HEINRICH HEINE
UNIVERSITÄT DÜSSELDORF

The control of the honeybee sex determination pathway by alternative splicing

Inaugural-Dissertation

zur Erlangung des Doktorgrades der

Mathematisch-Naturwissenschaftlichen Fakultät

der Heinrich-Heine-Universität Düsseldorf

vorgelegt von

Inga Nissen

aus Haan (Rheinland)

Düsseldorf, März 2012

aus dem Institut für Evolutionsgenetik

der Heinrich-Heine-Universität Düsseldorf

Gedruckt mit der Genehmigung der

Mathematisch-Naturwissenschaftlichen Fakultät

der Heinrich-Heine-Universität Düsseldorf

Referent: Prof. Dr. Martin Beye

Korreferent: Prof. Dr. Laura Rose

Tag der mündlichen Prüfung: 20.4.2012

Erklärung

Die hier vorgelegte Dissertation habe ich eigenständig und ohne unerlaubte Hilfsmittel angefertigt. Die Dissertation wurde in der vorgelegten oder in ähnlicher Form noch bei keiner anderen Institution eingereicht. Ich habe bisher keine erfolglosen Promotionsversuche unternommen.

Inga Nissen Düsseldorf, den 23. März 2012

Contents

Chapter I

Allgemeine Einleitung

Im Laufe der Evolution haben sich in der Natur viele verschiedene Mechanismen entwickelt, die die Geschlechtsdifferenzierung in den unterschiedlichen Organismen regulieren. Insbesondere die initialen Signale für die Einleitung der männlichen oder weiblichen Entwicklung weisen ein breites Spektrum an Determinationsfaktoren auf, wozu unter anderem Umwelteinflüsse als auch genetische Faktoren zählen. Während bei Reptilien beispielsweise die Temperatur der entscheidende Determinationsfaktor ist (Bull, 1980), findet man u.a. bei Säugetieren, Vögeln, der Taufliege *Drosophila melanogaster* und dem Fadenwurm *Caenorhabditis elegans* geschlechtschromosomal gekoppelte Systeme (Erickson and Quintero, 2007; Cline and Meyer, 1996; Graves, 1995; Schartl, 2004). In etwa 20% aller Tierarten wird die haplodiploide Geschlechtsbestimmung verwirklicht. Prinzipiell entwickeln sich dabei aus unbefruchteten Eizellen männliche Tiere, die nur einen Chromosomensatz aufweisen und aus befruchteten Eizellen weibliche Tiere mit zwei Chromosomensätzen (Bull, 1983).

I.I Die Geschlechtsbestimmung bei der Honigbiene *Apis mellifera*

Die Honigbiene weist eine besondere Variante der haplodiploiden Geschlechtsbestimmung auf: die haplodiploide komplementäre Geschlechtsdetermination. Hier entscheidet der allelische Zustand am geschlechtsbestimmenden Lokus (Sex Determining Locus - SDL) über das Geschlecht des Individuums (Whiting, 1943). Im Falle der Honigbiene handelt es sich dabei um das multiallelische Gen *complementary sex determiner* (*csd*), wobei man bisher 15 verschiedene Allele nachweisen konnte (Hasselmann et al., 2008; Hasselmann and Beye, 2004). Liegt *csd* am SDL im heterozygoten Zustand vor, entwickeln sich Weibchen, während bei hemi- oder homozygotem Allelzustand die männliche Entwicklung eingeleitet wird (Abb. 1). In der Natur kommen allerdings nur diploide Weibchen und haploide Männchen vor, da die unfruchtbaren diploiden Drohnen im frühen Larvenstadium von den Arbeiterinnen gefressen werden.

Wie durch RNA-Interferenz (RNAi) Experimente gezeigt werden konnte, ist die heterozygote Form der aktive Zustand von *csd*. In genetisch weiblichen Embryonen führt die Repression der *csd* Transkription zur Entwicklung von Männchen. Genetisch männliche Embryonen werden durch das Ausschalten des Gens *csd* in ihrer Entwicklung nicht beeinflusst. (Gempe et al., 2009; Hasselmann et al., 2008). Die einzelnen *csd*-Allele kodieren für SR-ähnliche

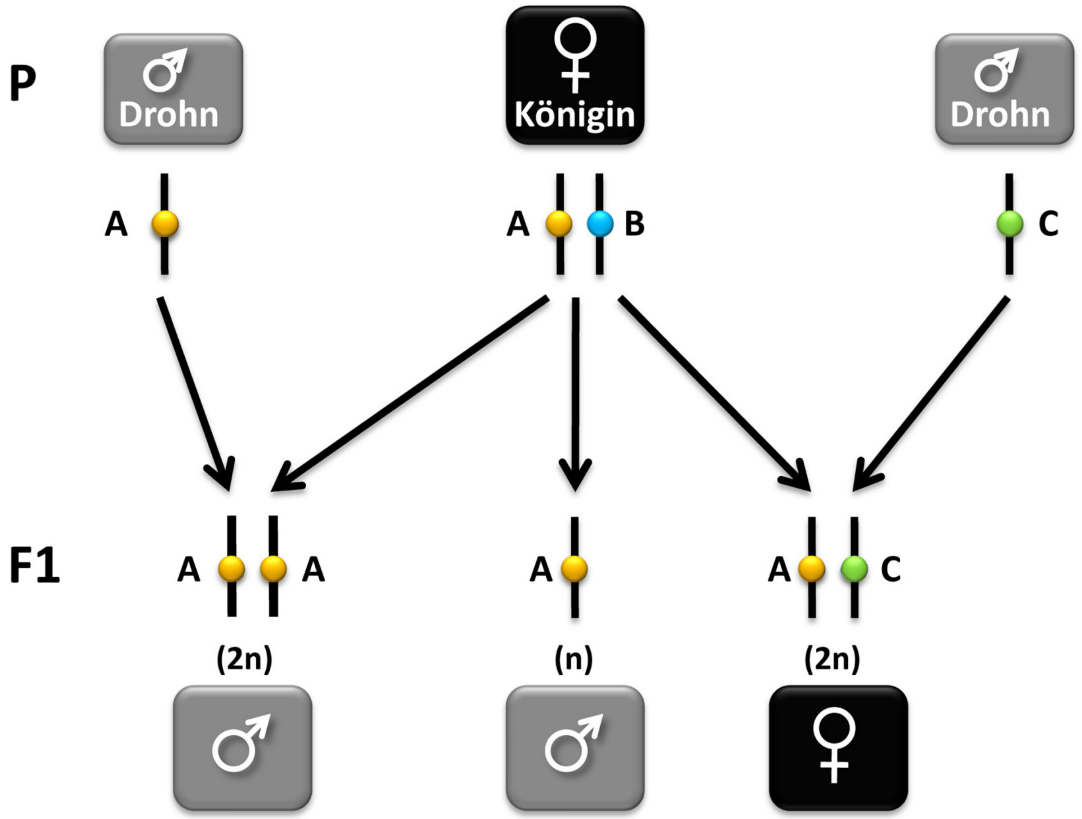

Abb. 1: Komplementäre Geschlechtsbestimmung der Honigbiene
In der Honigbiene entscheidet die allelische Zusammensetzung des *csd*-Gens am geschlechtsbestimmenden Lokus über das Geschlecht der Nachkommen in der F1-Generation. Diploide (2n) Nachkommen mit zwei verschiedenen *csd*-Allelen (heterozygoter Allelzustand; A/C oder B/C) sind weiblich. Haploide (n) Nachkommen mit nur einem *csd*-Allel (A oder B; hemizygoter Allelzustand) oder zwei identischen *csd*-Allelen (homozygoter Allelzustand; hier: A/A) sind männlich.
csd: complementary sex determiner; (2n) diploid; (n) haploid; P: Generation der Eltern; F1: Nachkommen der ersten Generation

Proteine, die sich in Sequenz und Molekulargewicht unterscheiden. Hauptverantwortlich für diese Unterschiede ist die sogenannte Hypervariable Region (HVR) am Carboxy-terminalen (C-terminalen) Ende von Csd. Zusätzlich zur HVR besitzen Csd-Proteine eine Prolinreiche Region (PR) und eine Arginin-/Serin-reiche Domäne im C-Terminus (Beye et al., 2003; Beye, 2004), die vermutlich für Protein-Protein-Interaktionen wichtig sind.

Das Gen *csd* ist durch eine Genduplikation aus dem Gen *feminizer (fem)* hervorgegangen und ist ihm in der geschlechtsbestimmenden Kaskade übergeordnet (Abb. 2A). Im aktiven Zustand beeinflusst Csd in seiner Eigenschaft als SR-ähnliches Protein das geschlechtsspezifische Spleißen des *fem*-Gens und leitet so den weiblichen Entwicklungsweg

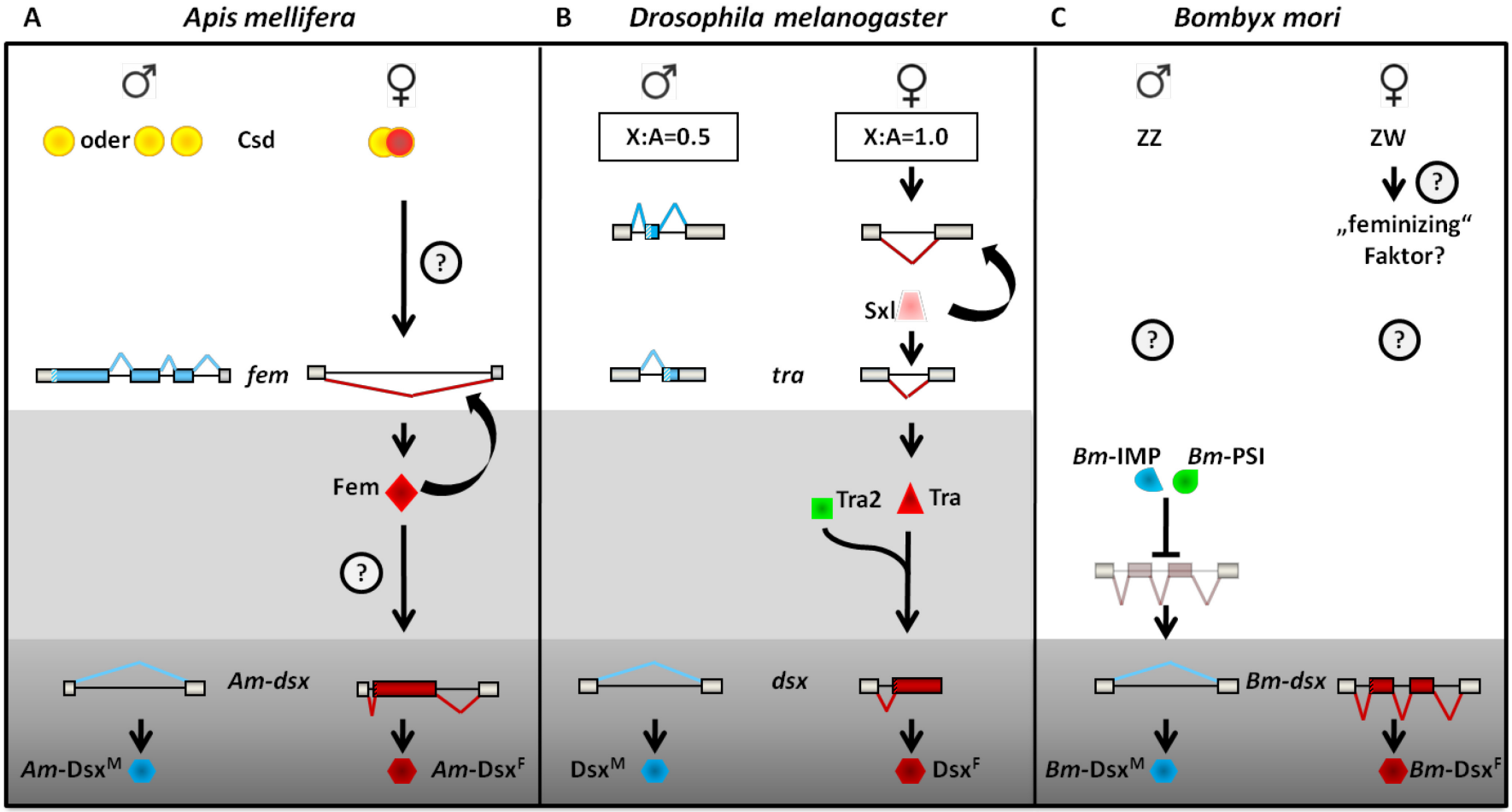

Abb. 2: Vergleichende Darstellung von Modellen der Geschlechtsdeterminationskaskaden von *Apis mellifera* (A), *Drosophila melanogaster* (B) und *Bombyx mori* (C). Die Exon-Intronstruktur der verschiedenen Gene zeigt nur einen Ausschnitt der tatsächlichen Transkripte. Männchen- bzw. weibchenspezifische Exons (Boxen) und Introns (Linien), sowie männchen- oder weibchenspezifisch exprimierte Proteine (dargestellt durch verschiedene Symbole) sind in blau (männl.), bzw. rot (weibl.) dargestellt. Exons die bei beiden Geschlechtern vorkommen sind grau dargestellt. Proteine, die in beiden Geschlechtern exprimiert werden, sind durch grüne Symbole gekennzeichnet. Fragezeichen zeigen Punkte in der Kaskade an, bei denen die Art der Regulation (direkt, indirekt, mögliche Beteiligung weiterer Faktoren, etc.) bislang unbekannt ist. Zur Erläuterung der einzelnen Stufen und (Suzuki, 2010; Suzuki et al., 2010)Elemente der Kaskaden siehe Text. Modell von *A. mellifera* modifiziert nach (Gempe, 2009), Modell von *D. melanogaster* modifiziert nach (Cline & Meyer, 1996); Modell von *B. mori* erstellt nach (Suzuki, 2010; Suzuki et al., 2010)

ein. Das männchenspezifische *fem*-Transkript setzt sich aus 12 Exons zusammen und kodiert für ein nicht funktionales Protein (Abb. 3A). In Weibchen wird ein alternativer Spleißdonor in

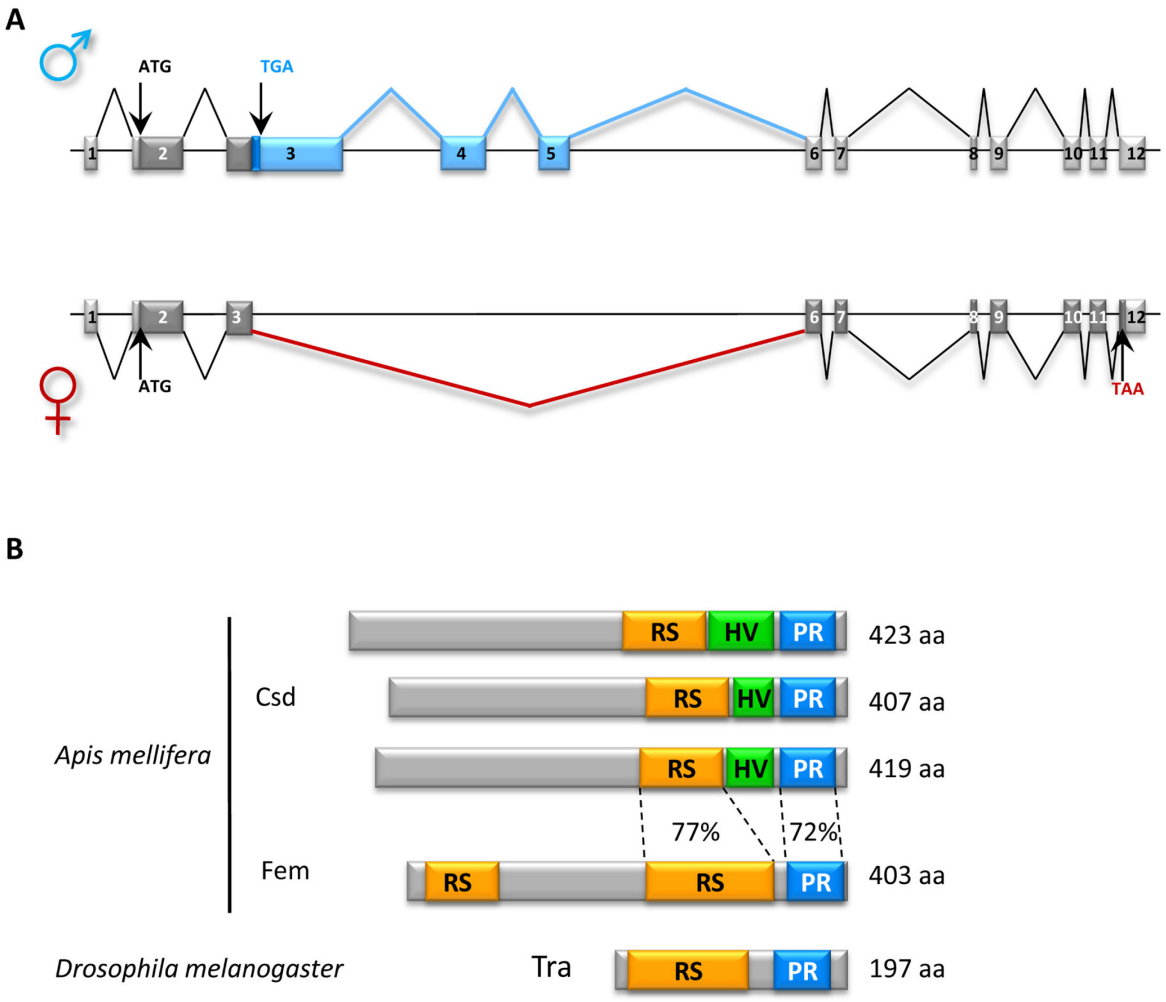

Abb. 3: Schematische Darstellung der alternativen Spleißvarianten des *feminizer* (*fem*) Gens der Honigbiene (A) und der Domänenstruktur der Proteine Complementary Sex Determiner (Csd), Feminizer (Fem) und Transformer (Tra) (B) (nach Hasselmann, 2008).
(A) *Das Gen fem* wird geschlechtsspezifisch gespleißt. Das Startkodon (ATG) in Exon 2 wird in beiden Geschlechtern benutzt. Männliches *fem*-Transkript (oben) besteht aus 12 Exons und umfasst einen offenen Leserahmen (ORF) von 510 bp, der für kein funktionales Protein kodiert. In Weibchen (unten) wird ein alternativer Spleißdonor in Exon 3 benutzt, wodurch direkt auf Exon 6 gespleißt wird. Durch die damit verbundene Entfernung des männchenspezifischen Stopkodons (TGA), wird ein Stopkodon am Anfang von Exon 12 benutzt (TAA). Das weibliche *fem*-Transkript hat einen ORF von 1206 bp und kodiert für das SR-ähnliche-Protein Fem. **(B)** Die Domänenstruktur des Fem-Proteins ist im Vergleich zu drei Csd-Proteinen und dem Protein Transformer (Tra) aus *D. melanogaster* dargestellt. Die Csd-Proteine besitzen eine RS-Domäne (Arginin/Serinreiche-Domäne), eine Hypervariable Region (HV) und eine prolinreiche Region (PR). Das Fem-Protein besitzt im Gegensatz zum Csd-Protein eine weitere RS-reiche Domäne, dafür aber keine Hypervariable Region. Das Protein Transformer (Tra) in *D. melanogaster* ist nur halb so groß wie das Fem-Protein und weist nur eine RS-Domäne und eine PR-Domäne auf. Arginin/Serinreiche Domänen (RS) sind in gelb dargestellt, prolinreiche Domänen (PR) in blau und die Hypervariable Region (HV) in grün. Exons sind durch Boxen, Introns durch Linien dargestellt. Kodierende Bereiche sind dunkelgrau, bzw. dunkelblau dargestellt, nicht kodierende hellgrau, bzw. hellblau.

Exon 3 benutzt; der hintere Teil von Exon 3, sowie die Exons 4 und 5 werden herausgespleißt. Dadurch wird auch das Stoppcodon entfernt, das in Männchen die vorzeitige Termination der Translation bewirkt (Abb. 3A). Das weibchenspezifische Transkript kodiert wie *csd* für ein SR-ähnliches Protein und nimmt eine wichtige Funktion in der weiteren Geschlechtsdeterminationskaskade ein (Hasselmann et al., 2008; Gempe et al., 2009). Ein Vergleich der Aminosäuresequenzen von Csd und Fem zeigt, dass sie über 70% Sequenzidentität aufweisen (Abb. 3B) (Hasselmann et al., 2008). Das Fem-Protein besitzt wie Csd-Proteine eine RS-reiche und eine PR-Domäne im C-terminalen Bereich. Das Protein weist keine HVR-Region auf und besitzt eine zusätzliche RS-reiche Domäne im N-terminalen Bereich (Gempe et al., 2009) (Abb. 3B). In Weibchen beeinflusst das Protein Fem das geschlechtsspezifische Spleißen der *Apis mellifera doublesex* (*Am-dsx*) prä-mRNA. Das weibchenspezifische *Am-dsx*-Transkript setzt sich aus 7 Exons zusammen, während in Männchen das weibchenspezifische Exon 5 übersprungen und direkt von Exon 4 auf Exon 6 gespleißt wird (Abb. 4A). Beide Transkripte kodieren für Transkriptionsfaktoren, die eine zinkfingerartige Domäne (DM-Domäne) besitzen und sich nur durch ihr C-terminales Ende unterscheiden (Cho et al., 2007) (Abb. 4B). Zusätzlich zu seiner Funktion als Spleißregulator von *Am-dsx* ist das Fem-Protein im Verlauf der Ontogenese für die Aufrechterhaltung des weibchenspezifischen Status verantwortlich, indem es autoregulativ für das Spleißen der eigenen prä-mRNA in den weibchenspezifischen Zustand sorgt (Gempe et al., 2009).

I.II Die Geschlechtsdetermination bei der Taufliege *Drosophila melanogaster*

Die genauen molekularen Mechanismen, die zur Einleitung der weibchen- oder männchenspezifischen Entwicklung des Honigbienenembryos führen, sind auf den drei Regulationsebenen *csd, fem* und *Am-dsx* noch weitgehend unbekannt. Am besten untersucht wurde die Geschlechtsdeterminationskaskade der Taufliege *Drosophila melanogaster* (Abb. 2B). Das initiale Signal ist hier das Verhältnis von Geschlechtschromosomen zu Autosomen (X:A), bzw. die Dosis an Numeratorgenen, die auf dem X-Chromosom lokalisiert sind, im Verhältnis zu Denominatorgenen, die auf den Autosomen liegen (Erickson and Quintero, 2007; Cline and Meyer, 1996; Schütt and Nöthiger, 2000). Je nach Dosierung der X-Chromosomen wird nun entweder die männliche (X:A = 0,5) oder die weibliche Entwicklung (X:A = 1) eingeleitet. Während der konstitutive

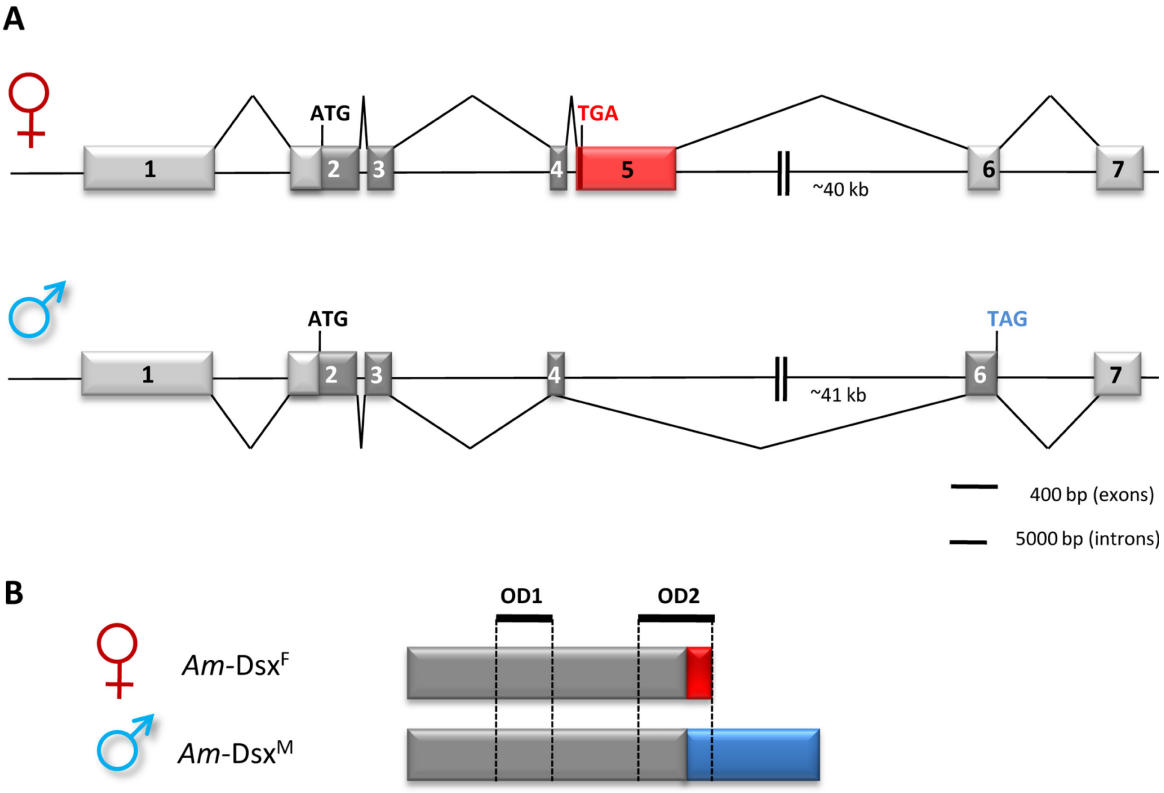

Abb. 4: Schematische Darstellung der alternativen Spleißvarianten des *Apis mellifera doublesex* (*Am-dsx*) Gens (A) und der *Am*-Dsx Proteine (B) (nach (Cho, 2007) und eigenen Daten (nicht publiziert)).

(A) Das Gen *Am-dsx* wird geschlechtsspezifisch gespleißt. Das Startkodon (ATG) in Exon 2 wird in beiden Geschlechtern benutzt. Weibliches *Am-dsx*-Transkript (oben) besteht aus 7 Exons und umfasst einen offenen Leserahmen (ORF) von 831 bp. Das weibchenspezifische Exon 5 (rot) wird in Männchen (unten) herausgespleißt, wodurch das weibchenspezifische Stopkodon (TGA) entfernt wird. Stattdessen wird ein Stopkodon (TAG) am Ende von Exon 6 benutzt, was zu einem ORF von 1008 bp führt. **(B)** Beide Transkripte kodieren für Transkriptionsfaktoren, die zwei Oligomerisierungsdomänen (OD; schwarze Balken) enthalten. Die aminoterminal gelegene OD1 ist in beiden Geschlechtern identisch und besteht aus einer atypischen Zinkfingerdomäne im Amino-Terminus, während sich die carboxyterminal gelegene OD2 zum Teil im geschlechtsspezifischen Proteinabschnitt (rot, bzw. blau dargestellt) befindet. Exons sind durch Boxen, Introns durch Linien dargestellt. Kodierende Bereiche sind dunkelgrau, bzw. dunkelrot dargestellt, nicht kodierende hellgrau, bzw. hellrot.

sex-lethal- (*sxl*) Promotor Sxl_{Pm} in beiden Geschlechtern aktiv ist und das komplette *sxl*-Transkript produziert wird, kann der weibchenspezifische Promotor Sxl_{Pe} nur beim richtigen Verhältnis von Numerator- zu Denominatorgenen aktiviert werden (Parkhurst and Ish-Horowicz, 1992). Dieser erzeugt eine verkürzte *sxl* prä-mRNA, der das Stoppcodon fehlt, das in Männchen zur vorzeitigen Termination der Translation führt. Das weibchenspezifische Transkript kodiert für das Sxl-Protein, das nach Deaktivierung des Promotors Sxl_{Pe} durch

positive Rückkopplung über Sxl_{Pm} die Synthese von weibchenspezifischen Transkript aufrechterhält (Keyes et al., 1992; Bell et al., 1991).

Sobald das Sxl-Protein in weiblichen Embryonen exprimiert wird, nimmt es die Rolle des Schlüsselregulators in der weibchenspezifischen Geschlechtsdetermination der Taufliege ein. In seiner Eigenschaft als RNA-Bindeprotein ist das Sxl-Protein nicht nur für das autoregulatorische Spleißen seiner eigenen prä-mRNA verantwortlich, sondern auch für das geschlechtsspezifische Spleißen der *male-specific lethal-2* (*msl-2*) und *transformer* (*tra*) prä-mRNAs.

MSL-2-Proteine sind in Männchen an der X-chromosomalen Dosiskompensation beteiligt, indem sie zusammen mit anderen Genen für die Hypertranskription der X-chromosomalen Gene sorgen (Gilfillan et al., 2004; Bashaw and Baker, 1995). Die Regulation erfolgt über das Spleißen von Intron 2 der *msl-2* prä-mRNA. In Weibchen bindet das Sxl-Protein zunächst an mehrere RNA-Bindestellen in Intron 2 und verhindert somit dessen Entfernung. Da Intron 2 Teil der 5'UTR ist, wird durch diesen Spleißprozess der offene Leserahmen (ORF – open reading frame) von *msl-2* nicht beeinflusst, jedoch die Proteinexpression im Cytoplasma: im Anschluss an den Spleißprozess bindet das Sxl-Protein an Bindestellen im 5'- und 3'UTR von *msl-2* Transkripten und inhibiert die Translation, wodurch im Weiteren die Hypertranskription X-chromosomaler Gene verhindert wird. In Männchen wird das Sxl-Protein nicht exprimiert, Intron 2 wird herausgespleißt und die Expression von MSL-2-Proteinen wird ermöglicht (Bashaw and Baker, 1997; Gebauer et al., 1999; Kelley et al., 1997).

Auch für die geschlechtsspezifische Expression von Tra-Protein und somit für die Geschlechtsdetermination, ist das Sxl-Protein von essentieller Bedeutung. In beiden Geschlechtern besteht die prä-mRNA von *tra* aus 3 Exons, jedoch unterscheidet sich das Spleißschema durch Verwendung unterschiedlicher Spleißakzeptoren in Exon 2. Dabei konkurriert Sxl in Weibchen mit dem Spleißfaktor U2AF um die Bindung an der männchenspezifischen Spleißstelle und wirkt somit als Repressor (Valcárcel et al., 1993). Daraufhin kommt es zur Verwendung einer schwächeren Spleißstelle in 3'-Richtung, wodurch das männchenspezifische Stoppcodon entfernt wird (Boggs et al., 1987). In Weibchen kodiert das *tra*-Transkript für ein SR-ähnliches Protein mit einer RS-Domäne und einer C-terminalen prolinreichen Region (Boggs et al., 1987). Tra ist ein Ortholog des zuvor

erwähnten Fem-Proteins der Honigbiene (Abb. 3B), und es wird angenommen, dass die beiden Proteine einen gemeinsamen evolutionären Vorfahren haben (Hasselmann et al., 2008).

Anders als Sxl wirkt Tra als aktivierender Spleißregulator beim weibchenspezifischen Spleißen der *doublesex* (*dsx*) prä-mRNA. In Abwesenheit von Tra kommt es zur Bildung des männchenspezifischen Transkripts dsx^M, das sich aus den Exons 1-3, sowie 5 und 6 zusammensetzt. In Weibchen wird durch die Expression von Tra ein schwacher Spleißakzeptor aktiviert, wodurch von Exon 3 direkt auf Exon 4 gespleißt wird. Am Ende von Exon 4 befindet sich eine Polyadenylierungsstelle, die die Transkription der weibchenspezifischen mRNA dsx^F beendet (Hedley and Maniatis, 1991; Burtis and Baker, 1989). Da Tra eine RNA-Bindedomäne fehlt, kann der Spleißregulator den weiblichen Spleißprozess nicht alleine einleiten und wird von einer Reihe von Spleißfaktoren unterstützt, darunter das SR-Protein Transformer2 (Tra2). Dabei assoziieren Tra und Tra2 zusammen mit anderen SR-Proteinen zu einem Spleißregulationskomplex und binden über die RNA-Bindedomäne von Tra2 an sechs Sequenzwiederholungen eines exonischen Spleißenhancers (ESE-exonic splicing enhancer) und ein purinreiches Element (PRE) in Exon 4 (Tian and Maniatis, 1993; Lynch and Maniatis, 1996, 1995). Dies führt zur Aktivierung des etwa 300 Nukleotide strangaufwärts gelegenen schwachen Spleißakzeptors von Exon 4 und dem Einschluss des weibchenspezifischen Exons durch die Spleißmaschinerie (Sciabica and Hertel, 2006). Doch sind die Tra/Tra2-Bindestellen nicht nur essentiell für die Initiation von weibchenspezifischem Spleißen, sondern auch für die Verwendung der weibchenspezifischen Polyadenylierungsstelle (Hedley and Maniatis, 1991; Burtis and Baker, 1989). Die *dsx*ESE Sequenzen bestehen aus einem charakteristischen, 13 Nukleotid langen Motiv (TC(T/A)(T/A)C(A/G)ATCAACA) (Nagoshi and Baker, 1990; Inoue et al., 1992), das mittlerweile ebenfalls in *dsx*-Genen anderer Dipteren identifiziert wurde (Hediger et al., 2004; Saccone et al., 2008; Lagos et al., 2005). Beim Seidenspinner *Bombyx mori* als auch bei der Honigbiene, ist dieses Motiv nicht vorhanden (Cho et al., 2007; Suzuki et al., 2001).

Das weibliche, wie auch das männliche *dsx*-Transkript kodieren für Transkriptionsfaktoren, die geschlechtsspezifisch die Aktivität der finalen, phänotypprägenden Gene der somatischen Geschlechtsbestimmungskaskade kontrollieren (Burtis et al., 1991). Dsx^F und Dsx^M besitzen zwei Oligomerisierungsdomänen (OD1 und OD2), die wichtige Schnittstellen

für DNA- und Proteininteraktionen sind. Die OD1 ist in beiden Proteinen identisch und besteht aus einer atypischen Zinkfinger-Domäne (DM-Domäne). Durch Verwendung der unterschiedlichen Spleißstellen unterscheiden sich die beiden Proteine jedoch im C-terminalen Bereich und somit in der Sequenz der OD2, wodurch die weibchen- und männchenspezifische Regulation der Zielgene gesteuert wird (An et al., 1996; Coschigano and Wensink, 1993; Burtis et al., 1991; Cho and Wensink, 1997; Sánchez et al., 2001; Keisman et al., 2001; Williams et al., 2008).

I.III Die Geschlechtsdetermination beim Seidenspinner *Bombyx mori*

In allen diesbezüglich untersuchten Lepidopteren hängt die Geschlechtsdetermination von der Zusammensetzung der Geschlechtschromosomen ab. Während Männchen homogametisch sind (ZZ), besitzen Weibchen entweder nur ein Geschlechtschromosom (Z0) oder zwei unterschiedliche Geschlechtschromosomen (ZW). Es wird angenommen, dass es sich bei dem Z0/ZZ–System um das ursprüngliche handelt (Traut and Marec, 1996), wobei vermutlich die unterschiedliche Dosis von einem oder mehreren Genen auf dem Z-Chromosom über das Geschlecht bestimmt. In *Bombyx mori* findet man das ZW/ZZ–System (Abb. 2C). Hier wird ein „feminizing"-Faktor auf dem W-Chromosom vermutet (Hashimoto, 1933), der jedoch noch nicht genauer bestimmt werden konnte. Durch Anwesenheit des W-Chromosoms werden regulative Vorgänge ausgelöst, die zur weiblichen Entwicklung führen. Am Ende dieser Kaskade steht wie bei *Apis mellifera* und *Drosophila melanogaster* das Gen *doublesex* (*Bm-dsx*), das ebenfalls geschlechtsspezifisch gespleißt wird und für Transkriptionsfaktoren mit einer DM-Domäne kodiert (Ohbayashi et al., 2001). Bei Expression eines *Bm-dsx* Minigens in HeLa-Zellextrakt tritt anders als bei *D. melanogaster,* das weibchenspezifische Spleißen als voreingestellter Weg auf (Funaguma et al., 2005). In Männchen wird das weibchenspezifische Spleißen durch die Spleißinhibitoren *Bm*PSI (*Bombyx mori* P-element somatic inhibitor) und *Bm*IMP (*Bombyx mori* IGF-II mRNA binding protein) verhindert, was sowohl durch RNAi Experimente in männlichen Zellen, als auch durch Ko-Expression der beiden Proteine in weiblichen Zellen gezeigt wurde (Suzuki et al., 2010, 2008). *Bmpsi* wird nicht geschlechtsspezifisch exprimiert und kodiert für ein Ortholog des zuerst in *D. melanogaster* identifizierten PSI-Proteins. Das Protein enthält wie PSI vier RNA-Bindemotive des KH- (hnRNP K-homology) Typs, sowie im C-Terminus zwei

hintereinander auftretende Wiederholungen eines 30 Aminosäure langen Motivs, der PSI-AB Domäne (Suzuki et al., 2008). Das Gen *Bmimp* ist auf dem Z-Chromosom lokalisiert und wird männchenspezifisch exprimiert. Es kodiert für ein Protein mit vier KH-Domänen und ein RNA-Bindemotiv des RRM-Typs (Suzuki et al., 2010). Sowohl *Bm*-PSI als auch *Bm*-IMP binden im weibchenspezifischen Exon der *Bm-dsx* prä-mRNA spezifisch an das CE1 Motiv, eine 20 Nukleotid lange, UA-reiche Sequenz, die drei Wiederholungen eines UAA-Motivs enthält (Suzuki et al., 2010, 2008). Der partielle Austausch dieser Region gegen ein Sequenzmotiv aus dem Gen P25 führte zu weibchenspezifischem Spleißen in männlichen Zellen (Suzuki et al., 2008). Neben ihrer Bindung an das CE1-Element, interagieren die Proteine *Bm*-IMP und *Bm*-PSI auch miteinander. Dabei wird durch Bindung des *Bm*-IMP-Proteins an *Bm*-PSI-Protein die Bindeaffinität von *Bm*-PSI an das CE1-Element erhöht, was im Endeffekt vermutlich das männchenspezifische Spleißen der *Bm-dsx* prä-mRNA fördert (Suzuki et al., 2010).

Diese Erkenntnisse, sowie die Abwesenheit von typischen Tra/Tra2-Bindestellen, lassen vermuten, dass das *Bombyx mori* Ortholog des Tra2-Proteins (*Bm*-Tra2) nicht in die Geschlechtsbestimmung involviert ist.

I.IV Tra2 – ein wichtiger Faktor der Geschlechtsdetermination

Orthologe des *tra2*-Gens wurden in den Insektenordnungen Diptera, Lepidoptera, Hymenoptera und Coleoptera isoliert oder annotiert (Burghardt et al., 2005; Salvemini et al., 2009; Sarno et al., 2010; Amrein et al., 1990; Mattox et al., 1996; Niu et al., 2005), unter anderem auch in der Honigbiene *Apis mellifera* (Dearden et al., 2006). Alle *tra2*-Transkripte kodieren durchweg für SR-Proteine und besitzen neben einer RNA-Bindedomäne (RBD) ein oder meist zwei RS-reiche Domänen (Amrein et al., 1990; Mattox et al., 1996). Bei mehreren Dipteren (*Drosophila melanogaster, Musca domestica, Ceratitis capitata, Anastrepha obliqua*) ist das Gen *Am-tra2* in die Geschlechtsdeterminationskaskade involviert (Burghardt et al., 2005; Salvemini et al., 2009; Sarno et al., 2010; Amrein et al., 1990; Mattox et al., 1996).

In *Drosophila melanogaster* wurden vier alternative Spleißprodukte von *tra2* beschrieben, die für drei Isoformen des Tra2-Proteins kodieren ($Tra2^{264}$, $Tra2^{226}$ und $Tra2^{179}$) (Amrein et al., 1990). Das Protein Tra2 ist in den geschlechtsspezifischen Spleißprozess der *doublesex*

prä-mRNA involviert und aktiviert dort den weibchenspezifischen Spleißakzeptor. Diese Aufgabe übernehmen insbesondere die längeren Isoformen $Tra2^{264}$ und $Tra2^{226}$, die konstitutiv in somatischem Gewebe und in weiblichen Keimzellen exprimiert werden (Amrein et al., 1990; Mattox et al., 1990, 1996). Neben seiner Aufgabe in der weibchenspezifischen Geschlechtsdetermination ist das Protein $Tra2^{226}$ essentiell für die Spermatogenese in männlichen Keimzellen. Dort wird es für die Prozessierung der *exuperantia* (*exu*) und der *alternative-testes-transcript* (*att*) mRNA benötigt (Mattox et al., 1996; Hazelrigg and Tu, 1994; Madigan et al., 1996). Da die männliche Fertilität auch von der richtigen Dosierung des $Tra2^{226}$-Proteins abhängt, kontrolliert das SR-Protein seine Expression in den männlichen Keimzellen über einen autoregulativen Mechanismus. Dabei reprimiert das $Tra2^{226}$- Protein die Entfernung des M1 Introns, was zur Entstehung eines Transkripts führt, das aufgrund des Fehlens der N-terminalen RS-Domäne für die verkürzte und vermutlich funktionslose Isoform $Tra2^{179}$ kodiert (Mattox et al., 1996; McGuffin et al., 1998; Mattox and Baker, 1991; Mattox et al., 1990; Amrein et al., 1990).

Eine weitere Rolle spielt das Tra2-Protein im Nervensystem von *D. melanogaster*, wo es zusammen mit Tra-Protein den weibchenspezifischen Spleißdonor der *fruitless* prä-mRNA aktiviert. Dieser Prozess wird durch Bindung von Tra- und Tra2-Proteinen an drei Wiederholungen desselben 13 Nukleotid langen Sequenzmotivs reguliert, das auch beim *doublesex* Spleißen essentiell für die Tra/Tra2-Bindung ist (Ryner et al., 1996; Heinrichs et al., 1998).

Auch in den Dipteren *Musca domestica, Ceratitis capitata* und *Anastrepha obliqua* wurde die Rolle des Gens *tra2* bei der Geschlechtsbestimmung untersucht. Durch Ausschalten der *tra2*-Transkription mittels RNAi wurde ein Einfluss auf das weibchenspezifische Spleißen der prä-mRNA von *tra* sowie von *dsx* beobachtet. Für alle drei Arten wurde postuliert, dass das Tra2-Protein sowohl eine direkte Rolle bei der (Auto-)regulation des Spleißprozesses der *tra* prä-mRNA spielt, als auch beim Spleißen der *dsx* prä-mRNA. Diese These wird durch die Anwesenheit von Tra/Tra2-Bindemotiven in den weibchenspezifischen Exons der jeweiligen *dsx* prä-mRNA unterstützt (Hediger et al., 2010; Burghardt et al., 2005; Salvemini et al., 2009; Sarno et al., 2010). Für das *tra2*-Gen von *Bombyx mori* (*Bm-tra2*) wurde bisher noch keine geschlechtsspezifische Funktion beschrieben, wobei der unterschiedliche Mechanismus des *Bm-dsx* Spleißens, sowie das

Fehlen von Tra/Tra2-Bindestellen in *Bm-dsx* gegen eine ähnliche Funktion wie der von *tra2* in *D. melanogaster* spricht (Suzuki et al., 2001, 2010, 2008; Niu et al., 2005). In der Honigbiene wurde ebenfalls ein *tra2*-Gen beschrieben (Dearden et al., 2006), allerdings ist bisher weder etwas über die Anzahl möglicher Spleißvarianten bekannt, noch über dessen Funktion. Die für *D. melanogaster* typischen Tra/Tra2-Bindestellen fehlen in *Am-dsx* allerdings genauso wie in *Bm-dsx* (Cho et al., 2007).

I.V Spleißen - ein wichtiger Mechanismus bei der Regulation der Geschlechtsdetermination

Das alternative Spleißen ermöglicht einem Gen für mehrere Transkripte und somit auch für verschiedene Proteine zu kodieren. Unter anderem spielt der Prozess des alternativen Spleißens eine äußerst wichtige Rolle in der Geschlechtsdetermination.

I.V.I Das Spleißen eukaryotischer prä-mRNAs

Beim Spleißen werden aus der prä-mRNA, die nicht kodierenden Introns entfernt und die Exons zu einer kodierenden Sequenz zusammengefügt. Die Entfernung des Introns erfolgt durch zwei aufeinanderfolgende Transesterreaktionen (Abb. 5). Im ersten Schritt attackiert die 2'OH-Gruppe des Adenosins der Verzweigungsstelle (branchpoint), die sich strangaufwärts vom Spleißakzeptor befindet, die Phosphodiesterbindung an der Exon/Intron-Grenze. Dadurch liegt das strangaufwärtsgelegene Exon frei vor, während das Intron eine Lariatstruktur ausbildet, die mit dem zweiten Exon verbunden ist. Im nächsten Schritt attackiert nun die 3'-OH Gruppe des freien Exons die Phosphodiesterbindung am Intron/Exon-Übergang des strangabwärts gelegenen Exons. Dadurch bildet sich eine neue Phosphodiesterbindung aus, die beiden Exons werden miteinander verbunden und das Intron-Lariat wird freigesetzt (Krämer, 1996; Staley and Guthrie, 1998).

Die Spleißreaktion wird im Zellkern von einem sogenannten Spleißosom, einem Ribonukleoprotein-Komplex, der aus zahlreichen Proteinen und fünf nukleären Ribonukleoproteinpartikeln (snRNPs) besteht, durchgeführt. snRNPs setzen sich aus Proteinen zusammen, die fest mit einer kleinen uracilreichen RNA (U snRNA) assoziiert sind (Zhou et al., 2002; Will and Lührmann, 1997; Krämer, 1996).

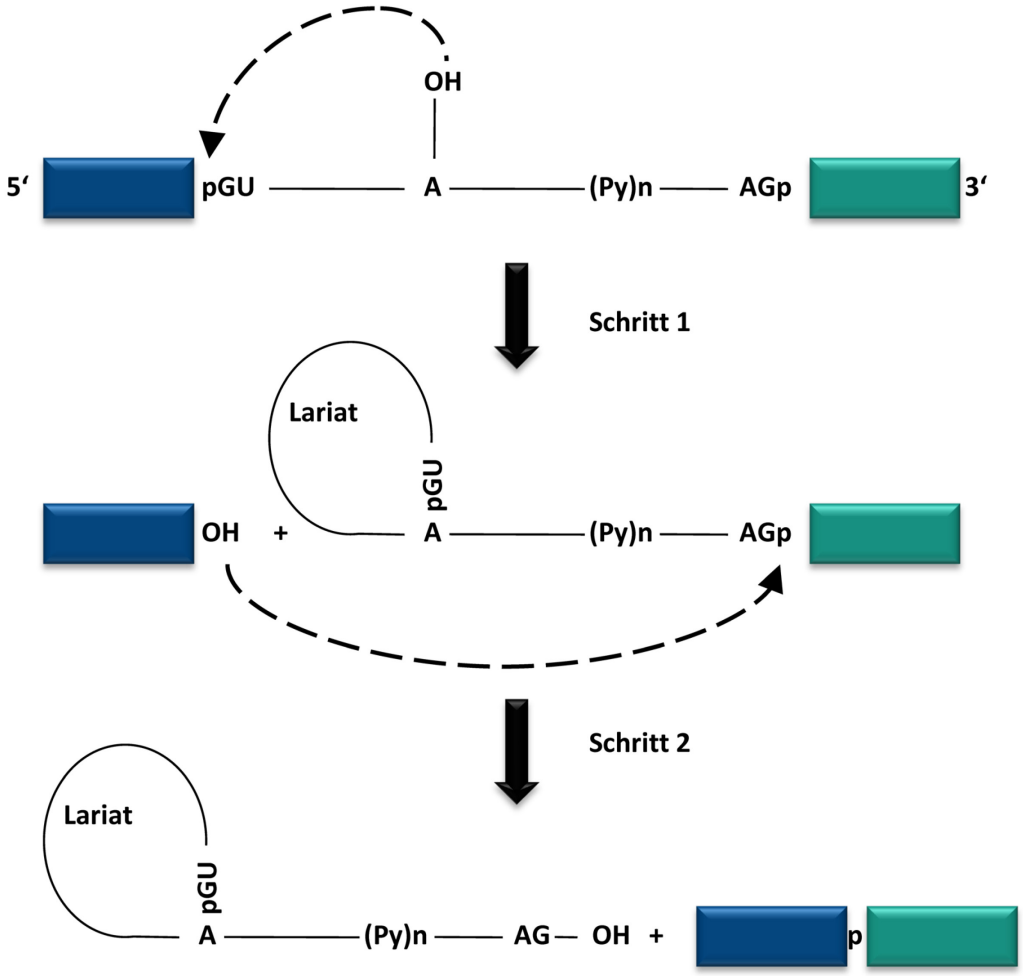

Abb. 5: Schematische Darstellung der Spleißreaktion (verändert nach Baralle, 2005). Das Spleißen findet in zwei Transesterreaktionen statt. Im ersten Schritt der Spleißreaktion attackiert die 2‘Hydroxylgruppe (2‘OH) des Adenosins (A) der Verzweigungsstelle durch einen nukleophilen Angriff die Phosphodiesterbindung an der 5‘ Spleißstelle. Dadurch entsteht eine 2‘-5‘-Phosphodiesterbindung an der Verzweigungsstelle und eine freie 3‘OH –Gruppe am 5‘ Exon (blau). Im zweiten Schritt der Umesterung attackiert die freie 3‘OH-Gruppe des 5‘Exons (blau) die 3‘Spleißstelle des strangabwärts gelegenen Exons (türkis). Dadurch wird das Intron-Lariat freigesetzt und die beiden Exons werden verknüpft.

Das Spleißosom erkennt Sequenzmotive, die das 5‘ und 3‘-Ende eines Introns markieren, die Spleißdonoren (5’Spleißstelle) und die Spleißakzeptoren (3’Spleißstelle) (Maniatis and Tasic, 2002). Der erste Schritt bei der Assemblierung des Spleißosoms ist die Bildung des sogenannten E-Komplexes (E=early) (Michaud and Reed, 1993, 1991), durch Bindung der U1snRNP an den Spleißdonor. Dabei kommt es zu Bildung eines RNA-Duplex zwischen der U1snRNA und 11 Nukleotiden des Spleißdonors (Freund et al., 2003; Kammler et al., 2001; Mount et al., 1983; Zhuang and Weiner, 1986), während die weiteren Komponenten des U1snRNP zur Stabilisierung des Duplex beitragen (Du and Rosbash, 2002). An den

Spleißakzeptor binden drei Proteine, die sich auch untereinander stabilisieren: SF1 (splice factor 1) bindet an die Verzweigungsstelle, während U2AF65 und U2AF35, die 65 kDa und die 35 kDa Untereinheit des U2AF (U2 auxiliary factor)-Proteins an den Polypyrimidintrakt, bzw. an das AG-Dinukleotid der 3'Spleißstelle binden (Berglund et al., 1998; Banerjee et al., 2004; Selenko et al., 2003). Im nächsten Schritt wird unter Umsetzung von ATP der A-Komplex gebildet (Michaud and Reed, 1991). Dabei wird SF1 an der Verzweigungsstelle durch U2snRNP ersetzt, wodurch sich die Verzweigungsstelle vorwölbt (Krämer, 1996; Staley and Guthrie, 1998). Im nächsten Schritt der Spleißosom-Assemblierung, der Bildung von Komplex B, kommt das U5•U4/U6 tri-snRNP hinzu, womit das Spleißosom vollständig ist (Boehringer et al., 2004). Als nächstes kommt es zu Wechselwirkungen von U5snRNP mit Exonsequenzen an der 5'Spleißstelle und an der 3'Spleißstelle (Newman, 1997). Erst nachdem die Spleißstellen korrekt angeordnet wurden, löst sich U6 snRNP (Lamond et al., 1988) während einer intramolekularen Umordnung von U4snRNP, wodurch die Basenpaarung von U6 snRNP mit U2 snRNP ermöglicht wird (Madhani and Guthrie, 1992; Ryan and Abelson, 2002). Das hat die Verdrängung von U1 durch U6snRNP zur Folge, wodurch letzteres mit dem 5'Ende des Introns in Wechselwirkung treten kann (Staley and Guthrie, 1999; McManus et al., 2007). Damit ist die erste Transesterreaktion abgeschlossen. Für die zweite Umesterung kommt es zu einer weiteren Umordnung der RNA im Spleißosom, durch die das freie 5'Exon neben dem 3'Exon zu liegen kommt. Dadurch kann die 3'OH-Gruppe des 5'Exons einen nukleophilen Angriff auf die 3'Spleißstelle durchführen, die beiden Exons werden miteinander verbunden und das Intron-Lariat wird freigesetzt (Staley and Guthrie, 1998).

I.V.II *Cis*-regulatorische Sequenzen helfen bei der Spleißstellenerkennung

Ein wichtiger Faktor beim prä-mRNA Spleißen ist die Erkennung der Exons bzw. der 5' und 3'Spleißstellen durch das Spleißosom. Zwar gibt es Merkmale, die den Intron/ Exon-Übergang charakteristisch markieren, insbesondere die ersten (GU) und letzten (AG) beiden Nukleotide eines Introns sind fast immer gleich, doch passt die Sequenz der einzelnen Spleißdonoren und –akzeptoren nicht immer zu ihrer Konsensussequenz. Hinzu kommt eine große Anzahl an kryptischen Spleißstellen, die zwar Ähnlichkeit zur Konsensussequenz zeigen, normalerweise aber nicht benutzt werden. Eine unterstützende Rolle bei der Spleißstellenerkennung nehmen kurze *cis*-regulatorische Elemente ein, die sich meist im näheren Bereich um die Spleißstellen befinden. Dabei unterscheidet man im Exon die

sogenannten exonischen spleißaktivierenden (ESE – exonic splicing enhancer) und spleißinhibierenden (ESS – exonic splicing silencer) Elemente, sowie im Intron die intronischen spleißinhibierenden (ISS – intronic splicing silencer), bzw. die intronischen spleißaktivierenden Elemente (ISE - intronic splice enhancer). Diese Sequenzen werden von den sogenannten nicht-snRNP-Proteinen gebunden, die per Definition zwar an der Spleißreaktion beteiligt sind, aber keine stabilen Komponenten von freien snRNP-Partikeln sind (Lamm and Lamond, 1993). Zu dieser Gruppe an Spleißfaktoren gehören unter anderem Proteine der SR- und der hnRNP-Familie (heterogenous nuclear ribonucleoprotein) (Smith and Valcárcel, 2000), die sowohl aktivierend als auch inhibierend wirken können.

I.V.III SR-Proteine und SR-ähnliche Proteine und ihre Rolle beim alternativen Spleißen

Das erste beschriebene SR-Protein war das Protein SF2/ASF (Krainer et al., 1990; Ge and Manley, 1990). Seitdem wurde eine große Anzahl weiterer Mitglieder dieser Proteinfamilie in Metazoen und auch in Pflanzen identifiziert. Eine Gemeinsamkeit dieser Proteine ist die Anwesenheit von mindestens einer RNA-Bindedomäne (RBD) im N-terminalen Bereich und der Arginin-Serin-reichen (RS-reichen) Domäne im C-Terminus (Graveley, 2000). Die 80-90 Aminosäuren lange RNA-Bindedomäne ist notwendig für die sequenzspezifische Bindung des Proteins an die RNA und bildet eine typische βαββαβ-Struktur aus, bestehend aus vier antiparallelen β-Faltblättern und zwei α-Helices (Nagai et al., 1990). Innerhalb der RBD befinden sich zwei hochkonservierte Bereiche, die als RNP-1 (K/R-G-F/Y-G/A-F/Y-V/I/L-X-F/Y) und RNP-2 (I/V/L-F/Y-I/V/L-X-N-L) bezeichnet werden (RNP – ribonucleoprotein consensus peptide) und Teil des β1- und des β3-Faltblatts sind. Die Aminosäuren im RNP-1 und RNP-2 Motiv sind direkt an der Erkennung der RNA beteiligt (Amrein et al., 1994; Nagai et al., 1990; Dreyfuss et al., 1988; Merrill et al., 1988), was unter anderem für das RNP-1 Motiv von Tra2 gezeigt wurde (Amrein et al., 1994). Die RS-reiche Domäne variiert in den verschiedenen SR-Proteinen in Länge und Anzahl der Arginine und Serine, wobei die Stärke einer RS-Domäne direkt mit der Anzahl an RS-Dipeptiden korreliert (Graveley, 2000; Graveley et al., 1998). Außerdem hängt die Fähigkeit eines SR-Proteins als Spleißaktivator zu wirken vom Phosphorylierungsstatus der RS-Domäne ab, da dieser Einfluss auf die subzelluläre Lokalisation sowie auf die Interaktion zwischen SR-Proteinen hat (Gui et al., 1994; Xiao and Manley, 1998; Yeakley et al., 1999; Wang et al., 1998; Colwill et al., 1996). Die RS-Domäne vermittelt die Interaktion mit anderen SR-Proteinen und den sogenannten SR-ähnlichen

Proteinen. SR-ähnliche Proteine besitzen wie die SR-Proteine eine RS-reiche Domäne, jedoch kein RNA-Bindemotiv. Beispiele aus der Geschlechtsdeterminationskaskade sind Tra in *Drosophila melanogaster* (Boggs et al., 1987), sowie Fem und Csd in der Honigbiene (Beye et al., 2003; Hasselmann et al., 2008). Während manche SR-Proteine und SR-ähnliche Proteine eine generelle Funktion als Spleißproteine haben, also in den Prozess des konstitutiven Spleißens involviert sind, spielen viele SR-Proteine und SR-ähnliche Proteine wie Tra und Tra2, insbesondere beim alternativen Spleißen eine große Rolle. Durch Bindung an ESE, ESS, ISE oder ISS Elemente können die SR-Proteine die Aktivierung der Spleißakzeptoren und Spleißdonoren beeinflussen. Meist wirken SR-Proteine durch Bindung an ESE-Sequenzen als Spleißaktivatoren, jedoch wurde auch eine inhibitorische Funktion beobachtet (Smith and Valcárcel, 2000). Tra2 wirkt beispielsweise in *Drosophila melanogaster* beim Spleißen der *doublesex* sowie der *fruitless* prä-mRNA einerseits als Spleißaktivator, hat aber beim autoregulativen Spleißen der eigenen prä-mRNA eine inhibitorische Wirkung.

I.VI Fragestellungen und Ziele der Arbeit

Die Entscheidung, ob sich aus einem Bienenei ein Drohn oder eine weibliche Honigbiene (*A. mellifera*) entwickelt, hängt von einer Reihe verschiedener Faktoren der Geschlechtsdeterminationskaskade ab. Bislang sind mit *csd, fem* und *Am-dsx* drei Gene bekannt, die in der haplodiploiden komplementären Geschlechtsdetermination eine wichtige Rolle spielen. Dabei ist der heterozygote Allelzustand des *csd*-Gens das initiale Signal, das die weibliche Entwicklung einleitet und das alternative Spleißen der *fem* prä-mRNA beeinflusst. In der Kaskade ist das Gen *fem*- dem *csd*-Gen untergeordnet und sorgt zum Einen in einem positiven autoregulativen Spleißprozess für die Aufrechterhaltung des weiblichen Signalwegs und zum Anderen für die Entstehung von weibchenspezifischem *Am-dsx*-Transkript. Ob Csd- und Fem-Proteine direkte Regulatoren des alternativen Spleißens der *fem*- und *Am-dsx*-prä-mRNA sind oder ob es noch weitere Regulationsebenen in der Geschlechtsdeterminationskaskade der Honigbiene gibt, ist bislang jedoch ungeklärt. Da weder Csd- noch Fem-Proteine eine RNA-Bindedomäne besitzen, die die Interaktion der Proteine mit der jeweiligen prä-mRNA vermitteln könnte, vermuten wir, dass noch mindestens ein weiteres Protein in die Spleißprozesse involviert ist. Ein potentieller Interaktionspartner von Csd- und Fem-Proteinen, der als Kofaktor die Interaktion mit der

fem- bzw. der *Am-dsx*-prä-mRNA vermitteln könnte, ist das Protein *Am*-Tra2. Beim *Am*-Tra2-Protein handelt es sich um das Ortholog des Tra2-Proteins der Taufliege *D. melanogaster*. Bei *D. melanogaster* vermittelt das Tra2-Protein zusammen mit dem Tra-Protein, einem Ortholog des Fem-Proteins, das weibchenspezifische Spleißen der *dsx*-prä-mRNA. Ein Ziel dieser Arbeit ist es deshalb, neben der Charakterisierung des *Am-tra2*-Gens, den Einfluss der Transkription des *Am-tra2*-Gens auf das geschlechtsspezifische Spleißen der *fem*- sowie der *Am-dsx*-prä-mRNA zu untersuchen. Dafür soll das *Am-tra2*-Gen in weiblichen Embryonen mittels RNA-Interferenz herunterreguliert werden. Im Weiteren habe ich mich auf die Aufklärung der *Am-dsx*-prä-mRNA Spleißregulation fokussiert. Um zu überprüfen, ob das Fem-Protein und weitere potentielle Spleißproteine direkt an der *Am-dsx* Spleißregulation beteiligt sind, soll in Schmetterlingszellen ein *in vivo* Spleißassay etabliert werden: Dabei kann mittels Koexpression der Proteine Fem und *Am*-Tra2 sowie des *Am-dsx* Minigenkonstrukts, der Einfluss dieser Proteine auf das Spleißen der *Am-dsx*-prä-mRNA untersucht werden. Abschließend soll der Mechanismus der *Am-dsx*-Spleißregulation untersucht werden.

Thesen

Folgende Thesen können im Rahmen dieser Arbeit aufgestellt werden:

1) Das Honigbienen-Ortholog des Gens *transformer 2* (*Am-tra2*) wird alternativ, aber geschlechtsunspezifisch gespleißt. Es existieren mindestens sechs Spleißvarianten, die für Isoformen des SR-Proteins *Am*-Tra2 kodieren.

2) Die Ergebnisse von RNA Interferenz Experimenten zeigen, dass die Transkription von *Am-tra2* in der geschlechtsbestimmenden Kaskade Einfluss auf das weibchen- und männchenspezifische Spleißen der *fem*-prä-mRNA sowie das weibchenspezifische Spleißen der *Am-dsx*-prä-mRNA hat. Während der Embryogenese ist die Transkription von *Am-tra2* außerdem essentiell für die Vitalität der Larve.

3) Mit einem *Am-dsx*-Minigenkonstrukt kann mittels eines *in vivo* Spleißassays das geschlechtsspezifische Spleißen der *Am-dsx*-prä-mRNA in *Sf*21-Zellen nachgestellt werden. Der männchenspezifische Spleißweg ist voreingestellt und unabhängig von der Expression honigbienenspezifischer Proteine. Für Bildung des weibchenspezifischen Transkripts ist die Expression von Fem-Protein notwendig. Die alleinige Expression von *Am*-Tra2-Protein ist nicht ausreichend um weibchenspezifisches Spleißen zu bewirken. In Kombination mit Fem-Protein hat es eine steigernde Wirkung auf die Bildung von weiblichem *Am-dsx*-Transkript.

4) Im Bereich vor dem weibchenspezifischen Spleißdonor der *Am-dsx*-prä-mRNA befinden sich zwei Wiederholungen eines Hexamer-Motivs (AAAGAG), die in Abhängigkeit von *Am*-Tra2- und Fem-Protein als exonische spleißsteigernde Sequenz (ESE) wichtig für die Aktivierung eines Spleißdonors sind.

5) Obwohl die einzelnen Komponenten der *dsx*-Spleißregulation von Honigbiene und Taufliege konserviert sind, ist der Mechanismus des *dsx*-Spleißens evolviert.

Chapter II: Manuscripts and additional experiments

Manuscript I

The *Am-tra2* gene is an essential cofactor of female splice regulation at two levels of the *complementary sex determination* hierarchy of the honeybee

Inga Nissen, Miriam Müller, Martin Beye

Department of Genetics, Heinrich Heine University Duesseldorf

40225 Duesseldorf, Germany

Manuscript currently under review at Genetics

Key words:

sex determination, insects, evolution of genetic pathways, *Apis mellifera*

Abstract

Heteroallelic and homo- or hemiallelic Csd (Complementary sex determiner) proteins determine sexual fate in the honeybee (*Apis mellifera*) by controlling the alternative splicing of the downstream gene *fem* (*feminizer*). Thus far, we have little understanding of how heteroallelic Csd proteins mediate the splicing of female *fem* mRNAs or how Fem proteins direct the splicing of honeybee *dsx* (*Am-dsx*) pre-mRNAs in the absence of an RNA binding domain (RBD). Here, we report that the honeybee *tra2* gene *(Am-tra2)*, which is a cofactor of female *dsx* splicing in *Drosophila melanogaster*, is an essential component of female splicing of the *fem* and *Am-dsx* transcripts in the honeybee. The *Am-tra2* transcripts are alternatively (but non-sex-specifically) spliced, and they are translated into six protein isoforms that all share the basic RBD/RS (arginine-/serine) domain structure. Knockdown studies showed that the *Am-tra2* gene is required to splice *fem* mRNAs into the productive female and non-productive male forms. We suggest that the *Am*-Tra2 protein is an essential cofactor of heteroallelic Csd proteins that can mediate the direct binding and splicing of *fem* pre-mRNAs that is implementing the female pathway. *Am*-Tra2 together with Fem proteins may also promote the splicing of more female *fem* mRNAs. This positive autoregulatory splicing loop maintains the female-determining state in honeybees, a process that has also been reported in some other dipteran insects. In males, the *Am*-Tra2 proteins may enhance the switch of *fem* transcripts into the non-productive male form when heteroallelic Csd proteins are absent. The *Am*-Tra2 protein is also an essential cofactor for *Am-dsx* female splice regulation, although the proposed canonical protein binding sites for this regulation are absent in the honeybee *Am-dsx* gene. We also showed that the *Am-tra2* gene has an essential function in honeybee embryogenesis that is unrelated to sex determination.

Introduction

In contrast to the well-studied sex chromosome system in *Drosophila melanogaster* (Cline and Meyer, 1996; Erickson and Quintero, 2007), sex in the honeybee (*Apis mellifera*) is determined by heterozygosity at the *complementary sex determiner* (*csd*) gene (Beye et al., 2003). Bees that are heterozygous at the *csd* locus develop into females, whereas bees that are homozygous or hemizygous at *csd* develop into males. Queens in honeybee colonies lay unfertilized eggs to produce fertile males (drones) and fertilized eggs to produce females that either differentiate into workers or queens; queen fate is determined by specific feeding of the queen larvae with royal jelly (Kucharski et al., 2008; Kamakura, 2011). Diploid males, homozygous for the *csd* gene, do not survive in a colony because they are eaten by worker bees shortly after they hatch from the egg. The *csd* gene translates into an SR-type protein that has at least 15 major allelic variants (Beye, 2004; Hasselmann et al., 2008) which differ at an average of approximately 6% of their amino acid residues in the putative specifying domain (PSD) (Hasselmann et al., 2008; Hasselmann and Beye, 2004). Females express heteroallelic Csd proteins that direct female splicing of *feminizer* (*fem)* pre-mRNAs (Hasselmann et al., 2008; Gempe et al., 2009). These female *fem* transcripts are also translated into SR-type proteins that are required for female differentiation. The Fem proteins promote the female splicing of the *Apis mellifera* dsx (*Am-dsx*) transcripts, which express a transcription factor of the DM type, a protein with a female-specific carboxy-terminal end (Dearden et al., 2006; Cho et al., 2007). In addition, Fem proteins direct splicing of their own pre-mRNAs into the productive female form, which generates an autoregulatory feedback loop that maintains the female state throughout development (Gempe et al., 2009). In the absence of Csd protein activity in males (homo- or hemiallelic Csd proteins), *fem* transcripts are spliced into the male form, which contains a translational stop codon in exon 3 that causes premature translation termination (Gempe et al., 2009). As a consequence, *Am-dsx* pre-mRNAs are spliced into the male variant expressing a Dsx protein (Gempe et al., 2009), which has a male-specific carboxy-terminal end as part of the oligomerization domain 2 (OD2) (Cho et al., 2007).

The *csd* gene is thus the primary signal of sex determination in the honeybee. The *csd* gene evolved recently in the honeybee lineage by gene duplication of an ancestral copy of the *fem* gene (Hasselmann et al., 2008). Although substantially diverged in sequence, the *fem* gene is the ortholog of the sex-determining gene *transformer* (*tra*) of *D. melanogaster*.

Thus far, we have little understanding of how the heteroallelic Csd proteins mediate *fem* splicing or how Fem proteins direct *Am-dsx* pre-mRNA splicing (Gempe and Beye, 2011). Both proteins lack an RNA-binding domain (RBD), which is suggestive of a cofactor that can directly interact with the respective RNA sequence.

In this study, we explored the role of the *tra2* gene of the honeybee (*Am-tra2* (Dearden et al., 2006)) in regulating the sex-specific splicing of the *fem* and *Am-dsx* transcripts. The honeybee *Am-dsx* pre-mRNAs lack the Tra/Tra2 protein canonical binding sites, suggesting that the sexual splice regulation of *dsx* transcripts has evolved.

In *D. melanogaster,* the RNA-binding protein Tra2 acts together with the Tra protein to promote the female splicing of *dsx* pre-mRNAs (Burtis and Baker, 1989; Inoue et al., 1992; Hedley and Maniatis, 1991; Lynch and Maniatis, 1996, 1995; Sciabica and Hertel, 2006; Amrein et al., 1990). *D. melanogaster* females express in somatic tissues two major protein isoforms of the Tra2 proteins ($Tra2^{264}$ and $Tra2^{226}$) that, together with the Tra proteins, activate a weak 3' splice acceptor site in *dsx* pre-mRNAs by binding to the six repeats of a 13-nucleotide exonic splicing enhancer (ESE) sequence and a single purine-rich element (PRE). This activation leads to the inclusion of the female exon 4 in female *dsx* mRNAs (Burtis and Baker, 1989; Inoue et al., 1992; Hedley and Maniatis, 1991; Lynch and Maniatis, 1996, 1995; Sciabica and Hertel, 2006). Splicing of honeybee *Am-dsx* has putatively evolved in that respect compared to *D. melanogaster*, as *Am-dsx* pre-mRNAs lack the canonical binding sites of Tra/Tra2 proteins.

The *tra2* gene has also been characterized in other dipteran species, including *Musca domestica, Anastrepha obliqua* and *Ceratitis capitata* (Burghardt et al., 2005; Salvemini et al., 2009; Sarno et al., 2010), as well as in the lepidopteran insect *Bombyx mori* (Niu et al., 2005). All Tra2 proteins share the same domain structure of a single RBD that is flanked by two arginine-/serine-rich (RS-rich) regions. In *D. melanogaster*, three Tra2 protein isoforms have been identified that vary in the length of the amino terminal RS-rich region (Amrein et al., 1990). The RBD consists of 80-90 amino acids that form a βαββαβ barrel-like topology. One side of the β-sheet surface (β1 and β3) of the RBD exposes two sequence elements, RNP-1 and RNP-2, that are directly involved in RNA recognition (Amrein et al., 1994; Nagai et al., 1990; Dreyfuss et al., 1988; Merrill et al., 1988).

In the dipteran insects *M. domestica, C. capitata* and *A. obliqua,* RNAi knockdown studies of the *tra2* gene showed that Tra2 proteins are also involved in female splicing of *tra* mRNAs (Burghardt et al., 2005; Hediger et al., 2010; Salvemini et al., 2009; Sarno et al., 2010). The authors suggest that Tra2 proteins act as cofactors in the autoregulatory splicing loop in which Tra/Tra2 proteins direct the female splicing of *tra* transcripts and thus the expression of Tra proteins. In the lepidopteran insect *B. mori*, the function of Tra2 proteins in sexual regulation of the *Bm-dsx* transcripts is not known. In this species, male splicing of *Bm-dsx* transcripts requires the splicing inhibitor (*Bm*-PSI) and the male-specific IMP (*Bm*-IMP) proteins. The activation of the female exon splicing is repressed in males by the binding of the *Bm*-PSI and the male-specific *Bm*-IMP proteins to the 20-nucleotide CE1 motif of the female exon (Suzuki et al., 2010, 2001, 2008).

In the male germline of *D. melanogaster*, the $Tra2^{226}$ protein isoform has an additional function in spermatogenesis in controlling the splicing of the *exuperantia* (*exu*) and *alternative-testes-transcript* (*att*) transcripts (Mattox et al., 1996; Hazelrigg and Tu, 1994; Madigan et al., 1996). In the testes, $Tra2^{226}$ proteins negatively affect their own expression by promoting the splicing of $tra2^{179}$ mRNAs, which produce no functional protein (Mattox et al., 1996; McGuffin et al., 1998; Mattox and Baker, 1991). This negative feedback loop controls the level of Tra2 expression, which is critical for proper spermatogenesis.

In this study, we report the cloning and functional analysis of the *A. mellifera Am-tra2* gene. Our study showed that the *Am-tra2* gene serves as a cofactor in female-specific splicing of *fem* and *Am-dsx* transcripts. Furthermore, we show that *Am-tra2* has a vital function in embryogenesis that differs from its reported functions in other species.

Materials and Methods

Bee Sources

Diploid female embryos were derived from the progeny of queens inseminated by semen from a single drone having a different sex allele than that of the queen. Haploid male embryos were collected from colonies that were headed by a virgin queen. These non-mated queens laid unfertilized male eggs that we induced by repeated CO_2 treatments.

RNA extraction, cDNA synthesis and PCR

Total RNA was extracted using the TRIzol protocol (GIBCO BRL Life Technologies, Germany). The first-strand cDNA from mRNA was generated by reverse transcription using an oligo-dT primer following the protocol of the supplier (Fermentas). We quantified the amount of cDNA in our samples in a NanoDrop ND-1000 spectral photometer and adjusted the amount of cDNA prior to PCR amplification. PCR was performed using GoTaq Flexi DNA Polymerase (RNAi experiments) according to the protocol of the supplier (Promega, Mannheim, Germany) and Taq polymerase (RACE and transcriptional analysis of *Am-tra2* throughout development). All RT-PCR fragments were resolved by agarose gel electrophoresis and stained with ethidium bromide. We amplified cDNA fragments of the elongation factor 1α (ef-1α) gene using oligonucleotides #EM033 and #EM034 for the semiquantitative studies across samples (supplementary data, table S1).

Characterization of *Am-tra2* gene

To determine the entire sequences of the *Am-tra2* transcripts, we performed 5’ and 3’ rapid amplification of cDNA ends (RACE) experiments following the manufacturer’s instructions (FirstChoice RLM-RACE kit; Ambion). The cDNAs were generated from male and female RNA samples of honeybee embryos. Gene-specific primers for RACE reactions were designed from the sequence of the *Am-tra2* gene model at the NCBI web site (NCBI Reference Sequence: XM_001121070.2) (supplementary data, table S1). All RACE products were cloned into the pGEM®-T vector (Promega), and both strands were sequenced. We translated the mRNA sequences into the amino acid sequence, and we predicted the protein domains by the similarity to domains in the PROSITE database (http://www.expasy.org/prosite/). The

GenBank accession numbers are JQ518311 (*Am-tra2*^{285}), JQ518314 (*Am-tra2*^{284}), JQ518312 (*Am-tra2*^{253}), JQ518314 (*Am-tra2*^{252}), JQ518313 (*Am-tra2*^{234}), JQ518316 (*Am-tra2*^{233}).

Transcriptional studies of the *Am-tra2* gene throughout development

Total RNA was extracted from male and female eggs (0-6 h, 9-24 h, 33-48 h, 72 h), larvae (L1 and L4 instar), pupae (3 days before hatching from comb), adult heads and germline tissue (testes of L4 larvae and ovaries of virgin queens). We amplified cDNA fragments using oligonucleotides #359 and #421 (supplementary data, table S3) that span the complete open reading frame (ORF) of all six *Am-tra2* splice variants. The identity of the amplicons of the male and female L1 larvae and pupae were verified by sequencing.

Functional studies of the *Am-tra2* gene

RNAi knockdown was induced in early embryogenesis at the syncytial stage (0–4 h after egg deposition) in females (Beye et al., 2003, 2002). *Am-tra2* dsRNA-1, encompassing the region from 322 to 767 bp (446 bp long), was generated using oligonucleotides #22M and #23M (supplementary data, table S3) from cloned cDNAs of the *Am-tra2*^{285} transcript following the protocol previously described (Hasselmann et al., 2008; Beye et al., 2003, 2002). *Am-tra2* dsRNA-2, encompassing the region from 108 to 499 bp (392 bp long), was generated using oligonucleotides #591 and #592 (supplementary data, table S1) from cloned cDNAs of transcript *Am-tra2*^{285}. The dsRNAs were dissolved in ddH_2O and injected at a concentration of 4-200 pg per embryo (supplementary data, table S1). In the control samples, we only injected ddH_2O (Roth).

We counted the number of embryos showing normal development approximately 70 hours after egg deposition and the number of hatched L1 larvae 77-80 hours after egg deposition. All embryos that showed malformed embryos, necrotic tissue within embryos or lack of segmentation pattern 70 h after injection were classified as aberrant.

To study the effect of *Am-tra2* knockdown on the splice patterns of *fem* and *Am-dsx,* we collected embryos and L1 larvae from dsRNA-treated and non-treated samples 77-80 hours after egg deposition. Treated embryos and controls were reared in the incubator at 35°C

until an age of 77-80 hours or until reaching the L1 larval stage, respectively, and were then frozen in liquid nitrogen.

Fragments corresponding to the female *fem* mRNAs were amplified using oligonucleotides #412 and #523 (supplementary data, table S3) that are composed of a part of exon 3 and exon 6 (size, 177 bp). We amplified fragments (size, 458 bp; exon 3-4 and part of exon 5) corresponding to the male *fem* mRNAs using oligonucleotides #410 and #566 (supplementary data, table S3). We only counted female *fem* splice products as female-specific if the fragments showed stronger ethidium bromide staining intensities than those in males (see for instance, Fig. 4B, lanes 31-35). We used oligonucleotides #417 and #418 (supplementary data, table S3) to amplify exons 4-5-6 (size, 1.2 kb) and exons 4-6 (size, 392 bp), which correspond to the female and male *Am-dsx* transcripts, respectively. The female *Am-dsx* transcripts were also specifically amplified by oligonucleotides #417 & #419 (supplementary data, table S3) (size, 188 bp) that encompass exon 4-5.

Phylogenetic and molecular evolutionary sequence analyses

We utilized nine *tra2* sequences from different insect species to compare the phylogenetic and molecular relationships: *Tribolium castaneum* (GenBank:XP_968550.2), *Acromyrmex echinatior* (GenBank: EGI70155.1), *Drosophila melanogaster* (NCBI Reference Sequence: NP_476764.1), *Bombyx mori* (GenBank: NP_001119705.1), *Nasonia vitripennis* (GenBank: XP_001601106.1), *Sciara ocellaris* (GenBank: CBX45935.1), *Ceratitis capitata* (GenBank: ACC68674.1), *Musca domestica* (GenBank: AAW34233.1) and *Anastrepha obliqua* (GenBank: CBJ17280.1).

Results

Genomic organization of the *tra2* gene in the honeybee

The existence of the *tra2* gene in the honeybee genome was predicted by the similarity of its RBD to those of other insects (Dearden et al., 2006). We isolated *Am-tra2* transcripts that included the 5' and 3' untranslated regions (UTRs) and three different polyadenylation sites (Fig. 1) using RACE experiments with cDNA preparations from both male and female embryos. Using cDNA preparations from embryos and pupae of males and females, we performed RT-PCRs to amplify the entire open reading frame (ORF) of *Am-tra2* with oligonucleotide primers that bound the 5' and the 3' ends of the ORF. As a result, we detected up to six splice variants (Fig. 1) that were not sex-specific in embryos or pupae. The six splice variants in embryos were *Am-tra2*^{285}, *Am-tra2*^{284}, *Am-tra2*^{253}, *Am-tra2*^{252}, *Am-tra2*^{234} and *Am-tra2*^{233} and the four transcript variants in pupae were *Am-tra2*^{285}, *Am-tra2*^{284}, *Am-tra2*^{253} and *Am-tra2*^{252} (Fig. 1). Three splice variants showed major sequence differences in exon 2. The other three are minor variants of the other three major variants that differ in three nucleotides in exon 4. All transcripts express essentially the same protein that harbors a RBD flanked by two RS-rich domains but differ in the length in the first RS domain (RS1) (Fig. 1 A-C). The RBD amino acid sequence has the strongest similarity (61 to 85% sequence identity) to the Tra2 proteins of a variety of insects (*Nasonia vitripennis*, *Tribolium castaneum*, *Acromyrmex echinatior*, *B. mori*, *Sciara ocellaris*, *D. melanogaster*, *M. domestica*, *C. capitata* and *A. obliqua*) (Burghardt et al., 2005; Salvemini et al., 2009; Sarno et al., 2010; Amrein et al., 1990; Mattox et al., 1996; Niu et al., 2005; The *Tribolium* Genome Sequencing Consortium, 2008; Nygaard et al., 2011), supporting our notion that we have identified the Tra2 ortholog of the honeybee (Fig. 2).

The largest transcript, *Am-tra2*^{285}, consists of 1401 nt and 5 exons and harbors an ORF coding for 285 amino acids (Fig. 1A). The other two major transcript variants (*Am-tra2*^{253} and *Am-tra2*^{234}) consist of 6 exons in which different parts of exon 2 are spliced out. These two variants are produced by sharing the same splice donor but using two alternative splice acceptors. They are 1248 nt and 1305 nt long and are putatively translated into 253- and 234-amino-acid proteins, respectively (Fig. 1B and C). All three major splice variants have minor splice variants that utilize an alternative splice acceptor site at exon 4, producing single-serine amino acid deletion in the RS2 domain of the putative protein (Fig. 1D).

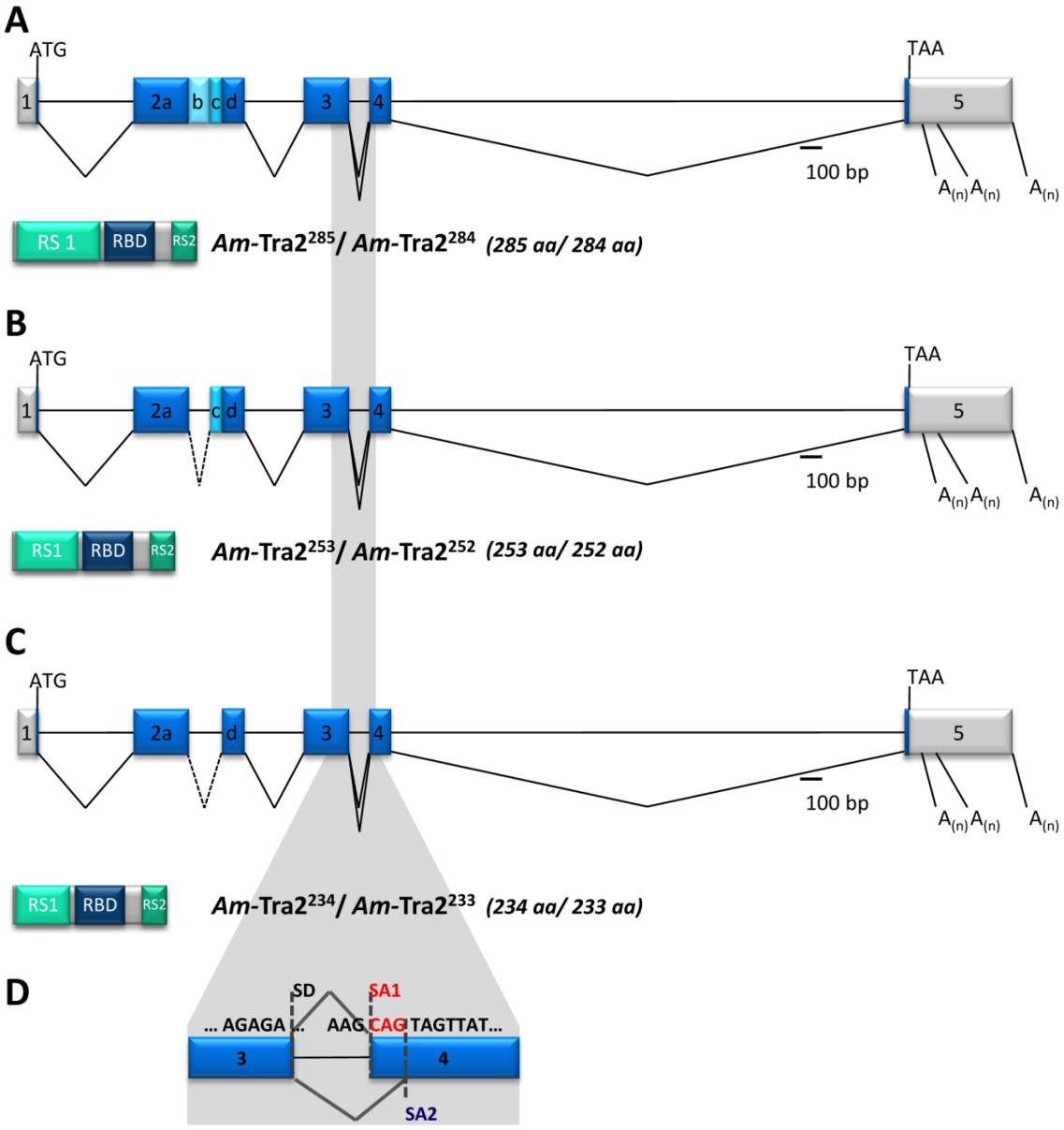

Figure 1: Genomic organization of the *Am-tra2* gene and the alternative spliced mRNAs.

Schematic representation of the intron and exon organization (presented as lines and boxes, respectively). The alternatively spliced transcripts are indicated by the connecting lines between exons. The three alternative polyadenylation sites are labeled $A_{(n)}$. The scale denotes the relative size of the introns and exons. The 5'- and 3'-UTRs are presented in grey and the ORF in blue boxes. Below the genomic organization, the domain structure and relative size of the predicted *Am*-Tra2 proteins are shown (RS, arginine/serine rich domain; RBD, RNA binding domain). Superscript of *Am*-Tra2 proteins denotes the number of amino acids in that particular protein isoform. **(A)** *Am*-Tra2^{285}, *Am*-Tra2^{284} **(B)** *Am*-Tra2^{253}, *Am*-Tra2^{252} **(C)** *Am*-Tra2^{234} and *Am*-Tra2^{233}. **(D)** Alternatively spliced variants in exon 4 producing tri-nucleotide and single amino acid differences. This splicing affected all three transcripts shown in A-C that are denoted as the *Am-tra2*^{284}, *Am-tra2*^{252} and *Am-tra2*^{233} transcripts**.** SA_1 and SA_2 label the alternative splice acceptors in exon 4.

Am-Tra2 has a typical Tra2 RBD that has evolved in the RNP-1 element

We compared our deduced amino acid sequence with those of dipteran (*D. melanogaster, M. domestica, S. ocellaris* and *A. obliqua*), coleopteran (*T. castaneum*), lepidopteran (*B. mori*) and other hymenopteran (*N. vitripennis* (wasp) and *A. echinatior* (ant)) insects to identify shared structural features of Tra2 proteins in holometabolous insects. We were only able to unambiguously align the amino acid sequences for the RBD and neighboring regions, but not for other parts of the protein (supplementary data, Fig. S1), suggesting that the RBD is evolutionary constrained. All Tra2 proteins in the different organisms share the two RS domains, but the arginine- and serine-rich sequence is highly diverged (supplementary data, Fig. S1), suggesting that these domains are faster-evolving and evolutionarily less constrained than the RBD.

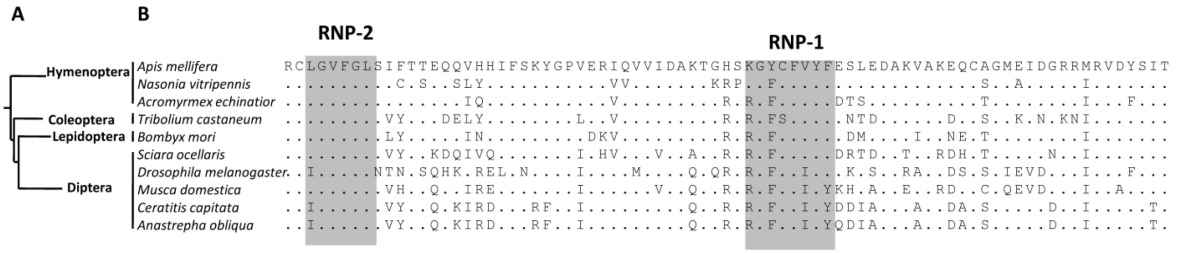

Figure 2: The RBD domain of the honeybee *Am*-Tra2 protein and its relation to the Tra2 RBD domain of other holometabolous insects. (A) The phylogenetic relationship of the members of the different insect orders used in this comparison (Savard et al., 2006). (B) Amino acid sequence alignment of Tra2 RBDs of *Apis mellifera, Nasonia vitripennis, Acromyrmex echinatior, Tribolium castaneum, Bombyx mori, Sciara ocellaris, Drosophila melanogaster, Musca domestica, Ceratitis capitata* and *Anastrepha obliqua*. Dots indicate amino acids identical to the predicted *Am*-Tra2 RBD of the honeybee. Boxes denote the RNP sequence elements.

In *D. melanogaster*, the RBD domain of Tra2 protein binds to ESEs, which is comprised of six nearly identical 13-nucleotide-long sequences. Similar motifs have also been detected in the *dsx* gene sequences of other dipteran insects, but are lacking in the *Am-dsx* pre-mRNAs of the honeybee. We next studied the sequence similarities within the RBDs of Tra2 proteins of different holometabolous insects (Fig. 2). The RBD amino acid sequence diverges in relation to phylogenetic distance. The RBD domains of the hymenopteran species honeybees, *N. vitripennis* and *A. echinatior,* show a pairwise sequence identity of 82 to 85%, whereas RBDs of the honeybee and dipteran species have pairwise sequence identity of 61 to 68%. Within the RBD of the *Am*-Tra2 protein, there are

two putative RNP (ribonucleoprotein consensus peptide) elements (Fig. 2) which directly interact with the RNAs (Amrein et al., 1994; Nagai et al., 1990; Dreyfuss et al., 1988; Merrill et al., 1988). In the eight-amino-acid-long honeybee RNP-1 sequence element, we detected 3 amino acid differences compared with that of *D. melanogaster*. A mutation in the first amino acid of RNP-1 has been shown in *D. melanogaster* to be essential for female *dsx* splicing (Amrein et al., 1994). Interestingly, this functional important first amino acid has especially evolved from amino acid R to K within the Hymenopteran lineage (Fig. 2). We propose that the target RNA sequences of RNP-1 element in the *Am*-Tra2 protein correspondingly evolved and thus differ from those found in *D. melanogaster.*

Am-tra2 transcript variants are not sex-specifically regulated and are transcribed throughout development

In *D. melanogaster, the tra2* gene produces three transcript variants that translate into three Tra2 protein isoforms ($Tra2^{264}$, $Tra2^{226}$ and $Tra2^{179}$) which differ in the length of the first RS domain (Amrein et al., 1990; Mattox et al., 1996). $Tra2^{264}$ and $Tra2^{226}$ are expressed in somatic tissues in both sexes and are involved in *dsx* regulation. To identify sex-specific transcripts and the relative abundance of transcripts throughout development in the honeybee, we simultaneously amplified all *tra2* transcript variants in 0-6 h, 9-24 h and 33-48 h old embryos; L1 and L4-L5 larvae; and pupae by semiquantitative RT-PCR (Fig. 3A). The transcripts are most abundant in the embryonic to larval stages in both sexes. However, the level of expression substantially decreases at the pupal stage. The *Am-tra2*^{285} and *Am-tra2*^{284} transcripts, which are translated into the largest RS domain-containing proteins of all the transcripts, have the highest level of expression (Fig. 3A). The *Am-tra2*^{253}/ *Am-tra2*^{252} transcripts are present in sizeable amounts in early embryos, whereas the *Am-tra2*^{234} / *Am-tra2*^{233} transcripts are apparently minor splice products. In *D. melanogaster,* there is a male-specific *tra2* transcript, *tra2*^{179}*,* whose splicing is mediated by $Tra2^{226}$ proteins derived from the general $tra2^{226}$ transcript. This negative feedback loop at the level of splicing regulates proper Tra2 protein expression in the germline, which is essential for correct sperm formation (McGuffin et al., 1998). To identify a similar role of the *Am-tra2* gene in sperm formation of the honeybee, we specifically searched for sex-specific *Am-tra2* transcripts in the female (ovary tissue) and male (testis tissue) germ lines (Fig. 3B). We detected no sex-

specific transcripts, suggesting that *Am-tra2* mRNAs are not sexually spliced in male germ cells.

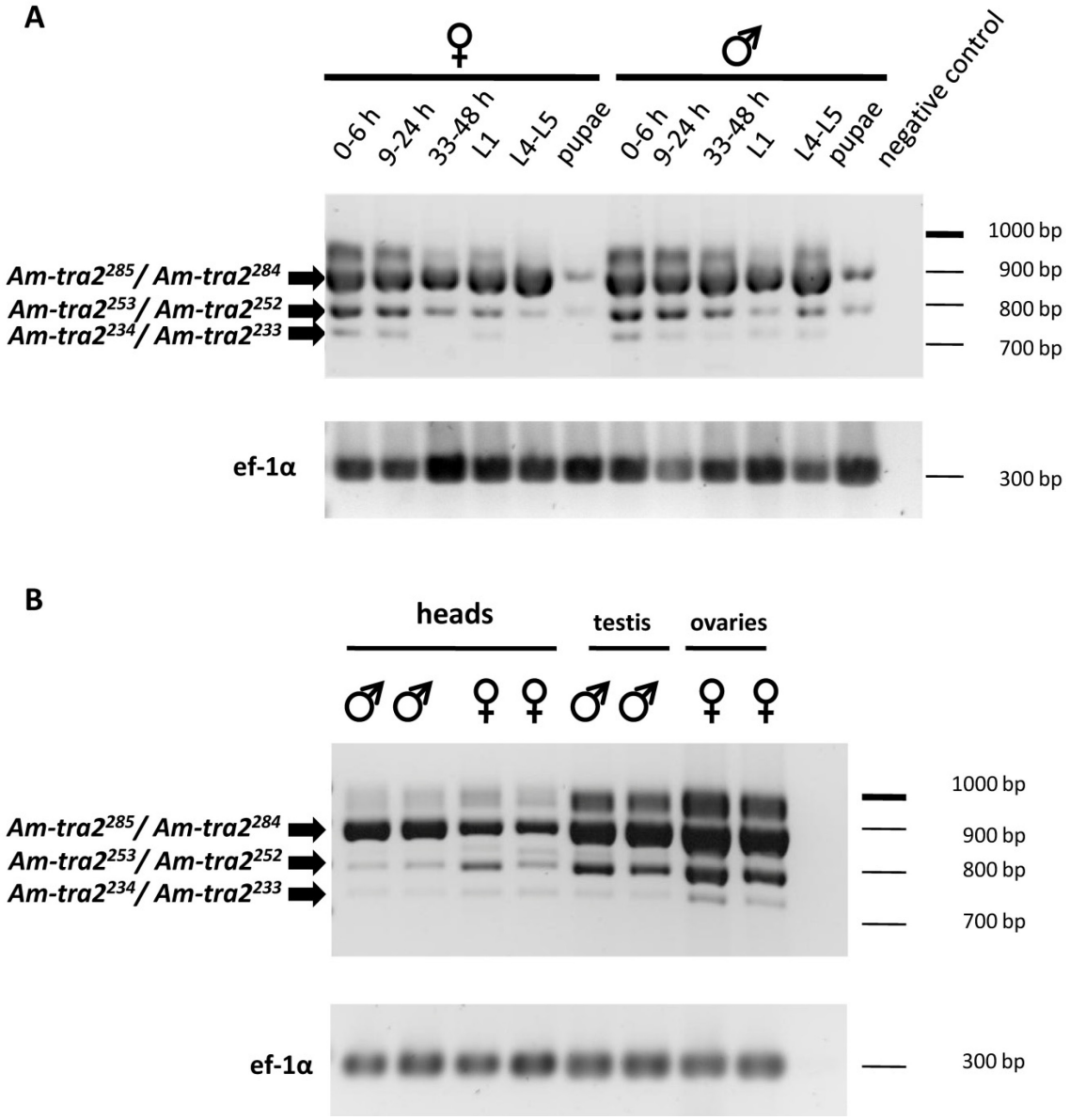

Figure 3: Semiquantitative transcriptional profile of *Am-tra2* transcripts throughout male and female development (A) and in somatic and gonadal tissue (B).

Fragments of *Am-tra2* transcripts spanning the entire open reading frame (ORF) were amplified by RT-PCR, resolved by agarose gel electrophoresis and stained with ethidium bromide. We used amplification of the cDNAs of the "housekeeping" gene *elongation factor 1α* (ef-1α) as a relative control to semiquantify *Am-tra2* transcripts across samples (Beye et al., 2003). All the fragments in 0-6-h-old embryos and pupae were cloned and sequenced. Ovaries were dissected from virgin queens and testes from L4 larvae; heads are from adult bees. L1, larval stage 1; L4-L5, larval stages 4 and 5. For the embryonic stages, the hours (h) after egg deposition are indicated.

Taken together, these results indicate that *Am-tra2* transcripts are not sex-specifically spliced, that *Am-tra2*^{285} and *Am-tra2*^{284} are the dominant transcripts and that the relative amount of *Am-tra2* decreases at the pupal stage.

Knockdown of the *Am-tra2* gene affects embryonic viability and female splicing of the *fem* and *Am-dsx* transcripts

Table 1: Development of *Am-tra2*-dsRNA treated female embryos.

Treatment/ amount of injected dsRNA per embryo	Embryos injected	Number of embryos showing normal development after ~70h		Hatched L1 larvae	
	N	N	%	N	%
non-treated	683	594	86.9	365	53.4
dd$H_2$0	406	182	44.8	97	23.9
***Am-tra2* dsRNA-1**					
200 pg	92	0	0	0	0
100 pg	25	0	0	0	0
75 pg	60	0	0	0	0
40 pg	20	0	0	0	0
20 pg	53	0	0	0	0
4 pg	40	5	12.5	2	5.0
***Am-tra2* dsRNA-2**					
224 pg	27	0	0	0	0
96 pg	100	15	15	0	0
67 pg	104	38	36.5	0	0
56 pg	105	14	13.3	4	3.8
33 pg	146	31	21.2	6	4.1
4 pg	155	29	18.7	9	5.8

The number of individuals (N) and the relative proportions (%) with respect to the total number of initially injected embryos are shown.

In the honeybee, the primary signal *csd* mediates, in the heteroallelic condition, the female splicing of *fem* transcripts. The downstream target of *csd*, the *fem* gene, is required to direct splicing of *Am-dsx* pre-mRNAs and its own *fem* transcripts (Gempe et al., 2009), the latter of which establishes a positive feedback loop of self-regulated *fem* female splicing (Gempe et al., 2009; Pane et al., 2002). Although substantially diverged in sequence, the *fem* gene and the *tra* gene of *D. melanogaster* are orthologs, whereas the *csd* gene was derived by gene duplication of an ancestral copy of the *fem* gene (Hasselmann et al., 2008). As Csd and Fem

proteins harbor no RBD but have similar sex-determining and splice regulation functions as the Tra protein in *D. melanogaster* (Gempe et al., 2009), we proposed that *Am*-Tra2 protein is the RNA-binding cofactor that is essential for *fem* and *Am-dsx* splicing. To study the role of the *Am-tra2* gene in female splicing, we induced *Am-tra2* knockdown by RNAi in 0-3-hour-old female embryos. We injected two dsRNAs that target either the region expressing the RS1 and RRM domains (dsRNA-1) or the entire RS1 domain (dsRNA-2).

Table 2: The production of female and male *fem* and *Am-dsx* transcripts in 72- to 80-hour-old individuals in response to embryonic *Am-tra2* dsRNA-2 treatment.

Treatment	**Number of embryos**	**Number of individuals with transcripts**					
		fem			***Am-dsx***		
		solely male	**solely female**	**male and female**	**solely male**	**solely female**	**male and female**
non-treated controls							
males	10	9	0	0	10	0	0
females	14	1	8	5	1	12	1
treated females							
ddH_20	15	0	8	15	1	13	1
dsRNA-2 (33 pg)	12	6	0	1	12	0	0
dsRNA-2 (4 pg)	11	0	4	7	1	1	9

No embryo reached larval stage L1 when we injected approximately 200 pg of dsRNA-1 or dsRNA-2 per embryo (Table 1). This is an amount that is substantially below the approximately 900 pg of dsRNA per embryo that we repeatedly used in previous studies, in which we observed no lethal effect (Beye et al., 2003; Hasselmann et al., 2008). Null mutant data suggest that the *tra2* gene in *D. melanogaster* is not essential for viability (Watanabe, 1975; Fujihara et al., 1978). Therefore, we further reduced the amount of *Am-tra2*-dsRNAs until we observed fully developed L1 larvae (Table 1). At concentrations of 56 pg of dsRNA-2 and 4 pg of dsRNA-1 per embryo, we obtained the first viable L1 larvae, but at a very low frequency (5% compared to 24% in our ddH_2O-treated controls). Even when we further reduced the dsRNA-2 concentration from 56 pg to 4 pg per embryo, the hatching rate did not substantially improve, suggesting that knockdown of *Am-tra2* caused some general lethality during embryogenesis. None of the hatched L1 larvae reached the L4 stage (data

not shown), suggesting that we cannot study the direct role of *Am-tra2* in morphological sexual differentiation. Taken together, our knockdown results suggest that the *Am-tra2* gene is essential for embryogenesis in the honeybee.

We proposed that in addition to a vital role for *Am-tra2* in embryogenesis, *Am-tra2* possibly has another function in sex determination, specifically in promoting female-specific splicing of the *fem* and *Am-dsx* transcripts. The knockdown of the *fem* and the *csd* genes, which regulate female splicing and sex determination, had no general effects on lethality, suggesting that the putative role of *Am-tra2* in activation of the female pathway did not cause the embryonic lethality (Beye et al., 2003; Gempe et al., 2009). To study the sex-determining role of the *Am-tra2* gene, we injected 4 pg or 33 pg of dsRNA-2 in embryos and studied after 77-80 hours the splice patterns, irrespective of whether the larvae hatched. If *Am-tra2* promotes female splicing, we expected that knockdown of this gene would induce male-like splice patterns in these females.

The injection of 4 pg of dsRNA-2 (Table 2) induced male splicing of *Am-dsx* mRNAs in females (Fig. 4A, lanes 1-10), which is entirely absent in the control embryos (Fig. 4A, lanes 21-30), which produce only the *Am-dsx* female splice product. This result suggests that the *Am-tra2* gene is essential to promote female splicing of *Am-dsx* transcripts. The 4 pg dsRNA-2 treatment, however, did not affect the splicing of female *fem* mRNAs (Fig. 4B, lanes 1-10).

The treatment of female embryos with the higher concentration of 33 pg of dsRNA-2 per embryo compromised female splicing of the *fem* mRNAs (Fig. 4B, lanes 11-20) compared to our control embryos (Fig. 4B, lanes 26-35; Table 2), indicating that *Am*-Tra2 protein is also required for *fem* female transcript splicing. In contrary to our expectation, the knock down of *Am-tra2* gene in females produced not the alternative, male splice form of the *fem* transcripts. We consistently observed in the 33 pg-treated females the absence of the male *fem* transcript (Fig. 4B, lanes 11-20). This result suggests that *Am*-Tra2 protein may also have a role in splicing the *fem* pre-mRNAs into the male form. Taken together, these results suggest that *Am*-Tra2 promotes female splicing of the productive female *fem* mRNAs and possibly also of the non-productive male *fem* mRNAs.

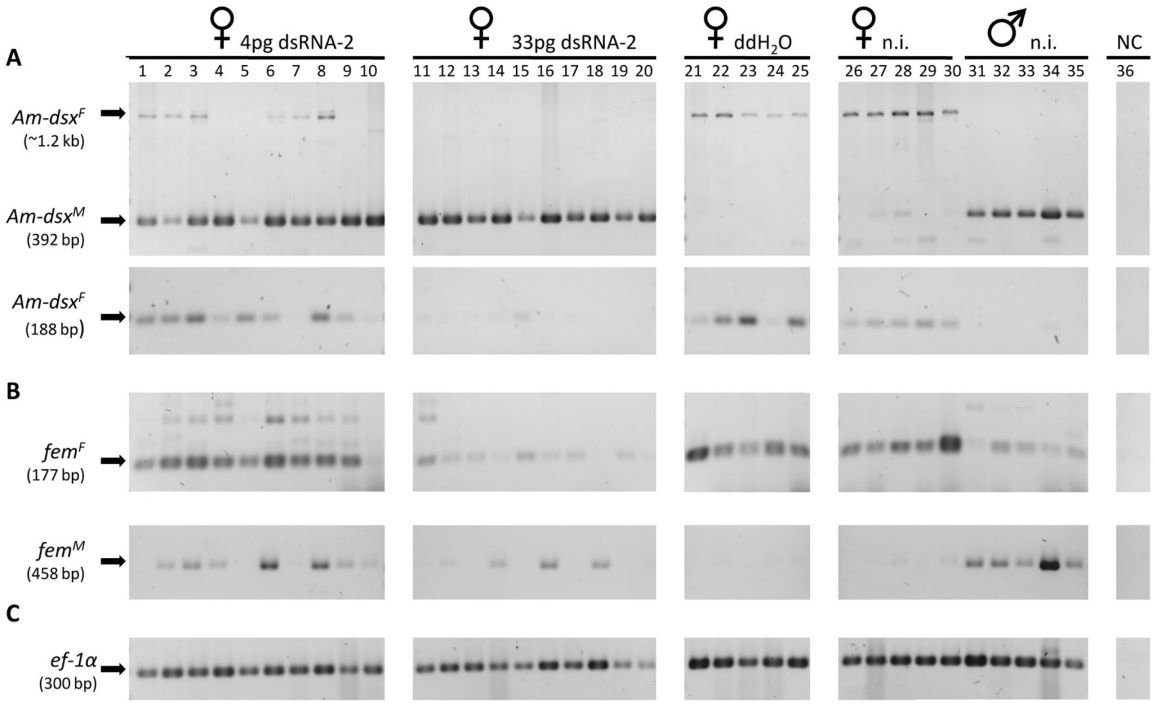

Figure 4: Sex-specific splicing of the *fem* and *Am-dsx* transcripts in *Am-tra2* dsRNA-2 treated embryos.

The male and female *Am-dsx* (A) and *fem* (B) mRNAs of individuals 77-80 hours after egg-laying were studied using semiquantitative RT-PCR. Early embryos were injected with 4 pg of *Am-tra2* dsRNA-2 (lanes 1-10), 33 pg of *Am-tra*2 dsRNA-2 (lanes 11-20) or ddH_20 (lanes 21-25). The untreated (labeled as n.i.) female and male controls are shown in lanes 26-30 and 31-35, respectively. NC denotes our control PCR in which no cDNA was added (lane 36). Fragments corresponding to the *fem* female (size of 188 bp) and male (size of 449 bp) mRNAs and the *Am-dsx* female (size of 1.2 kb and 188 bp), and male (size of 392 bp) mRNAs were resolved by agarose gel electrophoresis and stained with ethidium bromide. We used amplification of the cDNAs of the "housekeeping" gene *elongation factor 1α* (ef-1α) as a relative control to semiquantify *Am-tra2* transcripts across embryonic samples.

Discussion

Heteroallelic Csd proteins determine honeybee femaleness and set the downstream regulator of the sex determination cascade, *fem,* into the female mode by alternative splicing (Beye et al., 2003; Hasselmann et al., 2008). The Fem proteins in females maintain the female-determined state by promoting female splicing of the *fem* mRNAs (positive autoregulation) and direct female splicing of the *Am-dsx* transcripts (Hasselmann et al., 2008; Gempe et al., 2009). In this study, we showed that the *Am-tra2* gene, an ortholog of the *tra2* gene of *D. melanogaster*, is a component of the honeybee sex-determination hierarchy. The *Am-tra2* proteins are required to regulate female and male splicing of *fem* mRNAs and female splicing of the *Am-dsx* mRNAs. In addition, we showed that the *Am-tra2* gene has an essential role in embryogenesis that is not related to sex determination.

We characterized the *Am-tra2* gene in the honeybee and showed that the deduced *Am*-Tra2 proteins share the same domain structure as other Tra2 orthologs described thus far (Salvemini et al., 2009; Burghardt et al., 2005; Sarno et al., 2010; Mattox et al., 1996; Amrein et al., 1988; Goralski et al., 1989; Bandziulis et al., 1989). *Am*-Tra2 protein contains a RBD that is supposed to directly interact with the pre-mRNA and two flanking RS-rich domains that provide a potential surface for an interaction with other proteins, such as Tra proteins (Sciabica and Hertel, 2006; Graveley, 2000; Hoshijima et al., 1991; Amrein et al., 1988). We identified six splice variants of *Am-tra2* mRNAs that translate into proteins that differ in the length of the first RS domain and in the absence/presence of one amino acid (serine) in the second RS domain. The six splice variants are not sex-specifically regulated throughout development, suggesting that *Am*-Tra2 proteins are constitutively expressed. However, we observed that the level of *Am-tra2* transcription substantially decreases at the pupal stage, which possibly relates to *Am-tra2'*s function in sexual and embryonic differentiation. We also showed that the *Am-tra2* gene is not sex-specifically spliced in the gonadal tissues (Fig. 3B). This finding is in contrast to the germline specific control of *tra2* transcripts in males of *D. melanogaster*. Here, the $Tra2^{226}$ protein directs splicing of the $tra2^{179}$ transcript in the fruit fly germ line, thereby regulating the level of $Tra2^{226}$ protein expression that is critical for proper sperm formation (McGuffin et al., 1998; Mattox et al., 1990).

When we repressed *Am-tra2* gene by injecting 4 pg of dsRNAs into early embryos, we observed a sizeable reduction of the female *Am-dsx* but not of the female *fem* splice variants, suggesting a role for the *Am-tra2* gene in the regulation of female *Am-dsx* mRNA splicing (Fig. 5). Our result suggests a conserved role of the *Am*-Tra2 protein in *Am-dsx* regulation, although the canonical binding motif to which Tra2 proteins are supposed to bind is absent in the *Am-dsx* gene. This finding suggests that the Tra2 protein binding sites have evolved in the honeybee. In *D. melanogaster,* Tra2, together with the Tra proteins, binds to six repeats of a 13-nucleotide motif (TC(T/A)(A/T)C(A/G)ATCAACA) on the *dsx* pre-mRNA and promotes the activation of the weak female splice acceptor that directs the production of the female *dsx* transcripts. In other dipteran species (*M. domestica, C. capitata, Bactrocera oleae* and different *Anastrepha* species), the canonical Tra/Tra2 binding motifs are consistently present in the *dsx* genes (Hediger et al., 2004; Ruiz et al., 2005; Saccone et al., 2008; Ruiz et al., 2007; Lagos et al., 2005) and are apparently utilized to promote female splicing (Burghardt et al., 2005; Salvemini et al., 2009; Sarno et al., 2010). We conclude that the *Am*-Tra2 protein is an essential, constitutively expressed cofactor that, together with the female-specific Fem protein, directs the female processing of *dsx* transcript.

Consistent with an evolved binding site of the Fem/*Am*-Tra2 proteins, we identified several amino acid replacements in the RBD that affect the designated binding nucleotide sequence of the RNA. We found three amino acid sites in the RNP-1 sequence element that diverged with respect to the *D. melanogaster* sequence. The RNP-1 and RNP-2 sequence elements are part of the RBD, which has a βαββαβ barrel-like structure. The RNP sequence elements have exposed positions at the surface of β-sheets β1 and β3 and are used to bind directly to the ribonucleotide sequence (Amrein et al., 1994; Nagai et al., 1990; Dreyfuss et al., 1988; Merrill et al., 1988). Mutation of the first amino acid arginine of the RNP-1 sequence element in *D. melanogaster* abolishes the female processing of *dsx* pre-mRNAs (Amrein et al., 1994). This critical arginine amino acid residue in the RNP-1 sequence element is replaced in the honeybee by a lysine. These findings support our conclusion that the corresponding Fem/*Am*-Tra2 protein binding sites diverged from this of the fruitfly.

The role of the Tra2 protein in regulating female splicing of *dsx* transcripts in the silk moth *B. mori* is thus far not known. The canonical Tra/ Tra2 binding sites are also absent in the *Bm-dsx* gene*,* and female transcripts are constitutively produced, even when splicing is studied

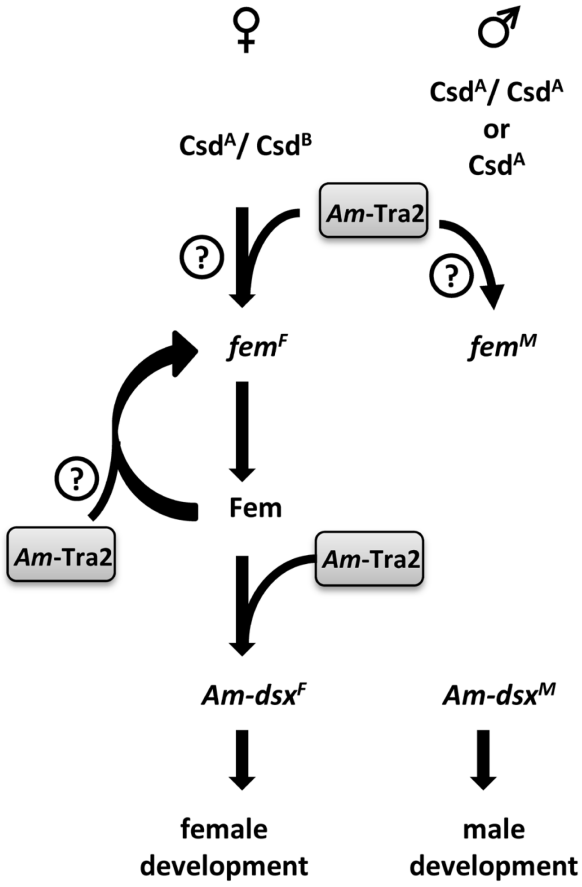

Figure 5: Model of the role of the *Am*-Tra2 protein in honeybee sex determination. *Am*-Tra2 protein is necessary for splicing of the productive female and *fem* and *Am-dsx* mRNAs. It is also required to splice the non-productive male *fem* splice form. In females, the *Am*-Tra2 protein either acts as a cofactor of heterozygous Csd proteins (Csd^A/Csd^B) or as a cofactor of Fem proteins (fem^F positive autoregulatory loop) to promote female processing of the *fem* pre-mRNAs (fem^F). *Am*-Tra2 protein is also a cofactor that directs together with Fem protein female-specific splicing of the *Am-dsx* pre-mRNA ($Am\text{-}dsx^F$). In males, *Am*-Tra2 protein directs splicing of male *fem* mRNAs in the absence of active Csd proteins (homoallelic Csd^A/Csd^A or Csd^A proteins) and/or Fem proteins.

in HeLa cell nuclear extracts (Suzuki et al., 2001). The latter result suggests that the constitutive female splicing does not require *Bm*-Tra2 activity, although this has not been shown thus far by functional studies. Factors that inhibit the female splicing and that determine maleness have been characterized. The male-specific *Bm*-IMP proteins and the *Bm*-PSI proteins bind together to the CE1 sequence at female exon 4 (Suzuki et al., 2010, 2001, 2008) to repress the activation of the female exon in the *Bm-dsx* pre-mRNA.

With the honeybees in the phylogeny at the base of holometabolous insects (Savard et al., 2006), our comparison across the different insect orders suggests that the role of Tra2 proteins in regulating female *dsx* splicing is the ancestral state in holometabolous insects.

When we repressed the *Am-tra2* gene by injecting 33 pg of dsRNAs, we also observed a reduction of the productive female and the non-productive male *fem* splice variants. This result indicates a function of the *Am-tra2* gene also on the level of splicing of the *fem* gene. We have previously shown that in females, Fem as well as heteroallelic Csd proteins direct the processing of *fem* pre-mRNAs

(Gempe et al., 2009). Fem proteins maintain the female state using a positive feedback loop, whereas the heteroallelic Csd proteins are the initial signal that determines femaleness. Our results presented here suggest that the *Am-Tra2* protein acts as an essential cofactor of female *fem* transcript splicing (Fig. 5). We suggest that the *Am*-Tra2 protein is a required cofactor of heteroallelic Csd proteins in females that mediates the binding and splicing of female *fem* pre-mRNAs. In males, where the active, heteroallelic Csd proteins are absent, the *Am*-Tra2 proteins may enhance the switch of the *fem* transcripts into the non-productive male form.

The *fem* gene is an ortholog of the *tra* gene of dipteran insects. In *D. melanogaster,* Tra2 proteins are not deployed in regulating female splicing of *tra* transcripts. Female *tra* mRNA processing in this species is regulated by the Sxl proteins (Bell et al., 1991; Valcárcel et al., 1993; Bell et al., 1988; Sosnowski et al., 1989). In contrast, Tra2 proteins in the dipteran insects *M. domestica* and *C. capitata* are, as in the honeybee, required to splice *tra* transcripts into the female form (Burghardt et al., 2005; Salvemini et al., 2009; Hediger et al., 2010). In these dipteran species, Tra2 proteins are evidently cofactors of an autoregulatory loop in females in which the maternally provided Tra proteins mediate female *tra* mRNAs. The reinforced production of more female *fem* mRNAs and Fem proteins implements female differentiation. The presence of a male-determining factor, M, apparently impairs this *tra* positive regulatory loop, resulting in male *tra* pre-mRNA splicing and male differentiation (Hediger et al., 2010; Pane et al., 2002). In addition, honeybee *Am*-Tra2 proteins are also required to produce the male *fem* transcripts. This finding suggests that Tra2 proteins in the honeybee can directly mediate the splicing of the male *fem* transcripts that are non-productive and are not required for male differentiation (Gempe et al., 2009). This *tra2* gene function has not been reported in other insects.

Our knockdown studies in early embryos also showed that the *Am-tra2* gene is essential and has a vital role in embryogenesis. We suggest that this lethal effect during embryogenesis is not caused by unspecific effects due to our dsRNA method, as we (i) targeted two different regions of the transcript with our dsRNAs, (ii) observed lethal effects with dsRNA concentrations that were substantially below that of previous experiments (4-40 times) that showed no lethal embryonic effects (Beye et al., 2003; Hasselmann et al., 2008) and (iii) the viability of embryos did not further increase above approximately 5% when we

decreased the dsRNA concentration by 1/10th (Table 1). Our previous results showed that knockdown of the components of the sex-determination hierarchy produced no lethal embryonic outcome (Gempe et al., 2009; Hasselmann et al., 2008; Beye et al., 2003), suggesting that the lethal effects of the *Am-tra2* gene are not related to the regulation of sex determination. This additional role in embryogenesis is absent in other dipteran insects (Watanabe, 1975; Fujihara et al., 1978; Salvemini et al., 2009; Burghardt et al., 2005), suggesting that this function is an evolved feature in insects.

Taken together, our results suggest that the *Am-tra2* gene is a non-sex-specifically expressed cofactor that is essential for generating the productive female and non-productive male *fem* transcripts. We suggest that the *Am*-Tra2 protein is a required cofactor of heteroallelic Csd proteins that mediates binding to *fem* pre-mRNAs and directs female *fem* splicing. This process initiates the primary signal *csd* by activating the female pathway. It may promote a positive regulatory loop in females in which Fem, together with the *Am*-Tra2 proteins, mediates the production of more female *fem* mRNAs, which is a role that has also been reported in some other dipteran insects. In males and in the absence of active, heteroallelic Csd proteins, *Am*-Tra2 may enhance the switch of *fem* transcripts into the non-productive male form. We also showed using knockdown experiments that the *Am*-Tra2 protein is an essential cofactor of female *Am-dsx* splice regulation. Although the canonical Tra/Tra2 protein binding sites are absent in the *Am-dsx* gene, this *Am*-Tra2 protein function is shared with the Tra2 proteins of other dipteran insects, suggesting an ancestral role. In addition, the *Am-tra2* gene has an essential function in honeybee embryogenesis that is unrelated to sex determination and has thus far not been reported in other insects.

Acknowledgements

We thank the reviewer for helpful comments on the manuscript; Eva-Maria Theilenberg, Marion Müller-Borg, Jessica Langer and Isabel Röös for technical support; Dalibor Titera for providing bee crosses; and Michael Griese for bee keeping support. This work was supported by grants from the Deutsche Forschungsgemeinschaft DFG.

Supplementary Data

Table S3: Sequences of oligonucleotides that were used

to synthesize dsRNA	
#22M: tra-2_ds_FOR	TAATACGACTCACTATAGGGCGAAGTCGTAGTCGCAGCCGTAGTCGTT
#23M: tra-2_ds_REV	TAATACGACTCACTATAGGGCGACTGGTGTTGGTGTATGAGCTCGTTG
#591	TAATACGACTCACTATAGGGAGTCGTAGTGCAAGTCCTAGAAGACC
#592	TAATACGACTCACTATAGGGATTTTCCCTGTTTCCAACATGAC
to analyze *Am-dsx* splicing	
#417	CTATTGGAGCACAGTAGCAAACTTG
#418	GGCTACGTATGTTTAGGAGGACC
#419	GAAACAATTTTGTTCAAAATAGAATTCC
to analyze *fem* splicing	
#412	CTGATTTTTCAATATTTACAGCTAAAACTGTAC
#523	CAACATCTGATGAACTTAAACGG
#410	TGAAGTTAATAACATATTTTTAATTCATCAATGAAG
#566	TGTACCATCTGAAGATTCTAATTTTTCG
to amplify *elongation factor-1α*	
#EM033	CGTTCGTACCGATCTCCGGATG
#EM034	GCTGCTGGAGCGAATGTTAC
in 5'RACE experiments	
5'RACE OLLI:	TGAACGGCTTCGTG
in 3'RACE experiments	
#33M (3'RACE J1 OUTER)	ACT CTC GCG AAT GTG ATA GGA CCA T
#34M (3'RACE J2 INNER)	TCA CAC TCC CGC AGT CCA ATG TCA T
3'RACE OLLI	AGAACAGTGTGCAG
to clone full ORF of *Am-tra2*	
#359	GATCGGATCCATGAGTGACATTGAGCGAAGTAGTAG
#421	TGACACGCGTTTAATATCGACGTGGTGAATAAGAGC

Figure S1: Multiple Sequence Alignment of the Tra2 proteins of *Apis mellifera (A.m.), Drosophila melanogaster (D.m.)* and *Bombyx mori (B.m)*. The RNA recognition motif (RRM) is highlighted in black, and the linker region is highlighted in dark grey. The arginine-serine rich motifs (RS1 and RS2) are shown in light grey. The RNP-1 and RNP-2 sequence elements are marked by black bars.

Reference list

Amrein, H., Gorman, M., and Nöthiger, R. (1988). The sex-determining gene *tra-2* of *Drosophila* encodes a putative RNA binding protein. Cell *55*, 1025–1035.

Amrein, H., Hedley, M. L., and Maniatis, T. (1994). The Role of Specific Protein-RNA and Protein-Protein Interactions in Positive and Negative Control of pre-mRNA Splicing by Transformer-2. Cell *76*, 735–746.

Amrein, H., Maniatis, T., and Nöthiger, R. (1990). Alternatively spliced transcripts of the sex-determining gene *tra-2* of *Drosophila* encode functional proteins of different size. The EMBO journal *9*, 3619–3629.

Bandziulis, R. J., Swanson, M. S., and Dreyfuss, G. (1989). RNA-binding proteins as developmental regulators. Genes & Development *3*, 431–437.

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*, 229–239.

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*, 1037–1046.

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: news and reviews in molecular, cellular and developmental biology *26*, 1131–1139.

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*, 419–429.

Beye, M., Härtel, S., Hagen, A., Hasselmann, M., and Omholt, S. W. (2002). Specific developmental gene silencing in the honey bee using a homeobox motif. Insect molecular biology *11*, 527–532.

Burghardt, G., Hediger, M., Siegenthaler, C., Moser, M., Dübendorfer, A., and Bopp, D. (2005). The *transformer2* gene in *Musca domestica* is required for selecting and maintaining the female pathway of development. Development genes and evolution *215*, 165–176.

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.

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*, 1733–1741.

Cline, T. W., and Meyer, B. J. (1996). Vive la différence: males vs females in flies vs worms. Annual review of genetics *30*, 637–702.

Dearden, P. K., Wilson, M. J., Sablan, L., Osborne, P. W., Havler, M., Mcnaughton, E., Kimura, K., Milshina, N. V., Martin, H., Gempe, T., et al. (2006). Patterns of conservation and change in honey bee developmental genes. Genome Research, 1376–1384.

Dreyfuss, G., Swanson, M. S., and Pinol-Roma, S. (1988). Heterogeneous nuclear ribonucleoprotein particles and the pathway of mRNA formation. Trend in biochemical sciences, 86–91.

Erickson, J. W., and Quintero, J. J. (2007). Indirect effects of ploidy suggest X chromosome dose, not the X:A ratio, signals sex in *Drosophila*. PLoS biology *5*, e332.

Fujihara, T., Kawabe, M., and Oishi, K. (1978). A sex-transformation gene in *Drosophila*. Journal of Heredity *69*, 229–236.

Gempe, T., Hasselmann, M., Schiøtt, M., Hause, G., Otte, M., and Beye, M. (2009). Sex determination in honeybees: two separate mechanisms induce and maintain the female pathway. PLoS biology *7*, e1000222.

Gempe, T., and Beye, M. (2011). Function and evolution of sex determination mechanisms, genes and pathways in insects. BioEssays : news and reviews in molecular, cellular and developmental biology *33*, 52–60.

Goralski, T. J., Edstrom, J.-E., and Baker, B. S. (1989). The sex determination locus *transformer-2* of *Drosophila* encodes a polypeptide with similarity to RNA binding proteins. Cell *56*, 1011–1018.

Graveley, B. R. (2000). Sorting out the complexity of SR protein functions. RNA (New York, N.Y.) *6*, 1197–1211.

Hasselmann, M., Gempe, T., Schiøtt, 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*, 519–522.

Hasselmann, M., and Beye, M. (2004). Signatures of selection among sex-determining alleles of the honey bee. Proceedings of the National Academy of Sciences of the United States of America *101*, 4888–4893.

Hazelrigg, T., and Tu, C. (1994). Sex-specific processing of the *Drosophila exuperantia* transcript is regulated in male germ cells by the *tra-2* gene. Proceedings of the National Academy of Sciences of the United States of America *91*, 10752–10756.

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

Hediger, M., Henggeler, C., Meier, N., Perez, R., Saccone, G., and Bopp, D. (2010). Molecular characterization of the key switch F provides a basis for understanding the rapid divergence of the sex-determining pathway in the housefly. Genetics *184*, 155–170.

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*, 579–586.

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*, 833–836.

Inoue, K., Hoshijima, K., Higuchi, I., Sakamoto, H., and Shimura, Y. (1992). Binding of the *Drosophila* transformer and transformer-2 proteins to the regulatory elements of *doublesex* primary transcript for sex-specific RNA processing. Proceedings of the National Academy of Sciences of the United States of America *89*, 8092–8096.

Kamakura, M. (2011). Royalactin induces queen differentiation in honeybees. Nature *473*, 478–483.

Kucharski, R., Maleszka, J., Foret, S., and Maleszka, R. (2008). Nutritional control of reproductive status in honeybees via DNA methylation. Science (New York, N.Y.) *319*, 1827–1830.

Lagos, D., Ruiz, M. F., Sánchez, 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*, 111–121.

Lynch, K. W., and Maniatis, T. (1996). Assembly of specific SR protein complexes on distinct regulatory elements of the *Drosophila doublesex* splicing enhancer. Genes & Development *10*, 2089–2101.

Lynch, K. W., and Maniatis, T. (1995). Synergistic interactions between two distinct elements of a regulated splicing enhancer. Genes & Development *9*, 284–293.

Madigan, S. J., Edeen, P., Esnayra, J., and McKeown, M. (1996). *att*, a target for regulation by tra2 in the testes of *Drosophila melanogaster*, encodes alternative RNAs and alternative proteins. Molecular and cellular biology *16*, 4222–4230.

Mattox, W., McGuffin, M. E., and Baker, B. S. (1996). A negative feedback mechanism revealed by functional analysis of the alternative isoforms of the *Drosophila* splicing regulator transformer-2. Genetics *143*, 303–314.

Mattox, W., Palmer, M. J., and Baker, B. S. (1990). Alternative splicing of the sex determination gene *transformer-2* is sex-specific in the germ line but not in the soma. Genes & Development *4*, 789–805.

Mattox, W., and Baker, B. S. (1991). Autoregulation of the splicing of transcripts from the *transformer-2* gene of *Drosophila*. Genes & Development *5*, 786–796.

McGuffin, M. E., Chandler, D., Somaiya, D., Dauwalder, B., and Mattox, W. (1998). Autoregulation of *transformer-2* alternative splicing is necessary for normal male fertility in *Drosophila*. Genetics *149*, 1477–1486.

Merrill, B. M., Stone, K. L., Cobianchi, F., Wilson, S. H., and Williams, K. R. (1988). Phenylalanines that are conserved among several RNA-binding proteins form part of a nucleic acid-binding pocket in the A1 heterogeneous nuclear ribonucleoprotein. The Journal of biological chemistry *263*, 3307–3313.

Nagai, K., Oubridge, C., Jessen, T. H., Li, J., and Evans, P. R. (1990). Crystal structure of the RNA-binding domain of the U1 small nuclear ribonucleoprotein A. Nature *348*, 515–520.

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 Biochimica et Biophysica Sinica *37*, 728–736.

Nygaard, S., Zhang, G., Schiøtt, M., Li, C., Wurm, Y., Hu, H., Zhou, J., Ji, L., Qiu, F., Rasmussen, M., et al. (2011). The genome of the leaf-cutting ant *Acromyrmex echinatior* suggests key adaptations to advanced social life and fungus farming. Genome Research *21*, 1339–1348.

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 (Cambridge, England) *129*, 3715–3725.

Ruiz, M. F., Eirín-López, J. M., Stefani, R. N., Perondini, A. L. P., Selivon, D., and Sánchez, L. (2007). The gene *doublesex* of *Anastrepha* fruit flies (Diptera, Tephritidae) and its evolution in insects. Development genes and evolution *217*, 725–731.

Ruiz, M. F., Stefani, R. N., Mascarenhas, R. O., Perondini, A. L. P., Selivon, D., and Sánchez, L. (2005). The gene *doublesex* of the fruit fly *Anastrepha obliqua* (Diptera, Tephritidae). Genetics *171*, 849–854.

Saccone, G., Salvemini, M., Pane, A., and Polito, L. C. (2008). Masculinization of XX *Drosophila* transgenic flies expressing the *Ceratitis capitata* DoublesexM isoform. The International journal of developmental biology *52*, 1051–1057.

Salvemini, M., Robertson, M., Aronson, B., Atkinson, P., Polito, L. C., and Saccone, G. (2009). *Ceratitis capitata transformer-2 gene* is required to establish and maintain the autoregulation of *Cctra*, the master gene for female sex determination. The International journal of developmental biology *53*, 109–120.

Sarno, F., Ruiz, M. F., Eirín-López, J. M., Perondini, A. L. P., Selivon, D., and Sánchez, L. (2010). The gene *transformer-2* of *Anastrepha* fruit flies (Diptera, Tephritidae) and its evolution in insects. BMC evolutionary biology *10*, 140.

Savard, J., Tautz, D., Richards, S., Weinstock, G. M., Gibbs, R. A., Werren, J. H., Tettelin, H., and Lercher, M. J. (2006). Phylogenomic analysis reveals bees and wasps (Hymenoptera) at the base of the radiation of Holometabolous insects. Genome Research *16*, 1334–1338.

Sciabica, K. S., and Hertel, K. J. (2006). The splicing regulators Tra and Tra2 are unusually potent activators of pre-mRNA splicing. Nucleic acids research *34*, 6612–6620.

Sosnowski, B. A., Belote, J. M., and McKeown, M. (1989). Sex-Specific Alternative Splicing of RNA from the *transformer* Gene Results from Sequence-Dependent Splice Site Blockage. Cell *58*, 449–459.

Suzuki, M. G., Imanishi, S., Dohmae, N., Asanuma, M., and Matsumoto, S. (2010). Identification of a male-specific RNA binding protein that regulates sex-specific splicing of *Bmdsx* by increasing RNA binding activity of *Bm*PSI. Molecular and cellular biology *30*, 5776–5786.

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. Molecular and cellular biology *28*, 333–343.

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 biochemistry and molecular biology *31*, 1201–1211.

The *Tribolium* Genome Sequencing Consortium (2008). The genome of the model beetle and pest *Tribolium castaneum*. Nature *452*, 949–955.

Valcárcel, 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*, 171–175.

Watanabe, T. K. (1975). A New Sex-Transforming Gene on the Second Chromosome of *Drosophila melanogaster*. Jap. J. Genet. *50*, 269–271.

The *Am-tra2* gene is an essential cofactor of female splice regulation at two levels of the *complementary sex determination* hierarchy of the honeybee

Journal: Genetics (under review)

Impact Factor: 4,242

1st author

Author's contribution: 80%

- Characterization of the *Am-tra2* gene (together with Miriam Müller)
- Transcriptional studies of the *Am-tra2* gene throughout development
- Functional studies of the *Am-tra2* gene (except injection of dsRNA-1)
- Phylogenetic and molecular evolutionary sequence analyses
- Statistical analysis
- Authoring the manuscript

Manuscript II

Fem and *Am*-Tra2 proteins direct female specific splicing of the *Am-dsx* pre-mRNA by activation of the female splice donor site in the honeybee *Apis mellifera*

Inga Nissen, Heiner Schaal, Martin Beye

In preparation for submission to Molecular and Cellular Biology

Abstract

Molecular analysis of the diversity of sex determination systems can contribute in our understanding of how and why pathways and gene functions evolved. Sex in the honey bee (*Apis mellifera*) is determined by heterozygosity of the *complementary sex determiner* (*csd*) gene. This sex-specific signal is realized by a cascade of genes that involves the *feminizer* (*fem*) and the *Apis mellifera-doublesex (Am-dsx)* gene. We have previously reported that Csd proteins control the female-specific processing of the *fem* transcript encoding the Fem protein, while Fem proteins are required to promote female specific splicing of *Am-dsx* transcripts. However, it is still unknown whether Fem protein is the direct regulator of female *Am-dsx* splicing and how this process is molecularly regulated. To analyze the molecular control of splice regulation of female and male *Am-dsx* transcripts, we have developed an *in vivo* splice assay in *Sf*21-cells by using *Am-dsx* minigene constructs to mimic sex-specific *Am-dsx* splicing. Beside Fem proteins, we co-expressed *Am*-Tra2 protein as a candidate protein that has the ability to directly interact with the pre-mRNAs and analyzed the *Am-dsx* splicing products. Here we report that both proteins, female specific Fem and non-sex-specific *Am*-Tra2 proteins are direct regulators of alternative splicing of the *Am-dsx* minigene transcripts. Fem and *Am*-Tra2 proteins activate the splice donor site of the female exon that requires exonic splice enhancer sequences (AAAGAG) within the region 20 nucleotides upstream of the splice donor site. In *Drosophila melanogaster* Tra and Tra2 also direct splicing of female *dsx* transcripts, but here the proteins are activating a splice acceptor site in the female exon by using other ESE sequences. This suggests that the underlying mechanism of producing the sex-specific transcription factor Dsx in these insects has diverged.

Introduction

Sex in the honeybee (*Apis mellifera*) is determined by the heteroallelic combination of the *complementary sex determiner* (*csd*) gene (Beye et al., 2003) in contrast to the well studied sex chromosome system of *Drosophila melanogaster* (Cline and Meyer, 1996; Erickson and Quintero, 2007). More than 15 allelic variants of this *complementary sex determining* gene are segregating in honeybee populations (Hasselmann et al., 2008; Hasselmann and Beye, 2004) that differ at an average of approximately 6% of their amino acid residues in the putative specifying domain (Hasselmann et al., 2008; Hasselmann and Beye, 2004). Bees heterozygous at the *csd* gene are females, whereas bees homozygote or hemizygote (haploid individuals) at *csd* are males. In the honeybee colonies queens lay unfertilized eggs to produce fertile males (drones) and fertilized eggs to produce females that either differentiate into workers or queens as determined by worker's differential feeding of larvae with royal jelly (Kucharski et al., 2008; Kamakura, 2011). Diploid males in a colony do not survive because they are eaten by worker bees shortly after they hatch from the egg. The *csd* gene translates into an SR-type protein (Beye, 2004) that in the heteroallelic constitution is required for female differentiation and directs the female splicing of *fem* pre-mRNAs (Hasselmann et al., 2008; Gempe et al., 2009). This has been shown by *fem* and *csd* gene knockdown studies. The female spliced *fem* mRNAs translate into Fem proteins that mediate the splicing of the female *Am-dsx* mRNAs. In males, the activity of the primary signal *csd* that newly evolved in the honeybee lineage (Hasselmann et al., 2008) and the downstream target gene *fem* are not required. In the presence of homoallelic or hemiallelic Csd proteins, non-productive *fem* mRNAs are generated by introducing a premature translational stop codon into exon 3 during the splice process (Gempe et al., 2009). In the absence of active Fem protein *Am-dsx* pre-mRNAs are spliced into the male mode, lacking female exon 5. The alternative spliced *Am-dsx* transcripts in males and females translate into transcription factors that differ in a part of their C-terminal sequence (Dearden et al., 2006; Cho et al., 2007).

Here, we study the splice control of *Am-dsx* transcripts in honeybees. Dsx orthologs are transcription factors of the DM type that have an atypical zinc-finger domain in common and are key players in aspects of sexual differentiation in insects, crustacean, nematodes and mammalians (Vincent et al., 2001; Keisman and Baker, 2001;

Christiansen et al., 2002; Keisman et al., 2001; Kopp et al., 2000; Sánchez et al., 2001; Sánchez and Guerrero, 2001; Williams et al., 2008; Raymond et al., 1998; Hediger et al., 2004; Suzuki et al., 2003; Kato et al., 2011; Matson et al., 2011) by integrating sex-specific information within the general developmental program. In dipteran species including *D. melanogaster*, *Ceratitis capitata*, *Anastrepha obliqua*, *Musca domestica*, *Bactrocera oleae* (Ruiz et al., 2005, 2007; Hediger et al., 2004; Saccone et al., 2008; Lagos et al., 2005; Burtis and Baker, 1989; Hoshijima et al., 1991), as well as in the lepidopteran insect *Bombyx mori* (Ohbayashi et al., 2001; Suzuki et al., 2001, 2003) and the hymenopterans *A. mellifera* and *Nasonia vitripennis* (Oliveira et al., 2009; Cho et al., 2007), *dsx* is sex-specific regulated by an alternative splicing process. These female and male *dsx* mRNAs translate Dsx proteins which have the N-terminal regions in common, but have sex-specific sequences in segments of the oligomerization domain OD2 (An et al., 1996; Oliveira et al., 2009; Ruiz et al., 2007, 2005; Hediger et al., 2004; Saccone et al., 2008; Lagos et al., 2005; Cho et al., 2007; Ohbayashi et al., 2001). The sex-specific peptides of Dsx proteins control the transcription of target genes which are necessary for somatic male and female differentiation by binding to *cis* regulatory elements (Burtis and Baker, 1989; Coschigano and Wensink, 1993; Sánchez et al., 2001; Keisman et al., 2001; Williams et al., 2008).

In females of *D. melanogaster*, the splicing of the *dsx* transcripts is directly regulated by the Tra and Tra2 proteins. The SR-type protein Tra together with Tra2 protein directly interacts with Exonic Splicing Enhancers (ESEs) and a purine rich element (PRE) located in exon 4 to activate a weak splice acceptor site (3´ splicing site) what leads to the inclusion of the female-specific exon 4 (Burtis and Baker, 1989; Inoue et al., 1992; Hedley and Maniatis, 1991; Lynch and Maniatis, 1996, 1995; Sciabica and Hertel, 2006; Tian and Maniatis, 1993). The ESEs are located ~300 nt downstream of the splicing acceptor site and consist of six repeats (*dsx*RE) of a 13-nucleotide (nt) sequence with the consensus motif (TC(T/A)(T/A)C(A/G)ATCAACA) (the so called *dsx* repeat elements (*dsx*RE)). This Tra/Tra2 binding motif has been observed in other dipteran species that suggests a canonical motif for Tra/Tra2 dependent *dsx* splice regulation (Salvemini et al., 2009; Hediger et al., 2004; Burghardt et al., 2005). The inclusion of the female *dsx* exon 4 results in the usage of a female specific polyadenylation site that causes cleavage of female transcripts ~570 nt downstream of the Tra/Tra2 binding sites. In contrast in males a polyadenylation signal terminates transcription after the male specific exon 5 (Hedley

and Maniatis, 1991). In *D. melanogaster* males, *dsx* pre-mRNAs are spliced into the male form in the absence of Tra proteins (Hoshijima et al., 1991; Burtis and Baker, 1989).

In the lepidopteran insect *B. mori* in which females are heterogametic (*ZW* chromosomes) and males homogametic (*ZZ*) the mechanism of sexual regulation of *dsx* splice control differs, although like in the *D. melanogaster dsx* gene, the female specific exons are retained. Female splicing is regulated by default and requires no additional signal from the sex-determining pathway. Unlike in *D. melanogaster,* the female-specific exon 5 of *Bm-dsx* gene has a strong 3' splice (acceptor site) site and the canonical Tra/Tra2 binding sites are apparently absent in the *Bm-dsx* gene. Male splicing of the *Bm-dsx* transcripts, however, requires the male-specific *B. mori* IGF-II mRNA binding (*Bm*-IMP) proteins (Suzuki et al., 2001, 2008, 2010). Together with P-element somatic inhibitor (*Bm*-PSI) proteins (Suzuki et al., 2001, 2008, 2010) they mediate the exclusion of the female-specific exon 4 by binding to the Exonic Splicing Silencer (ESS) element CE1. CE1 is a 20 nucleotides long sequence that is located in exon 4 and has reiterated UAA motif.

Here we study how female and male splicing of the *Am-dsx* transcripts is regulated in the honeybee in the absence of canonical *D. melanogaster* Tra/Tra2 binding sites. The Fem protein, like Tra in *D. melanogaster* is a female-specific factor that directs female splicing of *Am-dsx* pre-mRNAs. The Fem protein (403 aa) has substantially diverged in the amino acid sequence compared to the predicted *D. melanogaster* Tra ortholog (197 aa) and has an N-terminal region (212 aa) that is entirely absent in the Tra protein of the fruitfly. We have thus far no information how male *Am-dsx* mRNA processing is controlled and whether a male-specific factor is required.

In this study we report on another mechanism of *Am-dsx* splice regulation in the honeybee. The Fem/*Am*-Tra2 proteins activate the 5’ donor site of the female exon 5 that requires a new sequence motif leading to the inclusion of the female exon and the female transcript. Male *Am-dsx* splice variants are produced in the absence of an activity provided by the sex-determination pathway (default state) implying that the sexual switch at the level of *Am-dsx* is entirely controlled by absence and presence of Fem proteins.

Materials and Methods

Plasmid constructions

The pDO vector which was used for construction of vectors pDO Fem, pDO *Am*-Tra2, pDO *Am-dsx*^{mg} and pDO *hGH* is derived from the pIZ/V5-His vector (Invitrogen). We inserted in the multiple cloning site (MCS) a sequence that transcribes a Rubia-EGFP-fluorescence protein. Downstream of the second OpIE2 promoter we inserted a second multiple cloning site (MCS2) and a second OpIE2 polyA-site (for sequence of pDO vector see supplementary data Fig. S2 B).

The full length *feminizer* (Fem Srev238P; GenBank: EU100941.1) open reading frame (ORF) was amplified by PCR using primers #367 and #369 from vector pGEMT Fem Srev238P excluding the last three nucleotides. The full-length *Apis mellifera transformer2 (Am-tra2)* (splice variant *Am-tra2*^{285}; GenBank: JQ518311) ORF was amplified by PCR with primers #359 and #421. Both fragments were cloned into pDO vector using the *Bam*HI and *Mlu*I restriction sites by replacing the Rubia-eGFP fragment producing the pDO Fem and pDO *Am*-Tra2 expression plasmid.

To generate the *Am-dsx* minigene construct, a fragment containing *Am-dsx* exon 4, intron 4, exon 5 and first 225 nucleotides of intron 5, was amplified by PCR with primers #245 and #246 from genomic DNA. A second fragment comprising *Am-dsx* last 245 nucleotides of intron 5 and complete exon 6 was amplified by PCR using primers #247 and #248 from genomic DNA. These fragments were cloned into pGEMT vector (pGEMT-Kit, Promega) that produced the plasmid pGEMT dsx_Ex4-In5_225 and pGEMT dsx-In5_last245bp-Ex6. The second *Am-dsx* fragment was restricted from pGEMT dsx-In5_last245bp-Ex6 vector using *Spe*I and *Not*I restriction enzymes and cloned into the pGEMT dsx_Ex4-In5_225 plasmid using *Spe*I and *Not*I restriction sites to create the pGEMT *Am-dsx*^{mg} plasmid. *Am-dsx*^{mg} sequence was amplified by PCR using primers #317 and #320 and cloned into pDO vector using the *Avr*II and *Xho*I restriction sites, by replacing the *Avr*II/*Xho*I fragment downstream of the second OpIE2 promoter. This cloning produced the *Am-dsx* minigene plasmid pDO *Am-dsx*^{mg}. This minigene is identical from exon 4-6 to the genomic sequence (GenBank: GB18426), except for the intron 5 that we shortened from 39550 nt to about 470 nt (containing first 225 nt and last 245 nt of intron 5).

To create the pDO *hGH* plasmid, a fragment comprising last 10 nucleotides of exon 1 of *human Growth Hormone* (*hGH*) (GenBank: NT_010783.15) along with full intron 1, exon 2,

intron 2, exon 3, intron 3, exon 4, intron 4, exon 5, and first 198 nucleotides of exon 5 was amplified from vector pXGH5 using primers #496 and #497. This amplified product was cloned into pDO vector using the *Hind*III and *Xba*I restriction sites by replacing the Rubia-eGFP fragment. In *Sf*21-cells the transcribed *hGH* pre-mRNA is spliced into 313 nt long transcripts that contain exon 1, first 105 nt of Intron 1 and exon 5 (for sequence of spliced and unspliced *hGH* transcripts see supplementary data, Fig. S2 A).

To produce the pOPI *env*/pRR splice reporter gene plasmid, a fragment containing OpIE2 baculovirus promoter sequence (for sequence see supplementary data, Fig. S2 C) was amplified by PCR from vector pIZ/V5-His using primers #611 and #613. This amplified product was cloned into *SV-env/GAR egfp* splice reporter gene plasmid using the *Cla*I and *Eco*RI restriction sites by replacing the SV40 promoter sequence.

The splice reporter gene, the plasmid *SV-env/GAR egfp*, is an eGFP fused version of SV-*env*/GAR plasmid (Caputi et al., 2004) which carries *env* ((*viral*) *envelope glycoprotein*) sequence fused to *egfp (SV-env/GAR egfp).* The GAR sequence is a strong ESE element within the HIV-1 exon 5 that enhances splicing from the SD1 to the so called splice acceptor SA7 in human cells (Caputi et al., 2004). To identify putative ESE sequences targeted by Fem and *Am*-Tra2 we replaced the GAR element by putative regulative *Am-dsx* DNA fragments (pRRs) (figure 3 A + B). To generate the plasmid pOPI *env*/pRR1, pOPI *env*/pRR2, pOPI *env*/pRR3, pOPI *env*/pRR4, pOPI *env*/pRR1_mut1, pOPI *env*/pRR1_mut2, pOPI *env*/pRR1_mut3, pOPI *env*/pRR1_mut5 and pOPI *env*/pRR1_mut7 the GAR_ESE-*Eco*RI/*Sac*I fragment was replaced by the corresponding putative regulatory region (pRR) fragments (for sequence of pRRs see supplementary data, table S1).

Cell culture and transfection

*Sf*21 cells were maintained in Spodopan medium (PAN Biotech) supplemented with 20 µg/mL of gentamycin (Roth) following the manufacturer's instructions (Invitrogen). To assay *Am-dsx* minigene splicing or *env*/ pRR splice reporter gene splicing we seeded 1 x 10^6 cells/ well in six-well plates. We transiently transfected the cells with different amounts of plasmid DNA using 5µl Roti®-Insectofect reagent (Roth, Karlsruhe, Germany) per sample, following the procedure reported by manufacturer's instructions. 24 hours after transfection, we replaced the transfection reagent with Spodopan/gentamycin medium. Cells were harvested

72 hours after transfection. The transfection experiments were replicated at least three times.

RT-PCR analysis

Total RNA from *Sf*21cells was isolated using the TRIzol® protocol (GIBCO BRL Life Technologies, Germany). RNA was resolved in 11 µl of nuclease free water. The first strand cDNA from mRNA was generated by reverse transcription using 100 pmol oligo$(dT)_{18}$ primer and 200 U RevertAid^{TM} Reverse Transcriptase following the protocol of the supplier (Fermentas GmbH, St. Leon-Rot, Germany). To analyze splicing patterns of *Am-dsx* minigene mRNAs transcribed from pDO *Am-dsx*^{mg}, we amplified male and female specific fragments corresponding to the female and male transcripts by two separate PCRs using two sets of oligonucleotides (male: #417/#418; female #417/#419). To analyze transcripts of pOPI *env*/pRR splice reporter gene, PCR was performed using oligonucleotides #709 and #641. *hGH* mRNA fragments (217 nucleotides) were amplified with primers #465 and #597.

Semiquantitative RT-PCR was standardized in respect to the splice product of the co-transfected *hGH* control. All RT-PCR fragments were resolved by agarose gel electrophoresis (2%), stained with ethidium bromide and visualized with the UV-solo (Biometra). We show representative gel pictures of the three replicate RT-PCRs. Structure of splice products was determined by sequencing. The analysis of splice products was studied from three independent transfection assays and each of them was semiquantitatively analyzed by RT-PCRs in three replicates.

Results

Male-specific splicing is directed in the absence of Fem and *Am*-Tra2 proteins in *Sf*21 cells

Previous results showed that exon 5 of the *Am-dsx* pre-mRNA (Fig. 1A) is included in females in the presence of Fem activity and that exon 5 is excluded in its absence (Gempe et al., 2009). To study the mechanisms how the exon is selected and what trans-acting factors and binding sequence elements are required to direct sex-specific splicing we developed a splice assay in *Sf*21 cells. We used an *Am-dsx* minigene construct (Fig. 1B; cloned into plasmid pDO *Am-dsx*^{mg}) that we transfected into *Sf*21 cells to study sex-specific inclusion or exclusion of exon 5. To make the *Am-dsx* gene accessible to analysis we shortened the large intron 5 to the first 225 and the last 245 nucleotides. We co-transfected cells with pDO *hGH* transcribing pre-mRNA of the *hGH* gene (Selden et al., 1986) that served as transfection and splice control in our semiquantitative RT-PCR based assay.

We observed male-specific exclusion of exon 5 in cells that were transiently transfected solely with the *Am-dsx* minigene construct (Fig 1 C, lane 3). We detected no female splicing and no inclusion of exon 5 (Fig 1 C, lane 3). Beside the most abundant fragment corresponding to the male splice variant (*Am-dsx*$^{mg\text{-}M}$) we found low amounts of two alternative splice products (*Am-dsx*$^{mg\text{-}R1}$ and *Am-dsx*$^{mg\text{-}R2}$) in *Sf*21 cells that we did not detect in RT-PCR experiments from honeybee embryos. Splice product *Am-dsx*$^{mg\text{-}R1}$ includes full exon 4, the last 194 nucleotides of exon 5 and exon 6 (Fig. 1B +C; Fig. S2 D) suggesting that the female exon 5 splice donor is activated together with another splice acceptor (SA2; Fig. 1B) that is not used in honeybees. We were not able to sequence splice product *Am-dsx*$^{mg\text{-}R2}$. We next amplified fragments specifically of the female transcripts that include exon 4 and the beginning of exon 5 by using oligonucleotide primer pairs #417 and #419, but did not detect any PCR fragment (Fig 1B and Fig 1C, lane 3). These results suggest that our *Am-dsx* minigene can mimic male-specific splicing of *Am-dsx* pre-mRNAs in *Sf*21 cells.

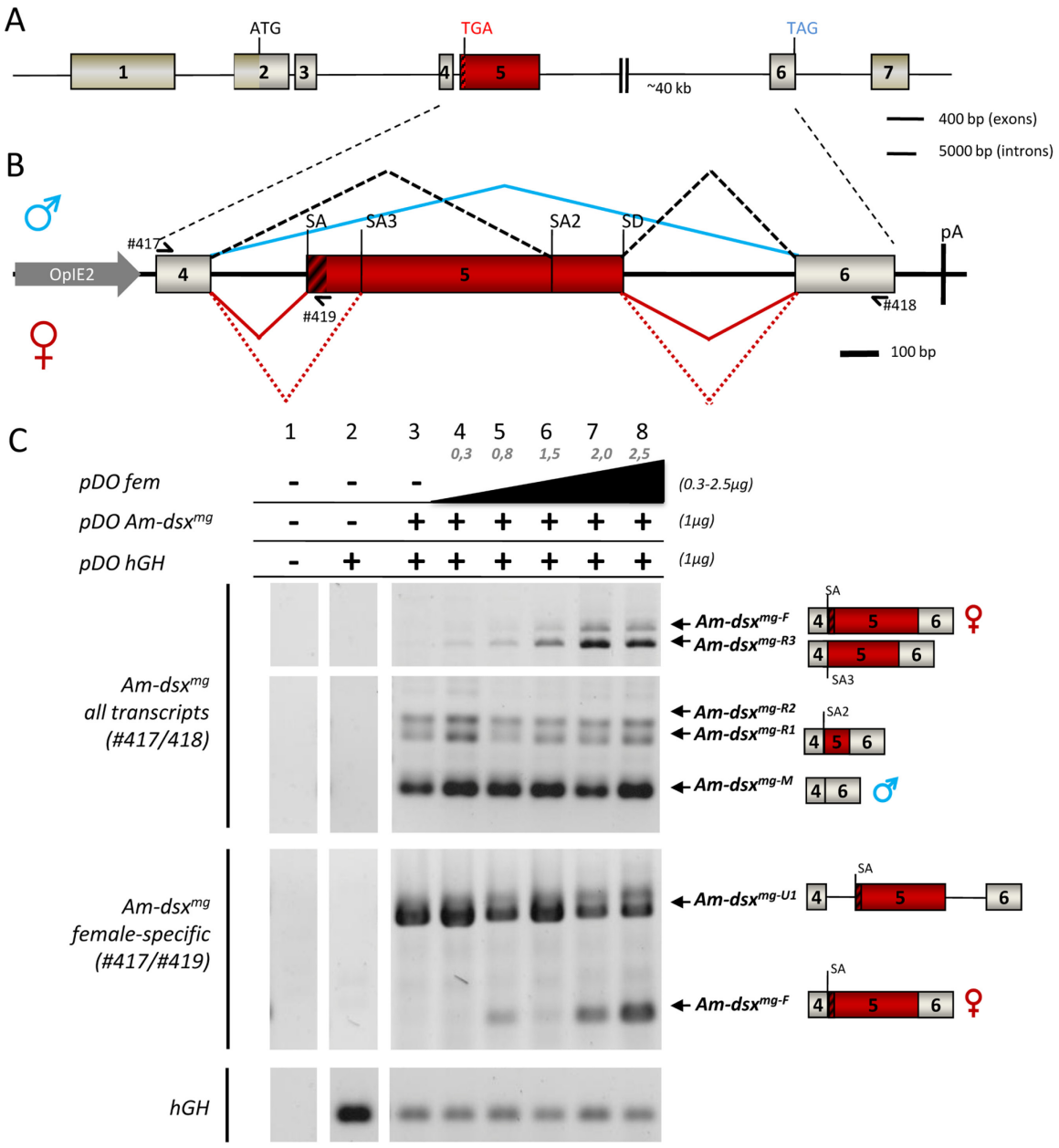

Fig. 1: RT-PCR analysis of splicing variants of the *Am-dsx* mini gene (pDO *Am-dsx*^{mg}) derived transcripts in the presence and absence of Fem protein in *Sf*21 cells.
(A) Diagram of the genomic organisation of the *Am-dsx* gene. The female-specific exon 5 is shown as red box; the common exons are shown as grey boxes. The common translational start codon is shown in black and the translational stop codons are shown in red (female) and blue (male). Please not that exons and introns are shown in different scales. Intron 5 has been shortened for better illustration (marked by two vertical bars). **(B) Intron and exon structure of the *Am-dsx*^{mg}** that was cloned into pDO vector under control of the viral OpIE2 promoter. pA indicates the polyadenylation signal that we derived from OpIE2 polyadenylation signal. The ~40 kb intron 5 of *Am-dsx* was shortened to 470 bp in the *minigene*. SA assigns the splice acceptor of the female exon and SD assigns the splice donor of the female exon. SA2 and SA3 indicate the alternatively used splice acceptors in the female exon that we observed when expressed in *Sf*21 cells. **(C) RT-PCR analysis of splicing variants of pDO *Am-dsx*^{mg} derived transcripts in the presence and absence of Fem protein in *Sf*21 cells.** 1x10^6 *Sf*21 cells were incubated with dd$H_2$0 (lane 1) or transiently transfected with 1µg of pDO *hGH* (lane 2-8), 1µg of

pDO *Am-dsx*^{mg} (lane 3-8) and increasing amounts of pDO *fem* (0.3, 0.8, 1.5, 2.0, 2.5 µg) (lane 4-8). Male and female specific fragments were amplified by using oligonucleotide primers as indicated in (B). Conditions for PCR using oligonucleotides #417/418 were optimized for amplification of small male *Am-dsx* fragments and large female *Am-dsx* fragments in two separate PCRs (shown in separate gel pictures). Semiquantitative RT-PCR was standardized in respect to the splice product of the cotransfected *hGH* control. Fragments were resolved in 2% agarose gels and stained with ethidium bromide. Fragments were cloned and sequenced to determine the exonic structure. The exonic structure of the amplified fragments is schematically shown to the right of the gel picture.

Fem proteins are sufficient to direct female-specific splicing in *Sf*21 cells

In a previous study we showed that knockdown of *fem* mRNAs in honeybees produced a switch from female to male splicing of *Am-dsx* pre-mRNAs (Gempe et al., 2009). To test whether Fem proteins are sufficient to direct female splicing in a dose dependent manner we transfected *Sf*21 cells with plasmids pDO *Am-dsx*^{mg}, pDO *fem* and pDO *hGH* that were expressing the *Am-dsx minigene*, the Fem proteins and the *hGH* splice control. We analyzed male- and female-specific splice products by RT-PCR which we semi-quantified across samples in respect of the splice product of the *hGH* control.

We detected female-specific inclusion of exon 5 with increasing amounts of Fem protein (*Am-dsx*$^{mg\text{-}F}$; Fig 1C, lane 4-8 upper panel). Beside the inclusion of entire female exon 5 in the presence of Fem protein we detected in the *Sf*21 cells *Am-dsx*$^{mg\text{-}R3}$ transcript (sequence, Fig. S2 E) that uses a splice acceptor site (SA3) that is located 140 nucleotides downstream of the female SA site (SA3, figure 1B and table S2).

The inclusion of female exon 5 leads to a substantial increase in size and thus in an underrepresentation of this fragment in PCRs when compared to the smaller male fragment (855 nucleotides difference). We confirmed the female splicing with oligonucleotide primers #417 and #419 that amplifies female transcripts that include the female exon 5, but no male or *Am-dsx*$^{mg\text{-}R3}$ transcripts. We detected again the female splice product (*Am-dsx*$^{mg\text{-}F}$) that increased with larger amounts of pDO *fem* plasmids that are expressing Fem proteins (Fig 1C, lane 4-8 lower panel). These results suggest that Fem proteins are sufficient to direct in a dose dependent manner female-specific splicing of *Am-dsx* transcripts.

We used in the subsequent experiments the oligonucleotide primers #417 and #419 to specifically amplify the female splice variants and the oligonucleotide primers #417 and #418 to analyze male splice products (Fig. 1B).

Expression of *Am*-Tra2 proteins enhances female-specific splicing in the presence of Fem proteins.

In a previous study we have identified the *tra2* ortholog gene, *Am-tra2*, in the honeybee genome (Dearden et al., 2006). Tra2 in *D. melanogaster* is not sex-specifically regulated in the soma (Mattox et al., 1990) and associates together with Tra protein as a part of a multiprotein complex to direct the female splicing of *dsx* pre-mRNAs (Tian and Maniatis, 1993; Lynch and Maniatis, 1996). We first studied whether expression of solely the *Am*-Tra2 proteins are sufficient to direct female splicing of the *Am-dsx* minigene. Previous studies on splice regulation of Tra2 proteins in *D. melanogaster* Kc cells showed that Tra2 proteins can induce low amount of *dsx* female-specific splicing, presumably due to the presence of endogenous expressed Tra proteins in this female cell line (Hoshijima et al., 1991).

To examine the role of *Am*-Tra2 in *Am-dsx* splicing we transfected *Sf*21 cells with increasing amounts of pDO *Am-tra2* plasmids expressing the *Am*-Tra2 proteins. We observed that *Am*-Tra2 proteins were not sufficient to induce female-specific splicing in *Sf*21 cells under different levels of expression (Fig. 2, lane 4-7) indicating that *Am*-Tra2 proteins cannot direct female splicing even when they are overexpressed. Next we examined whether *Am*-Tra2 protein expression can enhance female splicing when Fem proteins are expressed. We transiently cotransfected plasmids expressing Fem protein and increasing amounts of plasmids expressing *Am*-Tra2 proteins. Higher expression of *Am*-Tra2 proteins increased inclusion of exon 5 when Fem proteins are coexpressed (Fig. 2, lane 8-12), suggesting that the *Am*-Tra2 proteins enhance the female-specific splicing, but solely when Fem proteins are present.

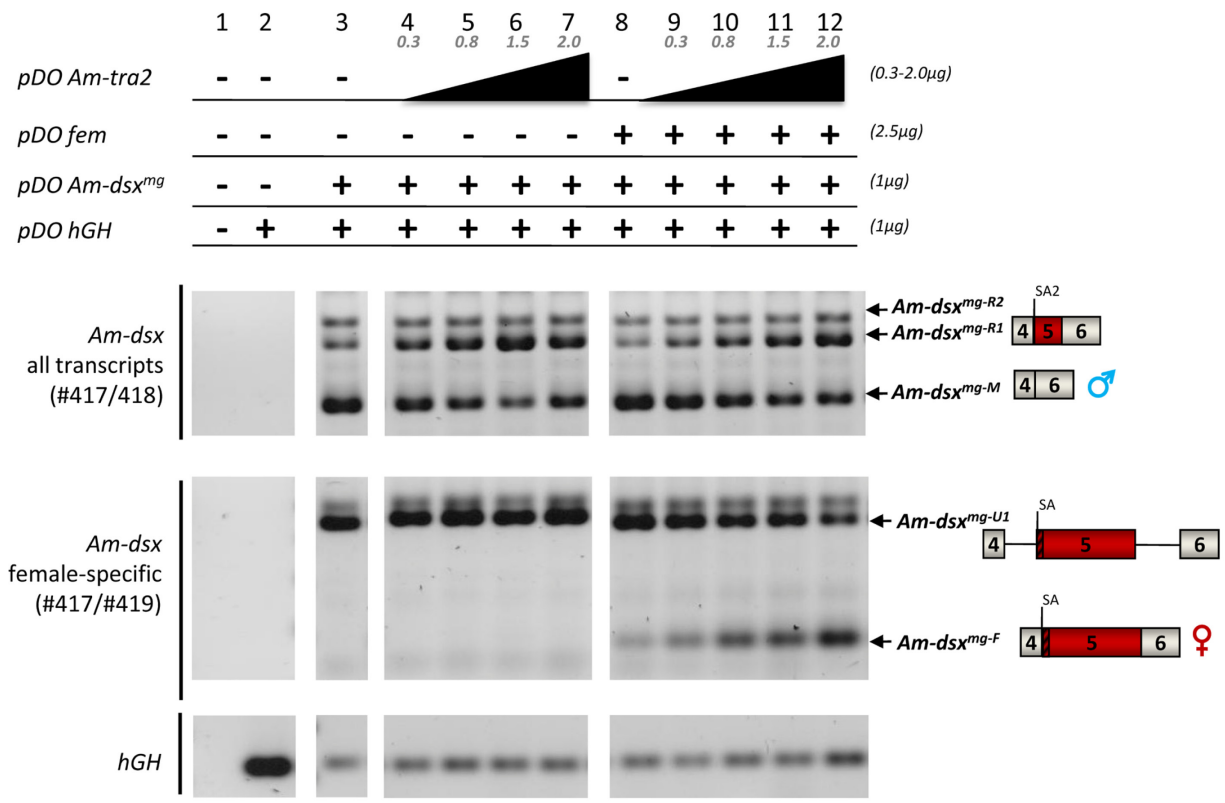

Fig. 2: RT-PCR analysis of splicing variants of pDO *Am-dsx*^{mg} derived transcripts in the presence and absence of Fem and *Am*-Tra2 protein in *Sf*21 cells. 1x10^6 *Sf*21 cells were incubated with dd$H_2$0 (lane 1) or transiently transfected with 1µg of pDO *hGH* (lane 2-12), 1µg of pDO *Am-dsx*^{mg} (lane 3-12), 2.5µg of pDO *fem* (lane 8-12) and increasing amounts of pDO *Am-tra2* (0.3, 0.8, 1.5, 2.0 µg) (lane 4-7 and 9-12). Male and female specific fragments were amplified by using oligonucleotide primers as indicated in figure 1B. Semiquantitative RT-PCRs were standardized in respect to the splice product of cotransfected hGH control. Conditions for the PCR using oligonucleotides #417/418 were optimized to amplify the small male *Am-dsx* fragments. Fragments were resolved in 2% agarose gels and stained with ethidium bromide. Fragments were cloned and sequenced to determine the exonic structure. The exonic structure of the amplified fragments is schematically shown to the right of the gel picture. SA assigns the splice acceptor of the female exon, SA2 and SA3 those of splice acceptors that are alternatively used in *Sf*21 cells.

Bioinformatic analyses of *Am-dsx* splice regulation

We closer inspect by bioinformatic tools how Fem and *Am*-Tra2 proteins can possibly direct the female-specific inclusion of exon 5. We first searched for canonical 13nt ESE (exonic splicing enhancer) sequence (TC(T/A)(T/A)C(A/G)ATCAACA) that recruit Tra2 and Tra proteins in *D. melanogaster*. In exon 4 to 6 and the connecting introns we did not identify sequences that match this motif.

Next, we evaluated the strength of splice acceptor and donor sites of the female exon 5 and compared them to those of the female *dsx* exon of *D. melanogaster* and to those of the constitutive spliced *Am-dsx* exons (Table S2 and S3). If the intron sequence upstream of the exon matches the consensus sequence well (the number of pyrimidines in the 12 nucleotide long tract 4 nucleotides upstream of the respective exon) the U2 snRNP Auxiliary Factor (U2AF) binds well and thus should activate the splice activator site without assistance from specific splice activators (Hoshijima et al., 1991; Tian and Maniatis, 1994; Graveley et al., 2001). We found that the female exon 5 of *Am-dsx* has a predicted strong acceptor site as we detected a long polypyrimidine tract (11 out of 12 nucleotides are pyrimidines) in intron 4 (Table S2). This acceptor site is predicted to be stronger compared to the other acceptor sites of constitutive spliced *Am-dsx* exons as they all have weaker matches to the consensus sequence of the U2AF binding site (number of pyrimidines ranging from 8 to 10 in the designated tract). The female exon in *D. melanogaster*, in contrast, has a weak acceptor site (only 6 out of 12 nucleotides are pyrimidines) (Hoshijima et al., 1991) therefore this 3' splice acceptor site is not used in males. This site is, however, activated in *D. melanogaster* females by Tra and Tra2 proteins that bind to the ESEs sequences and form a complex with other proteins that assist U2AF proteins in binding to the weak polypyrimidine tract (Graveley et al., 2001; Zuo and Maniatis, 1996; Lynch and Maniatis, 1996). We next studied the strength of splice donors. If the sequence at the splice donor site matches the consensus sequence well so that U1snRNA binds (this results in the formation of an RNA duplex between the free 5'end of the U1snRNA) (Freund et al., 2003; Kammler et al., 2001) then U1 proteins are recruited to the donor site without assistance of splice activators. The splice donor of the female exon 5 of the *Am-dsx* gene matches except of one position the consensus sequence of the predicted U1snRNA binding site (Table S3). All other constitutive spliced exons of *Am-dsx* gene have splice donors that have poorer matches to the consensus sequence indicating that the splice donor of female exon 5 is predicted to be used without assistance of a splice activator. In *D. melanogaster* the female exon lacks a donor splice site and a polyadenylation signal in this exon causes cleavage of the mRNAs at that point.

These comparisons suggest that female exon 5 of *Am-dsx* gene has strong acceptor and donor sites indicating that they are used without additional splice activators. Together with our results that the splicing of male *Am-dsx* transcripts is promoted in the absence of the Fem proteins both in the honeybee (Gempe et al., 2009) and in our splice assay here, we

propose that other transacting factors are required to repress the use of the strong acceptor and donor splice sites in female exon 5.

We further studied regulation of exon 5 splice donor site that is used in the presence of Fem and *Am*-Tra2 proteins (Fig. 3 and 4). We focused our study on the regulation of the splice donor site of the female exon, because our results showed that the use of this splice site substantially increased with various different splice acceptor sites by increasing amounts of Fem/*Am*-Tra2 proteins (Fig 1 and 2).

The exon sequence upstream of the female-specific splice donor site harbors several Fem/*Am*-Tra2 protein dependent ESEs

To further understand the control of exon 5 donor splice site activation by Fem and *Am*-Tra2 proteins, we cloned four overlapping 30 to 38 putative regulatory regions (pRR1-4) (Fig. 3A) that comprise 113 nt upstream of splice donor site of exon 5 into a splice reporter gene including a part of the *env* gene (Fig. 3B), a minigene that we modified from (Singh et al. 2010) in order to use it in *Sf*21 cells. If our sequence comprises Fem/*Am*-Tra2 protein dependent ESE sequences, the downstream splice donor ($SD7^{reporter}$, Fig. 3B) will be used. We focused our study on the first 113 nt upstream region, because it has been repeatedly reported that use of ESE decline with distance to splice sites (Kuo and Norton, 1999; Graveley et al., 1998; Shen and Green, 2006; Sciabica and Hertel, 2006).

We co-transfected increasing concentrations of our pOPI *env*/pRR plasmid containing the pRR together with control plasmid pDO *hGH* into *Sf*21 cells and studied the effects on activation of the reporter splice donor site ($SD1^{reporter}$) in respect to splicing of *hGH* control by semi-quantitative RT-PCR. We observed an increase of reporter splicing for pRR 1 to 4 in response to the expression of Fem and *Am*-Tra2 proteins (Figure 3C). For pRR1 – the sequence next upstream of the female splice donor – we found a strong dose dependent splice response with increasing amounts of Fem and *Am*-Tra2 proteins (Figure 3C; lane 3-6)

We did not detect such dose response for the pRR2 and 3 sequences (Figure 3C; lane 7-14). For these sequences we observed reporter gene splicing even in the presence of low amounts of Fem and *Am*-Tra2 proteins. We next studied whether splicing substrate was entirely converted even when low amounts of Fem/*Am*-Tra2 proteins are expressed using

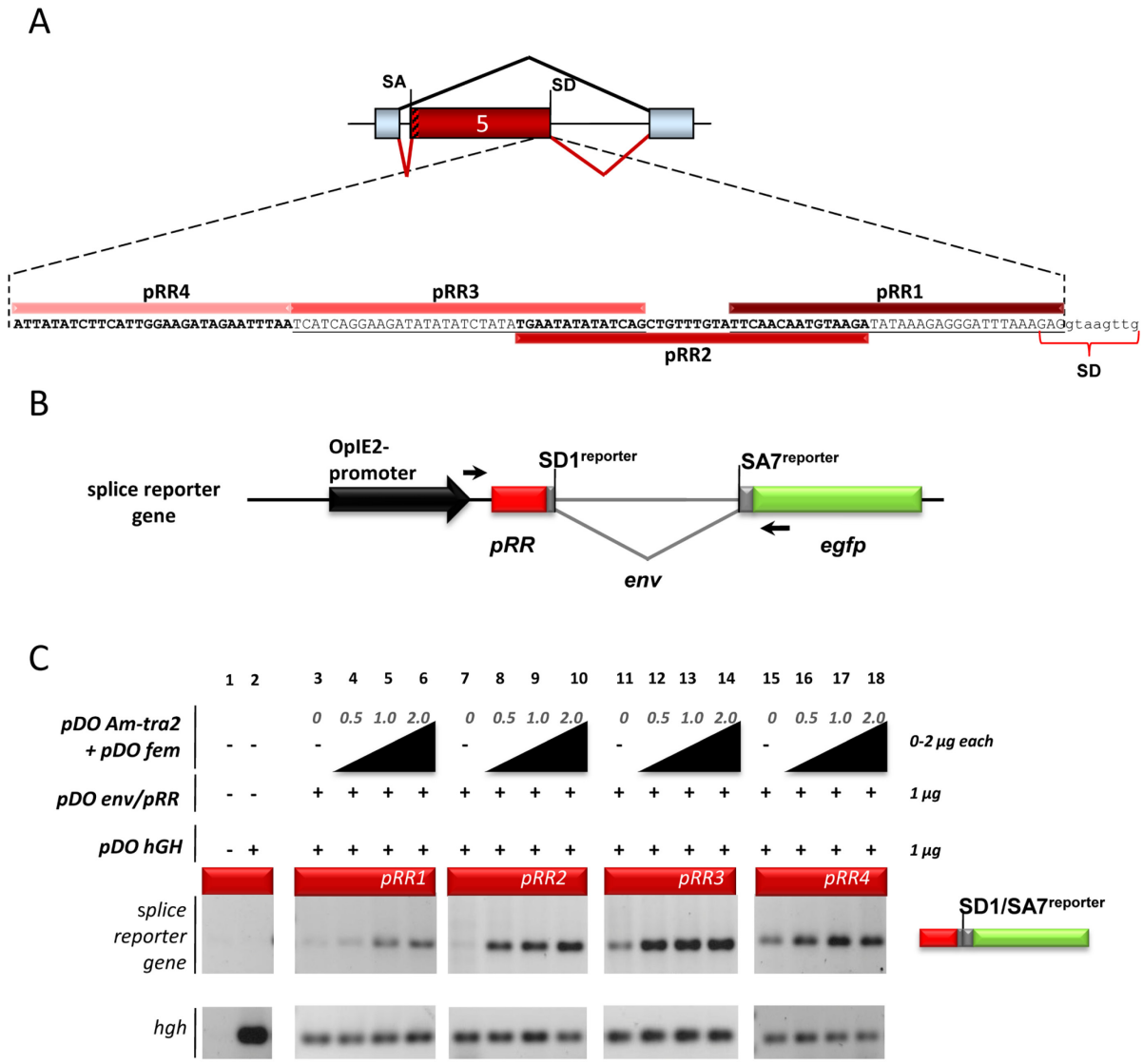

Fig. 3: Identification of Fem/*Am*-Tra2 protein dependent ESEs within the female *Am-dsx* exon. (A) Region of interest in exon 5 of *Am-dsx* and the sequences of fragments pRR1, pRR2, pRR3 and pRR 4. The female exon 5 is shown as a red box, the male and female common exons 4 and 6 are shown as grey boxes. Last 113 nucleotides of exon 5 (upper case letter) and first 9 nucleotides of intron 5 (lower case letter) are shown. SA indicates the splice acceptor, SD the splice donor of the female exon. **(B) Structure of the splice reporter gene construct (pOPI *env*/pRR) in order to identify splice enhancers.** The splice reporter gene is transcribed under control of the viral OpIE2 promoter. If pRR harbors a exonic splice enhancer (ESE) the reporter splice donor SD1^{reporter} is activated and directs splicing to splice acceptor SA7^{reporter} that we study by RT-PCR amplifications (arrows shows the designated binding sites of the oligonucleotide primers). *env* denotes a part of the gene *envelope*; *egfp* denotes the *enhanced green fluorescence protein* gene sequence. **(C) Semiquantitative analysis of Fem/*Am*-Tra2 dependent splicing in respect to the presence of the different pRR sequences in the splice reporter gene**. 1x10^6 *Sf*21 cells were incubated with dd$H_2$0 (lane 1) or transiently transfected with 1µg of pDO *hGH* (lane 2-18), 1µg of pOPI *env*/pRR1 (lane 3-6), 1µg of pOPI *env*/pRR2 (lane 7-10), 1µg of pOPI *env*/pRR3 (lane 11-14), 1µg of pOPI env/pRR4 (lane 15-18) and increasing amounts of pDO *fem* and pDO *Am-tra2* (0.5, 1.0, 2.0 µg each) (lane 4-6; 8-10; 12-14; 16-18). Transcripts were amplified by RT-PCRs using oligonucleotide primers as indicated in (B). Semiquantitative RT-PCRs were standardized in respect to the splice product of cotransfected *hGH* control. Fragments were resolved in 2% agarose gels and stained with ethidium bromide. Identity of splice products was verified by sequencing.

the example of pRR2. We tripled the amount of the transfected reporter gene plasmid pOPI *env*_pRR2 plasmid (supplementary data, Fig. S1) and observed again a strong dose dependent response implicating that the pRR2 and pRR3 sequences harbors very effective splice enhancers.

For pRR4, the sequence most distant to female exon 5 splice donor site, we sized only small differences in splicing in the absence and presence of Fem and *Am*-Tra2 proteins (Fig 3, lane 15-18) suggesting that this sequence is a not very effective in promoting the use of the splice donor site. Taken together these results indicate the first 100 nt upstream of the female splice donor site of *Am-dsx* harbors Fem/*Am*-Tra2 protein dependent ESEs that promote the use of a splice donor.

Defined mutations within the pRR1 sequence reduce splicing efficiency in presence of Fem and *Am*-Tra2.

We further characterized the sequence of Fem/*Am*-Tra2 ESEs in the pRR1 sequence that is located next upstream of the splice donor site of the female exon 5, because previous studies showed ESEs are most effective when located in close proximity of splice sites (Kuo and Norton, 1999; Graveley et al., 1998; Shen and Green, 2006).

To identify putative Fem/*Am*-Tra2 protein dependent ESEs within fragment pRR1 we predicted candidate sequences by the ESR (exonic splicing regulatory elements) search tool (http://esrsearch.tau.ac.il/). Two putative exonic splice enhancers (pESEs) were predicted (AAAGAG and AAGAGG; figure 4A) (Fairbrother et al., 2002) in which sequence AAAGAG appeared two times. To determine whether Fem/*Am*-Tra2 proteins target these sequences to promote splicing, we introduced a series of point mutations into the AAAGAG motif and studied whether the splicing efficiency is affected in our splice enhancer assay (figure 4A).

Point mutations that changed sequence AAAGAG to AAAGGG in either the first (pRR1_mut1) or the second (pRR1_mut2) hexamer repeat had no sizeable effect on the efficiency of splicing (Fig. 4B, lane 3-6 and 7-10). We detected a dosage dependent increase of splicing products of the reporter gene in the presence of Fem/*Am*-Tra2 proteins (figure 4B; lane 3-6 and lane 7-10). Next we examined the combined effect of point mutations in both hexameric repeats (pRR1_mut3). These double mutations abolished the Fem/*Am*-Tra2

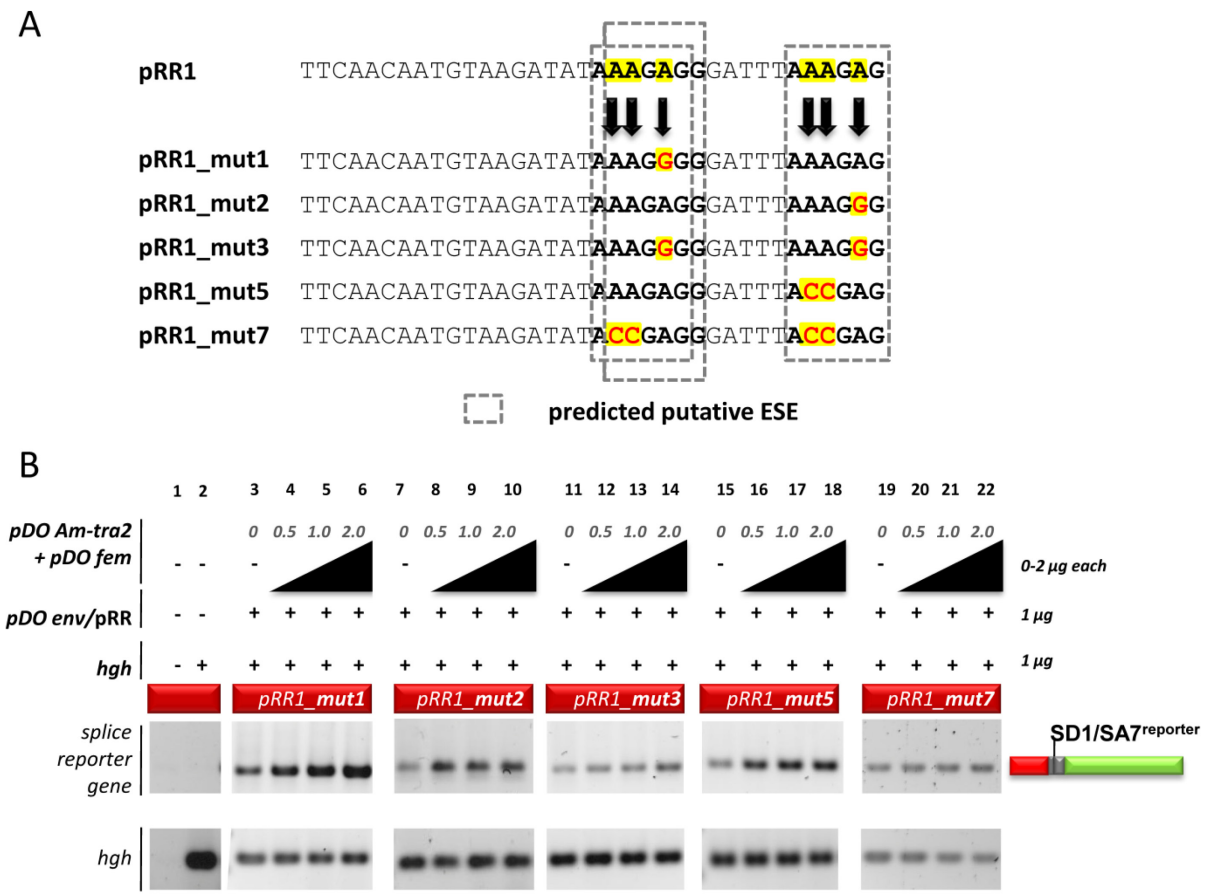

Fig. 4: Identification of Fem/*Am*-Tra2 protein dependent ESE sequences within the pRR1 fragment. (A) Sequences of mutated pRR1 fragments. Positions of changed nucleotides are marked in yellow. Dotted boxes indicate the predicted ESE sequences. **(B) The effect of point mutations in the pRR1 fragment on reporter gene splicing.** 1x10^6 *Sf*21 cells were incubated with dd$H_2$0 (lane 1) or transiently transfected with 1µg of pDO *hGH* (lane 2-18), 1µg of pOPI *env*/pRR1_mut1 (lane 3-6), 1µg of pOPI *env*/pRR1_mut2 (lane 7-10), 1µg of pOPI *env*/pRR1_mut3 (lane 11-14), 1µg of pOPI *env*/pRR1_mut5 (lane 15-18), 1µg of pOPI *env*/pRR1_mut7 (lane 19-22) and increasing amounts of pDO *fem* (0.5, 1.0, 2.0 µg) and of pDO *Am-tra2* (0.5, 1.0, 2.0 µg) (lane 4-6; 8-10; 12-14; 16-18; 20-22). Splice reporter gene transcripts were amplified in RT-PCRs by using oligonucleotide primers as indicated in figure 3B. Semiquantitative RT-PCRs were standardized in respect to the splice product of cotransfected *hGH* control. Fragments were resolved in 2% agarose gels and stained with ethidium bromide. Structure of splice products was verified by sequencing.

protein dosage dependent splicing (figure 4B; lane 11-14). To further characterize the sequences of the ESEs we analyzed two point mutations further upstream of the hexameric motif changing the sequence from AAAGAG to ACCGAG (pRR1_mut5 and pRR1_mut7; figure 4A). We observed essentially the same pattern. The mutation in one repeat had no sizeable effect on splicing efficiency while the mutations in both AAAGAG repeats abolished the Fem and *Am*-Tra2 protein dosage dependent splicing (figure 4B; lane 15-18 and lane 19-22). Our results showed that the AAAGAG motif in pRR1 fragment is essential to activate the splice donor in the presence of Fem/*Am*-Tra2 proteins.

Discussion

We have previously reported that the *fem* gene is required for female specific splicing of *Am-dsx* transcripts in the honeybee that implement female differentiation (Hasselmann et al., 2008; Gempe et al., 2009). However, a direct involvement of the Fem protein in the splicing process has thus far not been shown. The *fem* gene is a critical component of the complementary sex determining pathway as it maintains the femaleness by a positive regulatory loop and implements female differentiation by the primary signal *csd*. In the present study we provided evidences that female Fem protein is directly regulating the splicing of *Am-dsx* transcripts. We further showed that *Am*-Tra2 protein is a cofactor that enhances female splicing of *Am-dsx* pre-mRNAs in the presence of Fem proteins, but that the mechanism of splice regulation and the involved RNA binding sites differ from those in *D. melanogaster*. We established an *in vivo* splice assay in lepidopteran cells and were able to mimic sex-specific *Am-dsx* splicing using an *Am-dsx* minigene construct. In the absence of Fem protein male *Am-dsx* transcripts were generated, suggesting that male splicing is the default regulatory state in the absence of the complementary sex determining signal *csd*. In the presence of female Fem protein we were able to promote female specific splicing of *Am-dsx* transcripts. Taken together, our data suggests that Fem protein is sufficient to direct female *Am-dsx* splicing. *Am*-Tra2 protein further enhances female specific splicing in the presence of Fem protein, but alone it is not sufficient to direct female *Am-dsx* splicing in *Sf*21 cells. This finding supports our previous study in honeybees that shows that *Am-tra2* gene is essential for female *Am-dsx* splicing in the presence of female Fem proteins (Nissen, Müller, & Beye, under review).

This raises the question of how the SR-type protein Fem is able to promote, in the absence of an RBD and direct RNA binding, female *Am-dsx* splicing. We propose four explanations for this observation: (i) endogenously expressed *Sf*-Tra2 protein in *Sf*21 cells is able to replace the role of *Am*-Tra2 protein as co-factor of Fem protein in the splice process; (ii) other RNA-binding proteins are sufficient to mediate the interaction between Fem and the *Am-dsx* pre-mRNA; (iii) Fem protein has RNA binding abilities; (iv) Fem proteins when overexpressed, as it is performed here, can possibly direct the splicing. Our previous results, however, showed *Am-tra2* gene is essential to direct female *Am-dsx* pre-mRNA splicing (Nissen, Müller, & Beye, under review), suggesting that either overexpression or endogenous Tra2 of *Sf*21 cells is responsible.

We further studied how the Fem and *Am*-Tra2 proteins direct the use of the splice donor site. The splice acceptor and splice donor sites of the female *Am-dsx* exon are strong and are predicted to be recognized by the splicing machinery without assistance of particular splicing activators (Hoshijima et al., 1991; Tian and Maniatis, 1994; Graveley et al., 2001). The strong splice donor and acceptor sites that are not utilized neither in the honeybee (Cho, Huang, & Zhang, 2007; Nissen et al., under review) nor in *Sf*21 cells, suggesting that they are silenced by constitutive splicing repressor proteins that may bind to exonic or intronic splicing silencer sequences in the female exon or intron (Baralle and Baralle, 2005; Valcárcel et al., 1993; Suzuki et al., 2008; Qi et al., 2007). The female form of splicing in the response to Fem and *Am*-Tra2 proteins suggests that these proteins are splicing activators. We identified Fem/*Am*-Tra2 protein dependent exonic splicing enhancer (ESE) sequences in the first 100 nt upstream of the female splice donor site by using a reporter construct that can assay the impact of different ESE sequences on the activation of a general splice donor. Point mutations within the direct upstream sequence of the female splice donor site (sequence pRR1) showed that the two copies of the sequence AAAGAG are required for the activation of the splice donor by Fem and *Am*-Tra2 proteins. Mutation in only one of the AAAGAG sequence motifs did not reduce the efficiency of splicing. We suggest that both of these sequences are utilized to recruit Fem and *Am*-Tra2 proteins to the ribonucleotide sequence and to direct the use of the splice donor site of the female exon. Our analysis, however, cannot exclude that other sequence motifs further upstream are also involved to promote the use of the female specific splice donor or the splice acceptor site.

The predicted strength of the female splice donor and splice acceptor site further indicates that the sites are constructively silenced by splicing repressor proteins (Valcárcel et al., 1993; Suzuki et al., 2010, 2008) and are only activated in the presence of Fem protein in females. We propose two models how these splice repressor proteins and Fem proteins together control *Am-dsx* splicing. In the first model the splicing repressor proteins, which inhibit female splicing, bind in males to the same sequences as Fem and *Am*-Tra2 proteins. In females they are replaced by the Fem/*Am*-Tra2 protein complex possibly because they have a higher binding affinity to the RNA for instance to the AAAGAG sequence. In the second model Fem/*Am*-Tra2 protein complexes and the proteins repressing female splicing bind to other sequences of the *Am-dsx* pre-mRNA.

The role of female Fem protein as a direct component of sex-specific splice regulation suggests important implications for the control of the complementary sex determination pathway. Heteroallelic Csd protein is the primary signal that determines the female state and that directs the female processing of *fem* transcript that are translated into Fem proteins (Gempe et al., 2009; Hasselmann et al., 2008). Beside the role of Fem protein as a splice activator of female *Am-dsx* splicing, it is also responsible to maintain the female determined state throughout development by directing female splicing of its own pre-mRNAs into the productive female mode (Gempe et al., 2009). The Fem/*Am*-Tra2 proteins may also directly control the splicing of female *fem* transcripts. Indeed knockdown studies of the *Am-tra2* showed that *Am*-Tra2 protein is also required to promote *fem* transcript splicing into the female mode (Nissen et al., under review). This implies that *Am*-Tra2 has an important role on two levels of the cascade of sex determination, on the level of *Am-dsx* and on the level of *fem* pre-mRNA splicing.

Although the sex-specific splicing of *dsx* transcripts is highly conserved across different insects orders (Gempe and Beye, 2011) our results suggest that molecular control of this splice process can be very different. We compared *Am-dsx* regulation with those in *D. melanogaster* and the lepidopteran insect *B. mori* in which molecular details of the splicing process have been studied. In *D. melanogaster* the weak splice acceptor site is activated by Tra and Tra2 proteins (Inoue et al., 1992; Sciabica and Hertel, 2006; Tian and Maniatis, 1993). The binding of these proteins to six repeats of a 13 nt (ESE) sequence that is located ~300 nucleotides downstream of the weak acceptor splice site recruit the general splicing machinery (Inoue et al., 1992; Sciabica and Hertel, 2006) to promote the female splicing. This canonical 13 nt long binding motif is found at similar positions of the female exon in other dipteran species (Hediger et al., 2004; Saccone et al., 2008; Lagos et al., 2005) suggesting that an equivalent mechanisms by use of the splice acceptor is operating in these dipteran insects. In *D. melanogaster*, *M. domestica*, *C. capitata* and *Bactrocera oleae* the female exon is not spliced to further downstream exons and a polyadenylation signal terminates the female transcript in this exon. In contrast, Fem/*Am*-Tra2 proteins promote in the honeybee the use of the splice donor site by involving other ESEs in order to direct the inclusion of the female exon. In honeybees, the splice donor and acceptor sites of the female *Am-dsx* exon are very strong (Mount, 1982; Asang et al., 2008; Freund et al., 2005; Caputi et al., 2004; Hoshijima et al., 1991; Kammler et al., 2001; Norton, 1994) suggesting that some other

constitutive expressed splice repressors are involved in the honeybee to direct male splicing and exclusion of the female exon in the absence of the Fem proteins.

In *B. mori* the role of Tra or Tra2 proteins in *Bm-dsx* splice regulation is thus far not known. In this species the *Bm-dsx* transcripts are presumably spliced into the female form in the absence of male-determining splice regulator *Bm*-PSI and *Bm*-IMP proteins. The *Bm*-PSI proteins and the male-specific *Bm*-IMP proteins bind 15 nucleotide downstream of the female splice acceptor site to a 20 nucleotide long exonic splicing silencer sequence (CE1) (Suzuki et al., 2010, 2008) which is evidently silencing the female splicing. We suggest that honeybee orthologs of *Bm-imp* and *Bm-psi* genes are putative candidate genes that can control the male splicing of *Am-dsx* in the honeybee by silencing the strong female splice acceptor and donor sites in the female exon. Taken together, our comparison of the molecular control of *dsx* splicing in different orders of holometabolous insects implicate that generation of a sex-determining female and male *dsx* mRNAs is conserved (Gempe and Beye, 2011) while the mechanisms that produce the male and female transcripts have diverged.

In the near future, we will study the direct interaction of *Am*-Tra2 and Fem proteins with the *Am-dsx* pre-mRNA sequence by RNA binding shift experiments. By introducing mutations within the AAAGAG sequence of the *Am-dsx* minigene we will learn more about the impact of this sequence on female *Am-dsx* splicing and splice donor activation. In addition mutated splice acceptor and splice donor sites may help us to understand whether Fem/*Am*-Tra2 proteins also activate the splice acceptor site of the female exon in the honeybee.

Supplementary Data

Table S1: Sequences of oligonucleotides that were used

to clone pDO *fem*, pDO *Am-tra2*, pDO *hGH*	
#367 (fem_rev)	GATCACGCGTTTACATAGGTCCAAATCTAGGATTTGG
#369 (fem_fw)	GATCGGATCCATGAAACGGAATACAACAAATCATTC
#359 (Am-tra2_fw)	GATCGGATCCATGAGTGACATTGAGCGAAGTAGTAG
#421 (Am-tra2_rev)	TGACACGCGTTTAATATCGACGTGGTGAATAAGAGC
#496 (hGH_fw)	GATCAAGCTTATGGCTACAGGTAAGCGCCCCT
#497 (hGH_rev)	GATCTCTAGACTAGAAGCCACAGCTGCCCTCCACAG
to clone *Apis mellifera dsx* minigene (*Am-dsx^{mg}*)	
#245	GATCGGGCCCGTGAAAATGTGGAAATTCTATTGGAGCACA
#246	GATCACTAGTTATTAGAATTATAATATTATTTAATATTTTTTTAAATATA
#247	GATCACTAGTACGTGATATTATTGACACTATGATCTATCTTATAAAA
#248	GATCGCGGCCGCTGAGGCTACGTATGTTTAGGAGGACC
#317	GATCCCTAGGGTGAAAATGTGGAAATTCTATTGGAG
#320	GATCCTCGAGTGAGGCTACGTATGTTTAGGAGGACC
to clone OpIE2-promoter	
#611	GATC ATCGAT TCATGATGATAAACAATGTATGGTGCTAAT
#613	GATCGAATTCAAATTCGAACAGATGCTGTTCAACTGTG
to generate pRR linker	
#622 (pRR1_fw)	AATTC TTCAACAATGTAAGATATAAAGAGGGATTTAAAGAG GAGCT
#623 (pRR1_rv)	C CTCTTTAAATCCCTCTTTATATCTTACATTGTTGAA G
#618 (pRR2_fw)	AATTC TGAATATATATCAGCTGTTTGTATTCAACAATGTAAGA GAGCT
#619 (PRR2_rv)	C TCTTACATTGTTGAATACAAACAGCTGATATATATTCA G
#620 (pRR3_fw)	AATTC TCATCAGGAAGATATATATCTATATGAATATATATCAG GAGCT
#621 (pRR3_rv)	C CTGATATATATTCATATAGATATATATCTTCCTGATGA G
#729 (pRR4_fw)	AATTC ATTATATCTTCATTGGAAGATAGAATTTAA GAGCT
#730 (pRR4_rv)	C TTAAATTCTATCTTCCAATGAAGATATAAT G
#654 (pRR1_mut1_fw)	AATTC TTCAACAATGTAAGATATAAAGGGGGATTTAAAGAG GAGCT
#655 (pRR1_mut1_rv)	C CTCTTTAAATCCCCCTTTATATCTTACATTGTTGAA G
#648 (pRR1_mut2_fw)	AATTC TTCAACAATGTAAGATATAAAGAGGGATTTAAAGGG GAGCT
#649 (pRR1_mut2_rv)	C CCCTTTAAATCCCTCTTTATATCTTACATTGTTGAA G
#646 (pRR1_mut3_fw)	AATTC TTCAACAATGTAAGATATAAAGGGGGATTTAAAGGG GAGCT
#647 (pRR1_mut3_rv)	C CCCTTTAAATCCCCCTTTATATCTTACATTGTTGAA G
#724 (pRR1_mut5_fw)	AATTC TTCAACAATGTAAGATATAAAGAGGGATTTACCGAG GAGCT
#725 (pRR1_mut5_rv)	C CTCGGTAAATCCCCCTTTATATCTTACATTGTTGAA G
#735 (pRR1_mut7_fw)	AATTC TTCAACAATGTAAGATATACCGAGGGATTTACCGAG GAGCT
#736 (pRR1_mut7_rv)	C CTCGGTAAATCCCTCGGTATATCTTACATTGTTGAA G
to analyze *Am-dsx* splicing	
#417	CTATTGGAGCACAGTAGCAAACTTG
#418	GGCTACGTATGTTTAGGAGGACC
#419	GAAACAATTTTGTTCAAAATAGAATTCC
to analyze *hGH* splicing	
#465	GATCACGCGTCTAGAAGCCACAGCTGCCCTCC
#597	AGGTTTGGGGCTTCTGAATAG
to analyze splice product of the splice reporter gene (pOpI *env*/pRR)	
#709	CAGCATCTGTTCGAATTTGAATTC
#641	AGCACTGCACGCCGTAGGTC

Table S2: Splice acceptor sites in the *dsx* gene of the honeybee and in intron 3 of *D. melanogaster.*

	Intron/Exon	Acceptor sequence	Number of Pyrimidines
Am-dsx	4/5	UUUCUUUAUUCUCU**AG**	11
	1/2	UCUUUUUAUGUUGC**AG**	10
	2/3	GAUUCUUUCCUAAC**AG**	9
	3/4	UUUGAUUGUGUUAC**AG**	8
	5/6	UUGUUUGUUUCAAC**AG**	9
	6/7	UUUUUUCUGAUUCC**AG**	10
Dm-dsx	3/4	AUCUGAUCUAAACC**AG**	6

Acceptor sequences used in females are intron 4 of *Am-dsx* and intron 3 of *Dm-dsx*. The largest number of pyrimidines that can be found in the polypyrimidine tract is 12. Letters shown in bold, mark the end of introns.

Table S3: Similarity of putative U1snRNA binding sites of splice donors to the consensus sequence in the *Am-dsx* gene.

	Exon/Intron	U1 snRNA binding site	H-Bond Score
Am-dsx	5/5	GAG**GU**AAGUGU	20.2
	1/1	UGG**GU**AAGUUU	17.5
	2/2	GAG**GU**UAGUCU	15.7
	3/3	CCA**GU**GAGUGA	12.1
	4/4	AUG**GU**AAAGUA	12.0
	6/6	UCA**GU**AAGUCU	14.5
consensus sequence		CAG**GU**AAGUAU	23.8

The underlined part of the sequence indicates the intron. The hydrogen bond (H-Bond) score (http://www.uni-duesseldorf.de/rna/html/hbond_score.php) calculates the strength of binding between the pre-mRNAs and the U1snRNA which take the number of hydrogen bonds into account (Freund et al., 2003; Kammler et al., 2001). Consensus sequence shows the sequence which has optimal binding abilities to the U1snRNAs.

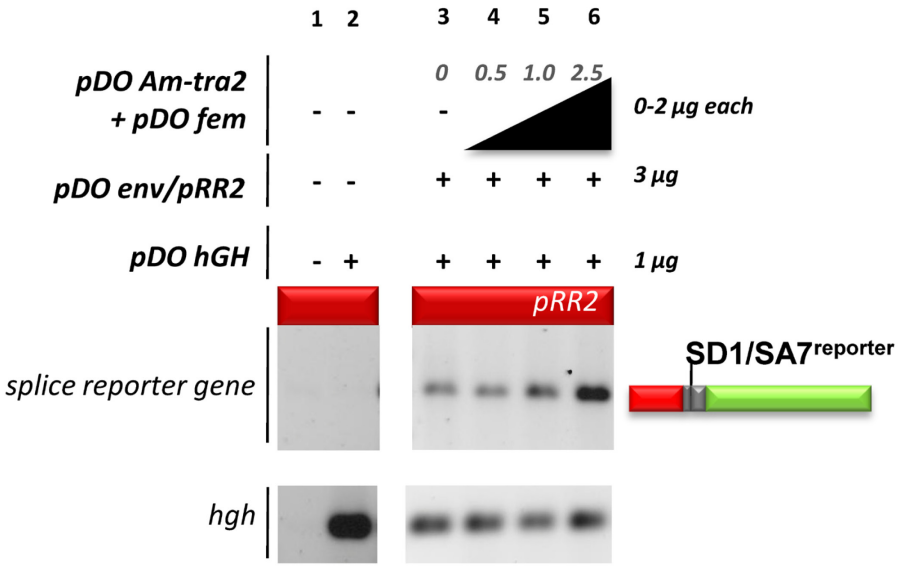

Fig. S1: Fem/*Am*-Tra2 protein dosage dependent splicing when excess of reporter gene with the pRR2 fragment (pOPI *env*/pRR2) is provided. 1x10^6 *Sf*21 cells were incubated with dd$H_2$0 (lane 1) or transiently transfected with 1µg of pDO *hGH* (lane 2-6), 3µg of pOPI *env*/pRR2 (triple amount compared to previous experiments; lane 3-6),) and increasing amounts of pDO *fem* and pDO *Am-tra2* (0.5, 1.0, 2.5 µg each) (lane 4-6). Transcripts were amplified by RT-PCRs using oligonucleotide primers as indicated in (Fig. 3B). Semiquantitative RT-PCRs of reporter transcripts were standardized in respect to the splice product of cotransfected *hGH* control. Fragments were resolved in 2% agarose gels and stained with ethidium bromide. Structure of splice products was determined by sequencing.

Fig. S2 Sequences of *hGH,* pDO vector, OpIE2 promoter, transcript *Am-dsx*$^{mg\text{-}R1}$ and transcript *Am-dsx*$^{mg\text{-}R3}$.

A) Sequence of human Growth Hormon (*hGH*) fragment, cloned into pDO vector (underlined sequence represents intron that is spliced out in *Sf*21 cells) :

ATGGCTACAGGTAAGCGCCCCTAAAATCCCTTTGGGCACAATGTGTCCTGAGGGGAGAGGCAGCGACCTGTA
GATGGGACGGGGGCACTAACCCTCAGGTTTGGGGCTTCTGAATGTGAGTATCGCCATGTAAGCCCAGTATTTG
GCCAATCTCAGAAAGCTCCTGGTCCCTGGAGGGATGGAGAGAGAAAAACAAACAGCTCCTGGAGCAGGGAG
AGTGCTGGCCTCTTGCTCTCCGGCTCCCTCTGTTGCCCTCTGGTTTCTCCCCAGGCTCCCGGACGTCCCTGCTCC
TGGCTTTTGGCCTGCTCTGCCTGCCCTGGCTTCAAGAGGGCAGTGCCTTCCCAACCATTCCCTTATCCAGGCTTT
TTGACAACGCTATGCTCCGCGCCCATCGTCTGCACCAGCTGGCCTTTGACACCTACCAGGAGTTTGTAAGCTCT
TGGGGAATGGGTGCGCATCAGGGGTGGCAGGAAGGGGTGACTTTCCCCCGCTGGGAAATAAGAGGAGGAG
ACTAAGGAGCTCAGGGTTTTTCCCGAAGCGAAAATGCAGGCAGATGAGCACACGCTGAGTGAGGTTCCCAGA
AAAGTAACAATGGGAGCTGGTCTCCAGCGTAGACCTTGGTGGGCGGTCCTTCTCCTAGGAAGAAGCCTATATC
CCAAAGGAACAGAAGTATTCATTCCTGCAGAACCCCCAGACCTCCCTCTGTTTCTCAGAGTCTATTCCGACACCC
TCCAACAGGGAGGAAACACAACAGAAATCCGTGAGTGGATGCCTTCTCCCCAGGCGGGGATGGGGGAGACCT

GTAGTCAGAGCCCCCGGGCAGCACAGCCAATGCCCGTCCTTCCCCTGCAGAACCTAGAGCTGCTCCGCATCTCC
CTGCTGCTCATCCAGTCGTGGCTGGAGCCCGTGCAGTTCCTCAGGAGTGTCTTCGCCAACAGCCTGGTGTACG
GCGCCTCTGACAGCAACGTCTATGACCTCCTAAAGGACCTAGAGGAAGGCATCCAAACGCTGATGGGGGTGA
GGGTGGCGCCAGGGGTCCCCAATCCTGGAGCCCCACTGACTTTGAGAGCTGTGTTAGAGAAACACTGCTGCCC
TCTTTTTAGCAGTCAGGCCCTGACCCAAGAGAACTCACCTTATTCTTCATTTCCCCTCGTGAATCCTCCAGGCCTT
TCTCTACACCCTGAAGGGGAGGGAGGAAAATGAATGAATGAGAAAGGGAGGGAACAGTACCCAAGCGCTTG
GCCTCTCCTTCTCTTCCTTCACTTTGCAGAGGCTGGAAGATGGCAGCCCCCGGACTGGGCAGATCTTCAAGCAG
ACCTACAGCAAGTTCGACACAAACTCACACAACGATGACGCACTACTCAAGAACTACGGGCTGCTCTACTGCTT
CAGGAAGGACATGGACAAGGTCGAGACATTCCTGCGCATCGTGCAGTGCCGCTCTGTGGAGGGCAGCTGTGG
CTTCTAG

B) Full sequence of pDO vector:

GGATCATGATGATAAACAATGTATGGTGCTAATGTTGCTTCAACAACAATTCTGTTGAACTGTGTTTTCATGTTT
GCCAACAAGCACCTTTATACTCGGTGGCCTCCCCACCACCAACTTTTTTGCACTGCAAAAAAACACGCTTTTGCA
CGCGGGCCCATACATAGTACAAACTCTACGTTTCGTAGACTATTTTACATAAATAGTCTACACCGTTGTATACGC
TCCAAATACACTACCACACATTGAACCTTTTTGCAGTGCAAAAAAGTACGTGTCGGCAGTCACGTAGGCCGGCC
TTATCGGGTCGCGTCCTGTCACGTACGAATCACATTATCGGACCGGACGAGTGTTGTCTTATCGTGACAGGAC
GCCAGCTTCCTGTGTTGCTAACCGCAGCCGGACGCAACTCCTTATCGGAACAGGACGCGCCTCCATATCAGCC
GCGCGTTATCTCATGCGCGTGACCGGACACGAGGCGCCCGTCCCGCTTATCGCGCCTATAAATACAGCCCGCA
ACGATCTGGTAAACACAGTTGAACAGCATCTGTTCGAATTTAAAGCTTGGTACCGAGCTCGGATCCACTAGTCC
AGTGTGGTGGAATTCATGGCCTCCTCCGAGGATGTCATCAAAGAGTTTATGAGATTTAAGGTCAAGATGGAGG
GAAGCGTCAACGGACACGAGTTCGAGATTGAGGGAGAAGGAGAAGGCCGGCCTTACGAGGGCACACAAACC
GCTAAGCTCAAGGTCACAAAAGGAGGACCCCTCCCCTTCTCCTGGGATATTCTGAGCCCTCAGTTCCAGTACGG
AAGCAAAGCCTATGTTAAACACCCTGCCGACATCCCTGACTATCTGAAGCTCTCCTTCCCTGAAGGCTTCAAGT
GGGAGAGATTCATGAACTTCGAGGACGGAGGCGTGGTGACAGTCACACAAGATAGCACCCTCCAGGACGGA
GAGTTTATTTATAAGGTGAAACTCAGAGGAACCAACTTCCCCTCCGATGGCCCTGTCATGCAAAAAAAAACAAT
GGGATGGGAAGCCTCCACCGAGAGAATGTATCCTGAGGATGGCGCTCTGAAAGGCGAAATTAAAATGAGACT
GAAACTCAAAGACGGAGGACACTACGATGCCGAGGTCAAAACAACCTACAAGGCCAAGAAACAAGTGCAGCT
GCCTGGCGCCTACATGACTGATATTAAACTCGACATTATCAGCCATAATGGGGACTACACCATCGTGGAACAAT
ATGAGAGAGCTGAGGGCAGACATAGCACAGGCGCTGGATCCGCGGCCGCGAGCTCACTAGTCATATGTTCTA
GAATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTA
AACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTC
ATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTACGGCGTGCAGTGCTT
CAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAG

CGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTG
GTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTAC
AACTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCC
GCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCC
CCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGA
TCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTAACCG
CGGTTCGAAGGTAAGCCTATCCCTAACCCTCTCCTCGGTCTCGATTCTACGCGTACCGGTCATCATCACCATCAC
CATTGAGTTTATCTGACTAAATCTTAGTTTGTATTGTCATGTTTTAATACAATATGTTATGTTTAAATATGTTTTT
AATAAATTTTATAAAATAATTTCAACTTTTATTGTAACAACATTGTCCATTTACACACTCCTTTCAAGCGCGTGGG
ATCGATGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGA
GCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCC
CTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGG
CGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTC
TCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCC
AAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTC
CAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGT
AGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGC
GCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTA
GCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTT
TCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGATGATAAACAATGTAT
GGTGCTAATGTTGCTTCAACAACAATTCTGTTGAACTGTGTTTTCATGTTTGCCAACAAGCACCTTTATACTCGG
TGGCCTCCCCACCACCAACTTTTTTGCACTGCAAAAAAACACGCTTTTGCACGCGGGCCCATACATAGTACAAA
CTCTACGTTTCGTAGACTATTTTACATAAATAGTCTACACCGTTGTATACGCTCCAAATACACTACCACACATTG
AACCTTTTTGCAGTGCAAAAAAGTACGTGTCGGCAGTCACGTAGGCCGGCCTTATCGGGTCGCGTCCTGTCAC
GTACGAATCACATTATCGGACCGGACGAGTGTTGTCTTATCGTGACAGGACGCCAGCTTCCTGTGTTGCTAACC
GCAGCCGGACGCAACTCCTTATCGGAACAGGACGCGCCTCCATATCAGCCGCGCGTTATCTCATGCGCGTGAC
CGGACACGAGGCGCCCGTCCCGCTTATCGCGCCTATAAATACAGCCCGCAACGATCTGGTAAACACAGTTGAA
CAGCATCTGTTCGAATTCGTTCAGCACCTAGGAGCAGCCTTAAGAGCAGCCTCGAGATCTTAGTTTGTATTGTC
ATGTTTTAATACAATATGTTATGTTTAAATATGTTTTTAATAAATTTTATAAAATAATTTCAACTTTTATTGTAACA
ACATTGTCCATTTACACACTCCTTTCAAGCGCGTGGCCGGGCTGCAGCACGTGTTGACAATTAATCATCGGCAT
AGTATATCGGCATAGTATAATACGACTCACTATAGGAGGGCCACCATGGCCAAGTTGACCAGTGCCGTTCCGG
TGCTCACCGCGCGCGACGTCGCCGGAGCGGTCGAGTTCTGGACCGACCGGCTCGGGTTCTCCCGGGACTTCGT
GGAGGACGACTTCGCCGGTGTGGTCCGGGACGACGTGACCCTGTTCATCAGCGCGGTCCAGGACCAGGTGGT
GCCGGACAACACCCTGGCCTGGGTGTGGGTGCGCGGCCTGGACGAGCTGTACGCCGAGTGGTCGGAGGTCG
TGTCCACGAACTTCCGGGACGCCTCCGGGCCGGCCATGACCGAGATCGGCGAGCAGCCGTGGGGGCGGGAG

TTCGCCCTGCGCGACCCGGCCGGCAACTGCGTGCACTTCGTGGCCGAGGAGCAGGACTGACCGACGCCGACC
AACACCGCCGGTCCGACGGCGGCCCACGGGTCCCAGGGGGGTCGACCTCGAAACTTGTTTATTGCAGCTTATA
ATGGTTACAAATAAAGCAATAGCATCACAAATTTCACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTT
TGTCCAAACTCATCAATGTATCTTATCATGTCT

C) Sequence of OpIE2 promoter cloned into pOPI *env*/ pRR expression vectors:

TCATGATGATAAACAATGTATGGTGCTAATGTTGCTTCAACAACAATTCTGTTGAACTGTGTTTTCATGTTTGCC
AACAAGCACCTTTATACTCGGTGGCCTCCCCACCACCAACTTTTTTGCACTGCAAAAAAACACGCTTTTGCACGC
GGGCCCATACATAGTACAAACTCTACGTTTCGTAGACTATTTTACATAAATAGTCTACACCGTTGTATACGCTCC
AAATACACTACCACACATTGAACCTTTTTGCAGTGCAAAAAAGTACGTGTCGGCAGTCACGTAGGCCGGCCTTA
TCGGGTCGCGTCCTGTCACGTACGAATCACATTATCGGACCGGACGAGTGTTGTCTTATCGTGACAGGACGCC
AGCTTCCTGTGTTGCTAACCGCAGCCGGACGCAACTCCTTATCGGAACAGGACGCGCCTCCATATCAGCCGCG
CGTTATCTCATGCGCGTGACCGGACACGAGGCGCCCGTCCCGCTTATCGCGCCTATAAATACAGCCCGCAACG
ATCTGGTAAACACAGTTGAACAGCATCTGTTCGAATTTA

D) Sequence of *Am-dsx* transcript *Am-dsx*$^{mg\text{-}R1}$ (underlined and not-underlined sequence represents exon junctions)

GTGAAAATGTGGAAATTCTATTGGAGCACAGTAGCAAACTTGTAGAACTTTTCCAATATCCTTGGGAAGCACT
GTTATTGATGTACATCAATTTAAAATATGCAGGGGCTAATCCGGAAGAAGTAGTGAGACGTATGGTTGATGAA
AGTGATCATTTATAAATGAAGATTATCAAATGATCACACGATTTATTATATTTATTCAACACGTGAAGATTTGTT
ACAATTATATCTTCATTGGAAGATAGAATTTAATCATCAGGAAGATATATATCTATATGAATATATATCAGCTGT
TTGTATTCAACAATGTAAGATATAAAGAGGGATTTAAAGAGCTAGCAACGAAATCCGTAACATGCACTTCTTGA
AAGCAATCAGAATGTCCCAACCTAGTCGTGCGTTCCGCTGCACCGCGGCATGTGCCGCACCCACGGGGCCGCC
AACGGGACCTCCGACATACGAGGGTGATGTGCCCTTCATAGGGGTTGGACCACCCCCGAATCCTATCCATTTC
AGGCCGTTTCTTCATCCCGAGAATGCTCATATACCGGCTACCAGACTACCGTCCAGCCCAGATGGTCCTCCTAA
ACATACGTAGCCTCA

E) Sequence of *Am-dsx* transcript *Am-dsx*$^{mg\text{-}R3}$ (underlined and not-underlined sequence represents exon junctions)

GTGAAAATGTGGAAATTCTATTGGAGCACAGTAGCAAACTTGTAGAACTTTTCCAATATCCTTGGGAAGCACT
GTTATTGATGTACATCAATTTAAAATATGCAGGGGCTAATCCGGAAGAAGTAGTGAGACGTATGGTTGATGAA
TTATATTGAAGAAAGATTCTAATATAACTTCATTATTAATGAACCGATTATTCTTATATTTTTGTTGGAAAAATG
ACAAGAGTAAATTTAATTTTGACGGTTTAATGATTAAAAGATAGATTTCTTTGTCGTGAAATAGTTGGCGCATG
TGGATAACGAAATTGGCAACAACATTCATATTTTAACAATATGTAATAAATTTTCTTTATATATCATGCACTTTTA
TTAAAGTAAAATGATGAAATTGTAGAACATTAATAAATTAAATTTTTTTGATAGTATATAAATCAAAAATTATAT
TACATATTATTCCAAAGAACTATTTCTATTATTCGGATTTTATTTCTGATAAAACATTATAAAATACAATGGTGCA
CTAAGTTTCTCGAAAATAAATTTAGAGTCGAGGACACAGGTGTTAAAACCATTACAATAAGATTTATGCTTCCA
GTATATATTTGTTCATGTCAATTTCGTTGTATTCGTCACCCTTAACAATAATATCGAATCAAAATTGCAGAAAGT
GATCATTTATAAATGAAGATTATCAAATGATCACACGATTTATTATATTTATTCAACACGTGAAGATTTGTTACA
ATTATATCTTCATTGGAAGATAGAATTTAATCATCAGGAAGATATATATCTATATGAATATATATCAGCTGTTTG
TATTCAACAATGTAAGATATAAAGAGGGATTTAAAGAGCTAGCAACGAAATCCGTAACATGCACTTCTTGAAA
GCAATCAGAATGTCCCAACCTAGTCGTGCGTTCCGCTGCACCGCGGCATGTGCCGCACCCACGGGGCCGCCAA
CGGGACCTCCGACATACGAGGGTGATGTGCCCTTCATAGGGGTTGGACCACCCCCGAATCCTATCCATTTCAG
GCCGTTTCTTCATCCCGAGAATGCTCATATACCGGCTACCAGACTACCGTCCAGCCCAGATGGTCCTCCTAAAC
ATACGTAGCCTCA

Reference list

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*. Molecular and cellular biology *16*, 3106–3111.

Asang, C., Hauber, I., and Schaal, H. (2008). Insights into the selective activation of alternatively used splice acceptors by the human immunodeficiency virus type-1 bidirectional splicing enhancer. Nucleic acids research *36*, 1450–1463.

Baralle, D., and Baralle, M. (2005). Splicing in action: assessing disease causing sequence changes. Journal of medical genetics *42*, 737–748.

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: news and reviews in molecular, cellular and developmental biology *26*, 1131–1139.

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*, 419–429.

Burghardt, G., Hediger, M., Siegenthaler, C., Moser, M., Dübendorfer, A., and Bopp, D. (2005). The *transformer2* gene in *Musca domestica* is required for selecting and maintaining the female pathway of development. Development genes and evolution *215*, 165–176.

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.

Caputi, M., Freund, M., Kammler, S., Asang, C., and Schaal, H. (2004). A Bidirectional SF2/ASF- and SRp40-Dependent Splicing Enhancer Regulates Human Immunodeficiency Virus Type 1 rev, env, *vpu* and *nef* Gene Expression. Journal of Virology *78*, 6517–6526.

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*, 1733–1741.

Christiansen, A. E., Keisman, E. L., Ahmad, S. M., and Baker, B. S. (2002). Sex comes in from the cold: the integration of sex and pattern. Trends in genetics : TIG *18*, 510–516.

Cline, T. W., and Meyer, B. J. (1996). Vive la différence: males vs females in flies vs worms. Annual review of genetics *30*, 637–702.

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

Dearden, P. K., Wilson, M. J., Sablan, L., Osborne, P. W., Havler, M., McNaughton, E., Kimura, K., Milshina, N. V., Martin, H., Gempe, T., et al. (2006). Patterns of conservation and change in honey bee developmental genes. Genome Research, 1376–1384.

Erickson, J. W., and Quintero, J. J. (2007). Indirect effects of ploidy suggest X chromosome dose, not the X:A ratio, signals sex in *Drosophila*. PLoS biology *5*, e332.

Fairbrother, W. G., Yeh, R.-F., Sharp, P. a, and Burge, C. B. (2002). Predictive identification of exonic splicing enhancers in human genes. Science (New York, N.Y.) *297*, 1007–1013.

Freund, M., Asang, C., S, K., C, K., Krummheuer, J., Hipp, M., Meyer, I., Gierling, W., Theiss, S., Preuss, T., et al. (2003). A novel approach to describe a U1 snRNA binding site. Nucleic Acids Research *31*, 6963–6975.

Freund, M., Hicks, M. J., Konermann, C., Otte, M., Hertel, K. J., and Schaal, H. (2005). Extended base pair complementarity between U1 snRNA and the 5' splice site does not inhibit splicing in higher eukaryotes, but rather increases 5' splice site recognition. Nucleic acids research *33*, 5112–5119.

Gempe, T., Hasselmann, M., Schiøtt, M., Hause, G., Otte, M., and Beye, M. (2009). Sex determination in honeybees: two separate mechanisms induce and maintain the female pathway. PLoS biology *7*, e1000222.

Gempe, T., and Beye, M. (2011). Function and evolution of sex determination mechanisms, genes and pathways in insects. BioEssays : news and reviews in molecular, cellular and developmental biology *33*, 52–60.

Graveley, B. R., Hertel, K. J., and Maniatis, T. (1998). A systematic analysis of the factors that determine the strength of pre-mRNA splicing enhancers. The EMBO journal *17*, 6747–6756.

Graveley, B. R., Hertel, K. J., and Maniatis, T. (2001). The role of U2AF35 and U2AF65 in enhancer-dependent splicing. RNA (New York, N.Y.) *7*, 806–818.

Hasselmann, M., Gempe, T., Schiøtt, 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*, 519–522.

Hasselmann, M., and Beye, M. (2004). Signatures of selection among sex-determining alleles of the honey bee. Proceedings of the National Academy of Sciences of the United States of America *101*, 4888–4893.

Hediger, M., Burghardt, G., Siegenthaler, C., Buser, N., Hilfiker-Kleiner, D., Dübendorfer, A., and Bopp, D. (2004). Sex determination in *Drosophila melanogaster* and *Musca domestica* converges at the level of the terminal regulator doublesex. Development genes and evolution *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*, 579–586.

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 (New York, N.Y.) *252*, 833–836.

Inoue, K., Hoshijima, K., Higuchi, I., Sakamoto, H., and Shimura, Y. (1992). Binding of the *Drosophila* transformer and transformer-2 proteins to the regulatory elements of *doublesex* primary transcript for sex-specific RNA processing. Proceedings of the National Academy of Sciences of the United States of America *89*, 8092–8096.

Kamakura, M. (2011). Royalactin induces queen differentiation in honeybees. Nature *473*, 478–483.

Kammler, S., Leurs, C., Freund, M., Krummheuer, J., Seidel, K., Tange, T. Ø., Lund, M. K., and Kjems, J. (2001). The sequence complementarity between HIV-1 5' splice site SD4 and U1 snRNA determines the steady-state level of an unstable env pre-mRNA. RNA *7*, 421–434.

Kato, Y., Kobayashi, K., Watanabe, H., and Iguchi, T. (2011). Environmental sex determination in the branchiopod crustacean *Daphnia magna*: deep conservation of a *Doublesex* gene in the sex-determining pathway. PLoS genetics *7*, e1001345.

Keisman, E. L., Christiansen, a E., and Baker, B. S. (2001). The sex determination gene *doublesex* regulates the A/P organizer to direct sex-specific patterns of growth in the *Drosophila* genital imaginal disc. Developmental cell *1*, 215–225.

Keisman, E. L., and Baker, B. S. (2001). The *Drosophila* sex determination hierarchy modulates *wingless* and *decapentaplegic* signaling to deploy dachshund sex-specifically in the genital imaginal disc. Development (Cambridge, England) *128*, 1643–1656.

Kopp, a, Duncan, I., Godt, D., and Carroll, S. B. (2000). Genetic control and evolution of sexually dimorphic characters in *Drosophila*. Nature *408*, 553–559.

Kucharski, R., Maleszka, J., Foret, S., and Maleszka, R. (2008). Nutritional control of reproductive status in honeybees via DNA methylation. Science (New York, N.Y.) *319*, 1827–1830.

Kuo, B. A., and Norton, P. A. (1999). Accurate selection of a 5' splice site requires sequences within fibronectin alternative exon B. Nucleic acids research *27*, 3945–3952.

Lagos, D., Ruiz, M. F., Sánchez, 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*, 111–121.

Lynch, K. W., and Maniatis, T. (1996). Assembly of specific SR protein complexes on distinct regulatory elements of the *Drosophila doublesex* splicing enhancer. Genes & Development *10*, 2089–2101.

Lynch, K. W., and Maniatis, T. (1995). Synergistic interactions between two distinct elements of a regulated splicing enhancer. Genes & Development *9*, 284–293.

Matson, C. K., Murphy, M. W., Sarver, A. L., Griswold, M. D., Bardwell, V. J., and Zarkower, D. (2011). DMRT1 prevents female reprogramming in the postnatal mammalian testis. Nature *476*, 101–104.

Mattox, W., Palmer, M. J., and Baker, B. S. (1990). Alternative splicing of the sex determination gene *transformer-2* is sex-specific in the germ line but not in the soma. Genes & Development *4*, 789–805.

Mount, S. M. (1982). A catalogue of splice junction sequences. Nucleic Acids Research *10*, 459–472.

Nissen, I., Müller, M., and Beye, M. The *Am-tra2* gene is an essential cofactor of female splice regulation at two levels of the complementary sex determination hierarchy of the honeybee. Genetics, under review.

Norton, P. A. (1994). Polypyrimidine tract sequences direct selection of alternative branch sites and influence protein binding. Nucleic Acids Research *22*, 3854–3860. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=308380&tool=pmcentrez&rendertype=abstract.

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., 145–158.

Oliveira, D. C. S. G., Werren, J. H., Verhulst, E. C., Giebel, J. D., Kamping, A., Beukeboom, L. W., and Zande, L. V. D. (2009). Identification and characterization of the *doublesex* gene of *Nasonia*. Insect molecular biology *18*, 315–324.

Qi, J., Su, S., and Mattox, W. (2007). The *doublesex* splicing enhancer components Tra2 and Rbp1 also repress splicing through an intronic silencer. Molecular and cellular biology *27*, 699–708.

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*, 691–695.

Ruiz, M. F., Eirín-López, J. M., Stefani, R. N., Perondini, A. L. P., Selivon, D., and Sánchez, L. (2007). The gene *doublesex* of *Anastrepha* fruit flies (Diptera, Tephritidae) and its evolution in insects. Development genes and evolution *217*, 725–731.

Ruiz, M. F., Stefani, R. N., Mascarenhas, R. O., Perondini, A. L. P., Selivon, D., and Sánchez, L. (2005). The gene *doublesex* of the fruit fly *Anastrepha obliqua* (Diptera, Tephritidae). Genetics *171*, 849–854.

Saccone, G., Salvemini, M., Pane, A., and Polito, L. C. (2008). Masculinization of XX *Drosophila* transgenic flies expressing the *Ceratitis capitata* DoublesexM isoform. The International journal of developmental biology *52*, 1051–1057.

Salvemini, M., Robertson, M., Aronson, B., Atkinson, P., Polito, L. C., and Saccone, G. (2009). *Ceratitis capitata transformer-2* gene is required to establish and maintain the autoregulation of *Cctra*, the master gene for female sex determination. The International journal of developmental biology *53*, 109–120.

Sciabica, K. S., and Hertel, K. J. (2006). The splicing regulators Tra and Tra2 are unusually potent activators of pre-mRNA splicing. Nucleic acids research *34*, 6612–6620.

Selden, R. F., Howie, K. B., Rowe, M. E., Goodman, H. M., and Moore, D. D. (1986). *Human growth hormone* as a reporter gene in regulation studies employing transient gene expression. Molecular and cellular biology *6*, 3173–3179.

Shen, H., and Green, M. R. (2006). RS domains contact splicing signals and promote splicing by a common mechanism in yeast through humans. Genes & Development *20*, 1755–1765.

Singh, K. K., Erkelenz, S., Rattay, S., Dehof, A. K., Hildebrandt, A., Schulze-Osthoff, K., Schaal, H., and Schwerk, C. (2010). Human SAP18 mediates assembly of a splicing regulatory multiprotein complex via its ubiquitin-like fold. *16*, 2442–2454.

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*. Development genes and evolution *213*, 345–354.

Suzuki, M. G., Imanishi, S., Dohmae, N., Asanuma, M., and Matsumoto, S. (2010). Identification of a male-specific RNA binding protein that regulates sex-specific splicing of *Bmdsx* by increasing RNA binding activity of *Bm*PSI. Molecular and cellular biology *30*, 5776–5786.

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. Molecular and cellular biology *28*, 333–343.

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 biochemistry and molecular biology *31*, 1201–1211.

Sánchez, L., Gorfinkiel, N., and Guerrero, I. (2001). Sex determination genes control the development of the *Drosophila* genital disc, modulating the response to Hedgehog, Wingless and Decapentaplegic signals. Development (Cambridge, England) *128*, 1033–1043.

Sánchez, L., and Guerrero, I. (2001). The development of the *Drosophila* genetical disc. BioEssays *23*, 698–707.

Tian, M., and Maniatis, T. (1993). A Splicing Enhancer Complex Controls Alternative Splicing of *doublesex* Pre-mRNA. Cell *74*, 105–114.

Tian, M., and Maniatis, T. (1994). A splicing enhancer exhibits both constitutive and regulated activities. Genes & Development *8*, 1703–1712.

Valcárcel, 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*, 171–175.

Vincent, S., Perkins, L. A., and Perrimon, N. (2001). Doublesex Surprises. Cell *106*, 399–402.

Williams, T. M., Selegue, J. E., Werner, T., Gompel, N., Kopp, A., and Carroll, S. B. (2008). The regulation and evolution of a genetic switch controlling sexually dimorphic traits in *Drosophila*. Cell *134*, 610–623.

Zuo, P., and Maniatis, T. (1996). The splicing factor U2AF35 mediates critical protein-protein interactions in constitutive and enhancer-dependent splicing. Genes & Development *10*, 1356–1368.

Fem and *Am*-Tra2 proteins direct female specific splicing of the *Am-dsx* pre-mRNA by activation of the female splice donor site in the honeybee *Apis mellifera*

Journal: Submission of manuscript in preparation to "Molecular and Cellular Biology"

Impact Factor: 6,057

1st author

Author's contribution: 95%

- Vector construction and cloning
- Experimental design
- Implementation of laboratory experiments
- Molecular evolutionary sequence analyses
- Authoring the manuscript

Heiner Schaal and Martin Beye: advisory function

The SV-*env*/GAR *egfp* vector was provided by Steffen Erkelenz.

Additional experiment: RNA Electro Mobility Shift Assay

Short introduction

In the study "Fem and *Am*-Tra2 proteins direct female specific splicing of the *Am-dsx* pre-mRNA by activation of the female splice donor site in the honeybee *Apis mellifera"* we present evidence for the direct involvement of Fem and *Am*-Tra2 proteins in sex-specific *Am-dsx* pre-mRNA splicing. Furthermore we showed that the last 36 nucleotides (putative regulatory region 1 - pRR1) of the female specific exon 5 contain exonic splicing enhancer (ESE) sequences that are affected by presence of Fem and *Am*-Tra2 proteins. We narrowed down the ESE motif on two repeats of the hexamer motif AAAGAG. To further analyze the *Am-dsx* splicing regulation, we studied binding of Fem and *Am*-Tra2 proteins to the ESE sequences in an RNA Electro Mobility Shift Assay (EMSA).

Results and Discussion

We labeled the putative *Am-dsx* ESE RNA sequences (last 23 nucleotides of *Am-dsx* exon 5) with biotin. We expressed Fem and *Am*-Tra2 proteins in TNT T7 coupled rabbit reticulocyte lysate. As positive control for protein-RNA interactions we used instead of Reticulocyte Lysate, Cytosolic liver extract containing iron-responsive protein (IRP) and as RNA target Iron responsive Element (IRE).

We incubated the putative *Am-dsx* ESE RNAs together with heterologous expressed Fem and *Am*-Tra2 proteins and resolved the RNAs/proteins by native polyacrylamide gel electrophoresis. To visualize the shifting of RNAs we transferred the free RNA and the RNA/protein-complexes onto a nylon membrane and analyzed the labeled RNA molecules by chemiluminescent detection. If RNA-protein complexes are formed the RNAs migrate slower than the free RNAs. By increasing the amount of unlabeled RNA we tested whether we would be able to reduce the number of protein/biotinylated RNA complexes what is evidence of specific RNA/protein interactions.

Our putative *Am-dsx* ESE RNAs did not migrate slower in response to the presence of the proteins Fem and *Am*-Tra2. Instead we observed a diffuse smear in the background (fig. 1A, lane 1-4) and only low amounts of free RNA probe (fig.1A, lane 2-5) suggesting a non-specific interaction of the biotin labeled RNA molecules with different proteins of the lysate. I suggest repeating this experiment with purified Fem and *Am*-Tra2 proteins and the absence

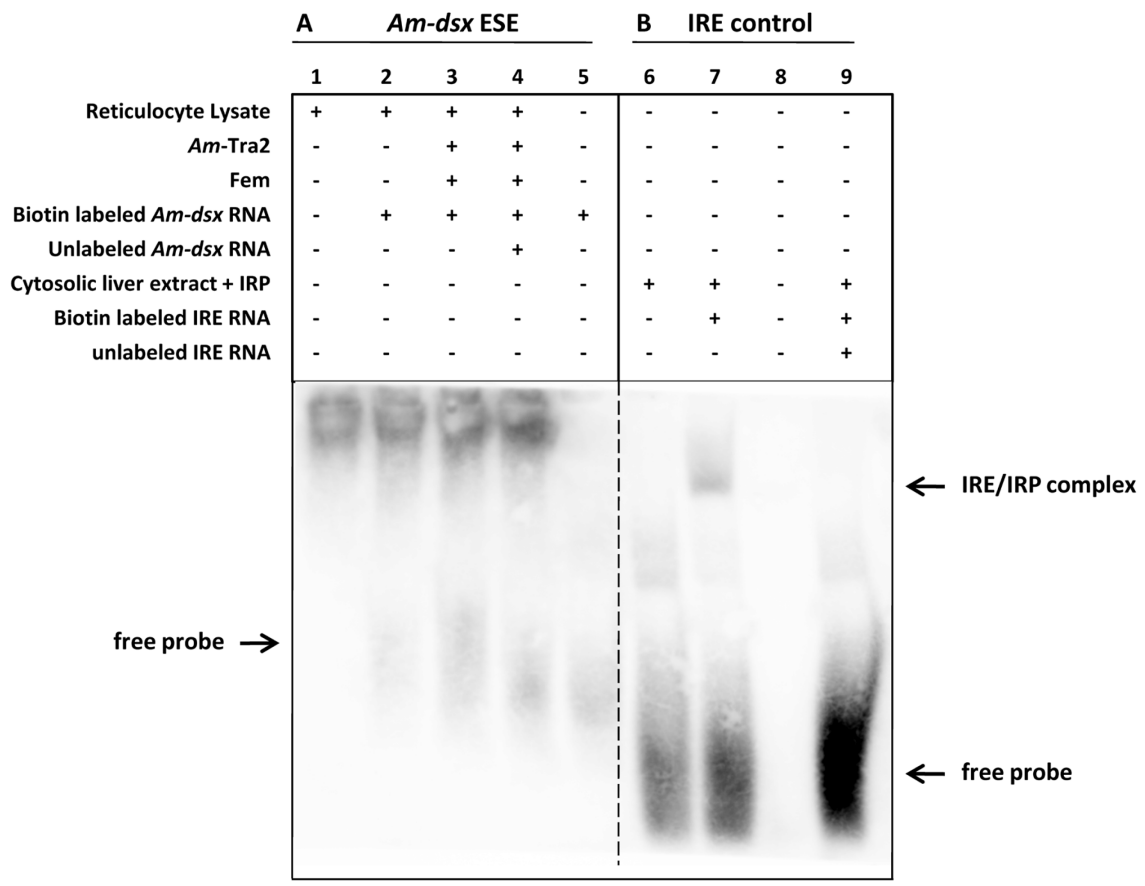

Fig. 1: RNA Electromobility Shift Assay (EMSA) studies of Fem and *Am*-Tra2 proteins with putative *Am-dsx* ESEs. RNA gel shift assays were performed for (A) of the putative *Am-dsx* ESE RNAs (last 23 nucleotides of *Am-dsx* exon 5) and (B) the iron responsive element (IRE) as positive control using the LightShift Chemiluminescent RNA EMSA Kit. Binding reactions (20µl) included 1x RNA EMSA Binding Buffer, 5% Glycerol, 2 µg of tRNA, 11.45 µl rabbit reticulocyte lysate (lane 1-4), 4 µg of Cytosolic liver extract containing iron-responsive protein (IRP; lane 6, 7, 9), 6.25 nM of the respective Biotin-labeled RNA (*Am-dsx* ESE RNA, lane 2-5; IRE RNA, lane 7-9) and 1 µM of the respective unlabeled RNA (*Am-dsx*, lane4; IRE, lane 9). Lane 8 is the negative control. Recombinant Fem and *Am*-Tra2 proteins were expressed in the TNT T7 coupled rabbit reticulocyte lysate. We used 0.5 µg of pGBKT7 Fem and 0.5 µg of pGBKT7 *Am*-Tra2 plasmids as template. In the controls we used ultrapure water instead of DNA. Reactions were resolved on a native 6% polyacrylamide gel in 0.5X TBE by gel electrophoresis and transferred to a nylon membrane. Band shifts were detected using the Pierce Chemiluminescent Nucleic Acid Detection Module.

of other proteins. To show that our method works well, we analyzed complex formation of the control protein Iron Responsive Protein (IRP) with the Iron Responsive Element (IRE) RNA. Here we observed a band shift (Fig. 1B, lane 6-7) which was reduced by addition of unlabeled free RNA (Fig. 1B, lane 9).

To optimize our experimental design for upcoming experiments, we will overexpress Fem and *Am*-Tra2 protein in *Escherichia coli* with subsequently purification. This should simplify the identification of the optimal amount of protein for the binding reaction and reduce the

background which was caused by non-specific binding of the putative *Am-dsx* ESE RNAs to other proteins.

Materials and Methods

Construction of T7 expression vectors

All vectors for the protein expression in reticulocyte lysate base on the pGBKT7 vector (Clontech) derived pGBKT7 Allel 16 Triple TAG3' vector (kindly provided by Marianne Otte). This vector lacks the multiple cloning site (MCS) and the c-Myc-epitope of the original pGBKT7 vector, but contains instead the coding sequence of allele D1-16 of the *complementary sex determiner* (*csd*) gene. The vector carries a Kanamycin resistence for expression in *E.coli* and a T7 promoter for *in vitro* transcription and translation.

To produce the protein expression plasmids, the full length *Apis mellifera transformer-2 (Am-tra2)* (splice variant *Am-tra2*^{285}; GenBank JQ518311) open reading frame was amplified by PCR using primers #660 and #370 from vector pDO *Am*-Tra2. In addition we restricted the coding sequence of *fem S2-38* (GenBank: AY569719.1) from pGBKT7 Myc Fem S2-38 vector (kindly provided by Marianne Otte). To generate the plasmid vectors pGBKT7 *Am*-Tra2 and pGBKT7 Fem S2-38, the amplified *Am-tra2* fragment and the restricted *fem S2-38* fragment were cloned into pGBKT7 Allel 16 Triple TAG3' vector using the *Eco*RI and *Sal*I restriction sites by replacing the csd Allel 16.

RNA/ protein binding reaction and RNA Electrophoretic Mobility Shift Assay (EMSA)

For the RNA binding shift assay we used the LightShift Chemiluminescent RNA EMSA Kit (Thermoscientific). The Kit includes a LightShift RNA EMSA Optimization and Control Kit which we used as positive control in the binding reaction following the manufacturer's instructions. We used the TNT T7 coupled Rabbit Reticulocyte Lysate of Promega to express *Am*-Tra2 and Fem proteins following the protocols provided by the manufacturer. To prepare small-scale reactions we used only one-half of the stated volume in the TNT Lysate Reactions. As template for *in vitro* transcription and translation we used 0.5 µg of pGBKT7 Fem and 0.5µg of pGBKT7 *Am*-Tra2 plasmids. In the controls we used ultrapure water

instead of DNA. After protein expression 11.45 µl of the reticulocyte lysate reaction were added to the respective binding reaction sample. 3'Biotin labeled and unlabeled target RNA was synthesized purified (HPLC) and lyophilised by Eurofins MWG Operon (Ebersberg, Germany) and resolved in ultrapure water. RNA was added in a final concentration of 6.25 nM (Biotin labeled RNA) and 1 µM (unlabeled RNA). To achieve optimal binding conditions we added 2 µl 50% Glycerol and 0.2 µl tRNA (10 mg/ml) to each sample. Positive controls and *Am-dsx* binding reactions were incubated 30 minutes at room temperature, immediately after this mixed with 5 µl REMSA Loading Buffer (5x) and loaded on a 6% native polyacrylamide gel (8 x 8 x 0.1 cm). Gels were run at 12.5 V/cm for 1h at 4°C in 0.5xTBE. Proteins and RNA were transferred to nylon membrane (Roth, Karlsruhe, Germany) using semi-dry transfer (Biorad) in cooled 0.5xTBE at 125 mA for 35 minutes. Immediately after transfer the membrane was irradiated with UV-light (120 mJ/cm^2 at 254 nm) for crosslinking. Chemiluminescent Nucleic Acid Detection was performed using the included Biotin labeled RNA detection module and the PhosphoImager following the manufacturer's instructions.

Table 1: Sequence of used DNA and RNA Oligonucleotide primers

Sequence of DNA Oligonucleotide primers to cloning of pGBKT7 *Am*-Tra2	
#660	GATCGAATTCATGAGTGACATTGAGCGAAGTAGTAG
#370	TGACGTCGACTTAATATCGACGTGGTGAATAAGAGC
Sequence of RNA Oligonucleotide for RNA binding assay (3'Biotin labeled and unlabeled)	
Am-dsx last 23nt	GAUAUAAAGAGGGAUUUAAAGAG

Author's contribution: "Experiment: RNA binding shift assay"

1st author: Inga Nissen

Author's contribution: 100%

Chapter III

Zusammenfassung

Die Entscheidung ob sich aus einem Embryo der Honigbiene (*Apis mellifera*) männliche oder weibliche Tiere entwickeln, wird im Rahmen der Geschlechtsdeterminationskaskade gefällt, in die die Gene *complementary sex determiner* (*csd*), *feminizer* (*fem*) und *Am-doublesex* (*Am-dsx*) involviert sind. In der vorliegenden Arbeit wurde mit *Am-transformer 2* (*Am-tra2*) ein weiterer Spleißfaktor charakterisiert, der auch in *Drosophila melanogaster* eine wichtige Rolle in der Geschlechtsdetermination spielt. Insgesamt sechs geschlechtsunspezifische Spleißvarianten kodieren für Isoformen des SR-Proteins *Am*-Tra2, die während der gesamten Entwicklung transkribiert werden. Dabei ist die Expression im Embryo am höchsten und nimmt im pupalen Stadium ab. Alle Isoformen weisen eine RNA-Bindedomäne (RBD), sowie zwei Arginin-/Serin reiche Domänen auf, die sich bei den Isoformen in der Länge unterscheiden. Die embryonale Repression der *Am-tra2*-Transkription über RNA-Interferenz zeigt, dass das Gen *Am-tra2* sowohl für das weibchenspezifische Spleißen der *fem*- und *Am-dsx*-prä-mRNA, als auch für die Produktion des männlichen *fem*-Transkripts wichtig ist. Zusätzlich hat die Repression der *Am-tra2*-Transkription eine erhöhte embryonale Letalitätsrate zur Folge, was auf eine vitale Funktion des Gens während der Embryogenese hindeutet. Die genaue Funktion von *Am-Tra2*-Proteinen auf Ebene der *fem*-Spleißregulation ist noch ungeklärt, da *Am*-Tra2-Proteine sowohl zusammen mit Csd-Heteromeren als auch mit Fem-Proteinen, das weibchenspezifische Spleißen der *fem*-prä-mRNA vermitteln könnten. Die Bedeutung von Fem- und *Am*-Tra2-Proteinen für das Spleißen der *Am-dsx*-prä-mRNA wurde mittels eines *in vivo* Spleißassays in Schmetterlingszellen (*Sf*21-Zellen) untersucht: Die Expression eines *Am-dsx* Minigenkonstrukts zeigte, dass der männliche Spleißweg von *Am-dsx* voreingestellt ist, während *Am*-Tra2-Protein zusammen mit Fem-Protein für die Produktion von weiblichem *Am-dsx* Transkript sorgt. Im Bereich direkt vor dem weibchenspezifischen Spleißdonor von *Am-dsx,* befinden sich exonische spleißsteigernde Sequenzen (ESEs), die bei Koexpression von *Am*-Tra2- und Fem-Proteinen zur Aktivierung des Spleißdonors beitragen. Mittels Insertion von Mutationen konnten innerhalb dieses Sequenzabschnitts zwei Wiederholungen eines Hexamer-Motivs (AAAGAG) identifiziert werden, die als ESE wichtig für die *Am*-Tra2-/Fem-Protein abhängige Spleißdonoraktivierung sind. Diese Befunde zeigen, dass die einzelnen Faktoren des *dsx*-Spleißprozesses zwischen *D. melanogaster* und *A. mellifera* konserviert sind, der Mechanismus der *dsx* Spleißregulation jedoch evolviert ist.

Summary

The decision of male or female development in the embryo of the honeybee (*Apis mellifera*) involves the genes *complementary sex determiner* (*csd*), *feminizer* (*fem*) and *Am-doublesex* (*Am-dsx*), which are components of the regulatory cascade of sex determination. In the present work, I characterized an additional splicing factor, *Am*-Tra2 protein, whose ortholog in *Drosophila melanogaster* has also an important function within sex determination. Six splice variants of the gene *Am-tra2* are transcribed throughout development and code the SR-protein *Am*-Tra2. The non-sex-specific transcripts are most abundant in embryonic to larval stages. Each isoform of the *Am*-Tra2 proteins has two arginine/serine-rich domains of differing lengths and an RNA binding domain (RBD). Embryonic repression of *Am-tra2* gene by RNA interference showed that the *Am-tra2* gene has an important role in female specific splicing of both *fem* and *Am-dsx* pre-mRNAs and the production of male *fem* transcripts. In addition, knockdown of the *Am-tra2* gene affects the viability of the embryo, suggesting that the *Am-tra2* gene has an essential function in embryogenesis of the honeybee. The role of *Am*-Tra2 protein on the level of *fem* pre-mRNA splicing is still unclear, as *Am*-Tra2 proteins may act as putative cofactors of Csd heteromeres or of Fem proteins that mediate splicing of the *fem* pre-mRNAs into the productive female form. The impact of Fem and *Am*-Tra2 proteins in *Am-dsx* pre-mRNA splicing was examined by an *in vivo* splice assay in lepidopteran cells (*Sf*21-cells). The expression of an *Am-dsx* mini gene showed, that male splicing of the *Am-dsx* pre-mRNA is the default state, whereas *Am*-Tra2 and Fem proteins direct together female specific *Am-dsx* splicing. The sequence directly upstream of the female specific splice donor site contains exonic splice enhancer sequences (ESEs) that are involved in the activation of the splice donor in presence of Fem and *Am*-Tra2 proteins. By introducing mutations within this sequence we identified two hexamer repeats (AAAGAG) that are important for the Fem/*Am*-Tra2 protein mediated splice donor activation. These findings suggest that the components which are involved in *dsx* pre-mRNA splicing are conserved between *D. melanogaster* and *A. mellifera*, while the underlying mechanism of producing the sex-specific *dsx* transcripts has evolved.

Literature

Amrein, H., Gorman, M., and Nöthiger, R. (1988). The sex-determining gene *tra-2* of *Drosophila* encodes a putative RNA binding protein. Cell *55*, 1025–1035.

Amrein, H., Hedley, M. L., and Maniatis, T. (1994). The Role of Specific Protein-RNA and Protein-Protein Interactions in Positive and Negative Control of pre-mRNA Splicing by Transformer-2. Cell *76*, 735–746.

Amrein, H., Maniatis, T., and Nöthiger, R. (1990). Alternatively spliced transcripts of the sex-determining gene *tra-2* of *Drosophila* encode functional proteins of different size. The EMBO journal *9*, 3619–3629.

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*. Molecular and cellular biology *16*, 3106–3111.

Asang, C., Hauber, I., and Schaal, H. (2008). Insights into the selective activation of alternatively used splice acceptors by the human immunodeficiency virus type-1 bidirectional splicing enhancer. Nucleic acids research *36*, 1450–1463.

Bandziulis, R. J., Swanson, M. S., and Dreyfuss, G. (1989). RNA-binding proteins as developmental regulators. Genes & Development *3*, 431–437.

Banerjee, H., Rahn, A., Gawande, B., Guth, S., Valcárcel, J., and Singh, R. (2004). The conserved RNA recognition motif 3 of U2 snRNA auxiliary factor (U2AF65) is essential in vivo but dispensable for activity in vitro. RNA *10*, 240–253.

Baralle, D., and Baralle, M. (2005). Splicing in action: assessing disease causing sequence changes. Journal of medical genetics *42*, 737–748.

Bashaw, G. J., and Baker, B. S. (1995). The *msl-2* dosage compensation gene of *Drosophila* encodes a putative DNA-binding protein whose expression is sex specifically regulated by Sex-lethal. Development (Cambridge, England) *121*, 3245–3258.

Bashaw, G. J., and Baker, B. S. (1997). The regulation of the *Drosophila msl-2* gene reveals a function for Sex-lethal in translational control. Cell *89*, 789–798.

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*, 229–239.

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*, 1037–1046.

Berglund, J. a, Abovich, N., and Rosbash, M. (1998). A cooperative interaction between U2AF65 and mBBP/SF1 facilitates branchpoint region recognition. Genes & development *12*, 858–867.

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 : news and reviews in molecular, cellular and developmental biology *26*, 1131–1139.

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*, 419–429.

Beye, M., Härtel, S., Hagen, A., Hasselmann, M., and Omholt, S. W. (2002). Specific developmental gene silencing in the honey bee using a homeobox motif. Insect molecular biology *11*, 527–532.

Boehringer, D., Makarov, E. M., Sander, B., Makarova, O. V., Kastner, B., Lührmann, R., and Stark, H. (2004). Three-dimensional structure of a pre-catalytic human spliceosomal complex B. Nature structural & molecular biology *11*, 463–468.

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*, 739–747.

Bull, J. J. (1983). Evolution of sex determining mechanisms (Benjamin/Cummings Pub. Co., Advanced Book Program (Menlo Park, Calif.)).

Bull, J. J. (1980). Sex Determination in Reptiles. The University of Chicago Press *55*.

Burghardt, G., Hediger, M., Siegenthaler, C., Moser, M., Dübendorfer, A., and Bopp, D. (2005). The *transformer2* gene in *Musca domestica* is required for selecting and maintaining the female pathway of development. Development genes and evolution *215*, 165–176.

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. The EMBO journal *10*, 2577–2582.

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.

Caputi, M., Freund, M., Kammler, S., Asang, C., and Schaal, H. (2004). A Bidirectional SF2/ASF- and SRp40-Dependent Splicing Enhancer Regulates Human Immunodeficiency Virus Type 1 rev, env, vpu and nef Gene Expression. Journal of Virology *78*, 6517–6526.

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*, 1733–1741.

Cho, S., and Wensink, P. C. (1997). DNA binding by the male and female doublesex proteins of *Drosophila melanogaster*. The Journal of biological chemistry *272*, 3185–3189.

Christiansen, A. E., Keisman, E. L., Ahmad, S. M., and Baker, B. S. (2002). Sex comes in from the cold: the integration of sex and pattern. Trends in genetics : TIG*18*, 510–516.

Cline, T. W., and Meyer, B. J. (1996). Vive la différence: males vs females in flies vs worms. Annual review of genetics *30*, 637–702.

Colwill, K., Pawson, T., Andrews, B., Prasad, J., Manley, J. L., Bell, J. C., and Duncan, P. I. (1996). The Clk/Sty protein kinase phosphorylates SR splicing factors and regulates their intranuclear distribution. The EMBO journal *15*, 265–275.

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

Dearden, P. K., Wilson, M. J., Sablan, L., Osborne, P. W., Havler, M., Mcnaughton, E., Kimura, K., Milshina, N. V., Martin, H., Gempe, T., et al. (2006). Patterns of conservation and change in honey bee developmental genes. Genome Research, 1376–1384.

Dreyfuss, G., Swanson, M. S., and Pinol-Roma, S. (1988). Heterogeneous nuclear ribonucleoprotein particles and the pathway of mRNA formation. Trend in biochemical sciences, 86–91.

Du, H., and Rosbash, M. (2002). The U1 snRNP protein U1C recognizes the 5’ splice site in the absence of base pairing. Nature *419*, 1–5.

Erickson, J. W., and Quintero, J. J. (2007). Indirect effects of ploidy suggest X chromosome dose, not the X:A ratio, signals sex in *Drosophila*. PLoS biology *5*, e332.

Fairbrother, W. G., Yeh, R.-F., Sharp, P. a, and Burge, C. B. (2002). Predictive identification of exonic splicing enhancers in human genes. Science (New York, N.Y.) *297*, 1007–1013.

Freund, M., Asang, C., S, K., C, K., Krummheuer, J., Hipp, M., Meyer, I., Gierling, W., Theiss, S., Preuss, T., et al. (2003). A novel approach to describe a U1 snRNA binding site. Nucleic Acids Research *31*, 6963–6975.

Freund, M., Hicks, M. J., Konermann, C., Otte, M., Hertel, K. J., and Schaal, H. (2005). Extended base pair complementarity between U1 snRNA and the 5’ splice site does not inhibit splicing in higher eukaryotes, but rather increases 5' splice site recognition. Nucleic acids research *33*, 5112–5119.

Fujihara, T., Kawabe, M., and Oishi, K. (1978). A sex-transformation gene in *Drosophila*. Journal of Heredity *69*, 229–236.

Funaguma, S., Suzuki, M. G., Tamura, T., and Shimada, T. (2005). The *Bmdsx* transgene including trimmed introns is sex-specifically spliced in tissues of the silkworm, *Bombyx mori*. Journal of insect science (Online) *5*, 17.

Ge, H., and Manley, J. L. (1990). A protein factor, ASF, controls cell-specific alternative splicing of SV40 early pre-mRNA in vitro. Cell *62*, 25–34.

Gebauer, F., Corona, D. F., Preiss, T., Becker, P. B., and Hentze, M. W. (1999). Translational control of dosage compensation in *Drosophila* by Sex-lethal: cooperative silencing via the 5’ and 3' UTRs of *msl-2* mRNA is independent of the poly(A) tail. The EMBO journal *18*, 6146–6154.

Gempe, T., Hasselmann, M., Schiøtt, M., Hause, G., Otte, M., and Beye, M. (2009). Sex determination in honeybees: two separate mechanisms induce and maintain the female pathway. PLoS biology *7*, e1000222.

Gempe, T., and Beye, M. (2011). Function and evolution of sex determination mechanisms, genes and pathways in insects. BioEssays : news and reviews in molecular, cellular and developmental biology *33*, 52–60.

Gilfillan, G. D., Dahlsveen, I. K., and Becker, P. B. (2004). Lifting a chromosome: dosage compensation in *Drosophila melanogaster*. FEBS letters *567*, 8–14.

Goralski, T. J., Edstrom, J.-E., and Baker, B. S. (1989). The sex determination locus *transformer-2* of *Drosophila* encodes a polypeptide with similarity to RNA binding proteins. Cell *56*, 1011–1018.

Graveley, B. R. (2000). Sorting out the complexity of SR protein functions. RNA (New York, N.Y.) *6*, 1197–1211.

Graveley, B. R., Hertel, K. J., and Maniatis, T. (1998). A systematic analysis of the factors that determine the strength of pre-mRNA splicing enhancers. The EMBO journal *17*, 6747–6756.

Graveley, B. R., Hertel, K. J., and Maniatis, T. (2001). The role of U2AF35 and U2AF65 in enhancer-dependent splicing. RNA (New York, N.Y.) *7*, 806–818.

Graves, J. A. M. (1995). The origin and function of the mammalian Y chromosome and the Y-borne genes- an evolving understanding. BioEssays *17*, 311–321.

Gui, J.-F., Lane, W. S., and Fu, X.-dong (1994). A Serine Kinase Regulates Intracellular Localization Of Splicing Factors in the Cell Cycle. Nature *369*, 678–682.

Hashimoto, H. (1933). The role of the W-chromosome in the sex determination of *Bombyx mori*. Jap. J. Genet. *8*, 245–258.

Hasselmann, M., Gempe, T., Schiøtt, 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*, 519–522.

Hasselmann, M., and Beye, M. (2004). Signatures of selection among sex-determining alleles of the honey bee. Proceedings of the National Academy of Sciences of the United States of America *101*, 4888–4893.

Hazelrigg, T., and Tu, C. (1994). Sex-specific processing of the *Drosophila exuperantia* transcript is regulated in male germ cells by the tra-2 gene. Proceedings of the National Academy of Sciences of the United States of America *91*, 10752–10756.

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

Hediger, M., Henggeler, C., Meier, N., Perez, R., Saccone, G., and Bopp, D. (2010). Molecular characterization of the key switch F provides a basis for understanding the rapid divergence of the sex-determining pathway in the housefly. Genetics *184*, 155–170.

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*, 579–586.

Heinrichs, V., Ryner, L. C., and Baker, B. S. (1998). Regulation of sex-specific selection of fruitless 5' splice sites by transformer and transformer-2. Molecular and cellular biology *18*, 450–458.

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 (New York, N.Y.) *252*, 833–836.

Inoue, K., Hoshijima, K., Higuchi, I., Sakamoto, H., and Shimura, Y. (1992). Binding of the *Drosophila* transformer and transformer-2 proteins to the regulatory elements of *doublesex* primary transcript for sex-specific RNA processing. Proceedings of the National Academy of Sciences of the United States of America *89*, 8092–8096.

Kamakura, M. (2011). Royalactin induces queen differentiation in honeybees. Nature *473*, 478–483.

Kammler, S., Leurs, C., Freund, M., Krummheuer, J., Seidel, K., Tange, T. Ø., Lund, M. K., and Kjems, J. (2001). The sequence complementarity between HIV-1 5' splice site SD4 and U1 snRNA determines the steady-state level of an unstable env pre-mRNA. RNA *7*, 421–434.

Kato, Y., Kobayashi, K., Watanabe, H., and Iguchi, T. (2011). Environmental sex determination in the branchiopod crustacean *Daphnia magna*: deep conservation of a *Doublesex* gene in the sex-determining pathway. PLoS genetics *7*, e1001345.

Keisman, E. L., Christiansen, a E., and Baker, B. S. (2001). The sex determination gene *doublesex* regulates the A/P organizer to direct sex-specific patterns of growth in the *Drosophila* genital imaginal disc. Developmental cell *1*, 215–225.

Keisman, E. L., and Baker, B. S. (2001). The *Drosophila* sex determination hierarchy modulates wingless and decapentaplegic signaling to deploy dachshund sex-specifically in the genital imaginal disc. Development (Cambridge, England) *128*, 1643–1656.

Kelley, R. L., Wang, J., Bell, L., and I., K. M. (1997). Sex lethal controls dosage compensation in *Drosophila* by a non-splicing mechanism. Nature *387*, 195–199.

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*, 933–943.

Kopp, a, Duncan, I., Godt, D., and Carroll, S. B. (2000). Genetic control and evolution of sexually dimorphic characters in *Drosophila*. Nature *408*, 553–559.

Krainer, a R., Conway, G. C., and Kozak, D. (1990). Purification and characterization of pre-mRNA splicing factor SF2 from HeLa cells. Genes & Development *4*, 1158–1171.

Krämer, A. (1996). The Structure and Function of Proteins Involved in Mammalian pre-mRNA Splicing. Annu. Rev. Biochem. *65*, 367–409.

Kucharski, R., Maleszka, J., Foret, S., and Maleszka, R. (2008). Nutritional control of reproductive status in honeybees via DNA methylation. Science (New York, N.Y.) *319*, 1827–1830.

Kuo, B. A., and Norton, P. A. (1999). Accurate selection of a 5' splice site requires sequences within fibronectin alternative exon B. Nucleic acids research *27*, 3945–3952.

Lagos, D., Ruiz, M. F., Sánchez, 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*, 111–121.

Lamm, G. M., and Lamond, A. I. (1993). Non-snRNP protein splicing factors. Biochimica et biophysica acta *1173*, 247–265.

Lamond, A. I., Konarska, M. M., Grabowski, P. J., and Sharp, P. A. (1988). Spliceosome assembly involves the binding and release of U4 small nuclear ribonucleoprotein. Proceedings of the National Academy of Sciences of the United States of America *85*, 411–415.

Lynch, K. W., and Maniatis, T. (1996). Assembly of specific SR protein complexes on distinct regulatory elements of the *Drosophila doublesex* splicing enhancer. Genes & Development *10*, 2089–2101.

Lynch, K. W., and Maniatis, T. (1995). Synergistic interactions between two distinct elements of a regulated splicing enhancer. Genes & Development *9*, 284–293.

Madhani, H. D., and Guthrie, C. (1992). A novel base-pairing interaction between U2 and U6 snRNAs suggests a mechanism for the catalytic activation of the spliceosome. Cell *71*, 803–817.

Madigan, S. J., Edeen, P., Esnayra, J., and McKeown, M. (1996). *att*, a target for regulation by tra2 in the testes of *Drosophila melanogaster*, encodes alternative RNAs and alternative proteins. Molecular and cellular biology *16*, 4222–4230.

Maniatis, T., and Tasic, B. (2002). Alternative pre-mRNA splicing and proteome expansion in metazoans. Nature *418*, 236–243.

Matson, C. K., Murphy, M. W., Sarver, A. L., Griswold, M. D., Bardwell, V. J., and Zarkower, D. (2011). DMRT1 prevents female reprogramming in the postnatal mammalian testis. Nature *476*, 101–104.

Mattox, W., McGuffin, M. E., and Baker, B. S. (1996). A negative feedback mechanism revealed by functional analysis of the alternative isoforms of the *Drosophila* splicing regulator transformer-2. Genetics *143*, 303–314.

Mattox, W., Palmer, M. J., and Baker, B. S. (1990). Alternative splicing of the sex determination gene *transformer-2* is sex-specific in the germ line but not in the soma. Genes & Development *4*, 789–805.

Mattox, W., and Baker, B. S. (1991). Autoregulation of the splicing of transcripts from the *transformer-2* gene of *Drosophila*. Genes & Development *5*, 786–796.

McGuffin, M. E., Chandler, D., Somaiya, D., Dauwalder, B., and Mattox, W. (1998). Autoregulation of *transformer-2* alternative splicing is necessary for normal male fertility in *Drosophila*. Genetics *149*, 1477–1486.

McManus, C. J., Schwartz, M. L., Butcher, S. E., and Brow, D. A. (2007). A dynamic bulge in the U6 RNA internal stem – loop functions in spliceosome assembly and activation. RNA *13*, 2252–2265.

Merrill, B. M., Stone, K. L., Cobianchi, F., Wilson, S. H., and Williams, K. R. (1988). Phenylalanines that are conserved among several RNA-binding proteins form part of a nucleic acid-binding pocket in the A1 heterogeneous nuclear ribonucleoprotein. The Journal of biological chemistry *263*, 3307–3313.

Michaud, S., and Reed, R. (1993). A functional association between the 5’ and 3' splice site is established in the earliest prespliceosome complex (E) in mammals. Genes & Development *7*, 1008–1020.

Michaud, S., and Reed, R. (1991). An ATP-independent complex commits pre-mRNA to the mammalian spliceosome assembly pathway. Genes & Development *5*, 2534–2546.

Mount, S. M. (1982). A catalogue of splice junction sequences. Nucleic Acids Research *10*, 459–472.

Mount, S. M., Pettersson, I., Hinterberger, M., Karmas, A., and Steitz, J. a (1983). The U1 small nuclear RNA-protein complex selectively binds a 5’ splice site in vitro. Cell *33*, 509–518.

Nagai, K., Oubridge, C., Jessen, T. H., Li, J., and Evans, P. R. (1990). Crystal structure of the RNA-binding domain of the U1 small nuclear ribonucleoprotein A. Nature *348*, 515–520.

Nagoshi, R. N., and Baker, B. S. (1990). Regulation of sex-specific RNA splicing at the *Drosophila doublesex* gene: *cis*-acting mutations in exon sequences alter sex-specific RNA splicing patterns. Genes & Development *4*, 89–97.

Newman, A. J. (1997). The role of U5 snRNP in pre-mRNA splicing. The EMBO journal *16*, 5797–5800.

Nissen, I., Müller, M., and Beye, M. The *Am-tra2* gene is an essential cofactor of female splice regulation at two levels of the complementary sex determination hierarchy of the honeybee. Genetics, under review.

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 Biochimica et Biophysica Sinica *37*, 728–736.

Norton, P. A. (1994). Polypyrimidine tract sequences direct selection of alternative branch sites and influence protein binding. Nucleic Acids Research *22*, 3854–3860. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=308380&tool=pmcentrez&rendertype=abstract.

Nygaard, S., Zhang, G., Schiøtt, M., Li, C., Wurm, Y., Hu, H., Zhou, J., Ji, L., Qiu, F., Rasmussen, M., et al. (2011). The genome of the leaf-cutting ant *Acromyrmex echinatior* suggests key adaptations to advanced social life and fungus farming. Genome Research *21*, 1339–1348.

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., 145–158.

Oliveira, D. C. S. G., Werren, J. H., Verhulst, E. C., Giebel, J. D., Kamping, A., Beukeboom, L. W., and Zande, L. V. D. (2009). Identification and characterization of the *doublesex* gene of *Nasonia*. Insect molecular biology *18*, 315–324.

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 (Cambridge, England) *129*, 3715–3725.

Parkhurst, S. M., and Ish-Horowicz, D. (1992). Common denominators for sex. Curr. Biol. *2*, 629–631.

Qi, J., Su, S., and Mattox, W. (2007). The *doublesex* splicing enhancer components Tra2 and Rbp1 also repress splicing through an intronic silencer. Molecular and cellular biology *27*, 699–708.

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*, 691–695.

Ruiz, M. F., Eirín-López, J. M., Stefani, R. N., Perondini, A. L. P., Selivon, D., and Sánchez, L. (2007). The gene *doublesex* of *Anastrepha* fruit flies (Diptera, Tephritidae) and its evolution in insects. Development genes and evolution *217*, 725–731.

Ruiz, M. F., Stefani, R. N., Mascarenhas, R. O., Perondini, A. L. P., Selivon, D., and Sánchez, L. (2005). The gene *doublesex* of the fruit fly *Anastrepha obliqua* (Diptera, Tephritidae). Genetics *171*, 849–854.

Ryan, D. E., and Abelson, J. (2002). The conserved central domain of yeast U6 snRNA: importance of U2-U6 helix Ia in spliceosome assembly. RNA (New York, N.Y.) *8*, 997–1010.

Ryner, L. C., Goodwin, S. F., Castrillon, D. H., Anand, A., Villella, A., Baker, B. S., Hall, J. C., Taylor, B. J., and Wasserman, S. a (1996). Control of male sexual behavior and sexual orientation in *Drosophila* by the *fruitless* gene. Cell *87*, 1079–1089.

Saccone, G., Salvemini, M., Pane, A., and Polito, L. C. (2008). Masculinization of XX *Drosophila* transgenic flies expressing the *Ceratitis capitata* DoublesexM isoform. The International journal of developmental biology *52*, 1051–1057.

Salvemini, M., Robertson, M., Aronson, B., Atkinson, P., Polito, L. C., and Saccone, G. (2009). *Ceratitis capitata transformer-2* gene is required to establish and maintain the autoregulation of *Cctra*, the master gene for female sex determination. The International journal of developmental biology *53*, 109–120.

Sarno, F., Ruiz, M. F., Eirín-López, J. M., Perondini, A. L. P., Selivon, D., and Sánchez, L. (2010). The gene *transformer-2* of *Anastrepha* fruit flies (Diptera, Tephritidae) and its evolution in insects. BMC evolutionary biology *10*, 140.

Savard, J., Tautz, D., Richards, S., Weinstock, G. M., Gibbs, R. A., Werren, J. H., Tettelin, H., and Lercher, M. J. (2006). Phylogenomic analysis reveals bees and wasps (Hymenoptera) at the base of the radiation of Holometabolous insects. Genome Research *16*, 1334–1338.

Schartl, M. (2004). Sex chromosome evolution in non-mammalian vertebrates. Current opinion in genetics & development *14*, 634–641.

Schütt, C., and Nöthiger, R. (2000). Structure, function and evolution of sex-determining systems in Dipteran insects. Development (Cambridge, England) *127*, 667–677.

Sciabica, K. S., and Hertel, K. J. (2006). The splicing regulators Tra and Tra2 are unusually potent activators of pre-mRNA splicing. Nucleic acids research *34*, 6612–6620.

Selden, R. F., Howie, K. B., Rowe, M. E., Goodman, H. M., and Moore, D. D. (1986). *Human growth hormone* as a reporter gene in regulation studies employing transient gene expression. Molecular and cellular biology *6*, 3173–3179.

Selenko, P., Gregorovic, G., Sprangers, R., Stier, G., Rhani, Z., Krämer, A., and Sattler, M. (2003). Structural basis for the molecular recognition between human splicing factors U2AF65 and SF1/mBBP. Molecular cell *11*, 965–976.

Shen, H., and Green, M. R. (2006). RS domains contact splicing signals and promote splicing by a common mechanism in yeast through humans. Genes & Development *20*, 1755–1765.

Singh, K. K., Erkelenz, S., Rattay, S., Dehof, A. K., Hildebrandt, A., Schulze-osthoff, K., Schaal, H., and Schwerk, C. (2010). Human SAP18 mediates assembly of a splicing regulatory multiprotein complex via its ubiquitin-like fold. *16*, 2442–2454.

Smith, C. W., and Valcárcel, J. (2000). Alternative pre-mRNA splicing: the logic of combinatorial control. Trends in biochemical sciences *25*, 381–388.

Sosnowski, B. A., Belote, J. M., and McKeown, M. (1989). Sex-Specific Alternative Splicing of RNA from the *transformer* Gene Results from Sequence-Dependent Splice Site Blockage. Cell *58*, 449–459.

Staley, J. P., and Guthrie, C. (1999). An RNA switch at the 5' splice site requires ATP and the DEAD box protein Prp28p. Molecular cell *3*, 55–64.

Staley, J. P., and Guthrie, C. (1998). Mechanical devices of the spliceosome: motors, clocks, springs, and things. Cell *92*, 315–326.

Suzuki, M. G. (2010). Sex determination: insights from the silkworm. Journal of genetics *89*, 357–363.

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*. Development genes and evolution *213*, 345–354.

Suzuki, M. G., Imanishi, S., Dohmae, N., Asanuma, M., and Matsumoto, S. (2010). Identification of a male-specific RNA binding protein that regulates sex-specific splicing of *Bmdsx* by increasing RNA binding activity of *Bm*PSI. Molecular and cellular biology *30*, 5776–5786.

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. Molecular and cellular biology *28*, 333–343.

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 biochemistry and molecular biology *31*, 1201–1211.

Sánchez, L., Gorfinkiel, N., and Guerrero, I. (2001). Sex determination genes control the development of the *Drosophila* genital disc, modulating the response to Hedgehog,

Wingless and Decapentaplegic signals. Development (Cambridge, England) *128*, 1033–1043.

Sánchez, L., and Guerrero, I. (2001). The development of the *Drosophila* genetical disc. BioEssays *23*, 698–707.

The Tribolium Genome Sequencing Consortium (2008). The genome of the model beetle and pest *Tribolium castaneum*. Nature *452*, 949–955.

Tian, M., and Maniatis, T. (1993). A Splicing Enhancer Complex Controls Alternative Splicing of *doublesex* Pre-mRNA. Cell *74*, 105–114.

Tian, M., and Maniatis, T. (1994). A splicing enhancer exhibits both constitutive and regulated activities. Genes & Development *8*, 1703–1712.

Traut, W., and Marec, F. (1996). Sex Chromatin In Lepidoptera. The Quarterly review of biology *71*, 239–256.

Valcárcel, 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*, 171–175.

Vincent, S., Perkins, L. A., and Perrimon, N. (2001). Doublesex Surprises. Cell *106*, 399–402.

Wang, H. Y., Lin, W., Dyck, J. a, Yeakley, J. M., Songyang, Z., Cantley, L. C., and Fu, X. D. (1998). SRPK2: a differentially expressed SR protein-specific kinase involved in mediating the interaction and localization of pre-mRNA splicing factors in mammalian cells. The Journal of cell biology *140*, 737–750.

Watanabe, T. K. (1975). A New Sex-Transforming Gene on the Second Chromosome of *Drosophila melanogaster*. Jap. J. Genet. *50*, 269–271.

Whiting, P. W. (1943). Multiple Alleles in Complementary Sex Determination of Habrobracon. Genetics *28*, 365–382.

Will, C. L., and Lührmann, R. (1997). Protein functions in pre-mRNA splicing. Current opinion in cell biology *9*, 320–328.

Williams, T. M., Selegue, J. E., Werner, T., Gompel, N., Kopp, A., and Carroll, S. B. (2008). The regulation and evolution of a genetic switch controlling sexually dimorphic traits in *Drosophila*. Cell *134*, 610–623.

Xiao, S. H., and Manley, J. L. (1998). Phosphorylation-dephosphorylation differentially affects activities of splicing factor ASF/SF2. The EMBO journal *17*, 6359–6367.

Yeakley, J. M., Tronchère, H., Olesen, J., Dyck, J. a, Wang, H. Y., and Fu, X. D. (1999). Phosphorylation regulates in vivo interaction and molecular targeting of serine/arginine-rich pre-mRNA splicing factors. The Journal of cell biology *145*, 447–455.

Zhou, Z., Licklider, L. J., Gygi, S. P., and Reed, R. (2002). Comprehensive proteomic analysis of the human spliceosome. Nature *419*, 182–185.

Zhuang, Y., and Weiner, a M. (1986). A compensatory base change in U1 snRNA suppresses a 5' splice site mutation. Cell *46*, 827–835.

Zuo, P., and Maniatis, T. (1996). The splicing factor U2AF35 mediates critical protein-protein interactions in constitutive and enhancer-dependent splicing. Genes & Development *10*, 1356–1368.

Danksagung

Ich danke ganz herzlich…

…Prof. Dr. Martin Beye für die Betreuung dieser Arbeit, die vielen hilfreichen Diskussionen und die Überlassung des Themas

…Prof. Dr. Laura Rose für die Übernahme des Korreferats

…Eva Theilenberg und Marion Müller-Borg für ihre unermüdliche Unterstützung im Labor

…Christina Schulte, Dr. Tanja Gempe, Björn Schmitt, Steve Przybilla und Jens Nissen fürs Korrekturlesen

…Dr. Marianne Otte für die Überlassung des einen oder anderen Vektors und nützliche Denkanstöße

…der ganzen Bienengruppe und allen ehemaligen Bienenleuten für die angenehme Arbeitsatmosphäre, die hilfreichen Diskussionen und die ständige Hilfsbereitschaft

…Prof. Dr. Heiner Schaal für seine Diskussionsbereitschaft und seine Ideen zu den Spleißexperimenten

…Steffen Erkelenz dafür, dass ich seine Vektoren verwenden durfte

…den Bienchen für das ja doch meist brave Eierlegen

…und meiner Familie und meinen Freunden für die Unterstützung, die Motivationshilfe und die Ablenkung wenn sie nötig war.