAAGAG
Aadsxf2   TGGGAGATGATGCCCCTGGTGTTGGTCATCCTAAATTATGCAGGCAGCGACCTTGAAGAG
Aadsxf4   TGGGAGATGATGCCCCTGGTGTTGGTCATCCTAAATTATGCAGGCAGCGACCTTGAAGAG
Aadsxf6   TGGGAGATGATGCCCCTGGTGTTGGTCATCCTAAATTATGCAGGCAGCGACCTTGAAGAG
          ************************************************************
           W  E  M  M  P  L  V  L  V  I  L  N  Y  A  G  S  D  L  E  E

Aadsxm    GCCTCGCGGAAAATTGACGAA-----------------------------------------
Aadsxf1   GCCTCGCGGAAAATTGACGAAGGAAAGATGATAATCAACGAATACGCGAGAAAACATAAT
Aadsxf3   GCCTCGCGGAAAATTGACGAAGGAAAGATGATAATCAACGAATACGCGAGAAAACATAAT
Aadsxf5   GCCTCGCGGAAAATTGACGAAGGAAAGATGATAATCAACGAATACGCGAGAAAACATAAT
Aadsxf2   GCCTCGCGGAAAATTGACGAAGGAAAGATGATAATCAACGAATACGCGAGAAAACATAAT
Aadsxf4   GCCTCGCGGAAAATTGACGAAGGAAAGATGATAATCAACGAATACGCGAGAAAACATAAT
Aadsxf6   GCCTCGCGGAAAATTGACGAAGGAAAGATGATAATCAACGAATACGCGAGAAAACATAAT
          *********************
           A  S  R  K  I  D  E  G  K  M  I  I  N  E  Y  A  R  K  H  N

Aadsxm    ------------------------------------------------------------
Aadsxf1   TTGAATGTCTTCGACGGTCTCGAGCTGAGGAACTCGACACGCCAGTACGGACTTTAATAG
Aadsxf3   TTGAATGTCTTCGACGGTCTCGAGCTGAGGAACTCGACACGCCAGTACGGACTTTAATAG
Aadsxf5   TTGAATGTCTTCGACGGTCTCGAGCTGAGGAACTCGACACGCCAGTACGGACTTTAATAG
Aadsxf2   TTGAATGTCTTCGACGGTCTCGAGCTGAGGAACTCGACACGCCAG---------------
Aadsxf4   TTGAATGTCTTCGACGGTCTCGAGCTGAGGAACTCGACACGCCAG---------------
Aadsxf6   TTGAATGTCTTCGACGGTCTCGAGCTGAGGAACTCGACACGCCAG---------------

           L  N  V  F  D  G  L  E  L  R  N  S  T  R  Q  Y  G  L  X

Aadsxm    ------------------------------------------------------------
Aadsxf1   ACCGGACGAAGGTGGCGAAATTCGAAATGTAATTTATCAAAAAAAAAGGTGACTCGAATT
Aadsxf3   ------------------------------------------------------------
Aadsxf5   ------------------------------------------------------------
Aadsxf2   ------------------------------------------------------------
Aadsxf4   ------------------------------------------------------------
Aadsxf6   ------------------------------------------------------------


Aadsxm    ------------------------------------------------------------
Aadsxf1   TCGAAAGCGGCCATGTGTTTAGTGTCAGTGTGCGGTATCGATATTATTTTATTTATTTTC
Aadsxf3   ------------------------------------------------------------
Aadsxf5   ------------------------------------------------------------
Aadsxf2   ------------------------------------------------------------
Aadsxf4   ------------------------------------------------------------
Aadsxf6   ------------------------------------------------------------

Aadsxm    ------------------------------------------------------------
Aadsxf1   TGTTTTTTGTAGG--AAAATGCTGGAAATTAATAATATAAGTGGTGTACTGTCGTCATCG
Aadsxf3   ---------------AAAATGCTGGAAATTAATAATATAAGTGGTGTACTGTCGTCATCG
Aadsxf5   ---------------AAAATGCTGGAAATTAATAATATAAGTGGTGTACTGTCGTCATCG
Aadsxf2   ---------------AAAATGCTGGAAATTAATAATATAAGTGGTGTACTGTCGTCATCG
Aadsxf4   ---------------AAAATGCTGGAAATTAATAATATAAGTGGTGTACTGTCGTCATCG
Aadsxf6   ---------------AAAATGCTGGAAATTAATAATATAAGTGGTGTACTGTCGTCATCG

                          K  M  L  E  I  N  N  I  S  G  V  L  S  S  S

Aadsxm    ------------------------------------------------------------
Aadsxf1   ATGAAGTTATTTTGCGAATGATACTTTGTTTTACGAGTGCCGTGGTTTTTGTGGACACAT
Aadsxf3   ATGAAGTTATTTTGCGAATGATACTTTGTTTTACGAGTGCCGTGGTTTTTGTGGACACAT
Aadsxf5   ATGGAGTTATTTTGCGAATGATACTTTGTTTTACGAGTGCCGTGGTTTTTGTGGACACAT
Aadsxf2   ATGAAGTTATTTTGCGAATGATACTTTGTTTTACGAGTGCCGTGGTTTTTGTGGACACAT
Aadsxf4   ATGAAGTTATTTTGCGAATGATACTTTGTTTTACGAGTGCCGTGGTTTTTGTGGACACAT
Aadsxf6   ATGAAGTTATTTTGCGAATGATACTTTGTTTTACGAGTGCCGTGGTTTTTGTGGACACAT

           M  K  L  F  C  E  X

Aadsxm    ------------------------------------------------------------
Aadsxf1   GCTGTGCGATGCTCTGTGTTGCGAATTTCAACGGACAATTGTGTTGTCGCTTCACTGGAC
Aadsxf3   GCTGTGCGATGCTCTGTGTTGCGAATTTCAACGGACAATTGTGTTGTCGCTTCACTGGAC
Aadsxf5   GCTGTGCGATGCTATGTGTTGCGAATTTCAACGGACAATTGTGTTGTCGCTTCACTGGAC
Aadsxf2   GCTGTGCGATGCTCTGTGTTGCGAATTTCAACGGACAATTGTGTTGTCGCTTCACTGGAC
Aadsxf4   GCTGTGCGATGCTCTGTGTTGCGAATTTCAACGGACAATTGTGTTGTCGCTTCACTGGAC
Aadsxf6   GCTGTGCGATGCTCTGTGTTGCGAATTTCAACGGACAATTGTGTTGTCGCTTCACTGGAC
```

```
Aadsxm    ------------------------------------------------------------
Aadsxf1   ATTTA---------------------------------------------------------
Aadsxf3   ATTTA---------------------------------------------------------
Aadsxf5   ATTTA---------------------------------------------------------
Aadsxf2   ATTTAGACCATTTCGCGACCTTCAAAGTCAATCAAAAGCGTCAAAGCAGAGACTTTACTA
Aadsxf4   ATTTA---------------------------------------------------------
Aadsxf6   ATTTA---------------------------------------------------------

Aadsxm    ------------------------GCCCACTGGGTGGTACATCAGTGGCGACTGTACGAG
Aadsxf1   ------------------------GCCCACTGGGTGGTACATCAGTGGCGACTGTACGAG
Aadsxf3   ------------------------GCCCACTGGGTGGTACATCAGTGGCGACTGTACGAG
Aadsxf5   ------------------------------------------------------------
Aadsxf2   CTCACCACAATGTCACTTATGAA-GCCCACTGGGTGGTACATCAGTGGCGACTGTACGAG
Aadsxf4   ------------------------GCCCACTGGGTGGTACATCAGTGGCGACTGTACGAG
Aadsxf6   ------------------------------------------------------------

                                  A  H  W  V  V  H  Q  W  R  L  Y  E

Aadsxm    CATTCGCTGTGCTCATTGCTCAAGCTGCAGGCGCAGAAGGGTACGTACGCCATGTGCTGC
Aadsxf1   CATTCGCTGTGCTCATTGCTCAAGCTGCAGGCGCAGAAGGGTACGTACGCCATGTGCTGC
Aadsxf3   CATTCGCTGTGCTCATTGCTCAAGCTGCAGGCGCAGAAGGGTACGTACGCCATGTGCTGC
Aadsxf5   ------------------------------------------------------------
Aadsxf2   CATTCGCTGTGCTCATTGCTCAAGCTGCAGGCGCAGAAGGGTACGTACGCCATGTGCTGC
Aadsxf4   CATTCGCTGTGCTCATTGCTCAAGCTGCAGGCGCAGAAGGGTACGTACGCCATGTGCTGC
Aadsxf6   ------------------------------------------------------------

          H  S  L  C  S  L  L  K  L  Q  A  Q  K  G  T  Y  A  M  C  C

Aadsxm    TCGCCTCGGTATGTGCTCGCTCCGGAGTACGCGCCCCACTTGCTGCCGATACCTCTGACC
Aadsxf1   TCGCCTCGGTATGTGCTCGCTCCGGAGTACGCGCCCCACTTGCTGCCGATACCTCTGACC
Aadsxf3   TCGCCTCGGTATGTGCTCGCTCCGGAGTACGCGCCCCACTTGCTGCCGATACCTCTGACC
Aadsxf5   ------------------------------------------------------------
Aadsxf2   TCGCCTCGGTATGTGCTCGCTCCGGAGTACGCGCCCCACTTGCTGCCGATACCTCTGACC
Aadsxf4   TCGCCTCGGTATGTGCTCGCTCCGGAGTACGCGCCCCACTTGCTGCCGATACCTCTGACC
Aadsxf6   ------------------------------------------------------------

          S  P  R  Y  V  L  A  P  E  Y  A  P  H  L  L  P  I  P  L  T

Aadsxm    ACGCAGCGCGCATCGCCGCCGCCGGCGCACTTGTAGCGATGCGATCACGCGCCGACTCAG
Aadsxf1   ACGCAGCGCGCATCGCCGCCGCCGGCGCACTTGTAGCGATGCGATCACGCGCCGACTCAG
Aadsxf3   ACGCAGCGCGCATCGCCGCCGCCGGCGCACTTGTAGCGATGCGATCACGCGCCGACTCAG
Aadsxf5   ------------------------------------------------------------G
Aadsxf2   ACGCAGCGCGCATCGCCGCCGCCGGCGCACTTGTAGCGATGCGATCACGCGCCGACTCAG
Aadsxf4   ACGCAGCGCGCATCGCCGCCGCCGGCGCACTTGTAGCGATGCGATCACGCGCCGACTCAG
Aadsxf6   ------------------------------------------------------------G
                                                                        *

          T  Q  R  A  S  P  P  P  A  H  L  X

Aadsxm    ACGGCGCAACAACCCACCGAGCCACCTTCGACATCGTCGGGGCGGAGCACCGCCACTTAA
Aadsxf1   ACGGCGCAACAACCCACCGAGCCACCTTCGACATCGTCGGGGCGGAGCACCGCCACTTAA
Aadsxf3   ACGGCGCAACAACCCACCGAGCCACCTTCGACATCGTCGGGGCGGAGCACCGCCACTTAA
Aadsxf5   ACGGCGCAACAACCCACCGAGCCACCTTCGACATCGTCGGGGCGGAGCACCGCCACTTAA
Aadsxf2   ACGGCGCAACAACCCACCGAGCCACCTTCGACATCGTCGGGGCGGAGCACCGCCACTTAA
Aadsxf4   ACGGCGCAACAACCCACCGAGCCACCTTCGACATCGTCGGGGCGGAGCACCGCCACTTAA
Aadsxf6   ACGGCGCAACAACCCACCGAGCCACCTTCGACATCGTCGGGGCGGAGCACCGCCACTTAA
          ************************************************************

Aadsxm    GACATCACCGGTGACAGAAGCGTGTGGGACATAAATACATGTGTTAAAAATATGAACTGG
Aadsxf1   GACATCACCGGTGACAGAAGCGTGTGGGACATAAATACATGTGTTAAAAATATGAACTGG
Aadsxf3   GACATCACCGGTGACAGAAGCGTGTGGGACATAAATACATGTGTTAAAAATATGAACTGG
Aadsxf5   GACATCACCGGTGACAGAAGCGTGTGGGACATAAATACATGTGTTAAAAATATGAACTGG
Aadsxf2   GACATCACCGGTGACAGAAGCGTGTGGGACATAAATACATGTGTTAAAAATATGAACTGG
Aadsxf4   GACATCACCGGTGACAGAAGCGTGTGGGACATAAATACATGTGTTAAAAATATGAACTGG
Aadsxf6   GACATCACCGGTGACAGAAGCGTGTGGGACATAAATACATGTGTTAAAAATATGAACTGG
          ************************************************************

Aadsxm    ATATAGTATTGTTAACCCTGTAAAAATATTACGTGTATGTGAAACGGTAGACTATAGTTT
Aadsxf1   ATATAGTATTGTTAACCCTGTAAAAATATTACGTGTATGTGAAACGGTAGACTATAGTTT
Aadsxf3   ATATAGTATTGTTAACCCTGTAAAAATATTACGTGTATGTGAAACGGTAGACTATAGTTT
Aadsxf5   ATATAGTATTGTTAACCCTGTAAAAATATTACGTGTATGTGAAACGGTAGACTATAGTTT
Aadsxf2   ATATAGTATTGTTAACCCTGTAAAAATATTACGTGTATGTGAAACGGTAGACTATAGTTT
Aadsxf4   ATATAGTATTGTTAACCCTGTAAAAATATTACGTGTATGTGAAACGGTAGACTATAGTTT
Aadsxf6   ATATAGTATTGTTAACCCTGTAAAAATATTACGTGTATGTGAAACGGTAGACTATAGTTT
          ************************************************************
```

```
Aadsxm    ATTTTGGTTTCTTACTATATCGTTCAATGATACCCAGTTTTGGTTATTGTGCGCCCATGT
Aadsxf1   ATTTTGGTTTCTTACTATATCGTTCAATGATACCCAGTTTTGGTTATTGTGCGCCCATGT
Aadsxf3   ATTTTGGTTTCTTACTATATCGTTCAATGATACCCAGTTTTGGTTATTGTGCGCCCATGT
Aadsxf5   ATTTTGGTTTCTTACTATATCGTTCAATGATACCCAGTTTTGGTTATTGTGCGCCCATGT
Aadsxf2   ATTTTGGTTTCTTACTATATCGTTCAATGATACCCAGTTTTGGTTATTGTGCGCCCATGT
Aadsxf4   ATTTTGGTTTCTTACTATATCGTTCAATGATACCCAGTTTTGGTTATTGTGCGCCCATGT
Aadsxf6   ATTTTGGTTTCTTACTATATCGTTCAATGATACCCAGTTTTGGTTATTGTGCGCCCATGT
          ************************************************************

Aadsxm    TATTTTAATATAGACTATGAATCGCGCCCAAACTAATATAATAAATCCGAATGTGTGTTT
Aadsxf1   TATTTTAATATAGACTATGAATCGCGCCCAAACTAATATAATAAATCCGAATGTGTGTTT
Aadsxf3   TATTTTAATATAGACTATGAATCGCGCCCAAACTAATATAATAAATCCGAATGTGTGTTT
Aadsxf5   TATTTTAATATAGACTATGAATCGCGCCCAAACTAATATAATAAATCCGAATGTGTGTTT
Aadsxf2   TATTTTAATATAGACTATGAATCGCGCCCAAACTAATATAATAAATCCGAATGTGTGTTT
Aadsxf4   TATTTTAATATAGACTATGAATCGCGCCCAAACTAATATAATAAATCCGAATGTGTGTTT
Aadsxf6   TATTTTAATATAGACTATGAATCGCGCCCAAACTAATATAATAAATCCGAATGTGTGTTT
          ************************************************************

Aadsxm    GTTTGTTACGCTTTCGCGTCTTAACTACTCAACCAACCACCGTAATCACATTGTCAGGGG
Aadsxf1   GTTTGTTACGCTTTCGCGTCTTAACTACTCAACCAACCACCGTAATCACATTGTCAGGGG
Aadsxf3   GTTTGTTACGCTTTCGCGTCTTAACTACTCAACCAACCACCGTAATCACATTGTCAGGGG
Aadsxf5   GTTTGTTACGCTTTCGCGTCTTAACTACTCAACCAACCACCGTAATCACATTGTCAGGGG
Aadsxf2   GTTTGTTACGCTTTCGCGTCTTAACTACTCAACCAACCACCGTAATCACATTGTCAGGGG
Aadsxf4   GTTTGTTACGCTTTCGCGTCTTAACTACTCAACCAACCACCGTAATCACATTGTCAGGGG
Aadsxf6   GTTTGTTACGCTTTCGCGTCTTAACTACTCAACCAACCACCGTAATCACATTGTCAGGGG
          ************************************************************

Aadsxm    TAAAAAAAAAAAAAAAAAAAAAAAAA
Aadsxf1   TAAAAAAAAAAAAAAAAAAAAAAAAA
Aadsxf3   TAAAAAAAAAAAAAAAAAAAAAAAAA
Aadsxf5   TAAAAAAAAAAAAAAAAAAAAAAAAA
Aadsxf2   TAAAAAAAAAAAAAAAAAAAAAAAAA
Aadsxf4   TAAAAAAAAAAAAAAAAAAAAAAAAA
Aadsxf6   TAAAAAAAAAAAAAAAAAAAAAAAAA
          **************************
```

Fig. 2.10 Nucleotide and deduced amino acid sequences of *Aadsx* transcripts. The deduced amino acid sequence is shown below the nucleotide sequence. Suffix ‘f’ or ‘m’ to *Aadsx* indicates female or male, respectively. Female-specific amino acids are underlined whereas male-specific amino acids are in italics. The amino acids in red background are specific to one type of female-specific protein (AaDsxF1) whereas the amino acids in grey background are specific to another female-specific protein (AaDsxF2). Start codon (ATG) and stop codon are highlighted in blue and red colours, respectively. X represents stop codon site.

In case of *A. mylitta*, single sequence of 229bp (upstream from the primer binding site) was obtained as a result of sequencing of 5'RACE-PCR product, in both the sexes whereas the sequencing of the 3'RACE-PCR products yielded two sequences (downstream to primer binding site) of 845bp and 860bp in female and one sequence of 590bp in male. Two female and one male-specific *Amdsx* transcripts were assembled from the sequences obtained through DOP-PCR and RACE experiments. The RT-PCR experiments using end to end common region primers (AaDfor_1 and Amdsx_R2) confirmed the RACE results (**Fig. 2.11**). The alignment of all the three *Amdsx* transcripts is shown in **Fig. 2.12**.

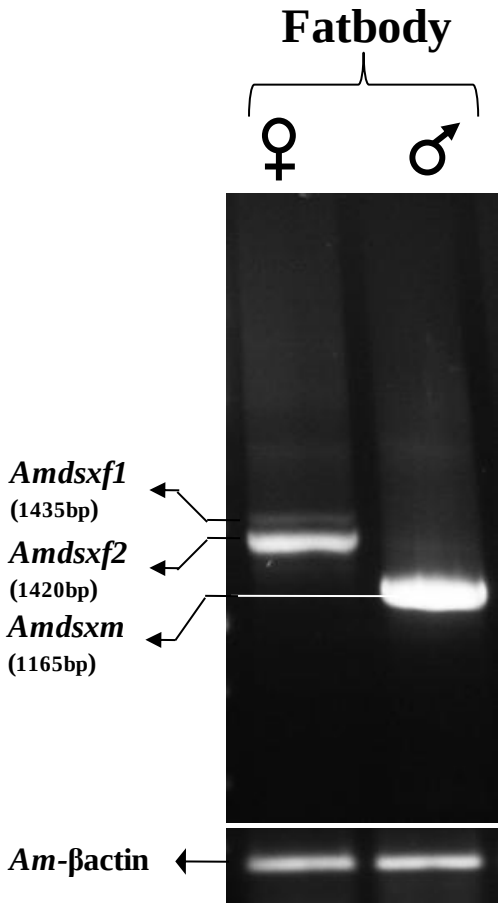


Fig. 2.11 Semi-quantitative RT-PCR using cDNA from *A. mylitta* fatbody as template and forward (AaDfor_1) and reverse primers (Amdsx_R2), designed in the common region of the first and last exons of *Amdsx* (see text). Two *Amdsx* bands are amplified in female whereas only single band is amplified in male. *Amdsx* splice variants and their respective sizes are indicated on the sides of the diagram.

```
Amdsxm      AAAGTCAGTCGCATCGTGACCGCGCGCCGTGTTGTTTGTATAACGTTCGCGCGCGCCGAC 60
Amdsxf2     AAAGTCAGTCGCATCGTGACCGCGCGCCGTGTTGTTTGTATAACGTTCGCGCGCGCCGAC 60
Amdsxf1     AAAGTCAGTCGCATCGTGACCGCGCGCCGTGTTGTTTGTATAACGTTCGCGCGCGCCGAC 60
            ************************************************************

Amdsxm      TTCGTCGCGATCACGACCGAGAACGAGGGCCCAAACATGGTATCCGTGGGCGCGTGGAGG 120
Amdsxf2     TTCGTCGCGATCACGACCGAGAACGAGGGCCCAAACATGGTATCCGTGGGCGCGTGGAGG 120
Amdsxf1     TTCGTCGCGATCACGACCGAGAACGAGGGCCCAAACATGGTATCCGTGGGCGCGTGGAGG 120
            ************************************************************
                                                M  V  S  V  G  A  W  R

Amdsxm      CGCCGCGCGCCCGACGACTGTGAGGAACGCTCCGATCCGGGCGCCTCCAGCTCGGCGGTG 180
Amdsxf2     CGCCGCGCGCCCGACGACTGTGAGGAACGCTCCGATCCGGGCGCCTCCAGCTCGGCGGTG 180
Amdsxf1     CGCCGCGCGCCCGACGACTGTGAGGAACGCTCCGATCCGGGCGCCTCCAGCTCGGCGGTG 180
            ************************************************************
             R  R  A  P  D  D  C  E  E  R  S  D  P  G  A  S  S  S  A  V

Amdsxm      CCGCGTGCGCCCCCGAACTGTGCCCGGTGCCGCAACCACAGACTTAAGGTCGAGCTGAAA 240
Amdsxf2     CCGCGTGCGCCCCCGAACTGTGCCCGGTGCCGCAACCACAGACTTAAGGTCGAGCTGAAA 240
Amdsxf1     CCGCGTGCGCCCCCGAACTGTGCCCGGTGCCGCAACCACAGACTTAAGGTCGAGCTGAAA 240
            ************************************************************
             P  R  A  P  P  N  C  A  R  C  R  N  H  R  L  K  V  E  L  K

Amdsxm      GGCCACAAACGGTACTGCAAATACCGATATTGCAACTGCGAGAAATGCCGGTTGACGGCA 300
Amdsxf2     GGCCACAAACGGTACTGCAAATACCGATATTGCAACTGCGAGAAATGCCGGTTGACGGCA 300
Amdsxf1     GGCCACAAACGGTACTGCAAATACCGATATTGCAACTGCGAGAAATGCCGGTTGACGGCA 300
            ************************************************************
             G  H  K  R  Y  C  K  Y  R  Y  C  N  C  E  K  C  R  L  T  A

Amdsxm      GACAGGCAGCGGGTGATGGCACTACAGACTGCTCTACGACGTGCTCAGGCGCAGGATGAG 360
Amdsxf2     GACAGGCAGCGGGTGATGGCACTACAGACTGCTCTACGACGTGCTCAGGCGCAGGATGAG 360
Amdsxf1     GACAGGCAGCGGGTGATGGCACTACAGACTGCTCTACGACGTGCTCAGGCGCAGGATGAG 360
            ************************************************************
             D  R  Q  R  V  M  A  L  Q  T  A  L  R  R  A  Q  A  Q  D  E

Amdsxm      GCGCGAGCAAGAGCTAAAGAGACAGGTGGTCATCTTCCAGGAGTAGAGTTAGATAGACCG 420
Amdsxf2     GCGCGAGCAAGAGCTAAAGAGACAGGTGGTCATCTTCCAGGAGTAGAGTTAGATAGACCG 420
Amdsxf1     GCGCGAGCAAGAGCTAAAGAGACAGGTGGTCATCTTCCAGGAGTAGAGTTAGATAGACCG 420
            ************************************************************
             A  R  A  R  A  K  E  T  G  G  H  L  P  G  V  E  L  D  R  P

Amdsxm      GAGCCGCCTATAGTGAAAGCACCGAGGAGTCCAGTGGTACCGCCGCCGCCGCCGCGGTCT 480
Amdsxf2     GAGCCGCCTATAGTGAAAGCACCGAGGAGTCCAGTGGTACCGCCGCCGCCGCCGCGGTCT 480
Amdsxf1     GAGCCGCCTATAGTGAAAGCACCGAGGAGTCCAGTGGTACCGCCGCCGCCGCCGCGGTCT 480
            ************************************************************
             E  P  P  I  V  K  A  P  R  S  P  V  V  P  P  P  P  R  S

Amdsxm      CTCGGTTCTGCGAGTTGCGATTCAGTTCCGGGTTCACCAGGAGTTTCTCCATATGCTCCA 540
Amdsxf2     CTCGGTTCTGCGAGTTGCGATTCAGTTCCGGGTTCACCAGGAGTTTCTCCATATGCTCCA 540
Amdsxf1     CTCGGTTCTGCGAGTTGCGATTCAGTTCCGGGTTCACCAGGAGTTTCTCCATATGCTCCA 540
            ************************************************************
             L  G  S  A  S  C  D  S  V  P  G  S  P  G  V  S  P  Y  A  P

Amdsxm      CCGCCGCCGTCGGTACCTCCTTTGCAAGCGATGGCACCGCTGATGCCACCACAACAGCCT 600
Amdsxf2     CCGCCGCCGTCGGTACCTCCTTTGCAAGCGATGGCACCGCTGATGCCACCACAACAGCCT 600
Amdsxf1     CCGCCGCCGTCGGTACCTCCTTTGCAAGCGATGGCACCGCTGATGCCACCACAACAGCCT 600
            ************************************************************
             P  P  P  S  V  P  P  L  Q  A  M  A  P  L  M  P  P  Q  Q  P

Amdsxm      GCAGTGTCACTAGAGACATTAGTTGAAAATTGTCACAAACTCCTGGAGAAGTTCCACTAC 660
Amdsxf2     GCAGTGTCACTAGAGACATTAGTTGAAAATTGTCACAAACTCCTGGAGAAGTTCCACTAC 660
Amdsxf1     GCAGTGTCACTAGAGACATTAGTTGAAAATTGTCACAAACTCCTGGAGAAGTTCCACTAC 660
            ************************************************************
             A  V  S  L  E  T  L  V  E  N  C  H  K  L  L  E  K  F  H  Y
```

```
Amdsxm      TCCTGGGAGATGATGCCCCTGGTGTTGGTTATCCTAAATTATGCGGGTAGCGACCTAGAA 720
Amdsxf2     TCCTGGGAGATGATGCCCCTGGTGTTGGTTATCCTAAATTATGCGGGTAGCGACCTAGAA 720
Amdsxf1     TCCTGGGAGATGATGCCCCTGGTGTTGGTTATCCTAAATTATGCGGGTAGCGACCTAGAA 720
            ************************************************************
                S  W  E  M  M  P  L  V  L  V  I  L  N  Y  A  G  S  D  L  E

Amdsxm      GAGGCCTCGCGGAAAATTGACGAA---------------------------------- 745
Amdsxf2     GAGGCCTCGCGGAAAATTGACGAAGGGAAGATGATAATCAACGAGTATGCCAGAAAACAT 780
Amdsxf1     GAGGCCTCGCGGAAAATTGACGAAGGGAAGATGATAATCAACGAGTATGCCAGAAAACAT 780
            ************************
                E  A  S  R  K  I  D  E  G  K  M  I  I  N  E  Y  A  R  K  H

Amdsxm      ------------------------------------------------------------
Amdsxf2     AATTTGAATGTCTTCGACGGTCTCGAGCTGAGGAACTCGACACGCCAG------------ 828
Amdsxf1     AATTTGAATGTCTTCGACGGTCTCGAGCTGAGGAACTCGACACGCCAGTACGGACTTTAA 840

                N  L  N  V  F  D  G  L  E  L  R  N  S  T  R  Q  Y  G  L  X

Amdsxm      ------------------------------------------------------------
Amdsxf2     ---AAAATGCTGGAAATTAATAATATAAGTGGTGTACTGTCGTCATCGATGAAGTTATTT 885
Amdsxf1     TAGAAAATGCTGGAAATTAATAATATAAGTGGTGTACTGTCGTCATCGATGAAGTTATTT 900

                K  M  L  E  I  N  N  I  S  G  V  L  S  S  S  M  K  L  F

Amdsxm      ------------------------------------------------------------
Amdsxf2     TGCGAATGATACTTTGTTTTACGAGTGCCGTGGTTTTTGTGGACACATGCTGTGCGTTGC 945
Amdsxf1     TGCGAATGATACTTTGTTTTACGAGTGCCGTGGTTTTTGTGGACACATGCTGTGCGTTGC 960

                C  E  X

Amdsxm      ------------------------------------------------------GCCCAC 751
Amdsxf2     TCTGTGTTGCGAATTTCAACGGACAATTGTGTTGTCGCTTCACTGGACATTTA-GCCCAC 1005
Amdsxf1     TCTGTGTTGCGAATTTCAACGGACAATTGTGTTGTCGCTTCACTGGACATTTA-GCCCAC 1020
                                                                  ******
                                                                   A  H

Amdsxm      TGTGTGGTGCATCAGTGGCGACTGTACGAGCGGTCGCTGTATTCGTTGCTCGAGCTGCAG 811
Amdsxf2     TGTGTGGTGCATCAGTGGCGACTGTACGAGCGGTCGCTGTATTCGTTGCTCGAGCTGCAG 1065
Amdsxf1     TGTGTGGTGCATCAGTGGCGACTGTACGAGCGGTCGCTGTATTCGTTGCTCGAGCTGCAG 1080
            ************************************************************
                C  V  V  H  Q  W  R  L  Y  E  R  S  L  Y  S  L  L  E  L  Q

Amdsxm      GCGCGGAAGGGCACGTACGCTATGTGCTGCTCGCCTCGGTATGTGCTCGCGCCGGAGTAT 871
Amdsxf2     GCGCGGAAGGGCACGTACGCTATGTGCTGCTCGCCTCGGTATGTGCTCGCGCCGGAGTAT 1125
Amdsxf1     GCGCGGAAGGGCACGTACGCTATGTGCTGCTCGCCTCGGTATGTGCTCGCGCCGGAGTAT 1140
            ************************************************************
                A  R  K  G  T  Y  A  M  C  C  S  P  R  Y  V  L  A  P  E  Y

Amdsxm      GCACCCCACTTGCTGCCGATCCCTCTCACCACACAGCGCGCATCGCCGCCGCCGGCGCAC 931
Amdsxf2     GCACCCCACTTGCTGCCGATCCCTCTCACCACACAGCGCGCATCGCCGCCGCCGGCGCAC 1185
Amdsxf1     GCACCCCACTTGCTGCCGATCCCTCTCACCACACAGCGCGCATCGCCGCCGCCGGCGCAC 1200
            ************************************************************
                A  P  H  L  L  P  I  P  L  T  T  Q  R  A  S  P  P  P  A  H

Amdsxm      TTGTAGCGATGCGACCACGCGCCGACTCAGACGGCGCAACAACCCACCGAGCCACCTTCA 991
Amdsxf2     TTGTAGCGATGCGACCACGCGCCGACTCAGACGGCGCAACAACCCACCGAGCCACCTTCA 1245
Amdsxf1     TTGTAGCGATGCGACCACGCGCCGACTCAGACGGCGCAACAACCCACCGAGCCACCTTCA 1260
            ************************************************************
                L  X

Amdsxm      ACATCGACGGGGCGGAGCACCGCCACTTAAGCCATCAACGGTGACAGAAGCGTGTGGGAC 1051
Amdsxf2     ACATCGACGGGGCGGAGCACCGCCACTTAAGCCATCAACGGTGACAGAAGCGTGTGGGAC 1305
Amdsxf1     ACATCGACGGGGCGGAGCACCGCCACTTAAGCCATCAACGGTGACAGAAGCGTGTGGGAC 1320
            ************************************************************
```

```
Amdsxm      ATAAATGCTGTGTTAAAAATATGAACTGGACATAGTAGTGTTAACCCTGTAAAAATGTTA 1111
Amdsxf2     ATAAATGCTGTGTTAAAAATATGAACTGGACATAGTAGTGTTAACCCTGTAAAAATGTTA 1365
Amdsxf1     ATAAATGCTGTGTTAAAAATATGAACTGGACATAGTAGTGTTAACCCTGTAAAAATGTTA 1380
            ************************************************************

Amdsxm      CGCGTACGTGAAACTGTAGATTATAGTTTATTTTCGTCTGTTACAATGTCGTCATTGATA 1170
Amdsxf2     CGCGTACGTGAAACTGTAGATTATAGTTTATTTTCGTCTGTTACAATGTCGTCATTGATA 1424
Amdsxf1     CGCGTACGTGAAACTGTAGATTATAGTTTATTTTCGTCTGTTACAATGTCGTCATTGATA 1440
            ************************************************************

Amdsxm      TCCAGTTTTGGTTATGGTGCGCCTATATTATCTATACTTTAAATTATTATAAAGCTGAAG 1230
Amdsxf2     TCCAGTTTTGGTTATGGTGCGCCTATATTATCTATACTTTAAATTATTATAAAGCTGAAG 1484
Amdsxf1     TCCAGTTTTGGTTATGGTGCGCCTATATTATCTATACTTTAAATTATTATAAAGCTGAAG 1500
            ************************************************************

Amdsxm      AGTTTGCTTGAACGCTTTTTTTTTTTCGAACTGACTACGCTATTGACCCAACAAGGTCTTA 1290
Amdsxf2     AGTTTGCTTGAACGCTTTTTTTTTTTCGAACTGACTACGCTATTGACCCAACAAGGTCTTA 1544
Amdsxf1     AGTTTGCTTGAACGCTTTTTTTTTTTCGAACTGACTACGCTATTGACCCAACAAGGTCTTA 1560
            ************************************************************

Amdsxm      ACAGCTTCACCGGAGTAAAAAAAAAAAAAAAAAAAAAAAAA 1332
Amdsxf2     ACAGCTTCACCGGAGTAAAAAAAAAAAAAAAAAAAAAAAAA 1586
Amdsxf1     ACAGCTTCACCGGAGTAAAAAAAAAAAAAAAAAAAAAAAAA 1602
            *****************************************
```

Fig. 2.12 ClustalW alignment of the *Amdsx* transcripts and their deduced amino acid sequences. The deduced amino acid sequence is shown below the nucleotide sequences. Suffix 'f' or 'm' to *Amdsx* indicates female or male, respectively. Female-specific amino acids are underlined whereas male-specific amino acids are in italics. The amino acids in red background are specific to one type of female-specific protein (AmDsxF1), whereas the amino acids in grey background are specific to another female-specific protein (AmDsxF2). Start codon (ATG) and stop codon are highlighted in blue and red colours, respectively. X represents stop codon sites.

2.3.2 Analysis of *Aadsx* and *Amdsx* transcripts

Alignment of *Aadsx* transcripts, RT-PCRs and genomic PCRs confirmed that *Aadsx* gene is made up of seven exons of variable lengths (**Table 2.4**). Exons 2-5 are skipped in case of males, leading to change in the reading frame for protein synthesis (**Fig. 2.13**). Exons 1,2,4 and 7 are common to all the female-specific *dsx* transcripts; exons 3 and 5 are specific to *Aadsxf1* and *Aadsxf2*, respectively and exon 6 is present in *Aadsxf1 Aadsxf2 Aadsxf3 Aadsxf4* whereas absent in *Aadsxf5* and *Aadsxf6* (**Table 2.5**). Tissue-specific expression profile as described in the experimental section showed no difference in the expression of different *dsx* splice forms in different tissues (**Fig. 2.9**) and in different developmental stages tested. RT-PCR using the combinations of a common forward primer (in the first exon) with different reverse primers (in different exons) (**Fig. 2.13**), gave the expected numbers and sizes of the amplicons in both male and female (**Fig. 2.14**). The primer combinations, the amplified transcript/s and the length/s of the amplicons in each RT-PCR experiments are summarized in the **Table 2.6**.

Table 2.4: Exons of *Aadsx* and their corresponding length

Serial No.	Exon No.	Exon length (bp)
1	Exon 1	741
2	Exon 2	99
3	Exon 3	133
4	Exon4	170
5	Exon 5	79
6	Exon 6	215
7	Exon 7	388

Table 2.5: *Aadsx* transcripts, their length and presence of different exons in female and male specific transcripts

Serial No.	Transcript	Length of the transcript (bp)	Exons present
1	*Aadsxf1*	1746	1,2,3,4,6,7
2	*Aadsxf2*	1676	1,2,4,5,6,7
3	*Aadsxf3*	1613	1,2,4,6,7
4	*Aadsxf4*	1598	1,2,4,6,7
5	*Aadsxf5*	1398	1,2,4,7
6	*Aadsxf6*	1383	1,2,4,7
7	*Aadsxm*	1344	1,6,7

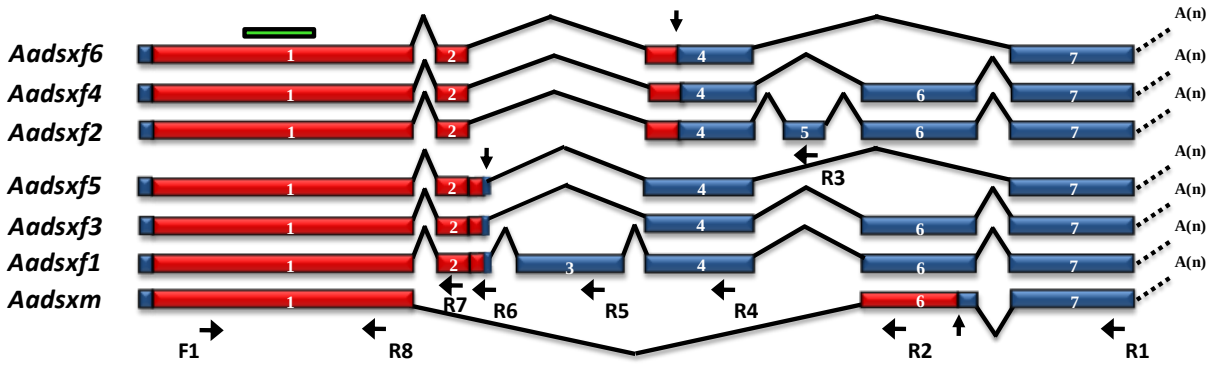


Fig. 2.13 Schematic representation of splice forms of *Aadsx* pre-mRNA, showing the primer positions for the RT-PCR experiments. Boxes are exons and lines are introns. Red coloured portion is the ORF whereas blue coloured regions are UTRs. Seven different splice products of *Aadsx* pre-mRNA, six female-specific (*Aadsxf1 Aadsxf2 Aadsxf3 Aadsxf4 Aadsxf5* and *Aadsxf6*) and one male-specific (*Aadsxm*), are produced. A(n) represents the poly-adenylation site. Vertical arrows represent stop codon sites whereas horizontal arrows represent primer positions. Green colour bar represents the probe region used for Southern hybridization.

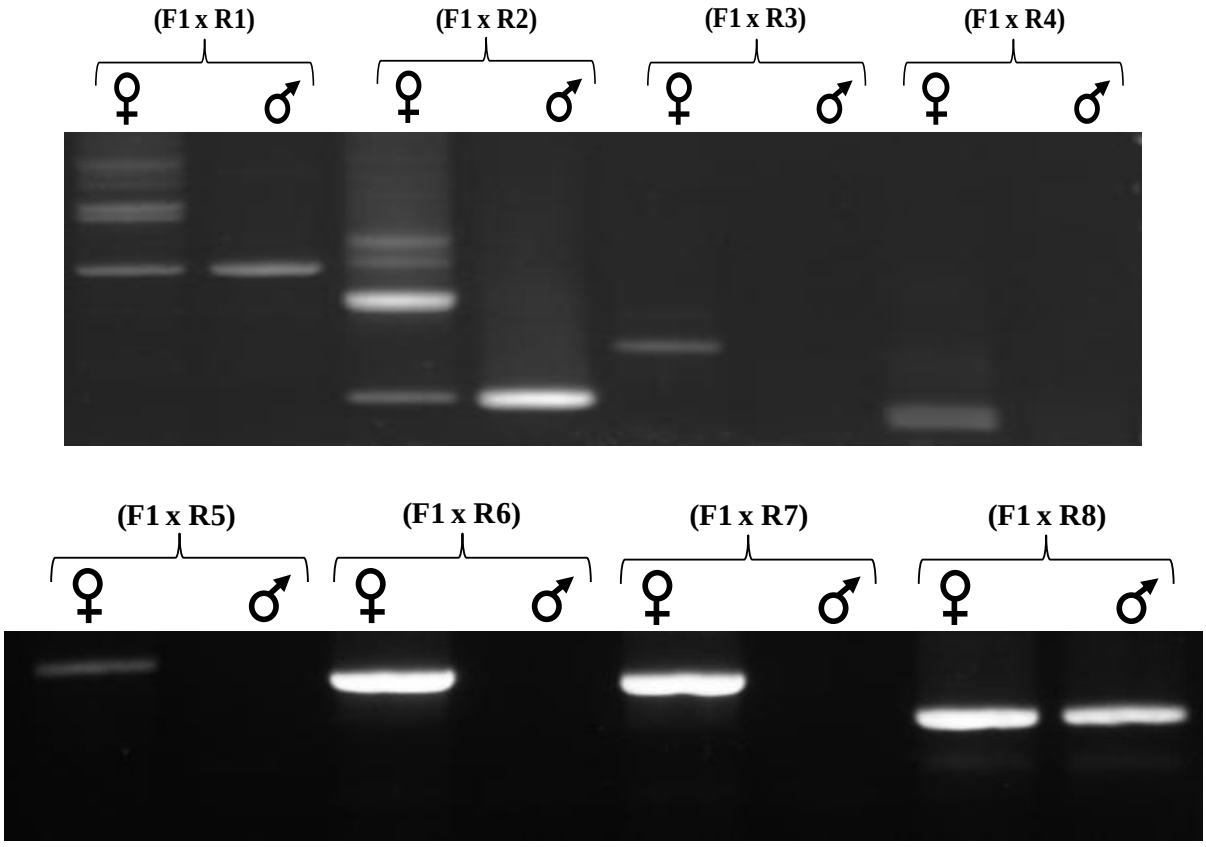


Fig. 2.14 Semi-quantitative RT-PCR using cDNA from *A. assama* fatbody as template and the combinations of one common forward (F1) and eight different reverse (R1-R8) primers (**Fig. 2.13**). Number of amplicons was consistent with the primer set used (**Table 2.6**).

Table 2.6: *Aadsx* transcripts amplicon size profile using different primer-pairs

Serial No.	Primer combinations	Transcripts amplified	Band size	
			Female	Male
1	F1 x R1	f1-f6 and m	1178, 1108, 1045, 1030, 830, 815	776
2	F1 x R2	f1, f2, f3, f4	803, 732, 669, 654	400
3	F1 x R3	f2	524	-
4	F1 x R4	f1-f6	498, 349, 364, 349, 364, 349	-
5	F1 x R5	f1	390	-
6	F1 x R6	f1, f3 and f5	309, 309	-
7	F1 x R7	f1-f6	275	-
8	F1 x R8	f1-f6 and m	122	122

* All the different *Aadsx* transcripts in this table are represented by f1-f6 (female specific transcripts) and m (male specific transcript).

In case of *Amdsx*, the full length female-specific transcripts are of 1587bp (*Amdsxf1*) and 1603bp (*Amdsxf2*) and the full length male-specific transcript (*Amdsxm*) is of 1333bp. The exon-intron boundaries could not be determined from the *Amdsx* transcript sequences because of lack of genomic sequence information.

2.3.3 *dsx* splicing pattern in *A. assama*

The sex-specific splicing pattern of *dsx* among dipterans, lepidopterans and hymenopterans are compared in the context of insect phylogeny. In *D. melanogaster*, the female-specific *dsx* mRNA connects exons 1, 2, 3, and 4, while in the male *dsx* mRNA exon 4 is skipped and exon 3 splices to exon 5 (Burtis and Baker, 1989) (**Fig. 2.15**). Thus, exons 1, 2 and 3 are common to both the sex-specific *dsx* transcripts; exon 4 is specific to females whereas exons 5 and 6 are specific to males. The same splicing pattern is observed in *A. gambiae* also. *B. mori dsx* (*Bmdsx*) has a different splicing pattern compared to *Drosophila dsx*: exons 1, 2, 5 and 6 are common to both male- and female-specific forms, whereas exons 3 and 5 are female-specific which are skipped in the male variant (*Bmdsxm*) (Suzuki et al., 2001) (**Fig. 2.15**). Thus the dipteran and lepidopteran *dsx* splicing pattern is different in the sense that, the final exon is not shared by the male and female splice forms in the former, whereas last two exons are shared by the male and female splice forms in the latter case. The

sex-specific splicing pattern of *Apis dsx* (*Am_dsx*) shares similarity, both with sex-specific splicing of dipteran and lepidopteran *dsx*. In *Apis mellifera*, the single female-specific exon in *Am_dsxf1* is skipped in the male variant (*Am_dsxm*) resulting in the final two exons being common to *Am_dsxf1* and *Am_dsxm* (**Fig. 2.15**). In *A. assama dsx* (*Aadsx*) there are four female-specific exons (exon 2 to exon 5) which are skipped in male variant (*Aadsxm*). Thus the sex-specific splicing of *Aadsx* is similar to that of *Bmdsx*. The major difference in the splicing of *Aadsx* with *Bmdsx* is the production of multiple female-specific splice forms in the former which can be grouped into two on the basis of their ORFs. The female-specific transcripts of each group differ from each other in their UTR regions, ultimately making only one kind of protein, thus together encoding two types of female-specific proteins. As in the case of *Bmdsx* here also there is no Tra-Tra2 binding motif related sequence. The fourth exon of *Aadsx* is > 90% identical to the fourth exon of *Bmdsx* (**Fig. 2.16**).

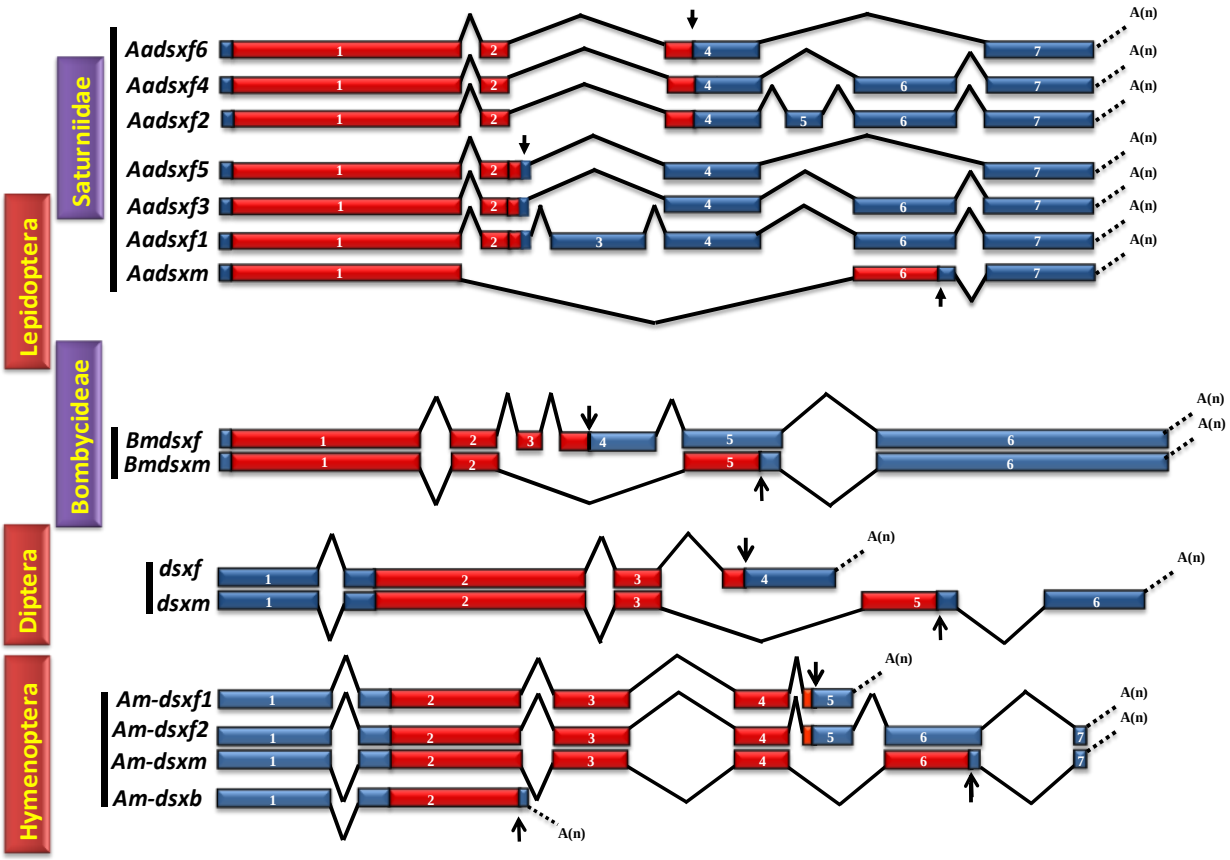


Fig. 2.15 Schematic representation of the *dsx* genes of representative insect species belonging to the orders, Lepidoptera (*A. assama* and *B. mori*), Diptera (*Drosophila*) and Hymenoptera (*A. mellifera*). Boxes represent exons and lines represent introns. Only exons are drawn to scale. Red coloured portion is the ORF whereas blue coloured regions are UTRs. Poly A tail is represented by A(n) and arrows represent the position of stop codon for the respective transcript.

2.3.4 Fourth exon of *Bmdsx*, *Aadsx* and *Amdsx* are conserved

As described earlier, 4th exon of *Bmdsx* contains the cis-acting element responsible for the sex-specific splicing of *Bmdsx* pre-mRNA and the skipping of exon 3 and exon 4 in *Bmdsx*m splice form. On analyzing the nucleotide sequence of *Aadsx* and *Amdsx* transcripts and then comparing them with the *Bmdsx* nucleotide sequences, we found the nucleotide sequence of 4th exon to be conserved in the *dsx* homologues from all the three silkmoths. This sequence was present in all the female splice variants of *Aadsx* (**Fig. 2.16**) and also in the female splice form of *Amdsx*, but absent in the male splice variants.

```
Aadsx      -AAAATGCTGGAAATTAATAATATAAGTGGTGTACTGTCGTCATCGATGAAGTTATTTTG  59
Amdsx      GAAAATGCTGGAAATTAATAATATAAGTGGTGTACTGTCGTCATCGATGAAGTTATTTTG  60
Bmdsx      GAAAATGCTGGAAATTAATAATATAAGTGGTGTACTGTCTTCGTCAATGAAGTTATTTTG  60
            ************************************  ** ** **************

Aadsx      CGAATGATACTTTGTTTTACGAGTGCCGTGGTTTTTGTGGACACATGCTGTGCGATGCTC  119
Amdsx      CGAATGATACTTTGTTTTACGAGTGCCGTGGTTTTTGTGGACACATGCTGTGCGTTGCTC  120
Bmdsx      CGAATGATACTTAGTTTTACAAGTGCCGTGGTGTGTGTTGACACTTGCTGTGCGATGCTG  120
           ************ ******* *********** * *** ***** ********* ****

Aadsx      TGTGTTGCGAATTTCAACGGACAATTGTGTTGTCGCTTCACTGGACATTTA-  170
Amdsx      TGTGTTGCGAATTTCAACGGACAATTGTGTTGTCGCTTCACTGGACATTTAG  172
Bmdsx      TG-----CGAATTTCAACGGAAATATTTGTTGTCGTAACATTGGATCTATGG  167
           **     ************** *  * ********   ** ****  * *
```

Fig. 2.16 ClustalW alignment of 4th exon of *Bmdsx* and *Aadsx* and the corresponding sequence of *Amdsx*. The *Bmdsx* nucleotide stretch in the yellow background is the cis-acting element responsible for the skipping of *Bmdsx* exon 3 and exon 4 in males.

2.3.5 Protein/s from the female nuclear extract binds to the *Aadsx* sequence identical to the exonic splicing silencer sequence of *Bmdsx* exon 4

Recently it is shown that fourth exon of *Bmdsx* contains a 20 nucleotide (UUAAUAAUAUAAGUGGUGUA) long Putative Exonic Splicing Silencer (PESS) responsible for the skipping of exon 3 and exon 4 in males (Suzuki et al., 2008). The PESS of *Bmdsx* is 100% identical to the similar sequence in the 4th exon of *Aadsx* and corresponding *Amdsx* sequence. Synthetic RNA oligo for PESS sequence was used as a probe for Gel Retardation Assay (GRA), using nuclear extracts prepared from the male and female fatbody tissues of *A. assama* larvae. We found the sex-specific binding of some protein/s from the female nuclear extract in contrast to what was observed in case of *B. mori* (**Fig. 2.17**). At this stage these results are preliminary and need further experimental support.

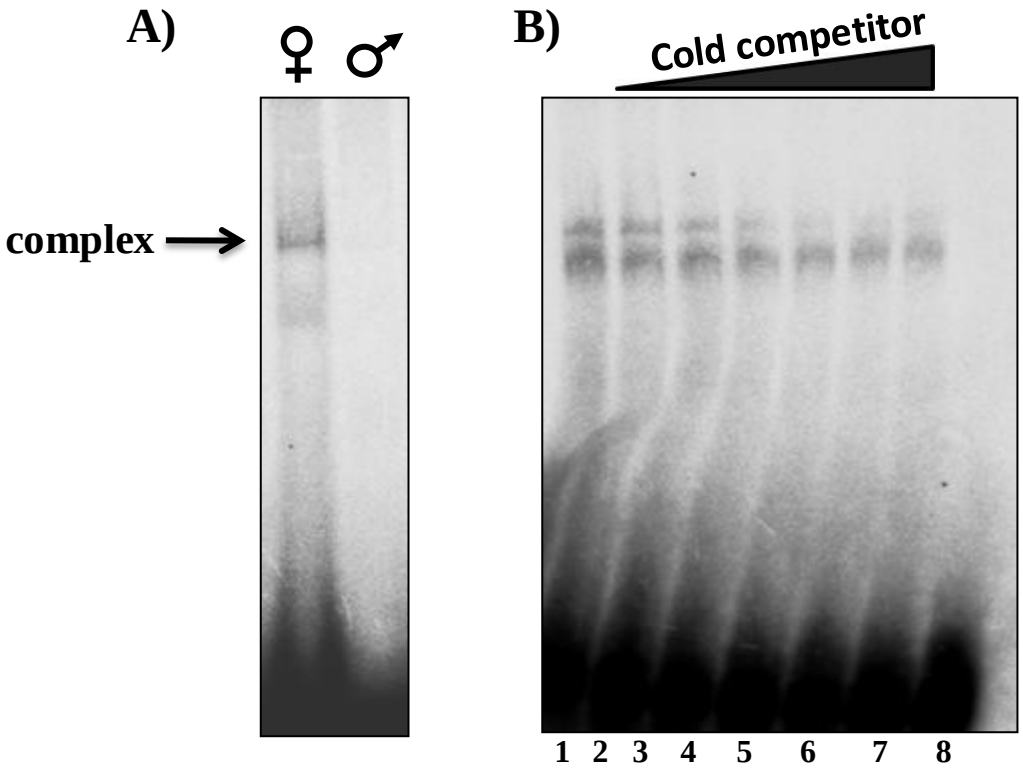


Fig. 2.17 A) Binding of nuclear protein/s from female fatbody to the putative PESE (Putative Exonic Splicing Enhancer) sequence of *Aadsx*. No binding is seen when male nuclear extract was used. **B**) The RNA-protein interaction between labelled RNA probe (.5pmol) and the fatbody nuclear (10µg) extract from female in the presence of increasing amount {1) 0µg 2) 10µg 3) 100µg 4) 1000µg 5) 10000µg 6) 10000µg} of unlabelled competitor. The last well (8th) represents the free probe.

2.3.6 Genomic copy number of *Aadsx*

To test whether the multiple sex-specific *Aadsx* transcripts are produced as a consequence of alternative splicing of pre-mRNA from a single gene or were transcribed by more than one gene, Southern hybridisation was done as described in the section 2.2.8. Genomic DNA from male and female were digested separately using two different Restriction enzymes (*Nde*I or *Spe*I). 370bp probe from the common region of *Aadsx* transcripts (**Fig. 2.13**) was used for hybridisation. Single band, common to both male and female hybridised to the probe (**Fig. 2.18**), indicating that *Aadsx* is present as a single copy per haploid genome.

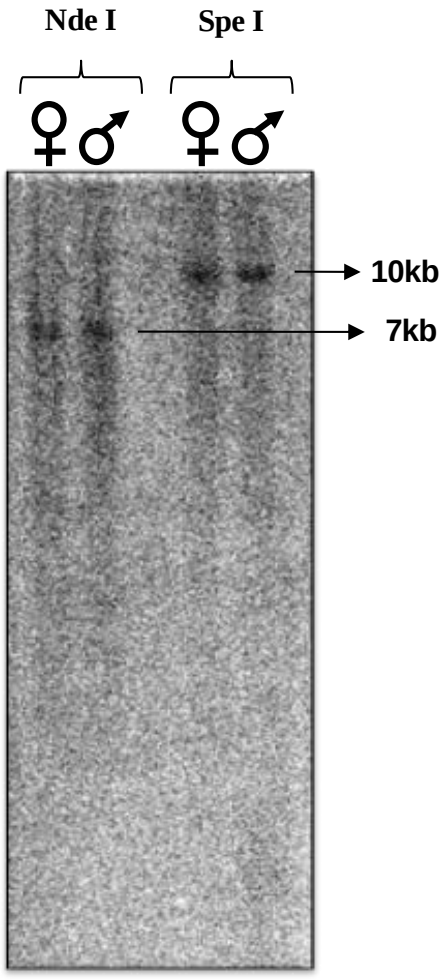


Fig. 2.18 *Aadsx* is present as a single copy per haploid genome. Southern hybridization was performed with a radio labelled probe common to both male and female *Aadsx* (**Fig. 2.13**), that has no restriction site recognized by any of the two enzymes used. The molecular sizes of the hybridized bands are indicated in the right side of the figure.

2.3.7 Sex-Specific proteins coded by *Aadsx* and *Amdsx* transcripts

AaDsx: On the basis of virtual translation of the six female-specific transcripts, the proteins coded by *Aadsx* can be grouped into two, three transcripts (*Aadsxf2*, *Aadsxf4* and *Aadsxf6*) coding for the female-specific protein, AaDsxF1 and the other three transcripts (*Aadsxf1*, *Aadsxf3* and *Aadsxf5*) coding for the second female-specific protein, AaDsxF2 (**Fig. 2.19**). The male-specific splice form, *Aadsxm*, on the other hand, codes for only one male-specific protein, AaDsxM (**Fig. 2.20**).

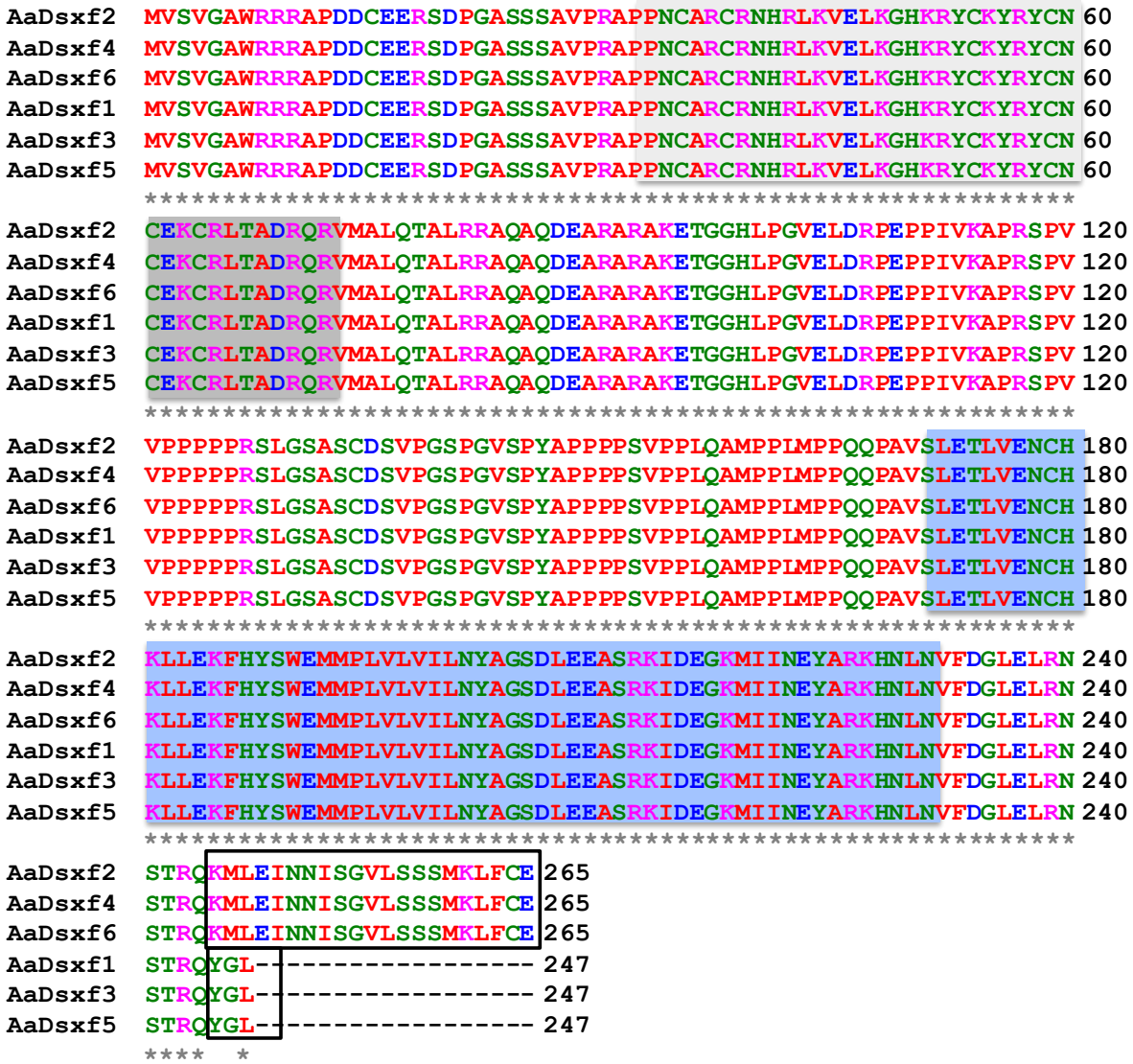


Fig. 2.19 ClustalW alignments of deduced amino acid sequences of female-specific *Aadsx* cDNAs. Shaded regions are DNA binding domain (OD1 domain) and the oligomerization domain (OD2 domain). Three splice forms (*Aadsxf2*, *Aadsxf4* and *Aadsxf6*) code for one kind of AaDsx protein i.e., AaDsxF1 whereas another three splice forms (*Aadsxf1*, *Aadsxf3* and *Aadsxf5*) code for AaDsxF2.

The deduced amino acid (aa) sequences from sex-specific transcripts can be divided into three parts: the region common to both the sexes (aa 1-216), the female-specific regions (aa 217-265 of AaDsxF1 and aa 217- 247 of AaDsxF2), and the male-specific region (aa number 217-279) (**Fig. 2.19, 2.20**). The female-specific (AaDsxF1 and AaDsxF2) and male-specific (AaDsxM) proteins are similar at their N-termini but differ at their C-termini, starting from the C-terminal region of DBD/OD2 domains. Both the female proteins (AaDsxF1 and AaDsxF2) differ at their C-termini by 21 aa (**Fig. 2.20**).

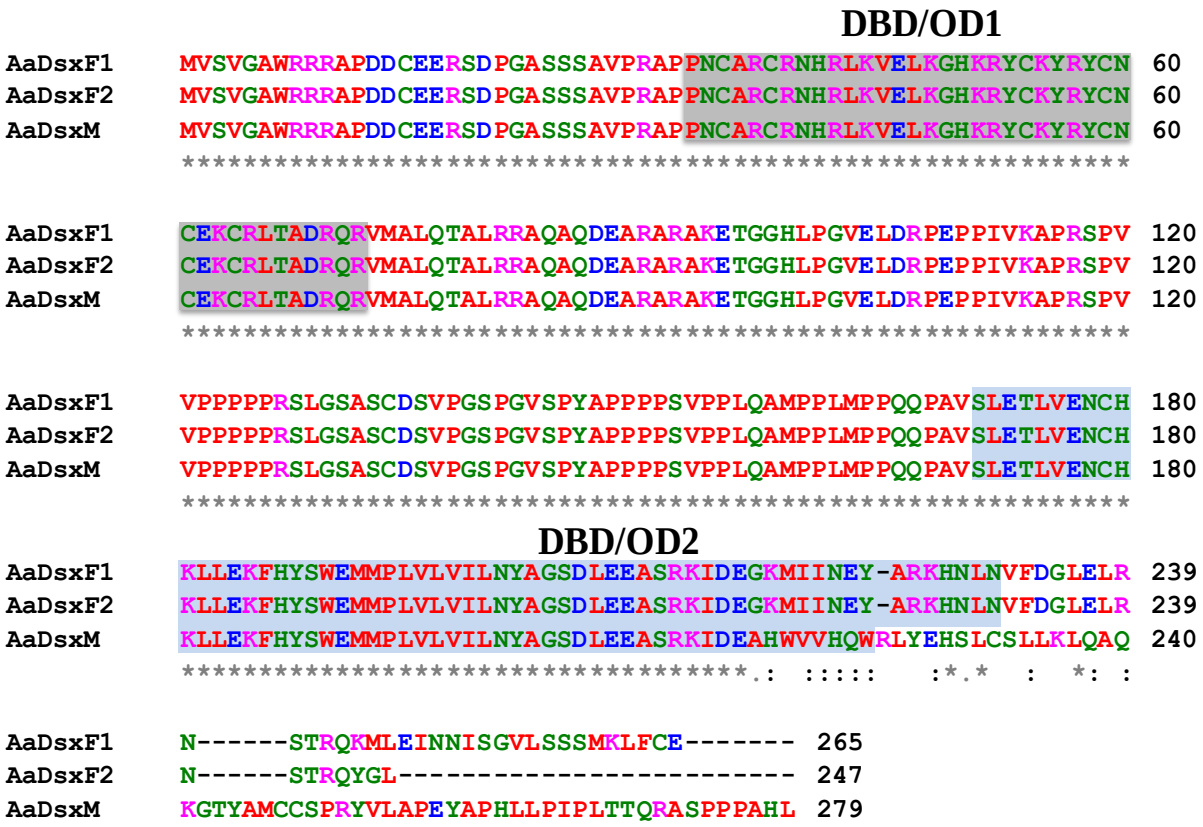


Fig. 2.20 ClustalW alignment of deduced amino acid sequences of sex-specific AaDsx proteins. Grey shaded regions represent the DNA binding domain (OD1 domain) whereas blue shaded regions represent the oligomerization domain (OD2 domain). Two female-specific proteins are AaDsxF1 and AaDsxF2, whereas male-specific protein is AaDsxM. The two female specific proteins differ in their extreme C-terminal regions but share common OD1 and OD2 domains. The male specific protein differs from the female specific proteins in its C-terminal OD2 domain.

AmDsx: The two female-specific *Amdsx* transcripts code for two female-specific putative proteins (AmDsxF1 and AmDsxF2) whereas the male-specific transcript code for male-specific putative protein (AmDsxM). Both the female-specific proteins share common OD1 and OD2 domains and as in case of AaDsx female-specific proteins, they also differ only at their extreme C-terminal end. The male-specific protein, AmDsxM differs from the female proteins at the C-terminal end starting from the C-terminal region of the OD2 domain. Alignment of these putative proteins is shown in **Fig. 2.21**.

```
AmDsxF1   MVSVGAWRRRAPDDCEERSDPGASSSAVPRAPPNCARCRNHRLKVELKGHKRYCKYRYCN 60
AmDsxF2   MVSVGAWRRRAPDDCEERSDPGASSSAVPRAPPNCARCRNHRLKVELKGHKRYCKYRYCN 60
AmDsxM    MVSVGAWRRRAPDDCEERSDPGASSSAVPRAPPNCARCRNHRLKVELKGHKRYCKYRYCN 60
          ************************************************************

AmDsxF1   CEKCRLTADRQRVMALQTALRRAQAQDEARARAKETGGHLPGVELDRPEPPIVKAPRSPV 120
AmDsxF2   CEKCRLTADRQRVMALQTALRRAQAQDEARARAKETGGHLPGVELDRPEPPIVKAPRSPV 120
AmDsxM    CEKCRLTADRQRVMALQTALRRAQAQDEARARAKETGGHLPGVELDRPEPPIVKAPRSPV 120
          ************************************************************

AmDsxF1   VPPPPPRSLGSASCDSVPGSPGVSPYAPPPPSVPPLQAMAPLMPPQQPAVSLETLVENCH 180
AmDsxF2   VPPPPPRSLGSASCDSVPGSPGVSPYAPPPPSVPPLQAMAPLMPPQQPAVSLETLVENCH 180
AmDsxM    VPPPPPRSLGSASCDSVPGSPGVSPYAPPPPSVPPLQAMAPLMPPQQPAVSLETLVENCH 180
          ************************************************************

AmDsxF1   KLLEKFHYSWEMMPLVLVILNYAGSDLEEASRKIDEGKMIINEYARKHNLNVFDGLELRN 240
AmDsxF2   KLLEKFHYSWEMMPLVLVILNYAGSDLEEASRKIDEGKMIINEYARKHNLNVFDGLELRN 240
AmDsxM    KLLEKFHYSWEMMPLVLVILNYAGSDLEEASRKIDEAHCVVHQWR-LYERSLYSLLELQA 239
          ***********************************.: :::::   :: .::. ***:

AmDsxF1   --------STRQKMLEINNISGVLSSSMKLFCE------- 265
AmDsxF2   --------STRQYGL------------------------ 247
AmDsxM    RKGTYAMCCSPRYVLAPEYAPHLLPIPLTTQRASPPPAHL 279
```

Fig. 2.21 ClustalW alignment of deduced amino acid sequences of sex-specific *Amdsx* proteins. Grey shaded regions represent the DNA binding domain (OD1 domain) whereas blue shaded regions represent the oligomerization domain (OD2 domain). Two female-specific proteins are AmDsxF1 and AmDsxF2, whereas male-specific protein is AmDsxM. The two female specific proteins differ in their extreme C-terminal regions but share common OD1 OD2 domains. The male specific protein differs from the female specific proteins in its C-terminal OD2 domain.

2.4 Discussion

Production of sex-specific transcripts is usual for any *doublesex* gene reported till date. In all the dipteran *dsx* reported so far, one male and one female-specific transcripts are produced. The same is also true for the lepidopteran, *B. mori dsx* gene (*Bmdsx*) (Ohbayashi et al., 2001). One exception is the *A. mellifera dsx* (*Am_dsx*) gene, which produces 4 alternatively spliced transcripts, two specific to female, one specific to male and one common to both the sexes. Both the female-specific transcripts share a common ORF, coding for the same type of putative protein (Cho et al., 2007). A generalized fact that emerges from the available data on *dsx* genes and their splice form in different organisms is, *dsx* produces sex-specific alternatively spliced transcripts to produce one male- and one female-specific Dsx proteins which differ only at their extreme C-termini.

In the study reported here we identified *dsx* genes, *Aadsx* and *Amdsx* from two species of wild silkmoths, *A. assama* and *A. mylitta*, respectively and showed that both *Aadsx* and *Amdsx* produce sex-specific transcripts. Unlike the case of *dsx* splicing in other insect species, the pre-mRNA of *Aadsx* gene sex-specifically splices to produce six splice variants in females and one in male. Analysis of different splice variants gave the idea of splice junctions which were further confirmed through the Genomic PCR results. All these results indicated the presence of 7 exons in *Aadsx* gene. As in the case of *Bmdsx*, there is no *dsx*RE (Tra/Tra2 binding sequence) and PRE in any of the exons of *Aadsx* gene. The *Aadsx* exon 4 which is specific to all the female splice forms shows 90% sequence similarity with the *Bmdsx* exon 4. The *Bmdsx* exon 4 contains the sequence which, on binding with BmPSI, acts as splicing suppressor and leads to the skipping of exons 3 and 4 in males (Suzuki et al., 2008). The sequences responsible for sex-specific splicing in *Bmdsx* were found to be conserved in *Aadsx* and *Amdsx* which points to the existence of common mechanism of sex-specific splicing of *dsx* in silkmoths. The six female-specific splice variants could be grouped into two on the basis of their ORF; three splice forms coding for putative protein, AaDsxF1 and the other three code for putative protein, AaDsxF2. The male-specific splice form codes for a male-specific protein, AaDsxM. Both the female-specific AaDsx proteins (AaDsxF1 and AaDsxF2) are having all the features of a functional Dsx protein. They share common DNA binding (OD1 domain) and oligomerization domain (OD2 domain) but differ in their amino acid composition at their extreme C-terminal ends; the longer of the two female-specific proteins (AaDsxF1) has additional 21 aa whereas the shorter one (AaDsxF2) has an

additional stretch of 3 aa. The difference in the female-specific transcripts of each group (having common ORF) lies in their 3' UTR region; this might have some regulatory significance since UTRs are the known target by the proteins or miRNAs to regulate translation (Ambros, 2004; Beckmann et al., 2005) . The male-specific AaDsx protein (AaDsxM) differs from AaDsxF at its C-terminus. The C-terminal region of OD2 domain is different in the sex-specific AaDsx proteins as in the case of sex-specific Dsx proteins in other insect species.

The *Amdsx pre-mRNA* splices to generate two female-specific and one male-specific splice forms. The two female splice forms differ from each other by the presence or absence of a stretch of 15nt. Because of this difference, the putative protein encoded by the female-specific transcripts differ at their C-termini; after the common shared region (aa 1-244), one of the female-specific putative proteins (AmDsxF1) contains an additional stretch of 21 aa whereas the other female-specific protein (AmDsxF2) contains an additional stretch of only 3 aa. Male splice form codes for a male-specific putative protein, which differs from the female specific proteins in its OD2 domain.

Existence of two female-specific Dsx proteins in wild silkmoths (*A. assama* and *A. mylitta*) raises a possibility that single female Dsx protein alone may not be able to execute the functions related to somatic sexual differentiation. Perhaps the activity of one Dsx protein needs the help of other Dsx protein to exert effect on their downstream genes involved in sexual differentiation. In *D. melanogaster* the Ix protein is required for the proper functioning of the female-specific Dsx protein. In *A. assama*, *ix* gene is found (Arunkumar, K.P and Nagaraju J, manuscript under preparation) but there is no sex-specific splicing of its pre-mRNA, as reported in the case of *B. mori ix* (*Bmix*) (Fujii and Shimada, 2007). So the possibility of *ix* being involved in sex-specific functions in silkmoths may be excluded. May be the other female-specific Dsx protein in wild silkmoths plays the same role what Ix does in case of *D. melanogaster*. This can only be proved by the functional analysis of the AaDsx proteins.

The conservation of PESS sequence of *Bmdsx* in the *Aadsx* and *Amdsx*, points towards the existence of common splicing mechanism of *dsx* gene in silkmoths. *Bmdsx* splices in the default female mode (due to the presence of strong female-specific splice acceptor sites), and male splice form is generated due to the binding of a trans-factor (BmPSI) to the cis-acting

PESS sequence present in the exon 4, which makes those splice site weak resulting in the skipping of exons 3 and 4 (Suzuki et al., 2008). In case of *Aadsx*, some protein(s) from the female nuclear extract rather than male nuclear extract, binds to the same sequence (UUAAUAAAUAUAAGUGGGUGUA). Considering the fact that sex-specific splicing of *Aadsx* shows similar pattern to that of *Bmdsx* splicing, it can be hypothesized that the skipping of exon 3 and exon 4 in male-specific *Aadsx* transcript is a default process (because of weak 3' splice site of exon 3) and binding of some protein(s) to exon 4 sequence in females makes the splice site strong enough to be recognised by the splicing machinery. Thus, in female forced splicing of *Aadsx* takes place so that exons 3 and 4 are included in *Aadsxf* (**Fig. 2.22**). This hypothesis further draws support by the presence of trace amount of male splice form of *Aadsx* in female tissues.

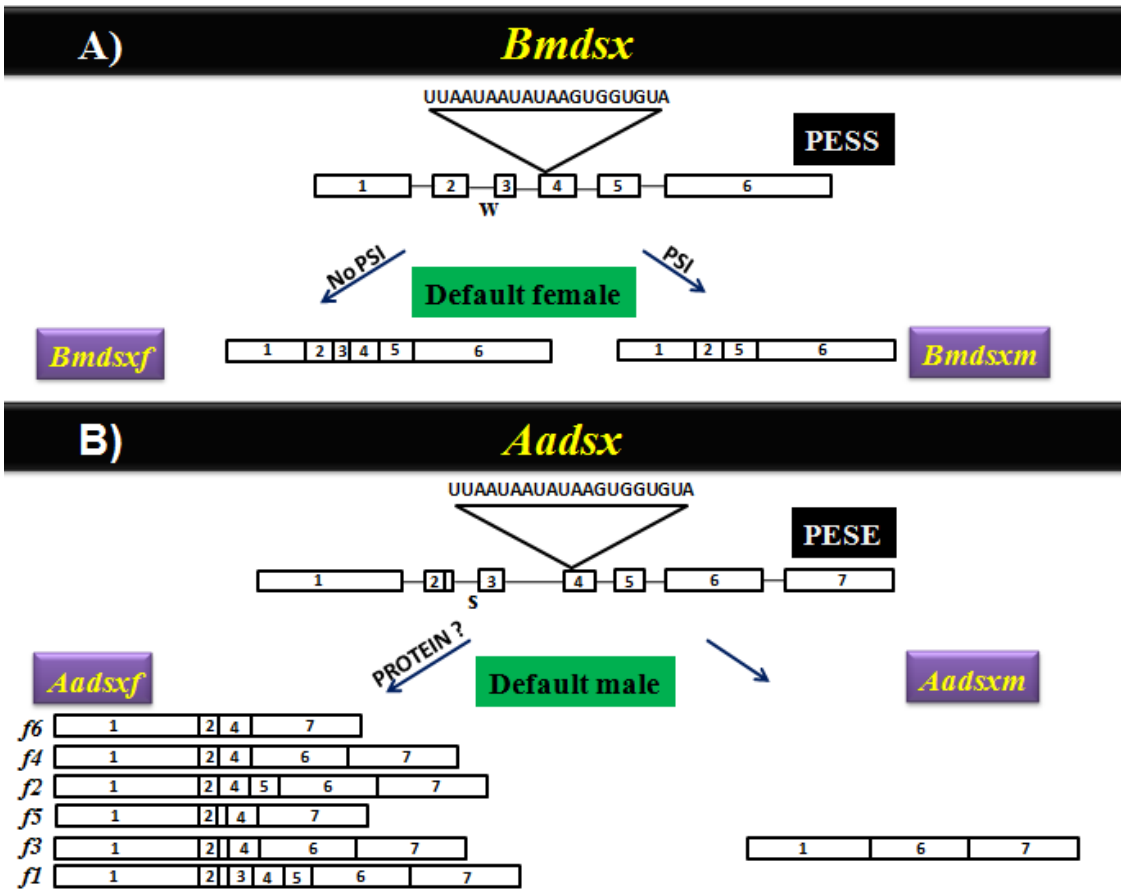


Fig. 2.22 Schematic diagram representing a hypothetical model for the *Aadsx* splicing (**A**) and its comparison with *Bmdsx* splicing (**B**). Boxes represent exons and lines represent introns, exons have been numbered. The presence of 3'strong splice-site-**s** (rather than weak-**w** in case of *Bmdsx*) in *Aadsx* leads to the default form as male, whereas binding of some protein in female to the **PESE** site results into forced female type splicing.

Chapter - III

Functional characterization of *Aadsx*

3.1 Introduction

In *Drosophila*, the sex-specific splice forms of *dsx* gene produce sex-specific proteins which differ only at their C-termini thus contributing to the sex-specific antagonistic functions of these proteins. The Dsx proteins act as transcriptional regulators which control the activity of genes responsible for the differentiation of sexually dimorphic traits. *Drosophila* yolk protein genes (*Yp1* and *Yp2*) are, by far the best characterized target genes of sex-specific Dsx proteins and are affected by any change in the sex determination pathway (Belote et al., 1985; Ota et al., 1981; Postlethwait et al., 1980). Dsx protein alters the basal level expression of these genes; DsxM represses basal transcription of these genes whereas DsxF enhances transcription of the same (An et al., 1996; Erdman et al., 1996). Several specific sites in the enhancer sequence of yolk protein encoding genes, designated as fat body enhancer elements (FBE) are utilized by sex-specific Dsx proteins to exert their antagonistic affects on these genes. DsxM and DsxF proteins bind to three sites within the 127-bp enhancer that directs sex- and tissue-specific transcription of yolk protein genes (Coschigano and Wensink, 1993). Binding of DsxF and DsxM to these enhancer elements has been shown by in-vitro studies as well (Burtis et al., 1991). Ectopic expression of DsxF in XY males results in appearance of some of the female characters such as enhanced expression of yolk protein genes and the production of female profile of pheromones (Baker et al., 1989; Waterbury et al., 1999). The female-specific differentiation in *D. melanogaster* is not achieved by the DsxF activity alone rather it requires to interact with the *intersex* (*ix*) gene product to execute its function (Waterbury et al., 1999). The *ix* gene is expressed in both males and females and functions together with the DsxF to regulate terminal differentiation (Garrett-Engele et al., 2002). Intersex protein is not required for the activity of DsxM. *Hermaphrodite* (*her*) is another gene expressed in a non sex-specific manner and acts together with the Dsx and Ix, to control female sexual differentiation (Li and Baker, 1998). The expression of Pheromone binding protein gene is also regulated in sex-specific manner by Dsx proteins.

Another known target gene of Dsx proteins is the '*takeout'* gene, expression of which is activated by DsxM and Fru (product of *fruitless* gene), and is repressed by DsxF in a Fru independent manner (Dauwalder et al., 2002). DsxM and Fru together govern the male-specific development of neural circuitry, a separate lineage of sex determination pathway.

Several studies have shown the function of Dsx protein of different insect species through it's over expression in transgenic background or through RNAi mediated knock-down of the *dsx* transcripts. Ectopic expression of female-specific *M. domestica* Doublesex protein, MdDsxF in the males of *Musca* and *Drosophila* activated their *vitellogenin* gene expression, whereas the ectopic expression of male-specific *M. domestica* Dsx protein, MdDsxM in *Drosophila* females produced male-like pigmentation of posterior tergites, suggesting that sex-specific splice forms of *Musca dsx* are not only structurally but also functionally conserved (Hediger et al., 2004).

The dsRNA mediated knock-down of female-specific Dsx protein in fruit fly *Bactrocera dorsalis*, hinders the yolk protein gene expression and affects its reproductive ability (Chen et al., 2008). Partial masculinisation of both somatic and germline tissues was observed on ectopic expression of male-specific Dsx protein of *Ceratitis capitata*, CcDsxM in *Drosophila* female (XX) flies (Saccone et al., 2008). Male-like pigmentation of the posterior tergites was observed in *Drosophila* female transgenic flies expressing CcDsxM. This pigmentation was more effective compared to that in the flies expressing MdDsxM (*M. domestica* male Dsx protein), probably because of higher sequence similarity of CcDsxM to DmDsxM and phylogenetically closer relationship of *Ceratitis* to *Drosophila* (both belong to Acalyptratae), than to *Musca* (Calyptratae) (Saccone et al., 2008).

Functional conservation of sex-specific *Anastrepha* Dsx proteins, DsxF and DsxM was established through over-expression of these proteins in *dsx* intersexual *Drosophila* flies (Alvarez et al., 2009). The sexual differentiation of *dsx* intersexual *Drosophila* flies followed female or male pathway on the expression of *Anastrepha* Dsx proteins, DsxF and DsxM, respectively. However, this transformation was not complete and the effect of heterologous protein expression on the differentiation of sexually dimorphic structures was varying. The expression of *Drosophila yolk protein* genes was also affected by the sex-specific *Anastrepha dsx* proteins similar to that affected by the sex-specific Dsx protein of *Drosophila* itself (Alvarez et al., 2009).

Apart from Doublesex protein of dipteran insects, studies have also been conducted to functionally characterize the sex-specific Dsx proteins of *B. mori*. Three different genes of *B. mori* i.e., *vitellogenin*, *Storage protein* (*SP1*) and *pheromone-binding protein* (*PBP*) express in sex-specific manner. *Vitellogenin* gene of *B. mori* is expressed in the fat body cells of

female during the larval-pupal ecdysis and the protein is released into the hemolymph subsequently to be taken up by the developing oocytes (Mine et al., 1983; Yano et al., 1994). SP1, one of the plasma proteins termed hexamerins, accumulates in the hemolymph of female larva of *B. mori* (Mine et al., 1983) and its expression is sex-specifically regulated at the level of transcription in the fat bodies of female (Izumi et al., 1988). The PBP is predominantly expressed in male antennae and accumulates in the pheromone-sensitive sensilla (Steinbrecht et al., 1995). PBP binds to the sex-attractant pheromone released by female moths of *B. mori* and conveys the signal to a membrane-bound receptor on a nerve cell (Pelosi and Maida, 1995).

Studies have shown that the ectopic expression of *Bmdsxf* from a ubiquitous promoter in the males of *B. mori*, activated the expression of female-specific genes, *vitellogenin* and *Sp1* whereas *PBP* gene was repressed (Suzuki et al., 2003). Gel mobility shift assays revealed the binding of BmDsx proteins to the enhancer sequence (ACATTGT) located between -95 and -89 nt relative to the transcriptional start site of the *vitellogenin* gene (Suzuki et al., 2003). Ectopic expression of *Bmdsxm* in the females of *B. mori* leads to the abnormal differentiation of some female-specific genital organs and caused partial male differentiation in female genitalia (Suzuki et al., 2005). The transgenic female expressing BmDsxM showed increased expression of PBP and decreased expression of *vitellogenin* (Suzuki et al., 2005).

To establish the functional conservation of *Antheraea assama* Dsx proteins and to elucidate the role of the two female-specific proteins that I have identified, (described in the preceding chapter) in the sexual differentiation of *A. assama*, dsRNA mediated knock-down of *Aadsx* transcripts was done and the expression of their downstream target genes was analyzed.

3.2 Materials and Methods

3.2.1 Target regions and double stranded RNA (dsRNA) synthesis

Three different regions of the *Aadsx* transcripts i.e., exon 1 (T1), 15bp region of exon 2 (T2) (present in splice forms *Aadsxf1*, *Aadsxf3* and *Aadsxf5*) and exon 6 (T3) (**Fig. 3.2**) were selected and amplified using the specific primers for these regions. Primer combinations to generate amplicons specific to T1, T2 and T3 regions of *Aadsx* transcripts are F1-R1, F2-R2 and F3-R3, respectively (**Table 3.1**). The amplicons were cloned in pCR®II-TOPO® cloning vector (Invitrogen, USA) harbouring dual promoter (T7 and Sp6). Orientations of cloned products were confirmed by sequencing. Two different primer combinations, one vector-specific (M13F or M13R) and the other gene-specific were taken to amplify the cloned *Aadsx* fragment including either of the promoter sequences (T7 or Sp6) present in the vector. In-vitro transcription was done with T7 and Sp6 Megascript kits as prescribed by the manufacturer (Ambion, USA) using template flanking T7 or Sp6 promoters. DNAse treatment was given to remove the DNA template from the in-vitro synthesized RNA followed by purification of the product by Trizol treatment (Invitrogen, USA) and isopropanol precipitation. Both the sense and antisense single stranded RNAs were dissolved in DEPC treated water, combined in equimolar amounts in 1X insect buffer saline (IBS, composition: NaCl-160 mM, KCl-10 mM, $CaCl_2$-4 mM) and annealed by heating to 95^{o}C and slow cooling overnight at room temperature. Similarly dsRNA specific to full length green fluorescent protein (GFP) was used as a non specific control. The dsRNA formation was confirmed by agarose gel electrophoresis and its concentration was determined spectrophotometrically. The annealed dsRNA samples were stored in -70^{o}C till further use.

3.2.2 Injections of dsRNA and Knock-down of *Aadsx* gene

70µg of dsRNA per larvae was injected into a set of 16 (8 males and 8 females) fifth instar first day larvae of *A. assama*, at their fourth abdominal legs using insulin syringe. GFP-dsRNA (non-specific dsRNA) or saline injected larvae (n=15) of the same developmental stage were maintained as experimental controls. The injected and the un-injected control larvae were reared on their natural host plants. Six days post injection, eight (4 males and 4 females) larvae from each group of injections were dissected and their gonads were observed under microscope using uniform scale. The fatbodies of the dissected larvae were collected

separately in eppendorf tubes in 'RNA-later' (Ambion, USA) and stored at-70^{o}C till the RNA isolation.

3.2.3 Primer designing

The primers for '*A. assama hexamerin* gene' were designed based on the template obtained (Unigene_Aa00258) from the EST database of *A. assama* (http://www.cdfd.org.in/wildsilkbase/home.php) through tblastn using Hexamerin protein (GI: 100134931) of *B. mori* as a query sequence. For designing of *A. assama vitellogenin* gene primers the nucleotide sequences of *vitellogenin* gene of six different lepidopteran insects (*Bombyx mori* - GI: 60391273, *B. mandarina* - GI: 32526657, *Antheraea yamamai* - GI: 123299275, *A. pernyi* - GI: 152002197, *Samia cynthia ricini* - GI: 12862882, *S.C. pryeri* - GI: 61651633) were aligned using the online program of ClustalW (http://www.ebi.ac.uk/Tools/clustalw2/index.html), and primers (Vitellogenin F2 and Vitellogenin R2) were designed from the region showing maximum conservation.

3.2.3 RNA isolation and RT-PCR

RNA from the stored fatbody tissues harvested from the injected and control larvae (4 males and 4 females) was isolated using Trizol method (Invitrogen, USA), according to the manufacturers' protocol. First strand cDNA synthesis by reverse transcription for any particular RNA sample was primed by oligo-T of 17-mers. Master mix containing 4µl 5X reverse transcriptase buffer, 2µl DTT (0.1 M), 2µl of all 4dNTP mix (250 µM), .5µl RNAse inhibitor and 1µl of SuperScript III (200 U/µl; Invitrogen, USA) was mixed with the heat denatured (at 75^{o}C) and snap-frozen RNA (1.4µg) containing 2µl oligo T (100 pmol/µl) and H_2O to 20µl and incubated at 42°C for 60 minutes. The reverse transcriptase was inactivated by incubation for 10 minutes at 70°C. The synthesized cDNA was used immediately for PCR amplification or stored in aliquots at -20°C for later use.

Primers for amplifying specific targets, their annealing temperatures (for RT-PCR) and the expected amplicon sizes are listed in **Table 3.1**. All the reactions were carried out in an Eppendorf PCR master cycler (Eppendorf, USA), using conditions- 94^{o}C for 2 minutes, 40 cycles of 94°C for 30 seconds, annealing temperature for 30 seconds, 72°C for 2 minutes and 1 cycle of 72°C for 10 minutes. The PCR reactions were run on 1.5% Agarose gels in TAE buffer or stored at -20°C till further use. PCR reaction components include 4.3µl H_2O,

1µl 10X PCR reaction buffer, 1µl dNTP mix (25 µM each), 1µl (5 pm/µl) of 5' and 3' primers (for specific targets), 1µl cDNA and 0.1 µl *Taq* DNA polymerase (5 U/µl). Reactions with RNA as templates were used as negative controls.

Table 3.1: Gene-specific primers used in the knock-down studies

Serial No.	Primer name	Primer sequence	Annealing temp (^{o}C)	Amplicon size (bp)
1	AaD F (**F1**)	5'-TAGATTGAAGGTCGAGCTGAAAGGCC-3'	58	574
2	Aad_Fem1 R (**R1**)	5'-GACATTCAAATTATGTTTTCTCGCG-3'		
3	RNAi F (**F2**)	5'-GACACGCCAGTACGGACTTTAATAG-3'	54	116
4	Aad_Fem6R (**R2**)	5'-ACACTGACACTAAACACATG-3'		
5	Aad_F7 (**F3**)	5'-ACTGTACGAGCATTCGCTGTGC-3'	59	175
6	Aad_3'm4 (**R3**)	5'-ATCGCATCGCTACAAGTGCG-3'		
7	Hex 4F	5'-GGCAAGAGGGCAAGATGAGG-3'	60	2046
8	Hex 4R	5'- GTAGCGATCCAGAGGGAAGC-3'		
9	Vitellogenin F2	5'-ATCCAAGGTGAAGAAGCTGC-3'	57	605
10	Vitellogenin R2	5'-ATTTCAGCCATTGCCTGCATC-3'		

3.3 Results

3.3.1 Functional analysis of *Aadsx* gene by RNA interference

In order to establish the functional conservation of *Aadsx* gene, the dsRNA specific to different regions (**Fig. 3.2**) of *Aadsx* transcripts were administered into the haemocoel of *A. assama* larvae as described in the experimental section 3.2.2. Four different parameters were checked for the effectiveness of RNAi experiments: a) Absence of target transcript/s, b) Expression profile of *Aadsx* downstream target genes, c) Change in the anatomy and morphology of sex organs and secondary sexual characters, and d) Mating, egg laying and hatching of the eggs. To check the downstream effect of *Aadsx* transcript/s knock-down, the expression of two *A. assama* genes i.e., *vitellogenin* and *hexamerin* were analyzed by semi-quantitative RT-PCR experiments, since the expression of these two genes are sex- and organ-specific (**Fig. 3.1**). The RT-PCR of the cDNA made from female fatbody RNA of experimental larvae showed complete abolishment of the target transcripts (**Fig. 3.3**) whereas in male organs *Aadsx* transcript remained unaffected (**Fig. 3.3**). Though, no significant morphological defect was observed in the moths eclosed from the RNAi-treated larvae, at molecular level expression of *vitellogenin* and *hexamerin* gene was completely abolished in all the experimental female larvae (**Fig. 3.3**). The gonads of the dsRNA injected female larvae were deformed and shrunken compared to the gonads of the control female larvae (**Fig. 3.4**). Though there was no knock-down effect seen at molecular level in the treated males they also showed slightly deformed gonads in all the cases where the *Aadsxm* was targeted.

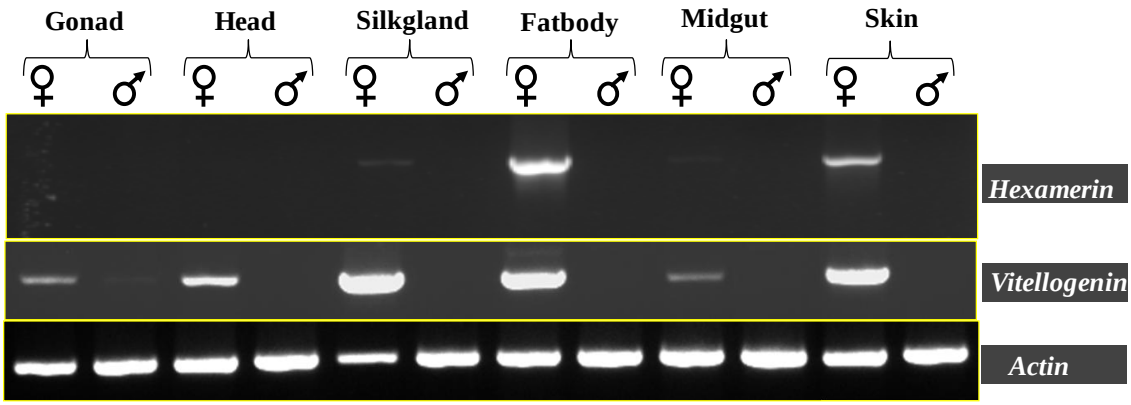


Fig. 3.1: Expression profiles of *hexamerin* and *vitellogenin* genes in different tissues at day 2 of the V instar larva. Both these genes are female-specific. *AaActin* is the loading control.

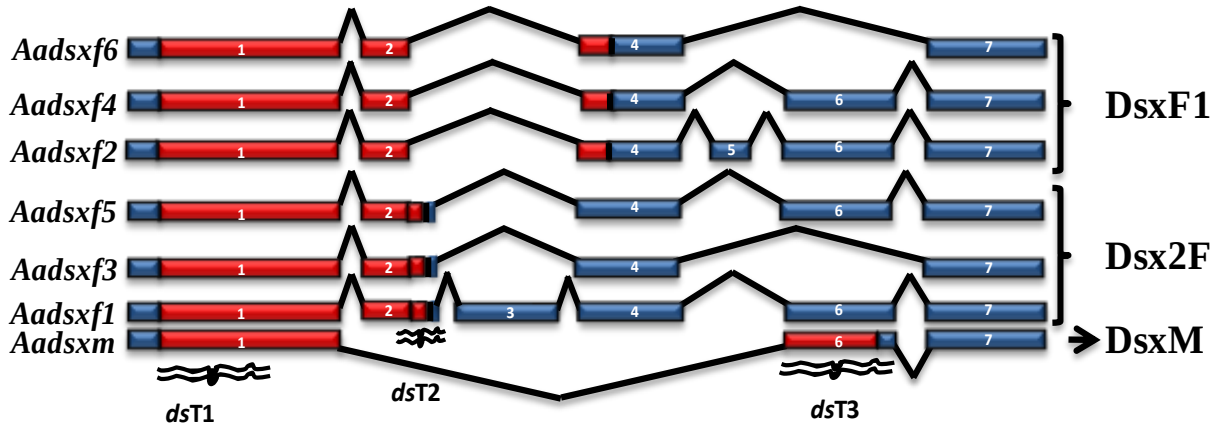


Fig. 3.2: Schematic representation of *Aadsx* transcripts and the regions chosen for making dsRNA for knock-down experiments. *Aadsx* splice forms are mentioned on the left side of the figure whereas the protein coded by them is mentioned on the right side. T1, T2 and T3 are the regions for which dsRNA was made.

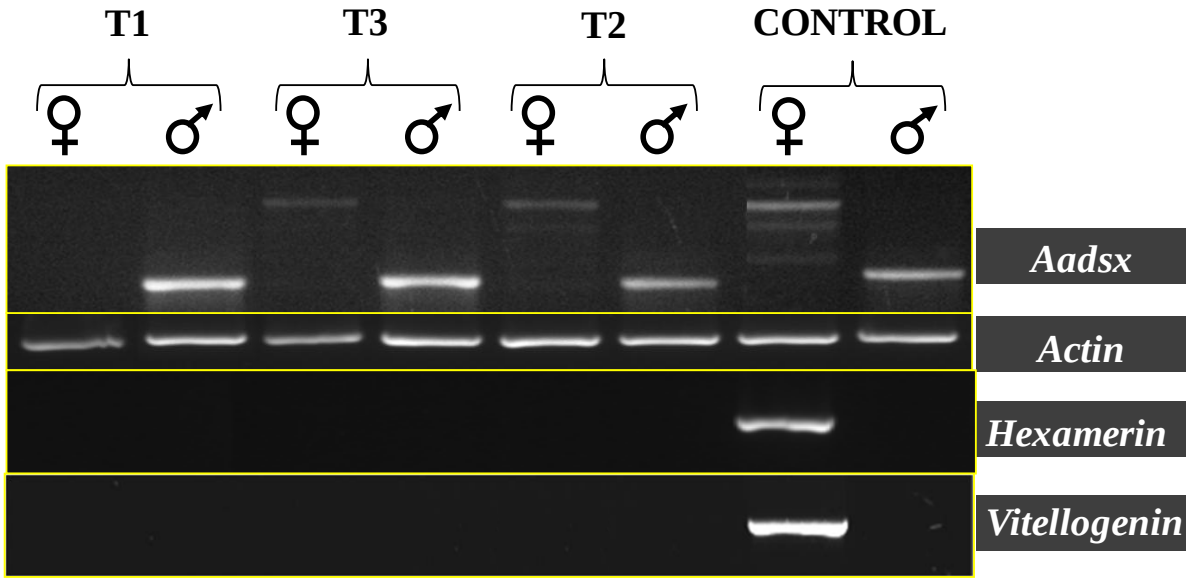
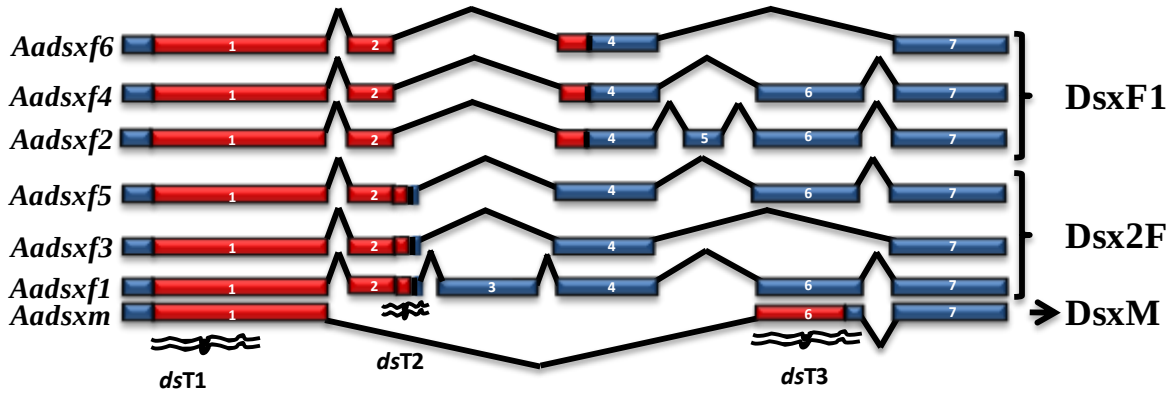


Fig. 3.3: Expression profile of *Aadsx*, *hexamerin* and *vitellogenin* genes in the fatbody tissue of knock-down and control larvae. Complete abolishment of *vitellogenin* and *hexamerin* expression is observed in females with the knock-down of *Aadsx* transcripts. Males were unaffected by the dsRNA injections. *A. assama* actin gene was used as a loading control.

A)



B)

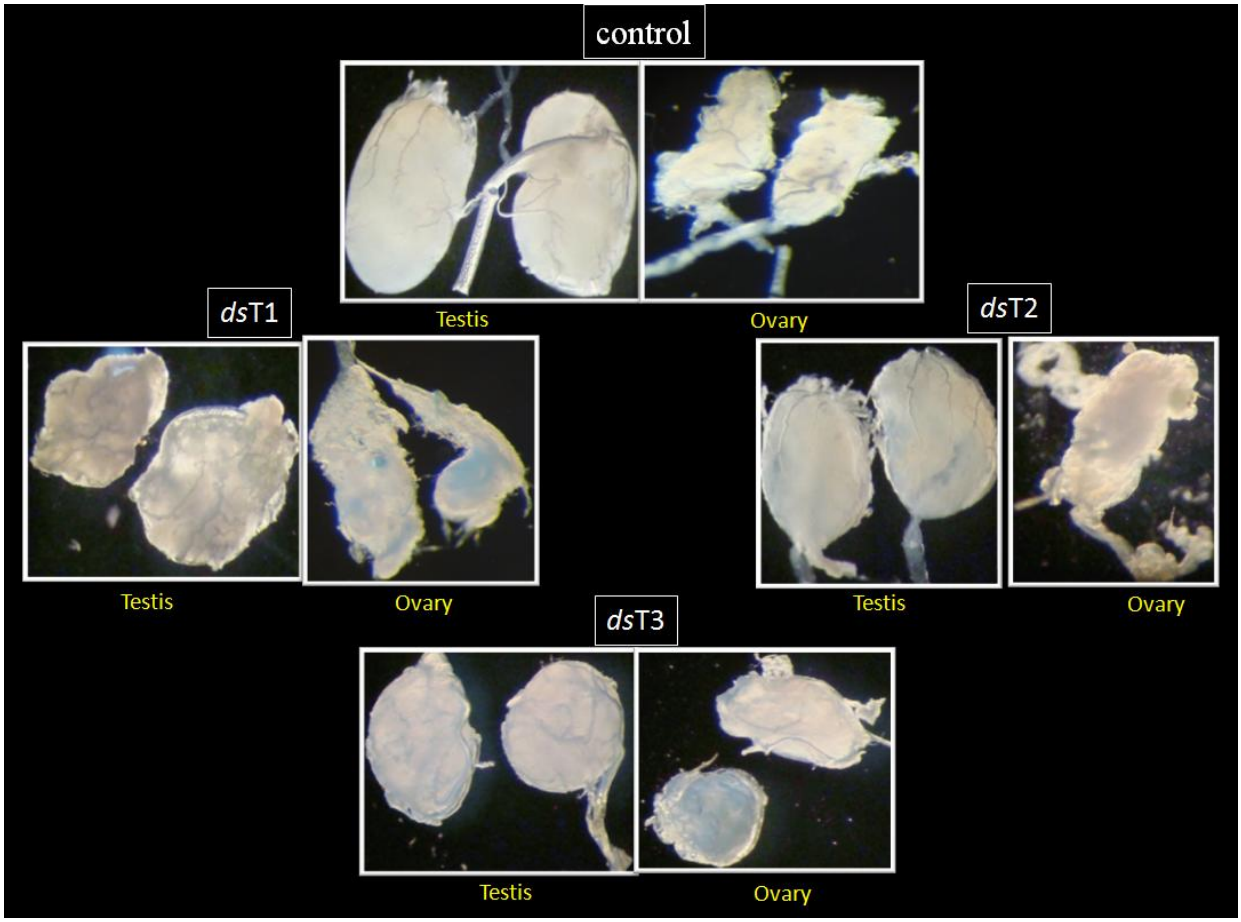


Fig. 3.4: A) Represents the *Aadsx* transcripts and the regions (T1, T2 and T3) selected for knock-down experiments. *Aadsx* splice forms are mentioned on the left and the corresponding protein is mentioned on the right side of the figure. **B)** Effect of *Aadsx* dsRNA injections on the larval gonads. Control gonads are from the larvae injected with dsGFP.

3.3.2 *Aadsx* knock-down effect in the next generation offspring

To see the effect of *Aadsx* knock-down in the next generation offspring, matings between female and male moths derived from larvae injected with *Aadsx* targets were performed; T1 female with T1 male, T2 female with T2 male and T3 female with T3 male and control females with control males. The numbers of eggs laid by the knock-down females mated with knock down males reduced drastically in all the *Aadsx* dsRNA injections compared to that of control females. dsGFP and saline injected larvae gave the same number of eggs as that of un-injected larvae. The eggs laid in different crosses initially looked like normal eggs, but in the later stages the eggs laid from the *Aadsx* knock-down females did not hatch. The un-injected females mated with *Aadsx*-dsRNA injected males recorded normal fecundity and hatchability (**Table 3.2**)

Table 3.2: Effect of dsRNA injections on egg laying and their hatching

Injection type	No. of eggs laid	No. of eggs hatched
T1	≈16	No hatching
T2	≈19	No hatching
T3	≈15	No hatching
CONTROL (dsGFP)	≈170	170 (all)
Un-injected	≈170	170 (all)

3.4 Discussion

The *Aadsx* gene is transcribed in all the organs throughout the larval development. The *Aadsx* gene produces sex-specific transcripts; six in female and one in male. Conceptual translation of these transcripts indicates the production of two female-specific and one male-specific Dsx proteins. In spite of the functional conservation of DM domain containing proteins among metazoans, the existence of multiple splice forms of *Aadsx* raises the question of their functional significance. Apart from that, female-specific transcripts having two different ORFs suggest that two AaDsx proteins may be essential for the female sexual differentiation. The function of *Aadsx in-vivo* was tested by transiently knocking down this using RNA interference (RNAi) technique. *Aadsx* transcripts were completely abolished in fatbody of dsRNA (for the common region) injected female larvae. The deformed gonads of the injected female larvae and the complete abolishment of the sex-specific expression of *vitellogenin* and *hexamerin* genes underlies the importance of *Aadsx* as a terminal regulatory gene in the hierarchy of regulatory genes controlling the sexual differentiation of *A. assama*. The failure of knock-down in male larvae may be because of the secondary structure of male-specific transcript *Aadsxm*, which does not allow the dicer components to access the transcript. Slightly deformed gonads in the dsRNA treated males may be attributed to partial knock-down of *Aadsxm* transcript in dsRNA (against *Aadsxm*) injected males because of which, the expression of *Aadsxm* in injected males was indistinguishable compared to the expression of *Aadsxm* in controls.

The effect of knock-down seen at the molecular level was consistent with the complete failure of the egg hatching in the treated batches; no visible morphological changes were noticed. It may be because of the involvement of some other factors in the sexual differentiation of the organism, governed by the common upstream signal but making a separate path other than the Dsx, to regulate all aspects of sexual differentiation. The knock-down of the transcripts having 15bp additional sequence (contributing to the change in the ORF) also gave the same results as that of the knock-down of all the *Aadsx* transcripts together. This raises two possibilities: 1) the group of transcripts (*Aadsxf2*, *Aadsxf4* and *Aadsxf6*) coding for the longer protein (AaDsF1) is not having any function so that the effect of knock-down of all the transcript at a time or only the knock-down of transcripts of second group (*Aadsxf1*, *Aadsxf3* and *Aadsxf5*) produces the similar results or 2) both the female-

specific AaDsx proteins require the presence of each other for the process of sexual differentiation. In view of the first possibility if the female AaDsx conceptual proteins are compared with the female-specific BmDsx protein, the differential portion of AaDsx protein coded by the longer ORF is present in female-specific BmDsx protein as well, which validates the genuineness of the AaDsxF1 protein. The second possibility could be more reasonable considering the existence of two female-specific transcripts having two different ORFs in another wild silkmoth, *A. mylitta*, similar to that of *Aadsx*. Further functional characterization of these proteins and their binding to the regulatory sequences of the downstream genes may shed light on this aspect. The presence of two female-specific Dsx proteins (through conceptual translation of *Aadsx* and *Amdsx* transcripts), in wild silkmoths whereas only one, as reported in the domesticated silkmoth, *B. mori* prompted us to revisit *Bmdsx* splice forms.

Chapter - IV

Identification of a novel female splice form of *Bmdsx* from silkworm, *Bombyx mori*

4.1 Introduction

Bmdsx, a homologue of *doublesex* gene of *Drosophila* is the bottom most gene of the sex determination cascade in *B. mori* (Ohbayashi et al., 2001). *Bmdsx* plays a very crucial role in silkworm sexual development. The pre-mRNA of *Bmdsx* gene sex-specifically splices to yield male-specific and female-specific mRNAs that encode male-specific (BmDsxM) and female-specific (BmDsxF) polypeptides (Ohbayashi et al., 2001; Suzuki et al., 2001). *Bmdsx* transcripts are expressed throughout the development in various tissues at larval, pupal, and adult stages in the silkworm. The pre-mRNA of *Bmdsx* gene is composed of 6 exons, the open reading frame (ORF) of male-specific transcript contains exons 1, 2 and 5 and the ORF of female specific transcript contains exons 1,2,3,4 and 5 (**Fig. 4.1**). *Bmdsx* pre-mRNA represents default mode of splicing in female and the female specific exons are devoid of putative Tra/Tra-2-binding sites (Suzuki et al., 2001). Recently it has been shown that the binding of BmPSI, a *Bombyx* homolog of P-element somatic inhibitor (PSI), to the exonic splicing silencer sequences (ESS) in exon 4 is required for the skipping of exon 3 and exon 4, thus maintaining proper male-specific splicing of *Bmdsx* pre-mRNA (Suzuki et al., 2008).

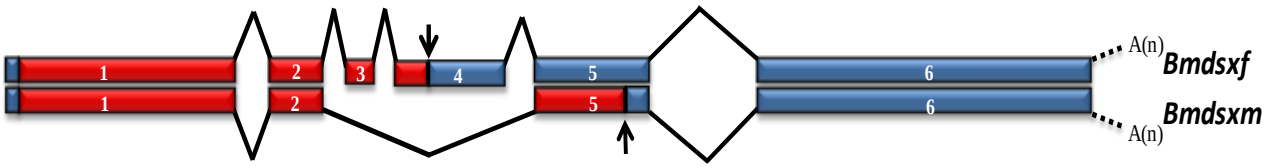


Fig. 4.1 Splice forms of *Bmdsx*. Boxes are exons and lines are introns. Red coloured portion is the ORF whereas blue coloured regions are UTRs. The *Bmdsx* pre-mRNA splices to generate 2 different splice products, one female-specific (*Bmdsxf*) and one male-specific (*Bmdsxm*). A(n) represents the poly-A tail and arrows represent the stop codon site.

In the present study, we identified a third splice form of *Bmdsx* gene. This isoform is female-specific (for the convenience we will refer to it as *Bmdsxf1*) and differs from the previously reported female-specific *dsx* isoform, *Bmdsxf*, by the presence of additional 15bp sequence (TACGGACTTTAATAG) towards the 3' end of exon 3. Due to the presence of this 15bp sequence the reading frame of the *Bmdsxf1* changes and the protein produced (through virtual translation) is short and contains only 246aa as compared to BmdsxF protein which is of 264aa.

4.2 Materials and Methods

4.2.1 Identification of *Bmdsx* EST having additional 15bp sequence (TACGGACTTTAATAG), from the database

The 15bp sequence (TACGGACTTTAATAG) stretch, which is a part of *A. assama* and *A. mylitta* female specific *doublesex* isoforms and responsible for generation of second type of female-specific Dsx protein in these silkmoths (information in chapter 2) was searched in the EST database (http://www.ncbi.nlm.nih.gov/dbEST) of *B. mori*, using the BLAST 2.1 program (http://www.ncbi.nlm.nih.gov/blast/). Virtual translation and alignment with the previously reported *Bmdsx* transcript confirmed that the obtained EST is a genuine *Bmdsx* transcript. Alignment of this EST sequence with *Bmdsx* genomic DNA sequence was done to find the exact location of this 15bp additional sequence, with respect to its positioning in the *Bmdsx* gene.

4.2.2 Designing of primers to confirm the presence of two female specific transcripts of *Bmdsx* gene

Forward and reverse primers were designed in the regions flanking the 15bp sequence which can amplify both of the female specific transcripts, one with the inclusion of 15bp and the other without this sequence. To amplify either of the female specific transcripts, keeping the forward primer constant, reverse primers were designed in such a manner that it can bind to specific female transcript. The primers used in this study, their annealing temperature and amplicon size they generate are listed in **Table 4.1**.

Table 4.1: Gene-specific primers used in the expression analysis of *Bmdsx*

Serial No.	Primer name	Primer sequence	Annealing temp (^{o}C)	Amplicon size (bp)
1	BdR2F (**F**)	5’-GCAGCGACTTGGATGAGG-3’	60	172 187
2	BmDsxR (**R1**)	5’-TTCATTGACGAAGACAGTACACC-3’		
3	BdR2F (**F**)	5’-GCAGCGACTTGGATGAGG-3’	60	131
4	BdR1R (**R2**)	5’-CAGCATTTTCTGGCGTGTC-3’		
5	BdR2F (**F**)	5’-GCAGCGACTTGGATGAGG-3’	60	150
6	BdR2R (**R3**)	5’-TTTCCAGCATTTTCTATTAAAGTCC-3’		

4.2.3 Insect species, RNA isolation and RT-PCR

A silkworm stock Pure Mysore (sex-limited), where male and female larvae could be distinguished by virtue of sex-limited larval markings (Nagaraja et al., 2005) was used as the source of RNA. The female larvae possess markings in the cephalic and thoracic segments while the male larvae lack the same and thus sex separation is made possible from early IV instar (**Fig. 4.2**). Tissues were dissected out from 5^{th} instar larvae, washed in PBS (135 mM NaCl, 2.7 mM KCl, 4.3 mM Na_2HPO4 and 1.4 mM KH_2PO_4), frozen immediately in liquid nitrogen and stored at -80^{o}C, till the time of RNA isolation. Total RNA was isolated from each tissue sample using TRIZOL reagent (Invitrogen, USA), according to manufacturer's protocol. DNAse treated total RNA was denatured at 75^{o}C for 10 minutes and immediately chilled on ice. First strand cDNA was synthesized for each RNA sample using SuperScript III reverse transcriptase (Invitrogen, USA), according to manufacturer's instructions, using 17-mers poly-T as primer. PCR was performed for each cDNA sample using Taq DNA polymerase (Fermentas), with the gene specific primers listed in **Table 4.1**. The PCR conditions were: 94^{o}C for 2 minutes, then 40 cycles of 94^{o}C for 30 seconds, 60^{o}C for 30 seconds and 72^{o}C for 30 seconds, followed by final extension at 72^{o}C for 10 minutes. PCR amplicons were analyzed by 2% agarose (50% Metaphor and 50% normal agarose) gel electrophoresis and the amplicons were confirmed by DNA sequencing.

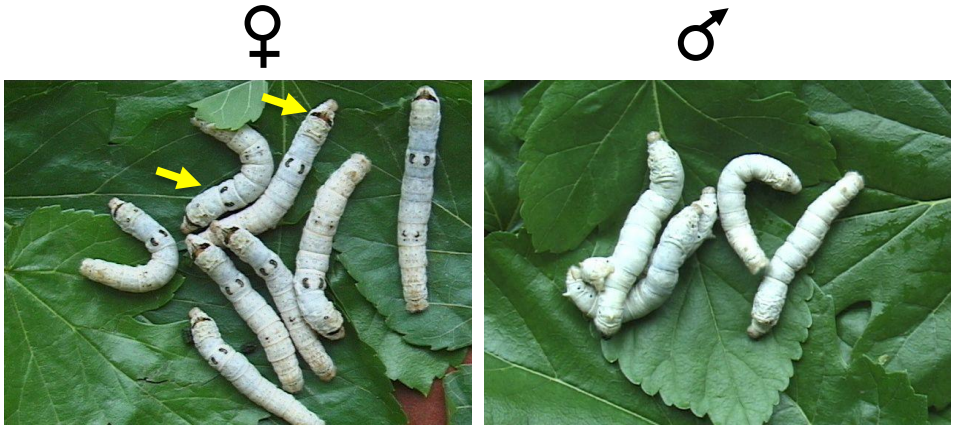


Fig. 4.2 Pure Mysore (sex limited) larvae. Female larvae possess crescent-shaped markings in the cephalic and thoracic segments (indicated by yellow arrows) while the male larvae are devoid of such markings.

4.3 Results

4.3.1 Identification of EST- BP121180

While looking for the 15bp conserved sequence of *A. assama* and *A. mylitta* female specific *dsx* forms in the NCBI EST database (http://www.ncbi.nlm.nih.gov/dbEST/), a single EST (BP121180) from *B. mori* showed the presence of this sequence in a stretch, positioned after 795 bp relative to the earlier reported female specific *Bmdsx* mRNA. Alignment of *Bmdsxf* ORF sequence with EST- BP121180 is shown in **Fig. 4.3**. The only difference between *Bmdsxf* and EST- BP121180 is the absence of 15bp sequence (TACGGACTTTAATAG) in the former. After aligning this EST sequence with *Bmdsx* mRNA and with *Bmdsx* genomic DNA sequence we found this sequence to be present immediately after exon 3 without any intronic sequence in between (**Fig. 4.4**). The presence of this 15bp sequence alters the reading frame of the female-specific *dsx* transcript to produce 18aa shorter protein as compared to the earlier reported BmDsxF. The new splice junction of exon 3 fulfils the GT-AG rule of splicing. Thus we believe that we have found a second BmDsxF in the silkworm, *B. mori*.

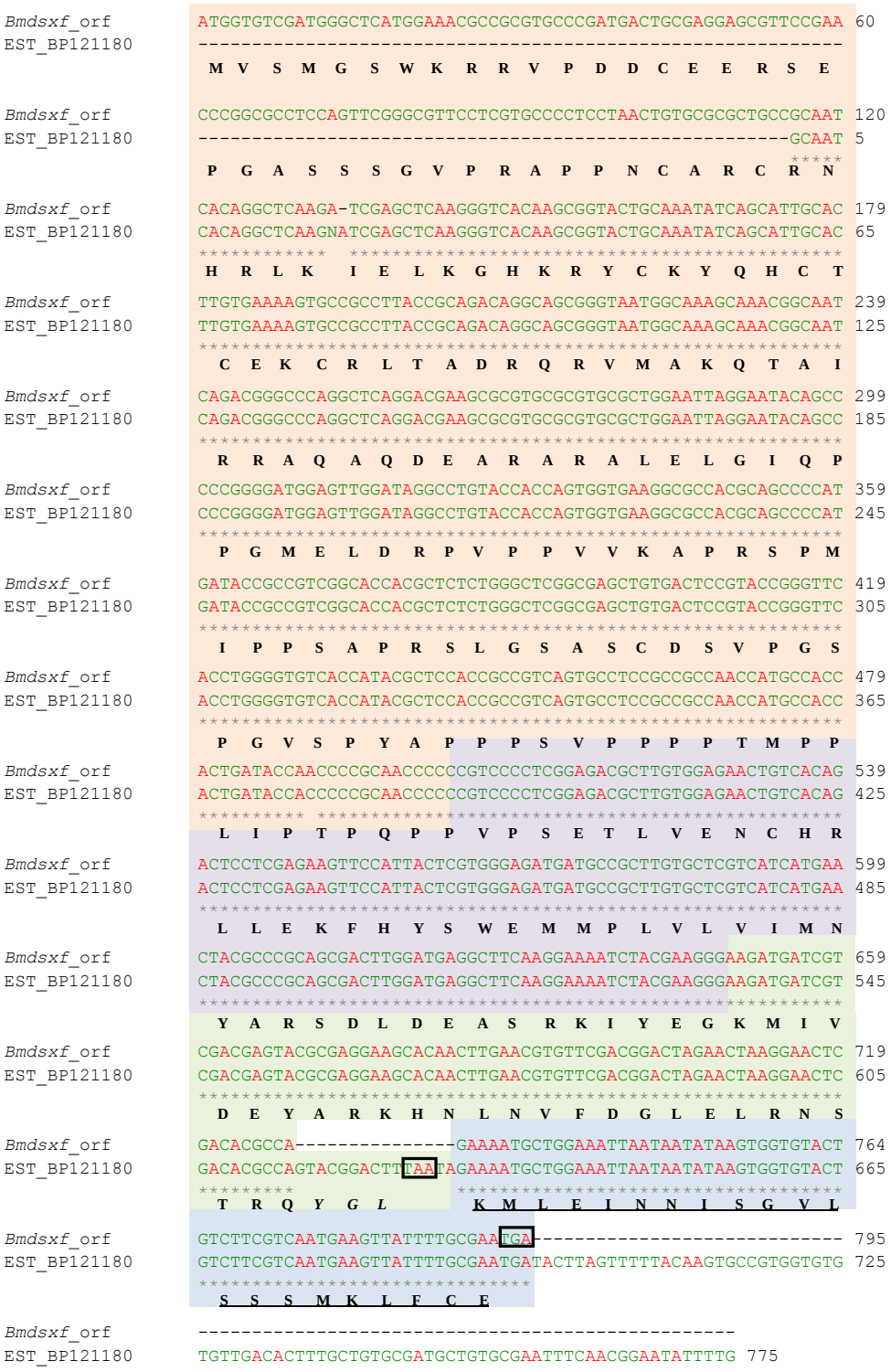


Fig. 4.3 ClustalW alignment of the *Bmdsxf*-ORF and EST-BP121180, showing the presence of additional 15bp sequence after exon 3 in the latter. The amino acids coded are given below the alignment; the amino acid in italics (*YGL*) is specific to *Bmdsxf1* and underlined amino acids are specific to BmDsxF. First four exons (shown with different colours in background) of *Bmdsx* gene are covered by ORF of *Bmdsxf*. The unshaded region after exon 3 shows the difference of 15bp sequence between *Bmdsxf* and EST-BP121180. Because of the presence of this 15bp sequence, the stop codon in the EST-BP121180 is located 66bp earlier compared to the stop codon of *Bmdsxf* splice form, as a result of which the new splice form of *Bmdsx* (named as *Bmdsxf1*) produces BmDsxF1 protein shorter by 18aa. The stop codons have been shown by the black boxes.

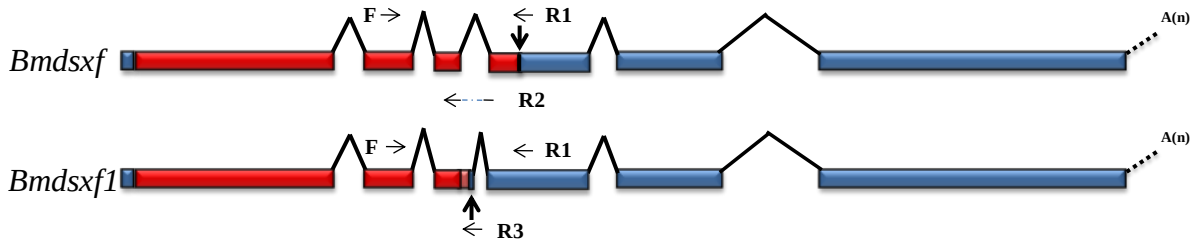


Fig. 4.4 Two female-specific splice forms of *Bmdsx* (*Bmdsxf* and *Bmdsxf1*). Boxes are exons and lines are introns; red colour represents the coding region whereas blue colour represents the UTR regions. *Bmdsxf* is the known and reported splice form. *Bmdsxf1* differs from *Bmdsxf* by the presence of 15bp sequence at the 3'end of exon3 (shown by a small box attached to exon 3). Stop codon of respective splice forms is shown by vertical arrows whereas primer positions are indicated by horizontal arrows. Priming region for primers F and R1 are present in both the female forms whereas the priming region for primers R2 and R3 is present in *Bmdsxf* and *Bmdsxf1* splice forms, respectively.

4.3.2 Expression analysis of female specific *Bmdsx* mRNA splice variants

We performed RT–PCR experiments to determine the expression patterns of the two female-specific *Bmdsx* splice variants, using one specific forward primer (F) and three reverse primers (R1, R2, and R3 as detailed in **Table 4.1**) (**Figs. 4.4 & 4.5**). Primers F and R2 amplified the female specific *Bmdsx* isoform, devoid of 15bp additional sequence (*Bmdsxf*), in a sex-specific manner whereas Primers F and R3 amplified female-specific *Bmdsx* transcript including 15bp sequence (*dsxf1*) in all the organs tested, of V instar larvae, but these primers also detected faint expression of these female specific splice variants in male tissues (**Fig. 4.6, B and C**). We confirmed that primers F and R1 amplified both the female specific *Bmdsx* isoforms in all the organs of 5^{th} instar larvae (**Fig. 4.6A**).

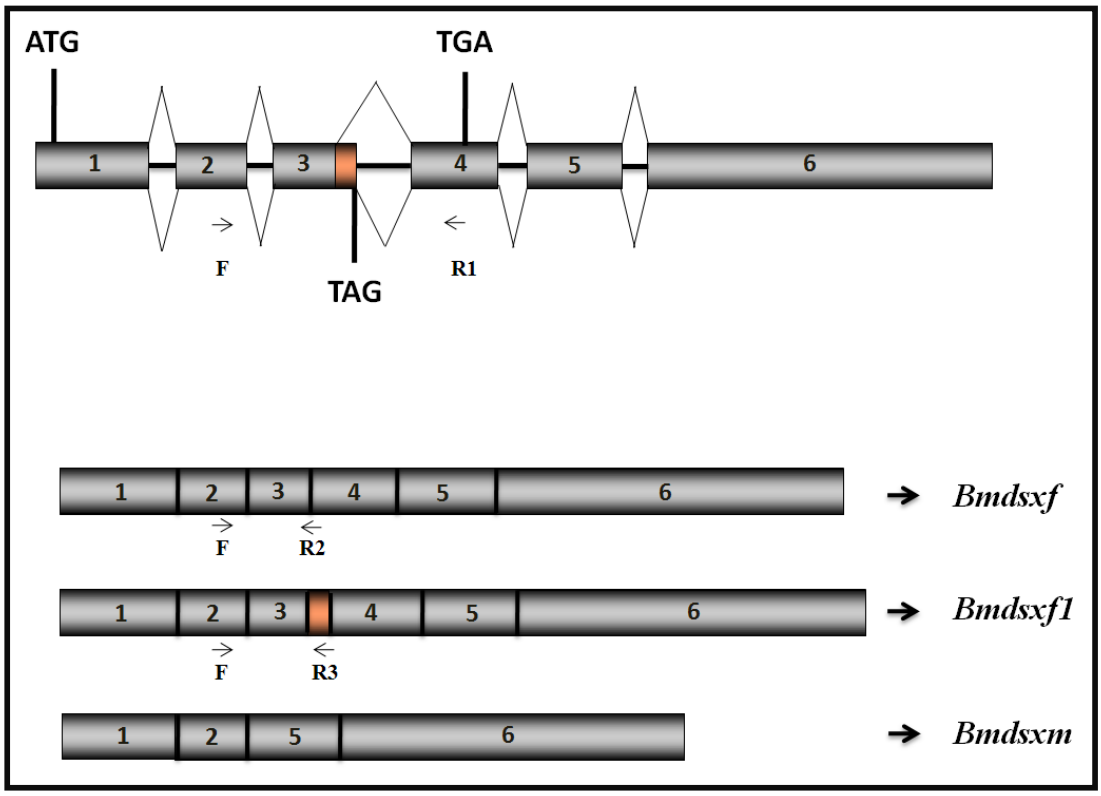


Fig. 4.5 Pre-mRNA of *Bmdsx* splices to produce three products i.e., *Bmdsxf*, *Bmdsxf1* and *Bmdsxm*. Boxes represent exons, whereas lines represent introns. Exons have been numbered. Arrows represent the primer positions used to confirm the presence of additional female splice form (*Bmdsxf1*) identified in the present study. The red coloured region in *Bmdsxf1* is the additional portion generated due to the alternative 5'splice-site selection.

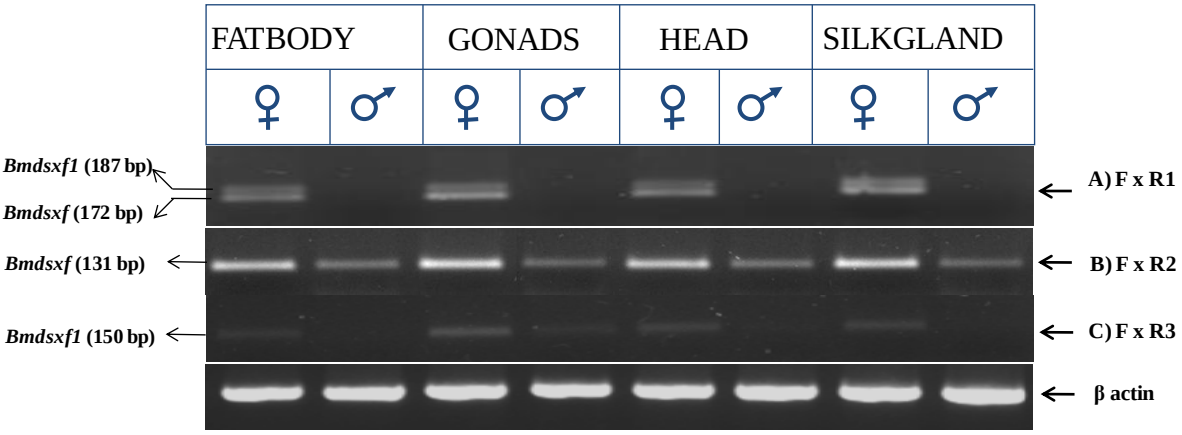


Fig. 4.6 Expression analysis of two female-specific splice forms of *Bmdsx* (*Bmdsxf* and *Bmdsxf1*). **A)** Amplification of both the female specific *Bmdsx* forms. **B)** Amplification of the *Bmdsxf* splice form. **C)** Amplification of the *Bmdsxf1* splice form. *Bmactin* (GenBank accession no. XM_623046) is the loading control. Male has not been included in the panel 'A'. Faint amplification in male panel could also be seen in B and C panels. Similar result for *Bmdsx* has also been reported earlier (Ohbayashi et al., 2001).

4.3.3 Proteins encoded by *Bmdsxf* and *Bmdsxf1* splice forms

The *Bmdsxf1* encoded protein (BmDsxF1) contains both OD1 (DM) and OD2 (Oligomerization) domains and is exactly similar to the BmDsxF protein reported earlier except for the difference in their extreme C-terminal regions. The inclusion of 15bp sequence at the end of 3rd exon in the *Bmdsxf1* splice form leads to change in the reading frame after amino acid 243, relative to the BmDsxF protein and early occurrence of the stop codon compared to the *Bmdsxf* splice form. As a result, the putative protein BmDsxF1 is of 246 aa compared to BmDsxF protein which is of 264 aa. Both the female proteins (AaDsxF and AaDsxF1) differ at their C-termini by 21 aa (**Fig. 4.7**).

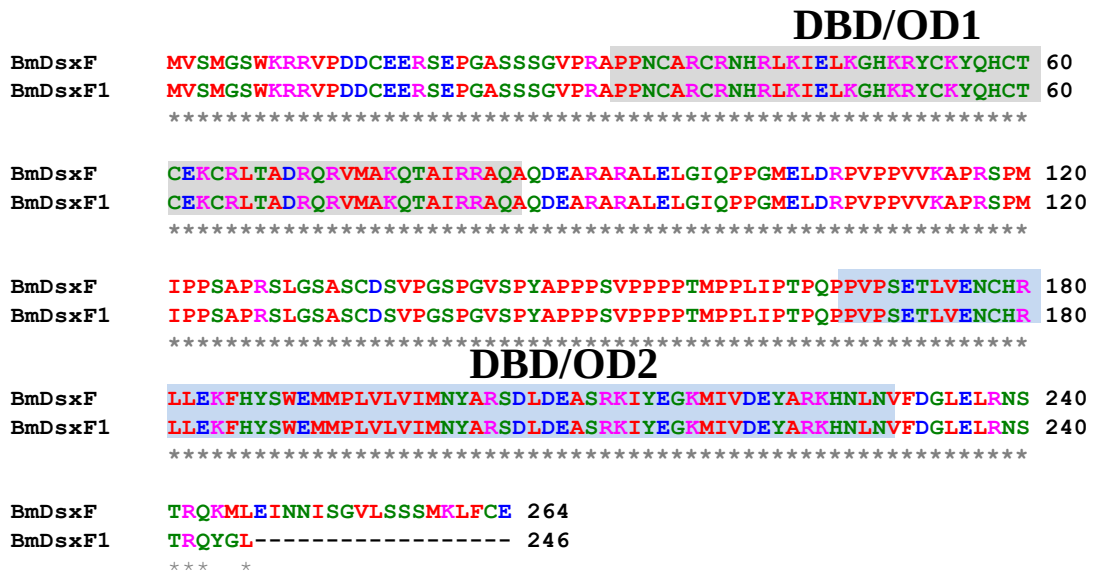


Fig. 4.7 ClustalW alignment of BmDsxF and BmDsxF1 proteins. BmDsxF1 is coded from female splice form of *Bmdsx* (*Bmdsxf1*) having 15bp insertion after exon3, relative to *Bmdsxf* transcript. Grey shaded region represents the DNA binding domain (OD1 domain) whereas blue shaded region represents the oligomerization domain (OD2 domain).

4.4 Discussion

Sex-specific BmDsx proteins, BmDsxF and BmDsxM are encoded by the transcripts produced by the alternative splice forms of *Bmdsx* pre-mRNA. We found a third splice variant, *Bmdsxf1* which is female-specific and the putative protein produced by this splice form is shorter than BmDsxF by 18 aa (**Fig. 4.7**).

In order to analyze the function of BmDsx proteins, Suzuki et al. constructed silkworm transgenic lines ectopically expressing $BmDsx^F$ or $Bmdsx^M$ (Suzuki et al., 2003; Suzuki et al., 2005). Though there was no change in the morphological characters of the male moths ectopically expressing $BmDsx^F$, they showed increased expression of *vitellogenin* (*vg*) and *Sp1* mRNAs compared to that in normal males (Suzuki et al., 2003). However, expression level of these genes was 20 fold and 200 fold less than those of normal female, respectively. On the other hand, female moths expressing transgenic $BmDsx^F$, showed enhanced expression of *vg* and *Sp1* which is consistent with the expected increase in the BmDsxF due to additional copy of the transgene $Bmdsx^f$ (two copies of endogenous *Bmdsxf* + single copy of transgene $Bmdsx^f$) (Suzuki et al., 2003). Transgenic males, having only single dose of $BmDsx^F$ encoded by $Bmdsx^f$ should show 50% expression of downstream target genes compared to normal females. But the increase was found only to be marginal. The reason attributed by authors for such a low level of expression of *vg* and *Sp1* in transgenic males expressing $BmDsx^F$ is the antagonistic effect of endogenous BmDsxM. But the effect of BmDsxM, in sexual differentiation of the organs and on the expression of *vg* and *Sp1* was more extensive in the transgenic females ectopically expressing $BmDsx^M$; *vg* expression was reduced by 25% as compared to that in normal females. In addition, female specific structures were repressed; instead certain male genital structures appeared in these transgenic females.

The less pronounced effect of ectopically expressed $BmDsx^F$ compared to ectopically expressed $BmDsx^M$ in the opposite sexes in regulating the sexual differentiation process and change in the morphology and expression levels of *vg*, *Sp1* and *PBP* genes, suggests the recruitment of additional factor by the BmDsxF protein to exert its effect fully on female sexual differentiation. Finding of a novel female-specific splice form of *dsx* in lepidopterans, which codes for another female specific putative protein, opens up a new avenue to unravel the sexual differentiation process governed by BmDsx proteins in the lepidopterans. A model

(**Fig. 4.7**) has been proposed to explain the pathway of female sexual differentiation governed by the two female-specific Dsx proteins.

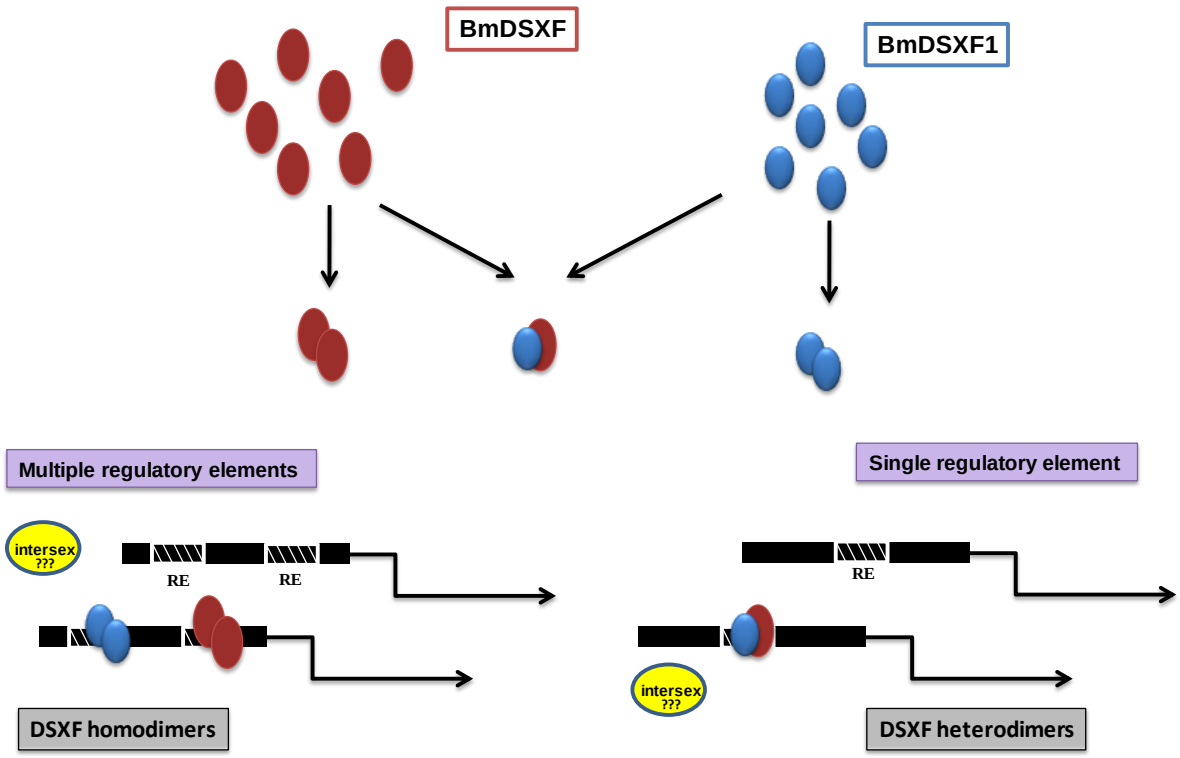


Fig. 4.7 Model explaining the mode of action of two Dsx proteins in females. **A**) Two different regulatory elements (RE) present in the promoters of downstream target genes (*vitellogenin*, *Sp1* and *PBP*). Homodimers of BmDsxF and BmDsxF1 bind seperatoly to these regulatory elements to independently regulate them. **B**) Alternatively, heterodimers of BmDsxF and BmDsxF1 bind to the single regulatory element present in the promoters of downstream target genes to regulate their expression.

Chapter - V

Concluding Remarks

Concluding remarks

Sex determination is the process of deciding the fate of the developing embryos to follow either the male or the female developmental pathway. These pathways which govern the differentiation of the "bipotential gonad", has been reasonably well characterized in *Drosophila melanogaster*.

Doublesex (*dsx*), the most downstream gene of the *Drosophila* sex determination cascade is the best studied example of sexual regulators conserved from worms (*mab3 in C. elegans*) to vertebrates (*Dmrt1*). The pre-mRNA of *dsx* splices to produce male- and female-specific transcripts which code for the male- and female-specific proteins, respectively. The characterization of the *dsx* homologues from different (many in the order Diptera, two in the order Hypmenoptera but only one in Lepidoptera) insect species led to the fact that the sex-specific splice forms of *dsx* pre-mRNA in all these species code for one male and one female-specific Dsx proteins, which regulate the downstream target genes responsible for the sex-specific characters.

In the present study we cloned and characterized the *dsx* homologues from two wild relatives (*Antheraea assama* and *Antheraea mylitta*) of the domesticated silkmoth *Bombyx mori*. Interestingly, we found the presence of multiple sex specific splice forms of *Aadsx* (in *A. assama*) and *Amdsx* (in *A. mylitta*) pre-mRNAs. On the basis of the proteins coded by these sex specific splice forms, they can be grouped into three categories: i) transcripts coding for the first type (DsxF1) of female specific protein ii) transcripts coding for the second type (DsxF2) of female specific protein, and iii) transcript coding for male specific protein (DsxM). Presence or absence of 15bp within the ORF of two groups of female-specific transcripts results in the production of two distinct female-specific Dsx proteins. The conservation of this 15bp sequence in all the female-splice forms of *Aadsx* and *Amdsx* and its absence in the previously reported splice forms of *Bombyx mori dsx* (*Bmdsx*), prompted us to reinvestigate the *Bmdsx* splice forms in the EST database of *B. mori*, which led to the identification and confirmation of existence of an another female splice form of *Bmdsx* (*Bmdsxf1*). Results obtained by the functional analysis of these female specific proteins in *A. assama*, suggest that both the female specific Dsx proteins are essential for the female sexual differentiation. At this stage we do not know whether the presence of two Dsx proteins in female is confined to silkmoths or is found in other insect orders as well. These

novel findings open up a new avenue to study the regulation of the downstream genes of sexual differentiation, which expands the area through which we observe the process of sex determination and sex differentiation governed by Dsx proteins in silkmoths (**Fig. 5.1**).

Apart from this, the conservation of the 4th exon sequence, especially the PESS sequence responsible for the sex-specific splicing of *Bmdsx* in the female specific transcripts of *Aadsx* and *Amdsx*, indicates the existence of common mechanism of sex-specific splicing of *dsx* homologues in silkmoths.

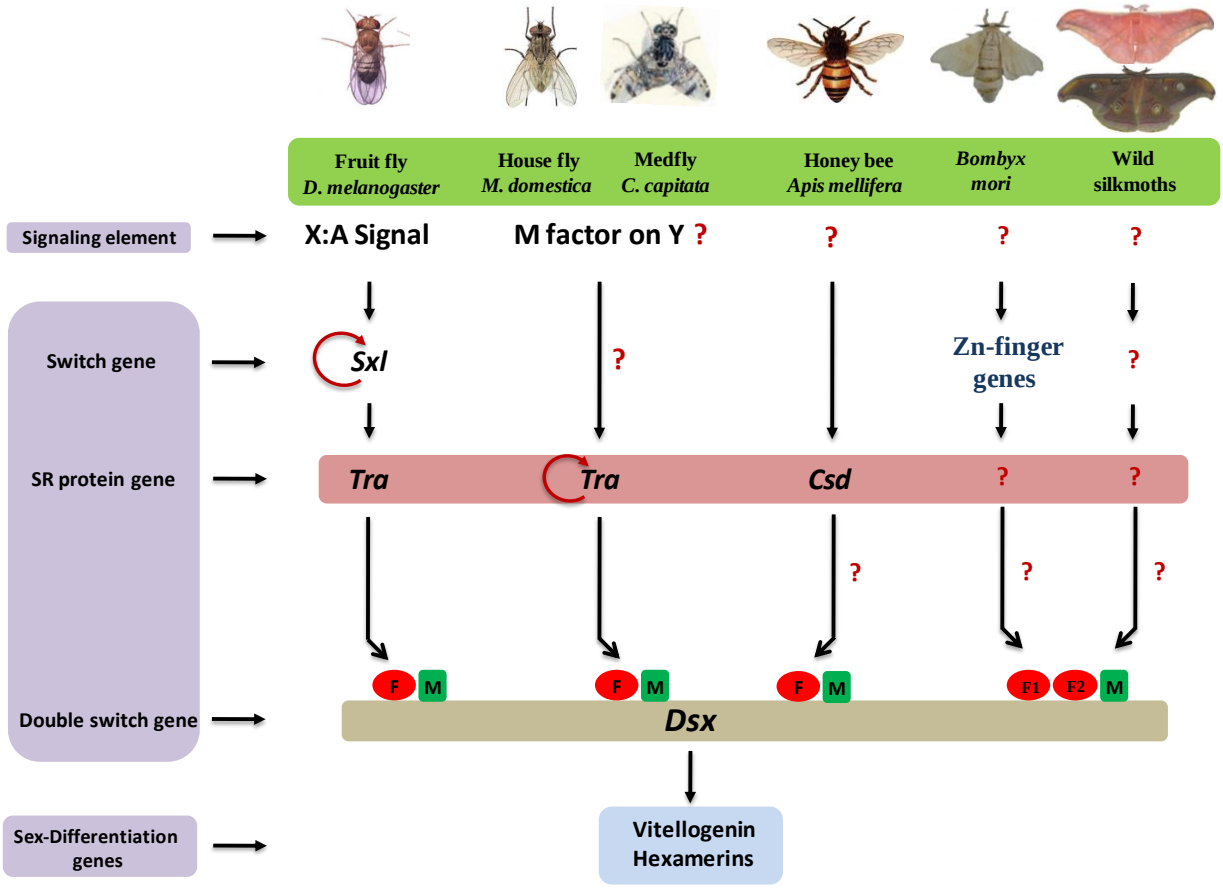


Fig. 5.1 Schematic diagram representing the players of sex determination cascade in the model insect species of Diptera (*D. melanogaster*, *M. domestica* and *C. capitata*), Hymenoptera (*A. mellifera*) and Lepidoptera (*B. mori*, *A. assama* and *A. mylitta*). Wild silkmoths include *A. assama* (upper) and *A. mylitta* (lower). Upstream genes of the sex determination cascade vary whereas *dsx*, the bottom most gene is conserved in all the insect species of different orders. In case of dipterans and hymenopterans, one male and one female specific Dsx proteins are produced. The study reported in the thesis concludes that lepidopterans produce two female specific (DsxF1 and DsxF2) and one male specific (DsxM) Dsx proteins. Both the female Dsx proteins are shown to be essential for the regulation of downstream sex differentiation genes.

References

Abe, H., Harada, T., Kanehara, M., Shimada, T., Ohbayashi, F. and Oshiki, T. (1998a). Genetic mapping of RAPD markers linked to the densonucleosis refractoriness gene, nsd-1, in the silkworm, Bombyx mori. *Genes Genet Syst* **73**, 237-42.

Abe, H., Kanehara, M., Terada, T., Ohbayashi, F., Shimada, T., Kawai, S., Suzuki, M., Sugasaki, T. and Oshiki, T. (1998b). Identification of novel random amplified polymorphic DNAs (RAPDs) on the W chromosome of the domesticated silkworm, Bombyx mori, and the wild silkworm, B. mandarina, and their retrotransposable element-related nucleotide sequences. *Genes Genet Syst* **73**, 243-54.

Abe, H., Ohbayashi, F., Shimada, T., Sugasaki, T., Kawai, S., Mita, K. and Oshiki, T. (2000a). Molecular structure of a novel gypsy-Ty3-like retrotransposon (Kabuki) and nested retrotransposable elements on the W chromosome of the silkworm Bombyx mori. *Mol Gen Genet* **263**, 916-24.

Abe, H., Ohbayashi, F., Shimada, T., Sugasaki, T., Kawai, S. and Oshiki, T. (1998c). A complete full-length non-LTR retrotransposon, BMC1, on the W chromosome of the silkworm, Bombyx mori. *Genes Genet Syst* **73**, 353-8.

Abe, H., Seki, M., Ohbayashi, F., Tanaka, N., Yamashita, J., Fujii, T., Yokoyama, T., Takahashi, M., Banno, Y., Sahara, K. et al. (2005). Partial deletions of the W chromosome due to reciprocal translocation in the silkworm Bombyx mori. *Insect Mol Biol* **14**, 339-52.

Abe, H., Sugasaki, T., Kanehara, M., Shimada, T., Gomi, S. J., Ohbayashi, F. and Oshiki, T. (2000b). Identification and genetic mapping of RAPD markers linked to the densonucleosis refractoriness gene, nsd-2, in the silkworm, Bombyx mori. *Genes Genet Syst* **75**, 93-6.

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.

Adams, M. D., Tarng, R. S. and Rio, D. C. (1997). The alternative splicing factor PSI regulates P-element third intron splicing in vivo. *Genes Dev* **11**, 129-38.

Ajimura., Ken, S., Hiroaki, A., Toshiki, T., Toru, S. and Mita., K. (2006). Are the zinc-finger motif genes, z1 and z20, located in the W chromosome involved in the sex-determination of the domesticated silkworm, *Bombyx mori*? *Abstracts from Seventh International Workshop on the Molecular Biology and Genetics of the Lepidoptera. J Insect Sci*, [will be available online http://www.insectscience.org/].

Akerib, C. C. and Meyer, B. J. (1994). Identification of X chromosome regions in Caenorhabditis elegans that contain sex-determination signal elements. *Genetics* **138**, 1105-25.

Albrecht, E. B. and Salz, H. K. (1993). The Drosophila sex determination gene snf is utilized for the establishment of the female-specific splicing pattern of Sex-lethal. *Genetics* **134**, 801-7.

Alvarez, M., Ruiz, M. F. and Sanchez, L. (2009). Effect of the gene doublesex of anastrepha on the somatic sexual development of Drosophila. *PLoS One* **4**, e5141.

Ambros, V. (2004). The functions of animal microRNAs. *Nature* **431**, 350-5.

An, W., Cho, S., Ishii, H. and Wensink, P. C. (1996). Sex-specific and non-sex-specific oligomerization domains in both of the doublesex transcription factors from Drosophila melanogaster. *Mol Cell Biol* **16**, 3106-11.

Arunkumar, K. P., Kifayathullah, L. and Nagaraju, J. (2008). Microsatellite markers for the Indian golden silkmoth, Antheraea assama (Saturniidae: Lepidoptera). *Molecular Ecology Resources* **9**, 268-270.

Bachiller, D. and Sanchez, L. (1991). Production of X0 clones in XX females of Drosophila. *Genet Res* **57**, 23-8.

Baker, B. S., Burtis, K., Goralski, T., Mattox, W. and Nagoshi, R. (1989). Molecular genetic aspects of sex determination in Drosophila melanogaster. *Genome* **31**, 638-45.

Baker, B. S. and Wolfner, M. F. (1988). A molecular analysis of doublesex, a bifunctional gene that controls both male and female sexual differentiation in Drosophila melanogaster. *Genes Dev* **2**, 477-89.

Baroiller, J. F. and D'Cotta, H. (2001). Environment and sex determination in farmed fish. *Comp Biochem Physiol C Toxicol Pharmacol* **130**, 399-409.

Bayrer, J. R., Zhang, W. and Weiss, M. A. (2005). Dimerization of doublesex is mediated by a cryptic ubiquitin-associated domain fold: implications for sex-specific gene regulation. *J Biol Chem* **280**, 32989-96.

Beckmann, K., Grskovic, M., Gebauer, F. and Hentze, M. W. (2005). A dual inhibitory mechanism restricts msl-2 mRNA translation for dosage compensation in Drosophila. *Cell* **122**, 529-40.

Belaid, B., Richard-Mercier, N., Pieau, C. and Dorizzi, M. (2001). Sex reversal and aromatase in the European pond turtle: treatment with letrozole after the thermosensitive period for sex determination. *J Exp Zool* **290**, 490-7.

Bell, L. R., Horabin, J. I., Schedl, P. and Cline, T. W. (1991). Positive autoregulation of sex-lethal by alternative splicing maintains the female determined state in Drosophila. *Cell* **65**, 229-39.

Bell, L. R., Maine, E. M., Schedl, P. and Cline, T. W. (1988). Sex-lethal, a Drosophila sex determination switch gene, exhibits sex-specific RNA splicing and sequence similarity to RNA binding proteins. *Cell* **55**, 1037-46.

Belote, J. M., Handler, A. M., Wolfner, M. F., Livak, K. J. and Baker, B. S. (1985). Sex-specific regulation of yolk protein gene expression in Drosophila. *Cell* **40**, 339-48.

Belote, J. M., McKeown, M., Boggs, R. T., Ohkawa, R. and Sosnowski, B. A. (1989). Molecular genetics of transformer, a genetic switch controlling sexual differentiation in Drosophila. *Dev Genet* **10**, 143-54.

Berta, P., Hawkins, J. R., Sinclair, A. H., Taylor, A., Griffiths, B. L., Goodfellow, P. N. and Fellous, M. (1990). Genetic evidence equating SRY and the testis-determining factor. *Nature* **348**, 448-50.

Beye, M. (2004). The dice of fate: the csd gene and how its allelic composition regulates sexual development in the honey bee, Apis mellifera. *Bioessays* **26**, 1131-9.

Beye, M., Hasselmann, M., Fondrk, M. K., Page, R. E. and Omholt, S. W. (2003). The gene csd is the primary signal for sexual development in the honeybee and encodes an SR-type protein. *Cell* **114**, 419-29.

Bhattacharya, A., Sudha, S., Chandra, H. S. and Steward, R. (1999). flex, an X-linked female-lethal mutation in Drosophila melanogaster controls the expression of Sex-lethal. *Development* **126**, 5485-93.

Billeter, J. C., Goodwin, S. F. and O'Dell, K. M. (2002). Genes mediating sex-specific behaviors in Drosophila. *Adv Genet* **47**, 87-116.

Boggs, R. T., Gregor, P., Idriss, S., Belote, J. M. and McKeown, M. (1987). Regulation of sexual differentiation in D. melanogaster via alternative splicing of RNA from the transformer gene. *Cell* **50**, 739-47.

Bopp, D., Bell, L. R., Cline, T. W. and Schedl, P. (1991). Developmental distribution of female-specific Sex-lethal proteins in Drosophila melanogaster. *Genes Dev* **5**, 403-15.

Bopp, D., Calhoun, G., Horabin, J. I., Samuels, M. and Schedl, P. (1996). Sex-specific control of Sex-lethal is a conserved mechanism for sex determination in the genus Drosophila. *Development* **122**, 971-82.

Botella, L. M., Donoro, C., Sanchez, L., Segarra, C. and Granadino, B. (1996). Cloning and characterization of the scute (sc) gene of Drosophila subobscura. *Genetics* **144**, 1043-51.

Bownes, M. (1994). The regulation of the yolk protein genes, a family of sex differentiation genes in Drosophila melanogaster. *Bioessays* **16**, 745-52.

Bridges, C. B. and Anderson, E. G. (1925). Crossing over in the X Chromosomes of Triploid Females of DROSOPHILA MELANOGASTER. *Genetics* **10**, 418-41.

Bull, J. J. (1983). Evolution of sex determining mechanisms. *Benjamin Cummings Publishing., Menlo Park, California.*

Bull, J. J. (1987). Sex determining mechanisms: an evolutionary perspective. *Experientia Suppl* **55**, 93-115.

Burghardt, G., Hediger, M., Siegenthaler, C., Moser, M., Dubendorfer, A. and Bopp, D. (2005). The transformer2 gene in Musca domestica is required for selecting and maintaining the female pathway of development. *Dev Genes Evol* **215**, 165-76.

Burtis, K. C. and Baker, B. S. (1989). Drosophila doublesex gene controls somatic sexual differentiation by producing alternatively spliced mRNAs encoding related sex-specific polypeptides. *Cell* **56**, 997-1010.

Burtis, K. C., Coschigano, K. T., Baker, B. S. and Wensink, P. C. (1991). The doublesex proteins of Drosophila melanogaster bind directly to a sex-specific yolk protein gene enhancer. *Embo J* **10**, 2577-82.

Carmi, I., Kopczynski, J. B. and Meyer, B. J. (1998). The nuclear hormone receptor SEX-1 is an X-chromosome signal that determines nematode sex. *Nature* **396**, 168-73.

Carr, J. L. and Bickham, J. W. (1981). Sex chromosomes of the Asian black pond turtle, Siebenrockiella crassicollis (Testudines: Emydidae). *Cytogenet Cell Genet* **31**, 178-83.

Casper, A. and Van Doren, M. (2006). The control of sexual identity in the Drosophila germline. *Development* **133**, 2783-91.

Cavaliere, D., Di Cara, F., Polito, L. C. and Digilio, F. A. (2009). Cloning and functional characterization of the intersex homologous gene in the pest lepidopteron Maruca vitrata. *Int J Dev Biol* **53**, 1057-62.

Chandler, D., McGuffin, M. E., Piskur, J., Yao, J., Baker, B. S. and Mattox, W. (1997). Evolutionary conservation of regulatory strategies for the sex determination factor transformer-2. *Mol Cell Biol* **17**, 2908-19.

Chase, B. A. and Baker, B. S. (1995). A genetic analysis of intersex, a gene regulating sexual differentiation in Drosophila melanogaster females. *Genetics* **139**, 1649-61.

Chen, S. L., Dai, S. M., Lu, K. H. and Chang, C. (2008). Female-specific doublesex dsRNA interrupts yolk protein gene expression and reproductive ability in oriental fruit fly, Bactrocera dorsalis (Hendel). *Insect Biochem Mol Biol* **38**, 155-65.

Cho, S., Huang, Z. Y. and Zhang, J. (2007). Sex-specific splicing of the honeybee doublesex gene reveals 300 million years of evolution at the bottom of the insect sex-determination pathway. *Genetics* **177**, 1733-41.

Cho, S. and Wensink, P. C. (1998). Linkage between oligomerization and DNA binding in Drosophila doublesex proteins. *Biochemistry* **37**, 11301-8.

Chu, D. S., Dawes, H. E., Lieb, J. D., Chan, R. C., Kuo, A. F. and Meyer, B. J. (2002). A molecular link between gene-specific and chromosome-wide transcriptional repression. *Genes Dev* **16**, 796-805.

Cline, T. W. (1978). Two closely linked mutations in Drosophila melanogaster that are lethal to opposite sexes and interact with daughterless. *Genetics* **90**, 683-98.

Cline, T. W. (1984). Autoregulatory functioning of a Drosophila gene product that establish es and maintains the sexually determined state. *Genetics* **107**, 231-77.

Cline, T. W. (1988). Evidence that sisterless-a and sisterless-b are two of several discrete "numerator elements" of the X/A sex determination signal in Drosophila that switch Sxl between two alternative stable expression states. *Genetics* **119**, 829-62.

Cline, T. W. and Meyer, B. J. (1996). Vive la difference: males vs females in flies vs worms. *Annu Rev Genet* **30**, 637-702.

Cline, T. W., Rudner, D. Z., Barbash, D. A., Bell, M. and Vutien, R. (1999). Functioning of the Drosophila integral U1/U2 protein Snf independent of U1 and U2 small nuclear ribonucleoprotein particles is revealed by snf(+) gene dose effects. *Proc Natl Acad Sci U S A* **96**, 14451-8.

Clinton, M. (1998). Sex determination and gonadal development: a bird's eye view. *J Exp Zool* **281**, 457-65.

Concha, C. and Scott, M. J. (2009). Sexual Development in Lucilia cuprina (Diptera, Calliphoridae) Is Controlled by the Transformer Gene. *Genetics*.

Cook, J. M. (1993). Sex Determination in the hymenoptera - a rerview of model and evidence. *Heredity* **71**, 421-435

Coschigano, K. T. and Wensink, P. C. (1993). Sex-specific transcriptional regulation by the male and female doublesex proteins of Drosophila. *Genes Dev* **7**, 42-54.

Cristino, A. S., Nascimento, A. M., Costa Lda, F. and Simoes, Z. L. (2006). A comparative analysis of highly conserved sex-determining genes between Apis mellifera and Drosophila melanogaster. *Genet Mol Res* **5**, 154-68.

Dauwalder, B., Tsujimoto, S., Moss, J. and Mattox, W. (2002). The Drosophila takeout gene is regulated by the somatic sex-determination pathway and affects male courtship behavior. *Genes Dev* **16**, 2879-92.

Dawes, H. E., Berlin, D. S., Lapidus, D. M., Nusbaum, C., Davis, T. L. and Meyer, B. J. (1999). Dosage compensation proteins targeted to X chromosomes by a determinant of hermaphrodite fate. *Science* **284**, 1800-4.

de Bono, M. and Hodgkin, J. (1996). Evolution of sex determination in caenorhabditis: unusually high divergence of tra-1 and its functional consequences. *Genetics* **144**, 587-95.

DeLong, L., Plenefisch, J. D., Klein, R. D. and Meyer, B. J. (1993). Feedback control of sex determination by dosage compensation revealed through Caenorhabditis elegans sdc-3 mutations. *Genetics* **133**, 875-96.

Demas, S., Duronslet, M., Wachtel, S., Caillouet, C. and Nakamura, D. (1990). Sex-specific DNA in reptiles with temperature sex determination. *J Exp Zool* **253**, 319-24.

Deodikar, G. B., Chowdhury, S. N., Bhuyan, B. N. and Kshirsagar, K. K. (1962). Cytogenetic studies in Indian silkworms. *Current Science* **31**, 247–248.

Dorizzi, M., Richard-Mercier, N. and Pieau, C. (1996). The ovary retains male potential after the thermosensitive period for sex determination in the turtle Emys orbicularis. *Differentiation* **60**, 193-201.

Dournon, C., Houillon, C. and Pieau, C. (1990). Temperature sex-reversal in amphibians and reptiles. *Int J Dev Biol* **34**, 81-92.

Dubendorfer, A., Hediger, M., Burghardt, G. and Bopp, D. (2002). Musca domestica, a window on the evolution of sex-determining mechanisms in insects. *Int J Dev Biol* **46**, 75-9.

Duffy, J. B. and Gergen, J. P. (1991). The Drosophila segmentation gene runt acts as a position-specific numerator element necessary for the uniform expression of the sex-determining gene Sex-lethal. *Genes Dev* **5**, 2176-87.

Dupont, S., Dennefeld, C., Krust, A., Chambon, P. and Mark, M. (2003). Expression of Sox9 in granulosa cells lacking the estrogen receptors, ERalpha and ERbeta. *Dev Dyn* **226**, 103-6.

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

Engel, W. and Schmid, M. (1981). H-Y antigen as a tool for the determination of the heterogametic sex in Amphibia. *Cytogenet Cell Genet* **30**, 130-6.

Ennis, T. J. (1976). Sex chromatin and chromosome numbers in Lepidoptera. *Can J Genet Cytol* **18**, 119-30.

Erdman, S. E. and Burtis, K. C. (1993). The Drosophila doublesex proteins share a novel zinc finger related DNA binding domain. *Embo J* **12**, 527-35.

Erdman, S. E., Chen, H. J. and Burtis, K. C. (1996). Functional and genetic characterization of the oligomerization and DNA binding properties of the Drosophila doublesex proteins. *Genetics* **144**, 1639-52.

Erickson, J. W. and Cline, T. W. (1991). Molecular nature of the Drosophila sex determination signal and its link to neurogenesis. *Science* **251**, 1071-4.

Erickson, J. W. and Cline, T. W. (1998). Key aspects of the primary sex determination mechanism are conserved across the genus Drosophila. *Development* **125**, 3259-68.

Evans, J. D., Shearman, D. C. and Oldroyd, B. P. (2004). Molecular basis of sex determination in haplodiploids. *Trends Ecol Evol* **19**, 1-3.

Ewert, M. A. and Nelson, C. E. (1991). Sex determination in turtles: diverse patterns and some possible adaptive values. *Copeia 1991*, 50–69.

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.

Flickinger, T. W. and Salz, H. K. (1994). The Drosophila sex determination gene snf encodes a nuclear protein with sequence and functional similarity to the mammalian U1A snRNP protein. *Genes Dev* **8**, 914-25.

Foster, J. W., Dominguez-Steglich, M. A., Guioli, S., Kwok, C., Weller, P. A., Stevanovic, M., Weissenbach, J., Mansour, S., Young, I. D., Goodfellow, P. N. et al. (1994). Campomelic dysplasia and autosomal sex reversal caused by mutations in an SRY-related gene. *Nature* **372**, 525-30.

Francis, R. C. and Barlow, G. W. (1993). Social control of primary sex differentiation in the Midas cichlid. *Proc Natl Acad Sci U S A* **90**, 10673-5.

Fujii, H. (1998). Genetical Stocks and Mutations of Bombyx mori: Important Genetic Resources. *Fukuoka, Japan: Kyushu University*.

Fujii, T. and Shimada, T. (2007). Sex determination in the silkworm, Bombyx mori: a female determinant on the W chromosome and the sex-determining gene cascade. *Semin Cell Dev Biol* **18**, 379-88.

Gailey, D. A., Billeter, J. C., Liu, J. H., Bauzon, F., Allendorfer, J. B. and Goodwin, S. F. (2006). Functional conservation of the fruitless male sex-determination gene across 250 Myr of insect evolution. *Mol Biol Evol* **23**, 633-43.

Garrett-Engele, C. M., Siegal, M. L., Manoli, D. S., Williams, B. C., Li, H. and Baker, B. S. (2002). intersex, a gene required for female sexual development in Drosophila, is expressed in both sexes and functions together with doublesex to regulate terminal differentiation. *Development* **129**, 4661-75.

Goldsmith, M. R. and Kafatos, F. C. (1984). Developmentally regulated genes in silkmoths. *Annu Rev Genet* **18**, 443-87.

Goldsmith, M. R., Shimada, T. and Abe, H. (2005). The genetics and genomics of the silkworm, Bombyx mori. *Annu Rev Entomol* **50**, 71-100.

Graham, P., Penn, J. K. and Schedl, P. (2003). Masters change, slaves remain. *Bioessays* **25**, 1-4.

Granadino, B., Campuzano, S. and Sanchez, L. (1990). The Drosophila melanogaster fl(2)d gene is needed for the female-specific splicing of Sex-lethal RNA. *Embo J* **9**, 2597-602.

Granadino, B., Penalva, L. O. and Sanchez, L. (1996). The gene fl(2)d is needed for the sex-specific splicing of transformer pre-mRNA but not for double-sex pre-mRNA in Drosophila melanogaster. *Mol Gen Genet* **253**, 26-31.

Graves, J. A. (1995). The evolution of mammalian sex chromosomes and the origin of sex determining genes. *Philos Trans R Soc Lond B Biol Sci* **350**, 305-11; discussion 311-2.

Grimaldi, D. A. and Engel, M. S. (2005). Evolution of the Insects. *New York: Cambridge University Press. 755p*.

Grutzner, F., Rens, W., Tsend-Ayush, E., El-Mogharbel, N., O'Brien, P. C., Jones, R. C., Ferguson-Smith, M. A. and 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**, 913-7.

Gutzke, W. H. and Paukstis, G. L. (1983). Influence of the hydric environment on sexual differentiation of turtles. *J Exp Zool* **226**, 467-9.

Hackstein, J. H. (1987). Spermatogenesis in Drosophila. *Results Probl Cell Differ* **15**, 63-116.

Hammes, A., Guo, J. K., Lutsch, G., Leheste, J. R., Landrock, D., Ziegler, U., Gubler, M. C. and Schedl, A. (2001). Two splice variants of the Wilms' tumor 1 gene have distinct functions during sex determination and nephron formation. *Cell* **106**, 319-29.

Hasselmann, M., Gempe, T., Schiott, M., Nunes-Silva, C. G., Otte, M. and Beye, M. (2008). Evidence for the evolutionary nascence of a novel sex determination pathway in honeybees. *Nature* **454**, 519-22.

Hediger, M., Burghardt, G., Siegenthaler, C., Buser, N., Hilfiker-Kleiner, D., Dubendorfer, A. and Bopp, D. (2004). Sex determination in Drosophila melanogaster and Musca domestica converges at the level of the terminal regulator doublesex. *Dev Genes Evol* **214**, 29-42.

Hedley, M. L. and Maniatis, T. (1991). Sex-specific splicing and polyadenylation of dsx pre-mRNA requires a sequence that binds specifically to tra-2 protein in vitro. *Cell* **65**, 579-86.

Hertel, K. J. and Maniatis, T. (1998). The function of multisite splicing enhancers. *Mol Cell* **1**, 449-55.

Hiaso, A. (1994). Principles of sericulture. 141.

Hilfiker, A., Amrein, H., Dubendorfer, A., Schneiter, R. and Nothiger, R. (1995). The gene virilizer is required for female-specific splicing controlled by Sxl, the master gene for sexual development in Drosophila. *Development* **121**, 4017-26.

Hodgkin, J. (1987). Primary sex determination in the nematode C. elegans. *Development* **101 Suppl**, 5-16.

Hodgkin, J., Horvitz, H. R. and Brenner, S. (1979). Nondisjunction Mutants of the Nematode CAENORHABDITIS ELEGANS. *Genetics* **91**, 67-94.

Hodgkin, J., Zellan, J. D. and Albertson, D. G. (1994). Identification of a candidate primary sex determination locus, fox-1, on the X chromosome of Caenorhabditis elegans. *Development* **120**, 3681-9.

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

Horabin, J. I. and Schedl, P. (1996). Splicing of the drosophila Sex-lethal early transcripts involves exon skipping that is independent of Sex-lethal protein. *Rna* **2**, 1-10.

Hoshijima, K., Inoue, K., Higuchi, I., Sakamoto, H. and Shimura, Y. (1991). Control of doublesex alternative splicing by transformer and transformer-2 in Drosophila. *Science* **252**, 833-6.

Huang, B., Wang, S., Ning, Y., Lamb, A. N. and Bartley, J. (1999). Autosomal XX sex reversal caused by duplication of SOX9. *Am J Med Genet* **87**, 349-53.

Hunter, C. P. and Wood, W. B. (1992). Evidence from mosaic analysis of the masculinizing gene her-1 for cell interactions in C. elegans sex determination. *Nature* **355**, 551-5.

Izumi, S., Sakurai, H., Fujii, T., Ikeda, W. and Tomino, S. (1988). Cloning of mRNA sequence coding for sex-specific storage protein of Bombyx mori. *Biochim Biophys Acta* **949**, 181-8.

Janzen, F. J. and Paukstis, G. L. (1991). Environmental sex determination in reptiles: ecology, evolution, and experimental design. *Q Rev Biol* **66**, 149-79.

Jimenez, R., Sanchez, A., Burgos, M. and Diaz de la Guardia, R. (1996). Puzzling out the genetics of mammalian sex determination. *Trends Genet* **12**, 164-6.

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

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

Kelley, R. L., Wang, J., Bell, L. and Kuroda, M. I. (1997). Sex lethal controls dosage compensation in Drosophila by a non-splicing mechanism. *Nature* **387**, 195-9.

Kent, J., Wheatley, S. C., Andrews, J. E., Sinclair, A. H. and Koopman, P. (1996). A male-specific role for SOX9 in vertebrate sex determination. *Development* **122**, 2813-22.

Keyes, L. N., Cline, T. W. and Schedl, P. (1992). The primary sex determination signal of Drosophila acts at the level of transcription. *Cell* **68**, 933-43.

Kimura, K., Ote, M., Tazawa, T. and Yamamoto, D. (2005). Fruitless specifies sexually dimorphic neural circuitry in the Drosophila brain. *Nature* **438**, 229-33.

Koike, Y., Mita, K., Suzuki, M. G., Maeda, S., Abe, H., Osoegawa, K., deJong, P. J. and Shimada, T. (2003). Genomic sequence of a 320-kb segment of the Z chromosome of Bombyx mori containing a kettin ortholog. *Mol Genet Genomics* **269**, 137-49.

Koopman, P., Gubbay, J., Vivian, N., Goodfellow, P. and Lovell-Badge, R. (1991). Male development of chromosomally female mice transgenic for Sry. *Nature* **351**, 117-21.

Koopman, P., Munsterberg, A., Capel, B., Vivian, N. and Lovell-Badge, R. (1990). Expression of a candidate sex-determining gene during mouse testis differentiation. *Nature* **348**, 450-2.

Kuhn, S., Sievert, V. and Traut, W. (2000). The sex-determining gene doublesex in the fly Megaselia scalaris: conserved structure and sex-specific splicing. *Genome* **43**, 1011-20.

Kulathinal, R. J., Skwarek, L., Morton, R. A. and Singh, R. S. (2003). Rapid evolution of the sex-determining gene, transformer: structural diversity and rate heterogeneity among sibling species of Drosophila. *Mol Biol Evol* **20**, 441-52.

Kurin, R. (2002). "The Silk Road: Connecting Cultures, Creating Trust" Talk Story, Fall 2002, No. 21. *Smithsonian Center for Folklife and Cultural Heritage*, 1-11.

Lagos, D., Koukidou, M., Savakis, C. and Komitopoulou, K. (2007). The transformer gene in Bactrocera oleae: the genetic switch that determines its sex fate. *Insect Mol Biol* **16**, 221-30.

Lagos, D., Ruiz, M. F., Sanchez, L. and Komitopoulou, K. (2005). Isolation and characterization of the Bactrocera oleae genes orthologous to the sex determining Sex-lethal and doublesex genes of Drosophila melanogaster. *Gene* **348**, 111-21.

Lance, V. A. (2003). Alligator physiology and life history: the importance of temperature. *Exp.Gerontol.* **38**, 801-805.

Lang, J. W., Andrews, H. and Whittaker, R. (1989). Sex determination and sex ratios in Crocodylus palustris. *American zoologist* **29**, 935-952.

Li, H. and Baker, B. S. (1998). hermaphrodite and doublesex function both dependently and independently to control various aspects of sexual differentiation in Drosophila. *Development* **125**, 2641-51.

Lifschytz, E. (1987). The developmental program of spermiogenesis in Drosophila: a genetic analysis. *Int Rev Cytol* **109**, 211-58.

Lundrigan, B. L. and Tucker, P. K. (1997). Evidence for multiple functional copies of the male sex-determining locus, Sry, in African murine rodents. *J Mol Evol* **45**, 60-5.

Lynch, K. W. and Maniatis, T. (1996). Assembly of specific SR protein complexes on distinct regulatory elements of the Drosophila doublesex splicing enhancer. *Genes Dev* **10**, 2089-101.

MacDougall, C., Harbison, D. and Bownes, M. (1995). The developmental consequences of alternate splicing in sex determination and differentiation in Drosophila. *Dev Biol* **172**, 353-76.

Madl, J. E. and Herman, R. K. (1979). Polyploids and sex determination in Caenorhabditis elegans. *Genetics* **93**, 393-402.

Manolakou, P., Lavranos, G. and Angelopoulou, R. (2006). Molecular patterns of sex determination in the animal kingdom: a comparative study of the biology of reproduction. *Reprod Biol Endocrinol* **4**, 59.

Marec, F., Tothova, A., Sahara, K. and Traut, W. (2001). Meiotic pairing of sex chromosome fragments and its relation to atypical transmission of a sex-linked marker in Ephestia kuehniella (Insecta: Lepidoptera). *Heredity* **87**, 659-71.

McAllister, B. F. and McVean, G. A. (2000). Neutral evolution of the sex-determining gene transformer in Drosophila. *Genetics* **154**, 1711-20.

Mehra, A., Gaudet, J., Heck, L., Kuwabara, P. E. and Spence, A. M. (1999). Negative regulation of male development in Caenorhabditis elegans by a protein-protein interaction between TRA-2A and FEM-3. *Genes Dev* **13**, 1453-63.

Meise, M., Hilfiker-Kleiner, D., Dubendorfer, A., Brunner, C., Nothiger, R. and Bopp, D. (1998). Sex-lethal, the master sex-determining gene in Drosophila, is not sex-specifically regulated in Musca domestica. *Development* **125**, 1487-94.

Miller, L. M., Plenefisch, J. D., Casson, L. P. and Meyer, B. J. (1988). xol-1: a gene that controls the male modes of both sex determination and X chromosome dosage compensation in C. elegans. *Cell* **55**, 167-83.

Mine, E., Izumi, S., Katsuki, M. and Tomino, S. (1983). Developmental and sex-dependent regulation of storage protein synthesis in the silkworm, Bombyx mori. *Dev Biol* **97**, 329-37.

Minniti, A. N., Sadler, C. and Ward, S. (1996). Genetic and molecular analysis of spe-27, a gene required for spermiogenesis in Caenorhabditis elegans hermaphrodites. *Genetics* **143**, 213-23.

Mita, K., Kasahara, M., Sasaki, S., Nagayasu, Y., Yamada, T., Kanamori, H., Namiki, N., Kitagawa, M., Yamashita, H., Yasukochi, Y. et al. (2004). The genome sequence of silkworm, Bombyx mori. *DNA Res* **11**, 27-35.

Morais da Silva, S., Hacker, A., Harley, V., Goodfellow, P., Swain, A. and 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**, 62-8.

Morrish, B. C. and Sinclair, A. H. (2002). Vertebrate sex determination: many means to an end. *Reproduction* **124**, 447-57.

Mrosovsky, N. and Pieau, C. (1991). Transitional change of temperature, pivotal temperature and thermosensitive stages for sex determiantion in reptiles. *Chelon. Cons. Biol.* **1**, 238-240.

Mukherjee, A. S. and Beermann, W. (1965). Synthesis of ribonucleic acid by the X-chromosomes of Drosophila melanogaster and the problem of dosage compensation. *Nature* **207**, 785-6.

Muller-Holtkamp, F. (1995). The Sex-lethal gene homologue in Chrysomya rufifacies is highly conserved in sequence and exon-intron organization. *J Mol Evol* **41**, 467-77.

Nagaraja, G. M., Mahesh, G., Satish, V., Madhu, M., Muthulakshmi, M. and Nagaraju, J. (2005). Genetic mapping of Z chromosome and identification of W chromosome-specific markers in the silkworm, Bombyx mori. *Heredity* **95**, 148-57.

Nagaraja, G. M. and Nagaraju, J. (1995). Genome fingerprinting of the silkworm, Bombyx mori, using random arbitrary primers. *Electrophoresis* **16**, 1633-8.

Nagaraju, J. and Goldsmith, M. R. (2002). Silkworm genomics - progress and prospects. *Current Science* **83**, 415-425.

Nanda, I., Schartl, M., Feichtinger, W., Epplen, J. T. and Schmid, M. (1992). Early stages of sex chromosome differentiation in fish as analysed by simple repetitive DNA sequences. *Chromosoma* **101**, 301-10.

Narendra, U., Zhu, L., Li, B., Wilken, J. and Weiss, M. A. (2002). Sex-specific gene regulation. The Doublesex DM motif is a bipartite DNA-binding domain. *J Biol Chem* **277**, 43463-73.

Niessen, M., Schneiter, R. and Nothiger, R. (2001). Molecular identification of virilizer, a gene required for the expression of the sex-determining gene Sex-lethal in Drosophila melanogaster. *Genetics* **157**, 679-88.

Nigon, V. (1949). Les modalités de la reproduction et le déterminisme du sexe chez quelques nematodes libres. *Ann. Sci. Nat. Zool. Biol. Anim.* **11**, 1-132.

Nigon, V. (1951). Polyploı¨die expe´rimentale chez un nematode libre, Rhabditis elegans. *Maupas. Bull. Biol. Fr. Belg.* **85**, 187–225.

Niimi, T., Sahara, K., Oshima, H., Yasukochi, Y., Ikeo, K. and Traut, W. (2006). Molecular cloning and chromosomal localization of the Bombyx Sex-lethal gene. *Genome* **49**, 263-8.

Niu, B. L., Meng, Z. Q., Tao, Y. Z., Lu, S. L., Weng, H. B., He, L. H. and Shen, W. F. (2005). Cloning and alternative splicing analysis of Bombyx mori transformer-2 gene using silkworm EST database. *Acta Biochim Biophys Sin (Shanghai)* **37**, 728-36.

O'Neil, M. T. and Belote, J. M. (1992). Interspecific comparison of the transformer gene of Drosophila reveals an unusually high degree of evolutionary divergence. *Genetics* **131**, 113-28.

Ohbayashi, F., Suzuki, M. G., Mita, K., Okano, K. and Shimada, T. (2001). A homologue of the Drosophila doublesex gene is transcribed into sex-specific mRNA isoforms in the silkworm, Bombyx mori. *Comp Biochem Physiol B Biochem Mol Biol* **128**, 145-58.

Oliveira, D. C., Werren, J. H., Verhulst, E. C., Giebel, J. D., Kamping, A., Beukeboom, L. W. and van de Zande, L. (2009). Identification and characterization of the doublesex gene of Nasonia. *Insect Mol Biol* **18**, 315-24.

Ortega, A., Niksic, M., Bachi, A., Wilm, M., Sanchez, L., Hastie, N. and Valcarcel, J. (2003). Biochemical function of female-lethal (2)D/Wilms' tumor suppressor-1-associated proteins in alternative pre-mRNA splicing. *J Biol Chem* **278**, 3040-7.

Ota, T., Fukunaga, A., Kawabe, M. and Oishi, K. (1981). Interactions between sex-transformation mutants of Drosophila melanogaster. I. Hemolymph vitellogenins and gonad morphology. *Genetics* **99**, 429-41.

Pane, A., De Simone, A., Saccone, G. and Polito, C. (2005). Evolutionary conservation of Ceratitis capitata transformer gene function. *Genetics* **171**, 615-24.

Pane, A., Salvemini, M., Delli Bovi, P., Polito, C. and Saccone, G. (2002). The transformer gene in Ceratitis capitata provides a genetic basis for selecting and remembering the sexual fate. *Development* **129**, 3715-25.

Parkhurst, S. M., Bopp, D. and Ish-Horowicz, D. (1990). X:A ratio, the primary sex-determining signal in Drosophila, is transduced by helix-loop-helix proteins. *Cell* **63**, 1179-91.

Pelosi, P. and Maida, R. (1995). Odorant-binding proteins in insects. *Comp Biochem Physiol B Biochem Mol Biol* **111**, 503-14.

Penalva, L. O., Ruiz, M. F., Ortega, A., Granadino, B., Vicente, L., Segarra, C., Valcarcel, J. and Sanchez, L. (2000). The Drosophila fl(2)d gene, required for female-specific splicing of Sxl and tra pre-mRNAs, encodes a novel nuclear protein with a HQ-rich domain. *Genetics* **155**, 129-39.

Penalva, L. O., Sakamoto, H., Navarro-Sabate, A., Sakashita, E., Granadino, B., Segarra, C. and Sanchez, L. (1996). Regulation of the gene Sex-lethal: a comparative analysis of Drosophila melanogaster and Drosophila subobscura. *Genetics* **144**, 1653-64.

Penalva, L. O. and Sanchez, L. (2003). RNA binding protein sex-lethal (Sxl) and control of Drosophila sex determination and dosage compensation. *Microbiol Mol Biol Rev* **67**, 343-59, table of contents.

Perry, M. D., Li, W., Trent, C., Robertson, B., Fire, A., Hageman, J. M. and Wood, W. B. (1993). Molecular characterization of the her-1 gene suggests a direct role in cell signaling during Caenorhabditis elegans sex determination. *Genes Dev* **7**, 216-28.

Pieau, C. (1996). Temperature variation and sex determination in reptiles. *Bioessays* **18[19], 26. Ref Type: Journal (Full)**.

Pieau, C., Dorizzi, M., Richard-Mercier, N. and 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**, 400-8.

Pienaar, J. and Greeff, J. M. (2003). Maternal control of offspring sex and male morphology in the Otitesella fig wasps. *J Evol Biol* **16**, 244-53.

Postlethwait, J. H., Bownes, M. and Jowett, T. (1980). Sexual phenotype and vitellogenin synthesis in Drosophila melanogaster. *Dev Biol* **79**, 379-87.

Powell, J. R., Jow, M. M. and Meyer, B. J. (2005). The T-box transcription factor SEA-1 is an autosomal element of the X:A signal that determines C. elegans sex. *Dev Cell* **9**, 339-49.

Pultz, M. A. and Baker, B. S. (1995). The dual role of hermaphrodite in the Drosophila sex determination regulatory hierarchy. *Development* **121**, 99-111.

Raymond, C. S., Parker, E. D., Kettlewell, J. R., Brown, L. G., Page, D. C., Kusz, K., Jaruzelska, J., Reinberg, Y., Flejter, W. L., Bardwell, V. J. et al. (1999). A region of human chromosome 9p required for testis development contains two genes related to known sexual regulators. *Hum Mol Genet* **8**, 989-96.

Raymond, C. S., Shamu, C. E., Shen, M. M., Seifert, K. J., Hirsch, B., Hodgkin, J. and Zarkower, D. (1998). Evidence for evolutionary conservation of sex-determining genes. *Nature* **391**, 691-5.

Rhind, N. R., Miller, L. M., Kopczynski, J. B. and Meyer, B. J. (1995). xol-1 acts as an early switch in the C. elegans male/hermaphrodite decision. *Cell* **80**, 71-82.

Romer, U. and Beisenhertz, W. (1996). Environmental determination of sex in Apistogramma (Cichlidae) and two other fresh water fishes (Teleostei). *Journal of Fish Biology* **48**, 714-725 Ref Type: Journal (Full).

Ruiz, M. F., Eirin-Lopez, J. M., Stefani, R. N., Perondini, A. L., Selivon, D. and Sanchez, L. (2007a). The gene doublesex of Anastrepha fruit flies (Diptera, Tephritidae) and its evolution in insects. *Dev Genes Evol* **217**, 725-31.

Ruiz, M. F., Goday, C., Gonzalez, P. and Sanchez, L. (2003). Molecular analysis and developmental expression of the Sex-lethal gene of Sciara ocellaris (Diptera order, Nematocera suborder). *Gene Expr Patterns* **3**, 341-6.

Ruiz, M. F., Milano, A., Salvemini, M., Eirin-Lopez, J. M., Perondini, A. L., Selivon, D., Polito, C., Saccone, G. and Sanchez, L. (2007b). The gene transformer of anastrepha fruit flies (Diptera, tephritidae) and its evolution in insects. *PLoS ONE* **2**, e1239.

Ruiz, M. F., Stefani, R. N., Mascarenhas, R. O., Perondini, A. L., Selivon, D. and Sanchez, L. (2005). The gene doublesex of the fruit fly Anastrepha obliqua (Diptera, Tephritidae). *Genetics* **171**, 849-54.

Ryner, L. C. and Baker, B. S. (1991). Regulation of doublesex pre-mRNA processing occurs by 3'-splice site activation. *Genes Dev* **5**, 2071-85.

Saccone, G., Peluso, I., Artiaco, D., Giordano, E., Bopp, D. and Polito, L. C. (1998). The Ceratitis capitata homologue of the Drosophila sex-determining gene sex-lethal is structurally conserved, but not sex-specifically regulated. *Development* **125**, 1495-500.

Saccone, G., Salvemini, M., Pane, A. and Polito, L. C. (2008). Masculinization of XX Drosophila transgenic flies expressing the Ceratitis capitata DoublesexM isoform. *Int J Dev Biol* **52**, 1051-7.

Salz, H. K. (1992). The genetic analysis of snf: a Drosophila sex determination gene required for activation of Sex-lethal in both the germline and the soma. *Genetics* **130**, 547-54.

Salz, H. K. and Flickinger, T. W. (1996). Both loss-of-function and gain-of-function mutations in snf define a role for snRNP proteins in regulating Sex-lethal pre-mRNA splicing in Drosophila development. *Genetics* **144**, 95-108.

Salz, H. K., Maine, E. M., Keyes, L. N., Samuels, M. E., Cline, T. W. and Schedl, P. (1989). The Drosophila female-specific sex-determination gene, Sex-lethal, has stage-, tissue-, and sex-specific RNAs suggesting multiple modes of regulation. *Genes Dev* **3**, 708-19.

Samuels, M., Deshpande, G. and Schedl, P. (1998). Activities of the Sex-lethal protein in RNA binding and protein:protein interactions. *Nucleic Acids Res* **26**, 2625-37.

Samuels, M. E., Bopp, D., Colvin, R. A., Roscigno, R. F., Garcia-Blanco, M. A. and Schedl, P. (1994). RNA binding by Sxl proteins in vitro and in vivo. *Mol Cell Biol* **14**, 4975-90.

Sanchez, L. (2008). Sex-determining mechanisms in insects. *Int J Dev Biol* **52**, 837-56.

Sanchez, L. and Nothiger, R. (1983). Sex determination and dosage compensation in Drosophila melanogaster: production of male clones in XX females. *Embo J* **2**, 485-491.

Sarre, S. D., Georges, A. and Quinn, A. (2004). The ends of a continuum: genetic and temperature-dependent sex determination in reptiles. *Bioessays* **26**, 639-45.

Satish, V., Shukla, J. N. and Nagaraju, J. (2006). CCCH-type zinc finger genes: candidate regulators of sex determination pathway in the silkworm, Bombyx mori. *Abstracts from Seventh InternationalWorkshop on the Molecular Biology and Genetics of the Lepidoptera. J Insect Sci*, [will be available online http://www.insectscience.org/].

Scali, C., Catteruccia, F., Li, Q. and Crisanti, A. (2005). Identification of sex-specific transcripts of the Anopheles gambiae doublesex gene. *J Exp Biol* **208**, 3701-9.

Schauer, I. E. and Wood, W. B. (1990). Early C. elegans embryos are transcriptionally active. *Development* **110**, 1303-17.

Schmahl, J., Eicher, E. M., Washburn, L. L. and Capel, B. (2000). Sry induces cell proliferation in the mouse gonad. *Development* **127**, 65-73.

Schutt, C. and Nothiger, R. (2000). Structure, function and evolution of sex-determining systems in Dipteran insects. *Development* **127**, 667-77.

Sekido, R. and Lovell-Badge, R. (2007). Mechanisms of gonadal morphogenesis are not conserved between chick and mouse. *Dev Biol* **302**, 132-42.

Sekido, R. and Lovell-Badge, R. (2008). Sex determination involves synergistic action of SRY and SF1 on a specific Sox9 enhancer. *Nature* **453**, 930-4.

Serna, E., Gorab, E., Ruiz, M. F., Goday, C., Eirin-Lopez, J. M. and Sanchez, L. (2004). The gene Sex-lethal of the Sciaridae family (order Diptera, suborder Nematocera) and its phylogeny in dipteran insects. *Genetics* **168**, 907-21.

Shearman, D. C. and Frommer, M. (1998). The Bactrocera tryoni homologue of the Drosophila melanogaster sex-determination gene doublesex. *Insect Mol Biol* **7**, 355-66.

Siegal, M. L. and Baker, B. S. (2005). Functional conservation and divergence of intersex, a gene required for female differentiation in Drosophila melanogaster. *Dev Genes Evol* **215**, 1-12.

Sievert, V., Kuhn, S., Paululat, A. and Traut, W. (2000). Sequence conservation and expression of the sex-lethal homologue in the fly Megaselia scalaris. *Genome* **43**, 382-90.

Sievert, V., Kuhn, S. and Traut, W. (1997). Expression of the sex determining cascade genes Sex-lethal and doublesex in the phorid fly Megaselia scalaris. *Genome* **40**, 211-4.

Sinclair, A., Smith, C., Western, P. and McClive, P. (2002). A comparative analysis of vertebrate sex determination. *Novartis Found Symp* **244**, 102-11; discussion 111-4, 203-6, 253-7.

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. and 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**, 240-4.

Singh, L. and 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**, 205-16.

Skipper, M., Milne, C. A. and Hodgkin, J. (1999). Genetic and molecular analysis of fox-1, a numerator element involved in Caenorhabditis elegans primary sex determination. *Genetics* **151**, 617-31.

Slee, R. and Bownes, M. (1990). Sex determination in Drosophila melanogaster. *Q Rev Biol* **65**, 175-204.

Solari, A. J. (1994). Sex chromosomes and sex determination in vertebrates. *Boca Raton: CRC press*.

Stahlhut, J. K. and Cowan, D. P. (2004). Single-locus complementary sex determination in the inbreeding wasp Euodynerus foraminatus Saussure (Hymenoptera: Vespidae). *Heredity* **92**, 189-96.

Steinbrecht, R. A., Laue, M. and Ziegelberger, G. (1995). Immunolocalization of pheromone-binding protein and general odorant-binding protein in olfactory sensilla of the silk moths *Antheraea* and *Bombyx*. *Cell Tissue Res* **282**, 203–217.

Steinmann-Zwicky, M., Amrein, H. and Nothiger, R. (1990). Genetic control of sex determination in Drosophila. *Adv Genet* **27**, 189-237.

Suzuki, M. G., Funaguma, S., Kanda, T., Tamura, T. and Shimada, T. (2003). Analysis of the biological functions of a doublesex homologue in Bombyx mori. *Dev Genes Evol* **213**, 345-54.

Suzuki, M. G., Funaguma, S., Kanda, T., Tamura, T. and Shimada, T. (2005). Role of the male BmDSX protein in the sexual differentiation of Bombyx mori. *Evol Dev* **7**, 58-68.

Suzuki, M. G., Imanishi, S., Dohmae, N., Nishimura, T., Shimada, T. and Matsumoto, S. (2008). Establishment of a novel in vivo sex-specific splicing assay system to identify a trans-acting factor that negatively regulates splicing of Bombyx mori dsx female exons. *Mol Cell Biol* **28**, 333-43.

Suzuki, M. G., Ohbayashi, F., Mita, K. and Shimada, T. (2001). The mechanism of sex-specific splicing at the doublesex gene is different between Drosophila melanogaster and Bombyx mori. *Insect Biochem Mol Biol* **31**, 1201-11.

Tacke, R. and Manley, J. L. (1999). Functions of SR and Tra2 proteins in pre-mRNA splicing regulation. *Proc Soc Exp Biol Med* **220**, 59-63.

Taylor, B. J. and Truman, J. W. (1992). Commitment of abdominal neuroblasts in Drosophila to a male or female fate is dependent on genes of the sex-determining hierarchy. *Development* **114**, 625-42.

Tian, M. and Maniatis, T. (1992). Positive control of pre-mRNA splicing in vitro. *Science* **256**, 237-40.

Torres, M. and Sanchez, L. (1991). The sisterless-b function of the Drosophila gene scute is restricted to the stage when the X:A ratio determines the activity of Sex-lethal. *Development* **113**, 715-22.

Torres, M. and Sanchez, L. (1992). The segmentation gene runt is needed to activate Sex-lethal, a gene that controls sex determination and dosage compensation in Drosophila. *Genet Res* **59**, 189-98.

Traut, W. and Marec, F. (1996). Sex chromatin in lepidoptera. *Q Rev Biol* **71**, 239-56.

Traut, W., Niimi, T., Ikeo, K. and Sahara, K. (2006). Phylogeny of the sex-determining gene Sex-lethal in insects. *Genome* **49**, 254-62.

Traut, W., Sahara, K. and Marec, F. (2007). Sex chromosomes and sex determination in Lepidoptera. *Sex Dev* **1**, 332-46.

Trent, C., Purnell, B., Gavinski, S., Hageman, J., Chamblin, C. and Wood, W. B. (1991). Sex-specific transcriptional regulation of the C. elegans sex-determining gene her-1. *Mech Dev* **34**, 43-55.

Tucker, P. K. and Lundrigan, B. L. (1993). Rapid evolution of the sex determining locus in Old World mice and rats. *Nature* **364**, 715-7.

Uller, T., Pen, I., Wapstra, E., Beukeboom, L. W. and Komdeur, J. (2007). The evolution of sex ratios and sex-determining systems. *Trends Ecol Evol* **22**, 292-7.

Valcarcel, J., Singh, R., Zamore, P. D. and Green, M. R. (1993). The protein Sex-lethal antagonizes the splicing factor U2AF to regulate alternative splicing of transformer pre-mRNA. *Nature* **362**, 171-5.

Valenzuela, N. and Lance, V. A. (2004). Temperature dependent sex determination in vertebrates. *Washington D.C.: Smithsonian Books.*

Valleley, E. M., Cartwright, E. J., Croft, N. J., Markham, A. F. and Coletta, P. L. (2001). Characterisation and expression of Sox9 in the Leopard gecko, Eublepharis macularius. *J Exp Zool* **291**, 85-91.

Vidal, V. P., Chaboissier, M. C., de Rooij, D. G. and Schedl, A. (2001). Sox9 induces testis development in XX transgenic mice. *Nat Genet* **28**, 216-7.

Villella, A. and Hall, J. C. (1996). Courtship anomalies caused by doublesex mutations in Drosophila melanogaster. *Genetics* **143**, 331-44.

Villeneuve, A. M. and Meyer, B. J. (1990). The regulatory hierarchy controlling sex determination and dosage compensation in Caenorhabditis elegans. *Adv Genet* **27**, 117-88.

Wagner, T., Wirth, J., Meyer, J., Zabel, B., Held, M., Zimmer, J., Pasantes, J., Bricarelli, F. D., Keutel, J., Hustert, E. et al. (1994). Autosomal sex reversal and campomelic dysplasia are caused by mutations in and around the SRY-related gene SOX9. *Cell* **79**, 1111-20.

Ward, S. and Carrel, J. S. (1979). Fertilization and sperm competition in the nematode Caenorhabditis elegans. *Dev Biol* **73**, 304-21.

Waterbury, J. A., Horabin, J. I., Bopp, D. and Schedl, P. (2000). Sex determination in the Drosophila germline is dictated by the sexual identity of the surrounding soma. *Genetics* **155**, 1741-56.

Waterbury, J. A., Jackson, L. L. and Schedl, P. (1999). Analysis of the doublesex female protein in Drosophila melanogaster: role on sexual differentiation and behavior and dependence on intersex. *Genetics* **152**, 1653-67.

West, S. A., Reece, S. E. and Sheldon, B. C. (2002). Sex ratios. *Heredity* **88**, 117-24.

Western, P. S., Harry, J. L., Graves, J. A. and Sinclair, A. H. (1999a). Temperature-dependent sex determination in the American alligator: AMH precedes SOX9 expression. *Dev Dyn* **216**, 411-9.

Western, P. S., Harry, J. L., Graves, J. A. and Sinclair, A. H. (1999b). Temperature-dependent sex determination: upregulation of SOX9 expression after commitment to male development. *Dev Dyn* **214**, 171-7.

Wilhelm, D. (2007). R-spondin1--discovery of the long-missing, mammalian female-determining gene? *Bioessays* **29**, 314-8.

Wilhelm, D., Palmer, S. and Koopman, P. (2007). Sex determination and gonadal development in mammals. *Physiol Rev* **87**, 1-28.

Wilkins, A. S. (1995). Moving up the hierarchy: a hypothesis on the evolution of a genetic sex determination pathway. *Bioessays* **17**, 71-7.

Willis, J. H., Wilkins, A. S. and Goldsmith, M. R. (1995). A brief history of Lepidoptera as model systems. *In Molecular Model Systems in the Lepidoptera, ed.* **Cambridge/ New York: Cambridge Univ. Press** pp. 1–20.

Xia, Q., Zhou, Z., Lu, C., Cheng, D., Dai, F., Li, B., Zhao, P., Zha, X., Cheng, T., Chai, C. et al. (2004). A draft sequence for the genome of the domesticated silkworm (Bombyx mori). *Science* **306**, 1937-40.

Yano, K., Sakurai, M. T., Izumi, S. and Tomino, S. (1994). Vitellogenin gene of the silkworm, Bombyx mori: structure and sex-dependent expression. *FEBS Lett* **356**, 207-11.

Yao, H. H. and Capel, B. (2005). Temperature, genes, and sex: a comparative view of sex determination in Trachemys scripta and Mus musculus. *J Biochem* **138**, 5-12.

Yasuda, J. (1996). [TA cloning: a rapid and efficient method for cloning of the PCR products]. *Tanpakushitsu Kakusan Koso* **41**, 518-21.

Yasukochi, Y. (1998). A dense genetic map of the silkworm, Bombyx mori, covering all chromosomes based on 1018 molecular markers. *Genetics* **150**, 1513-25.

Yi, W., Ross, J. M. and Zarkower, D. (2000). Mab-3 is a direct tra-1 target gene regulating diverse aspects of C. elegans male sexual development and behavior. *Development* **127**, 4469-80.

Younger-Shepherd, S., Vaessin, H., Bier, E., Jan, L. Y. and Jan, Y. N. (1992). deadpan, an essential pan-neural gene encoding an HLH protein, acts as a denominator in Drosophila sex determination. *Cell* **70**, 911-22.

Zarkower, D. (2001). Establishing sexual dimorphism: conservation amidst diversity? *Nat Rev Genet* **2**, 175-85.

Zarkower, D. and Hodgkin, J. (1992). Molecular analysis of the C. elegans sex-determining gene tra-1: a gene encoding two zinc finger proteins. *Cell* **70**, 237-49.

Zhang, P. and Stankiewicz, R. L. (1998). Y-Linked male sterile mutations induced by P element in Drosophila melanogaster. *Genetics* **150**, 735-44.

Zhou, Y., Gu, H. and Dorn, S. (2006). Single-locus sex determination in the parasitoid wasp Cotesia glomerata (Hymenoptera: Braconidae). *Heredity* **96**, 487-92.

Zhu, L., Wilken, J., Phillips, N. B., Narendra, U., Chan, G., Stratton, S. M., Kent, S. B. and Weiss, M. A. (2000). Sexual dimorphism in diverse metazoans is regulated by a novel class of intertwined zinc fingers. *Genes Dev* **14**, 1750-64.