Abstract
Insects exhibit a variety of sex determining mechanisms including male or female heterogamety and haplodiploidy. The primary signal that starts sex determination is processed by a cascade of genes ending with the conserved switch doublesex that controls sexual differentiation. Transformer is the doublesex splicing regulator and has been found in all examined insects, indicating its ancestral function as a sex determining gene. Despite this conserved function, the variation in transformer nucleotide sequence, amino acid composition and protein structure can accommodate a multitude of upstream sex determining signals. Transformer regulation of doublesex and its taxonomic distribution indicate that the doublesex – transformer axis is conserved among all insects and that transformer is the key gene around which variation in sex determining mechanisms has evolved.
Sexual development, one of the most important and widespread developmental processes, essentially entails one simple choice: becoming male or female. Although this suggests a common underlying genetic mechanism, an astoundingly diverse array of pathways regulates sex determination. Sánchez (2008) reviewed current knowledge of sex determining mechanisms in insects with a focus on primary signals. In flies (Diptera) the gene *doublesex (dsx)* acts as a conserved major switch at the bottom of the sex determining cascade (Dübendorfer *et al.*, 2002; Saccone *et al.*, 2002; Sánchez, 2008). The part of the sex determining cascade where the primary signal is transmitted to *dsx* has, until recently, received less attention. Data from Hymenoptera enabled comparison of sex determination mechanisms at a wider level within the insect class. This has directed focus towards *transformer (tra)* as a central player in the evolution of sex determination in insects. In this review, we describe how *tra* translates different primary signals into one of two sex specific pathways and consider how its function may serve as the key process around which insect sex determination mechanisms have evolved.

**Drosophila sex determination: the reference**

Insect sex determination has been extensively examined in *Drosophila melanogaster* (Nöthiger and Steinmann-Zwicky, 1985; Baker, 1989; Schütz and Nöthiger, 2000) and has served as a reference for all other insects (Graham *et al.*, 2002; Shearman, 2002; Pane *et al.*, 2005). In *Drosophila*, the upstream genomic region of *Sxl* contains two promoters: *p_early* and *p_maintenance*, which is the late promoter. The primary signal is based on the concentration of X-linked signal elements (XSE) that activate the early Sexlethal (*Sxl*) promoter in diploid XX individuals only (Erickson and Quintero, 2007) (see Fig. 6.1). Transcription from the early promoter of *Sxl* yields a transcript that is spliced to encode a functional early SXL protein. This splice pattern depends on the use of the 5' splice site from the early exon E1, whereas in later stages the 5' splice site of late exon 2 is used (Zhu *et al.*, 1997). It results in the default exclusion of exon 3 that contains in-frame stop codons in the early transcript. This early protein enables the production of a functional late SXL protein which further maintains female specific *Sxl* splicing by auto regulation. SXL also directs cryptic splicing of *tra* by binding to a polypyrimidine tract in the first *tra* intron and forces the general splicing factor U2AF to use the female specific 3' splice site in exon 3 instead of the non-sex-specific 3' splice site in exon 2. This *tra* transcript yields a functional TRA protein (Boggs *et al.*, 1987; Inoue *et al.*, 1990; Valcarcel *et al.*, 1993), which interacts with the non-sex-specific transformer-2 protein (TRA-2) (Amrein *et al.*, 1988) and binds to the *dsx* transcript in the middle of exon 4, called the *dsx* repeat element (*dsxRE*). This *dsxRE* contains six copies of the 13 nucleotide sequence
TC(T/A)(T/A)C(A/G)ATCAACA (Tian and Maniatis, 1993). Located between repeat element 5 and 6 of the\textit{dsxRE} is a purine-rich enhancer element (PRE) which is required for the specific binding of TRA-2 to the\textit{dsxRE} (Lynch and Maniatis, 1995). The binding of TRA/TRA-2 to the\textit{dsxRE} and PRE sites retains exon 4 in the\textit{dsx} pre-mRNA resulting in female specific splicing of\textit{dsx} at the bottom of the cascade (Hedley and Maniatis, 1991; Hoshijima et al., 1991; Ryner and Baker, 1991), generating a female specific DSX protein.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{Figure6.1.png}
\caption{The sex determination cascade in \textit{Drosophila melanogaster}. Boxes with numbers indicate transcripts with exon number and relative exon size. Dark gray transcripts are full length ORFs and yield a functional protein. Light gray transcripts contain early in-frame stop codons and give truncated non-functional proteins. Transcripts are designated by their gene name in italic. Proteins are designated by their capital gene name. Superscript F and M stand for female specific transcript or protein or male specific transcript or protein, respectively. \textit{p\textsubscript{e}Sxl} indicates\textit{Sxl} transcript from the early promoter, \textit{p\textsubscript{m}Sxl} indicates \textit{Sxl} from the late promoter. DSX\textsubscript{M} directs primarily male morphology but also interacts a little with FRU\textsubscript{M} to direct male behavior, which is indicated by a smaller arrow. The gray bottom half of the sex determining cascade shows the conserved part of the cascade.

In XY males the level of XSEs is insufficient for early \textit{Sxl} transcription and no early SXL protein is synthesized, preventing the autoregulatory loop from establishing. As a result, \textit{Sxl} pre-mRNA from the late promoter is male specifically spliced by default, yielding a truncated non-functional SXL protein. The absence
of SXL leads to the “default” splicing of the tra pre-mRNA and a non-functional TRA protein. Without TRA, dsx pre-mRNA is spliced by default generating a male-specific DSX protein.

The TRA/TRA2 complex also regulates female specific splicing of fruitless (fru) which yields a non-functional FRU protein (Ryner et al., 1996), while absence of TRA leads to male specific fru splicing and a functional FRU protein. Fru is not part of the (morphological) sex determination pathway but seems conserved in insects (and reviewed in Billeter et al., 2006; Gailey et al., 2006; Bertossa et al., 2009). It is conserved in both gene structure and its function as determiner of male sexual behavior.

Conservation of sex determining genes in insects

There is a common pattern in insect sex determining cascades: at the bottom is dsx, which has been identified for all examined dipteran (Saccone et al., 1996; Shearman and Frommer, 1998; Kuhn et al., 2000; Ohbayashi et al., 2001; Hediger et al., 2004; Lagos et al., 2005; Ruiz et al., 2005; Scali et al., 2005) and hymenopteran insect species (Cho et al., 2007; Oliveira et al., 2009). DSX has two characteristic domains: a DNA binding domain (DM or OD1) and an oligomerization domain (dsx dimer or OD2). Oliveira et al. (2009) showed for several insect species that amino acid alignment of these domains followed the established phylogeny, suggesting their importance in sexual differentiation. Conservation of dsx is in agreement with Wilkins’ theory (Wilkins, 1995) stating that regulatory elements are recruited into sex determining pathways, causing divergence towards the top, while dsx remains conserved at the bottom. However, as more and more sex determining cascades are elucidated, it appears that conservation is not only at the level of dsx, but also at the regulation of its sex specific splicing.

Transformer

After the initial identification in Ceratitis capitata (Pane et al., 2002), also in other insect species (Anastrepha sp., Bactrocera oleae, Lucilia cuprina, Musca domestica, Apis mellifera and Nasonia vitripennis), dsx splicing regulator genes have been identified that all appear to be D. melanogaster tra orthologs (Lagos et al., 2007; Ruiz et al., 2007; Hasselmann et al., 2008; Concha and Scott, 2009; Hediger et al., 2010; Verhulst et al., 2010). A tra ortholog has not (yet) been identified in Lepidoptera, perhaps due to the strong sequence divergence that characterizes tra evolution. In Bombyx mori no tra ortholog has been found based on the lack of dsxRE or PRE binding sites on Bmdsx and the presumed default mode of female specific splicing (Suzuki et al., 2001; Suzuki et al., 2008). However, dsxRE/PRE binding sites have only been identified in dipterans based on homology to Drosophila and are probably so diverged that recognition of these sites in other orders is difficult. Cho et al. (2007) reported the absence of dsxRE/PRE binding sites in the hymenopteran A. mellifera and suggested that Amdsx follows default female specific splicing, similar to B. mori. Nevertheless, a func-
tional *tra* ortholog has recently been found in *A. mellifera*, termed *feminizer*, which is functionally and structurally similar to *tra* (Hasselmann et al., 2008; Gempe et al., 2009). Interestingly, a comparison between *dsx* of hymenopterans *A. mellifera* and *N. vitripennis* revealed putative *dsxRE/PRE* binding sites that indeed have severely diverged from those of Diptera (Bertossa et al., 2009). Similar *dsxRE/PRE* binding sites have been identified in the *Nasonia fruitless* gene (Bertossa et al., 2009) and in *Nutra* (Verhulst et al., 2010). Hence, the illustrious feminizing factor on the W chromosome in *B. mori* (Fujii and Shimada, 2007; Traut et al., 2007), may be an unconfirmed ortholog of *tra* that also functions as active feminizing factor. In the mosquito *Anopheles gambiae* and the phorid fly *Megaselia scalaris*, only *dsx* has been found to date, but *tra* is surmised to be the regulating splice factor of *Agdsx* and *Msdsx* since *dsxRE/PRE* binding sites have been identified (Kuhn et al., 2000; Scali et al., 2005).

The functional importance of these *dsx* splicing regulators in female development has been shown by RNA interference (RNAi) in early embryos, which resulted in male specific *dsx* splicing (Pane et al., 2002; Lagos et al., 2007; Concha and Scott, 2009; Gempe et al., 2009; Hediger et al., 2010; Verhulst et al., 2010). The subsequent transformation of otherwise female offspring was not always complete and resulted in intersexes with various stages of masculinization, while male development remained unaffected. Although these *tra* genes differ largely in their nucleotide and amino acid composition, their function as the sex specific splicing regulator of *dsx* appears conserved (Pane et al., 2005). Strikingly, a conserved pattern of *tra* regulation in all insect species is the sex specific alternative splicing that produces transcripts in males that contain early in-frame

![Figure 6.2: Classification of insects and the conservation of *dsx* and *tra*](image_url)
stop codons and yield no protein. Only the female specific splicing of \textit{tra} premRNA yields a full length transcript and leads to TRA protein production. This active TRA protein directs female specific splicing of \textit{dsx}, implying a functional conservation in insect sex determination.

\textit{Tra} regulation of \textit{dsx} apparently constitutes the axis of insect sex determination. It likely acquired its function in the early ancestors of the insects, as \textit{tra} orthologs are found throughout the insect class including Diptera, Hymenoptera and Coleoptera (Fig. 6.2), but apparently has no sex determining function in the crustacean \textit{Daphnia} (Kato \textit{et al.}, 2010). The large sequence divergence indicates that \textit{tra} conservation is predominantly at the functional and less at the structural level. This becomes apparent when TRA protein sequences are compared among species. Comparison of the insect phylogeny to a phylogeny based on the TRA protein sequence reveals that its evolution has followed species divergence confirming that conservation lies in function rather than sequence (Fig. 6.3). Strikingly, an alignment of TRA orthologs shows that only the proline and Arg/Ser rich regions are conserved throughout the examined insect species, reflecting their function as splice factor. One additional domain is conserved in Hymenoptera only, a second domain is conserved in all species except \textit{Drosophila} (Kato \textit{et al.}, 2010) and a third domain is conserved in all Diptera (Fig. 6.3). The second domain may function in \textit{tra} autoregulation that is absent in \textit{D. melanogaster} and replaced by \textit{Sxl}. The other two domains are apparently not involved in \textit{tra} splicing but may have other unknown functions.

\textbf{Doublesex}

\textit{Dsx} belongs to a class of DM domain containing genes that are conserved outside the insect class, and regulates sex determination in both vertebrates (Suzuki \textit{et al.}, 2001; Bratus and Slota, 2006; Matsuda \textit{et al.}, 2007; Yoshimoto \textit{et al.}, 2008; Cao \textit{et al.}, 2009; Raghuveer and Senthilkumaran, 2009) and invertebrates (Hodgkin, 2002; Large and Mathies, 2007; Kato \textit{et al.}, 2008). \textit{Tra}, on the other hand, has been identified as a \textit{dsx} splicing factor in insects only (Kato \textit{et al.}, 2010). A comparison of sex determining cascades in different insect groups reveals that diversity essentially starts at the level of regulation of \textit{tra} splicing and that \textit{tra} acts as receptor for various primary signals. These signals are very diverse and include X-chromosome dose (Erickson and Quintero, 2007), a male determining factor on the Y chromosome and/or autosome (Schmidt \textit{et al.}, 1997; Pane \textit{et al.}, 2002), and a feminizing factor on the W chromosome (Fujii and Shimada, 2007) in diploids, as well as complementary sex determination and genomic imprinting sex determination in haplodiploids (Beye \textit{et al.}, 2003; Beukeboom \textit{et al.}, 2007b). Therefore, the conserved part of the insect sex determining cascade must be extended to include \textit{tra}, and from an evolutionary perspective, \textit{tra} appears to serve as the gene around which flexibility in sex determination is manifested.
Figure 6.3: Amino acid alignment and protein sequence tree of tra/fem proteins identified in insect species from three different orders. Upper part: Insect classification on the left is redrawn from Sánchez et al. (2008). On the right is the UPGMA consensus tree of the transformer protein sequences with bootstrap values indicated at the nodes (UPGMA cluster with a Jones-Taylor-Thornton matrix, 1000 bootstraps). In the middle is the protein sequence alignment showing all conserved areas. Green box indicates conserved domain in Diptera, yellow box indicates conserved domain in all species except D. melanogaster and purple box indicates conserved domain in Hymenoptera. Red box indicates shared Arg/Ser domains and blue box the common Pro-rich region. Lower part: Alignment of the conserved domains with similar colors as in the complete protein alignment (upper part). aa = amino acids. From top to bottom organisms and GenBank accession no.: Drosophila melanogaster (AAP49441); Lucilia cuprina (ACS4689); Anastrepha obliqua (ABW04165); Ceratitis capitata (AAM86673); Bactrocera oleae (CAG29243); Musca domestica (ACY40709); Nasonia vitripennis (NP_001128299); Apis mellifera (AVB56235); Bombus terrestris (ABY74329); Melipona compressipes (ABV79891) and Tribolium castaneum (XP_001809947).

Evolution of tra regulation...

In all tra (or fem) containing insects except D. melanogaster, female specific splicing of tra involves an auto regulatory loop, in which the TRA protein is required for female specific splicing of tra pre-mRNA (Pane et al., 2002; Lagos et al., 2007; Concha and Scott, 2009; Gempe et al., 2009; Salvemini et al., 2009; Hediger et al., 2010; Verhulst et al., 2010). Maternal input of tra mRNA or protein into the eggs has been demonstrated for all examined species except D. melanogaster and A. mellifera and has been surmised to start tra auto regulation. In the haplodiploid N. vitripennis it has been shown for the first time that sufficient levels of maternally provided tra mRNA in eggs is required for female development. Knockdown of tra in mothers leads to a diminished amount of maternally provided tra mRNA in eggs and results in diploid males (Verhulst et al., 2010). In Drosophila, Sxl has
been recruited upstream of tra and is female specifically regulated through its own autoregulatory loop. However, Siera and Cline (2008) showed that tra auto regulation may also be ancestral in Drosophila since a positive feedback loop of tra still operates through Sxl, which in turn regulates tra splicing. Tra regulation by X chromosome dose may occur outside the Drosophilidae but most likely in the absence of Sxl. How this is accomplished remains to be investigated. In A. mellifera a duplication of fem has been recruited into the sex determining cascade and initiates female specific splicing of fem transcripts (Hasselmann et al., 2008). Overall, the maternal provision of tra to eggs appears to be an ancestral regulatory mechanism, as all deviations from this system are of recent origin.

Two intriguing questions are how variations in tra regulation can account for the large variety in sex determining mechanisms in insects and how turnovers in signals and genes controlling tra can occur during evolution. A comparison between diploid and haplodiploid sex determination is particularly illustrative as a “flip-over” of tra regulation may lie at the basis of the difference between these two modes of sex determination.

... in diploid insects

The principle of tra regulation in diploid insects is that the paternally inherited genome inhibits female splicing of tra in a variety of ways. A diverse array of primary signals directly or indirectly regulates sex specific splicing of tra. A common theme in a number of dipteran insects is a masculinizing (M) factor on the Y chromosome that is transmitted through males only. M actively blocks the transcription or translation of tra, preventing the autoregulatory loop from establishing in ways that are not yet well understood (Pane et al., 2002; Lagos et al., 2007; Concha and Scott, 2009; Hediger et al., 2010). Thus, the paternally inherited M factor actively inhibits female development in XY individuals. In Drosophila the presence of twice as much X signal elements in XX animals directs female specific transcription of sxl and starts the female specific path of the sex determining cascade (Erickson and Quintero, 2007).

A special case is Lepidoptera in which females are the heterogametic sex (ZW females, ZZ males) (Traut et al., 2007)). As only females contribute a W chromosome containing a feminizing factor, males passively promote male development. In their theoretical treatise on the evolution of sex determination, Pomiankowski et al. (2004) inferred how, based upon initial allelic variation for dsx (dsxM: masculizing factor and dsx+: feminizing factor) conversion to tra regulation can evolve. Assuming that TRA splices only dsx+ into a female form but not dsxM, a mutation creating a stop codon in the tra exon 2 (traS) would be favorable for traS/traS males, as no female DSX is produced. Simulations showed that this could eventually lead to elimination of dsxM and to evolution of female heterogamy for traS/tra+ (Pomiankowski et al., 2004).

Two general rules emerge from comparing the different primary signals of diploid insects. First, a paternally derived genome is always necessary for male development and second, actively or passively, it always prevents the activation
of \textit{tra} (or \textit{Sxl}).

... in haplodiploid insects

A number of insect groups, including thrips (Thysanoptera), beetles (Coleoptera) and all Hymenoptera, have haplodiploid sex determination: males are haploid, develop from unfertilized eggs and only inherit a maternal genome, whereas females are diploid, develop from fertilized eggs and inherit a paternal and a maternal genome. It is therefore impossible for the paternally inherited genome to have a masculinizing effect as in diploids. Instead, the paternal genome must have acquired a complete reversal in sex determining function, i.e. be feminizing rather than masculinizing.

Until recently, knowledge about primary signals in haplodiploid species was limited to complementary sex determination (CSD) in which gender is determined by the allelic state of the \textit{complementary sex determiner (csd)} gene. Although CSD has been inferred for over 60 hymenopterans (van Wilgenburg \textit{et al.}, 2006), the \textit{csd} gene has been characterized in the honey bee only (Beye \textit{et al.}, 2003). Females are heterozygous and males hemizygous at this locus, but the biochemical details of CSD function are not yet completely known (Hasselmann \textit{et al.}, 2008; Gempe \textit{et al.}, 2009). Interestingly, the CSD primary signal can also be based on multiple loci (ml-CSD) (de Boer \textit{et al.}, 2008). However, in all CSD cases the paternally contributed genome provides the second \textit{csd} allele that is required for female development.

In another hymenopteran, \textit{Nasonia}, CSD has been ruled out as the primary signal (Werren \textit{et al.}, 2010). In \textit{Nasonia} female specific \textit{tra} is maternally provided to eggs. In embryos from fertilized eggs early zygotic expression of \textit{tra} is higher than in embryos from unfertilized eggs, which initiates an autoregulatory loop of \textit{tra} and results in female \textit{dsx} splicing (Verhulst \textit{et al.}, 2010). In embryos from unfertilized eggs no early zygotic expression of \textit{tra} occurs and the autoregulatory loop does not establish, leading to male specific \textit{tra} and \textit{dsx} splicing. The difference in zygotic \textit{tra} expression cannot be explained by masculinizing factors. Instead, the \textit{tra} gene, or a trans acting factor that influences \textit{tra} expression, on the maternal genome is rendered inactive by maternal imprinting. In an unfertilized egg, only this maternally imprinted gene is present that prevents \textit{tra} transcription and precludes the autoregulatory loop. In a fertilized egg, both a maternal and a paternal genome are present. The paternal set has an active, non silenced gene so that \textit{tra} will be transcribed enabling the maternally provided \textit{tra} mRNA to start autoregulation of \textit{tra}, eventually leading to female development (Verhulst \textit{et al.}, 2010). A mutant strain of \textit{N. vitripennis} that produces gynandromorphs and females from haploid unfertilized eggs (Beukeboom \textit{et al.}, 2007a; Kamping \textit{et al.}, 2007) may be explained by incomplete imprinting in the maternal germ line.

The honeybee and \textit{Nasonia} results indicate that, in contrast to diploids, the paternally inherited genome is always necessary for female development in haplodiploids and, actively or passively, promotes the activation of \textit{tra}. 

92
Conclusions and outlook

Much has been learned about sex determination in mammals (Wallis et al., 2008) and plants (Tanurdzic and Banks, 2004; Ming et al., 2007; Martin et al., 2009), but comparative work on a variety of insect species has been particularly fruitful for understanding how sex determination regulation evolves. Twenty-five years ago Nöthiger and Steinmann-Zwicky (1985) proposed that sex determination in all insects is based on a single principle. We can conclude that these authors were partly correct. The sex specific regulation of dsx splicing by tra appears to constitute a conserved gene axis in all insects. Clearly, the central gene around which diversity evolves is not Sxl, as was suggested from studies in Drosophila, but tra. A striking example of the central role of tra in the evolution of insect sex determination is the complete reversal in the paternal regulation of tra upon the separation of Hymenoptera and Diptera.

Although Drosophila has XX-XY sex determination, its processing of the primary signal and regulation of tra is different from all other flies with this mode of sex determination. A number of other insect groups likely rely on sex chromosome dose as primary signal, such as species with XX-XO sex determination (e.g. grasshoppers (Orthoptera), ZO-ZZ sex determination (e.g. some Lepidoptera, (Traut et al., 2007)) and paternal X-chromosome inactivation (e.g. coccids (Homoptera) and Sciarid flies (Sánchez, 2008). Whether and how tra regulation occurs in these groups remains an interesting unanswered question. In general, a broader taxonomic screen of how primary signals are processed by tra would be worthwhile as our current knowledge is virtually restricted to Diptera and Hymenoptera. Exploiting next generation sequencing technology will greatly expedite such an endeavor.

Acknowledgments

We thank Albert Kamping and Daniel Bopp for fruitful discussions on insect sex determination.