The dice of fate: the *csd* gene and how its allelic composition regulates sexual development in the honey bee, *Apis mellifera*

Martin Beye

Summary

Perhaps 20% of known animal species are haplodiploid: unfertilized haploid eggs develop into males and fertilized diploid eggs into females. Sex determination in such haplodiploid species does not rely on a difference in heteromorphic sex chromosome composition but the genetic basis has been elucidated in some hymenopteran insects (wasps, sawflies, ants, bees). In these species, the development into one sex or the others depends on an initial signal whether there is only one allele or two different alleles of a single gene, the *complementary sex determiner* (*csd*), in the zygotic genome. The gene has been most-recently identified in the honey bee and has been found to encode an arginine serine-rich (SR) type protein. Heterozygosity generates an active protein that initiates female development while hemizygosity/homozygosity results in a non-active CSD protein and default male development. I will discuss plausible models of how the molecular decision of male and female is made and implemented. Comparison to hierarchies of dipteran insects suggests that SR-type protein has facilitated the differentiation of sex-determining systems and hierarchies. *BioEssays* 26:1131–1139, 2004.

Introduction

Sex determination is a fundamental phenomenon of life. The pronounced differences in morphology, physiology and behavior that characterize male and female development have fascinated scientists for centuries. Mechanisms that underlie sex determination appear to share some general features in all species. In particular, the primary decision involves a cascade of genes that are required to produce the distinct male and female phenotypes. This choice between two alternative pathways, that of male and female development, is a paradigm for developmental decision making.

Martin-Luther-Universität Halle/Wittenberg, Biozentrum, Institut für Zoologie, Weinberg Weg 22, 06120 Halle, Germany.
E-mail: beye@zoologie.uni-halle.de
Funding agencies: Deutsche Forschungsgemeinschaft (2194/3, 2194/5) and the Kultusminister of Sachsen-Anhalt (3266A/0020L).
DOI 10.1002/bies.20098
Published online in Wiley InterScience (www.interscience.wiley.com).

Although sex determination is a very general principal, the initial signal varies considerably between species, suggesting that at least some of the underlying genetic components differ and are not conserved throughout evolution. Initial signals have been dissected to their molecular and genetic components in animals that have sex chromosomes, including the fruitfly *Drosophila melanogaster* and the nematode *Caenorhabditis elegans*, which both use the X:A ratio and a polygenic signal (genic balance),[1] and many mammalian species that use a single dominant male factor on the Y chromosome, *Sry*.[2]

Given the detailed molecular understanding of sexual regulation in these model organisms is there any serious motivation to study sex determination in the honey bee? In fact research on sex determination in bees has a long tradition that started more than 150 years ago and has continued to be a puzzle in genetic terms for more than 100 years. The difference in sexual regulation between the honey bee and organisms that have sex chromosomes is quite obvious: the sexual ratio between males and females in sex chromosomal systems is about 50% while, in the honey bee, only a few hundred males exist at specific times of the year, with the majority of individuals being females either ten thousands of sterile worker bees or a single reproductive queen. The sexual fate of an egg in bees depends on whether the queen fertilizes an egg, using the sperm stored in her spermatheca, while the caste differences between females depends on the food (royal jelly) that they receive during larval development. This haplodiploidy in which unfertilized haploid eggs develop into males and fertilized diploid eggs into females is actually widespread in the animal kingdom.

Whiting[3] presented evidence in a wasp ~60 years ago that sex is not governed by the fertilization process itself, but is regulated by a single genetic locus that appears in different versions or specificities. While I was still a student in biology, I decided that I wanted to learn more about the underlying genetic principal, the molecular nature and function of the so-called complementary sex determination in the honey bee. Part of my motivation was that I had kept honey bees since my childhood a hobby that was initiated by my grandfather. When I became a student in molecular genetics and biology, I wanted to apply the molecular techniques and biological concepts that

I had learned directly to the honey bee system. I didn't realize, however, how much effort it would be to identify the responsible gene in a genetic non-model organism.

As part of my graduate studies and subsequent post-doc work, we identified the relevant genomic region and mapped it to the chromosome,[4,5] but were far from identifying the gene. Greg Hunt's and Robert Page's laboratories used similar approaches.[6,7] Combining our different expertises, we were able to more precisely map the sex-determining region.[8] The final step of isolation required several years of work to develop and improve methodologies and technologies for the honey bee system. Together with my graduate student Martin Hasselmann, and with collaborators in the US (Robert Page's lab) and Norway (Stig Omholt's lab), we finally succeeded in identifying the gene responsible.[9]

In this article, I will review the progress that we have made over recent years in understanding the molecular basis of the complementary initial signal. I will shed some light on the historical aspects of haplodiploid sex determination that spans 150 years of research,[10] from the discovery that unfertilised eggs develop into males[11] to the isolation of the underlying genetic element of complementary sex determination.[9] I will discuss the current state of ideas of how the initial signal is implemented in different allelic combinations and of how this may relate to the activation of a downstream pathway. Finally, I will discuss how our findings relate to what is know about the sex-determining cascade of *D. melanogaster* and the available data from other dipteran insects that should help to elucidate how regulatory hierarchies have evolved over the last 270 million years.

Complementary sex determination: a history of 150 years of research

The mechanism of sex determination has been debated since at least the time of Aristotles. Even up till the early 19th century, it was still widely accepted that sexual fate is regulated by external conditions, probably because no alternative genetic concept existed. In 1845, however, Johann Dzierzon, a parish priest from the former Prussian province of Silesia, described how male honey bees develop from unfertilized eggs, which was the first rigorous report of a sex-determining mechanism.[11] His finding was based on the observation that a non-mated queen of the honey bee can only produce males. Later, cytological studies in the honey bee showed that designated male eggs are not fertilized[12] and are genetically haploid while eggs that develop into females are diploid.[13] This seemingly odd mechanism of male and female development is actually quite widespread and has independently evolved ~15 times in insects and mites alone[14] and occurs in perhaps 20% of all known animal species.[15] Haplodiploidy is found in different groups of insects (Hymenoptera, thrips, some bark beetles, some scale insects, white flies), the ticks/mites, and a group of rotifera.[16]

The first description of sex-specific chromosomes in grasshoppers and other insects[17,18] was followed by the elucidation of the precise nature of the chromosomal basis of sex determination in *D. melanogaster*.[19] Sex chromosomes, however, cannot provide the explanation for haplodiploid sex determination. The haploid, uniparental eggs receive a random half of the maternal diploid genome and develop into males irrespective of the chromosomal composition. This problem was most apparent to Bridges after his seminal publication of the genic balance hypothesis[19] in which he proposed that sex in *Drosophila* resulted from a balance of male- and female-determining genes on the X chromosome and the autosomes. Bridges proclaimed, "To me sex determination in the bee is the outstanding unsolved puzzle . . .. "[20]

The underlying genetic basis of haplodiploid sex determination in the insect order Hymenoptera was elucidated by the discovery of diploid males. Inbreeding crosses in the wasp *Bracon hepetor* resulted in 50% fertilized eggs becoming diploid male offspring with biparental origin.[21] This finding suggested that the process of fertilization is not itself the initial signal of sexual regulation. Subsequently, several allelic specificities were found to segregate in populations and led to the hypothesis of complementary sex determination.[3] Individuals that are heterozygous at the sex-determining locus develop into females, while individuals that are haploid (hemizygous) are males (Fig. 1A). Diploid males occur if the diploid fertilized eggs carry the same alleles and are homozygous (Fig. 1B). However, the diploid males (if fertile) produce triploid sterile offspring, because spermatogenesis is mitotic. In the case of the honey bee, these are eaten by worker bees shortly after they hatch from the egg. For the honey bee, up to 19 sex-determining alleles have been estimated to segregate in natural populations, based on the number of diploid males found in a population.[22] Males are haploid throughout the insect order Hymenoptera (sawflies, ants, bees, wasps), which includes over 200,000 species.[23] Diploid males have been documented in different species throughout the hymenopteran phylogeny, which is consistent with the view that these species follow the complementary mode of sexual regulation (superfamilies with diploid males are shown in bold in Fig. 2). No diploid males have been documented in several parasitic living species with a mating system that promotes extreme inbreeding (i.e. brother, sister mating are usual). This holds, for example, for the parasitic wasps of the superfamily Chrysidoidea, Chalcidoidea and Cynipoidea (underlined superfamilies in Fig. 2). Species that belong to some other eleven superfamilies, however, have not so far been tested for the mode of complementary sex determination (Fig. 2). Complementary sex determination appears to reside in the major and ancient branches of hymenopteran phylogeny, the Symphyta/Apocrita and Aculeata/Parasitica and most likely has not evolved independently in these branches (Fig. 2). Thus, complementary sex determination may be the ancestral

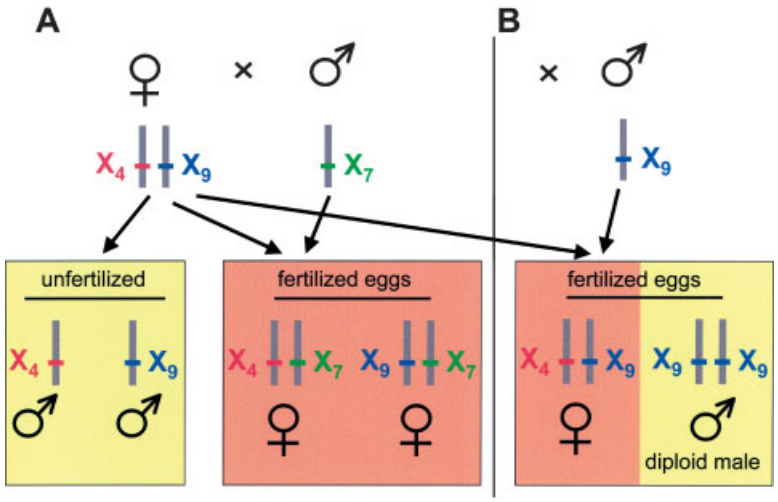

Figure 1. Genotypes and sexual fate under the system of complementary sex determination found in many hymenopteran species (ants, bees, wasps, sawflies). **A:** Males derive from unfertilized eggs and have only one allele (e.g. sex-determining allele X_4 or X_9). Fertilized eggs develop into females that are diploid and carry two different sex-determining alleles (heterozygous). **B:** Diploid males arise when the eggs have the same sex-determining alleles (homozygous). This occurs for example under inbreeding conditions in which the father has an allele in common with the mother.

mode of reproduction throughout the hymenopteran insects.[23] In species in which the parasitic life cycle promotes inbreeding, another mechanism has possibly evolved to overcome the serious costs of diploid male production. Experiments carried out in the parasitic wasp *Nasonia*[24] are consistent with a different model of sexual regulation, the paternal imprinting mechanism.[25]

Compared to the beautifully worked out cascade of sex determination of *D. melanogaster*,[26] it is surprising how little is known about the molecular genetic basis of haplodiploid sex determination. Complementary sex determination is in fact an attractive system to study because the initial signal is restricted to one locus instead of the more complex polygenic signaling of the X:A ratio in the fruitfly. Moreover, the molecular decision making on the basis of complementary alleles is a fascinating mechanism to study. The lack of sophisticated molecular genetic tools, techniques, resources and the need for a hymenopteran genetic model system can partly explain this obvious gap in information.

Positional cloning identifies the *complementary sex determiner*

Because the biochemical nature of the initial signal of complementary sex determination was totally unknown, we used a positional cloning and fine-scale mapping approach[27] to isolate and identify the genomic region of the *complementary sex determiner* (*csd*).[9] The natural occurring trait of female and diploid males in an inbred cross (Fig. 1B) were used to

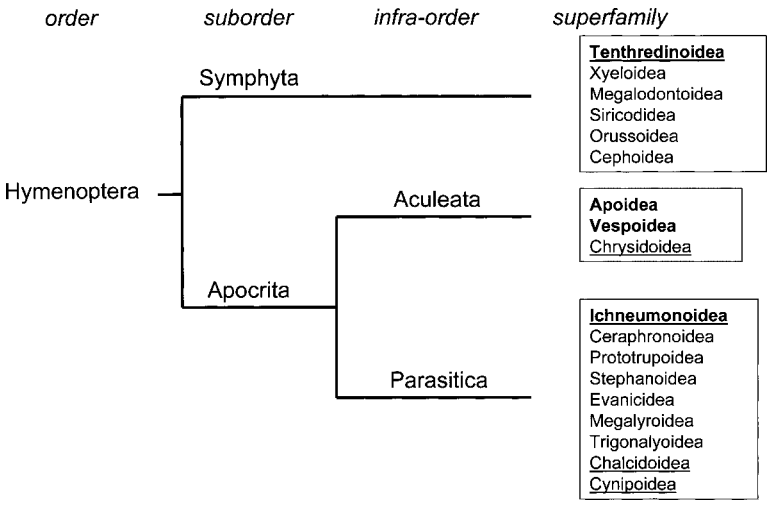

Figure 2. Phylogenetic view of sex determination in the insect order Hymenoptera.[23,56] Superfamilies in which species with diploid males have been documented are shown in bold. The occurrence of diploid males is consistent with the view that sex is regulated by the complementary mode of sex determination in these species. Superfamilies in which species were found that show no diploid male production are underlined. The lack of diploid males in these species gives evidence of a non-complementary mode of sex determination. These species have a parasitic life cycle that promotes inbreeding and thus have possibly evolved a different mode of sexual regulation more recently to overcome the serious costs of diploid male production.[23,25]

establish markers that were co-segregating with diploid male and female development. Two sex-linked markers were identified from two different groups, the Q marker from Hunt and Page[6] and the Z maker from Beye and co-workers.[4,5] The two laboratories joined forces and found that the two markers were flanking the sex-determining region at a distance of less than 360 kb.[8] The finding that 1 cM corresponds to about 50 kb in the honey bee facilitated the subsequent fine-scale mapping (by comparison, 1 cM corresponds to 600 kb in *D. melanogaster*). The Q marker with a genetic distance of 1 cM[6] was the starting point to initiate the chromosomal walk. Cloning the sex-determining genomic region, however, was the major obstacle: the corresponding genomic sequences were not cloned in several different genomic libraries that were constructed by various cloning vectors and approaches.[28] Finally, the genomic region was cloned using a partial genomic library that was enriched for the sex locus genomic fragment. This fragment had been previously identified by physical mapping of the genetic markers with pulsed field gel electrophoresis (PFGE).[8] 70 kb continuous DNA was subsequently isolated by overlapping clones. Consistent with a single locus complementary model, a 12 kb region between two genetic markers (Fig. 3A, genetic marker 1 and 2) was identified that was always heterozygous in females. The estimated precision of genetic mapping in the designated region was 5 kb per 1000 meiosis analyzed, suggesting that at least part of the sex-determining gene would be found in this region.

Exon prediction algorithms and subsequent analysis of transcripts identified a single gene in the sequence of the designated region, which we named the *complementary sex determiner* (*csd*) (Fig. 3A). The gene consists of nine exons and spans a genomic region of about 9 kb. *csd* transcripts are the same in male and female embryos, and no sequence or splicing differences were found between the sexes. Consistent with different allelic specificities, a variety of single nucleotide differences and insertions and deletions were found between alleles in the predicted gene coding and transcript sequences affecting the predicted amino acid sequence. Analysis of the open reading frame predicts a protein belonging to a broad family of proteins that are characterized by regions of reiterated arginine (R) and serine (S) amino acids that constitute the so-called RS domain.[29,30] Besides the RS domain, the protein also has a proline-rich region at its C terminus (Fig. 3B). Between these domains is a hypervariable region that differs

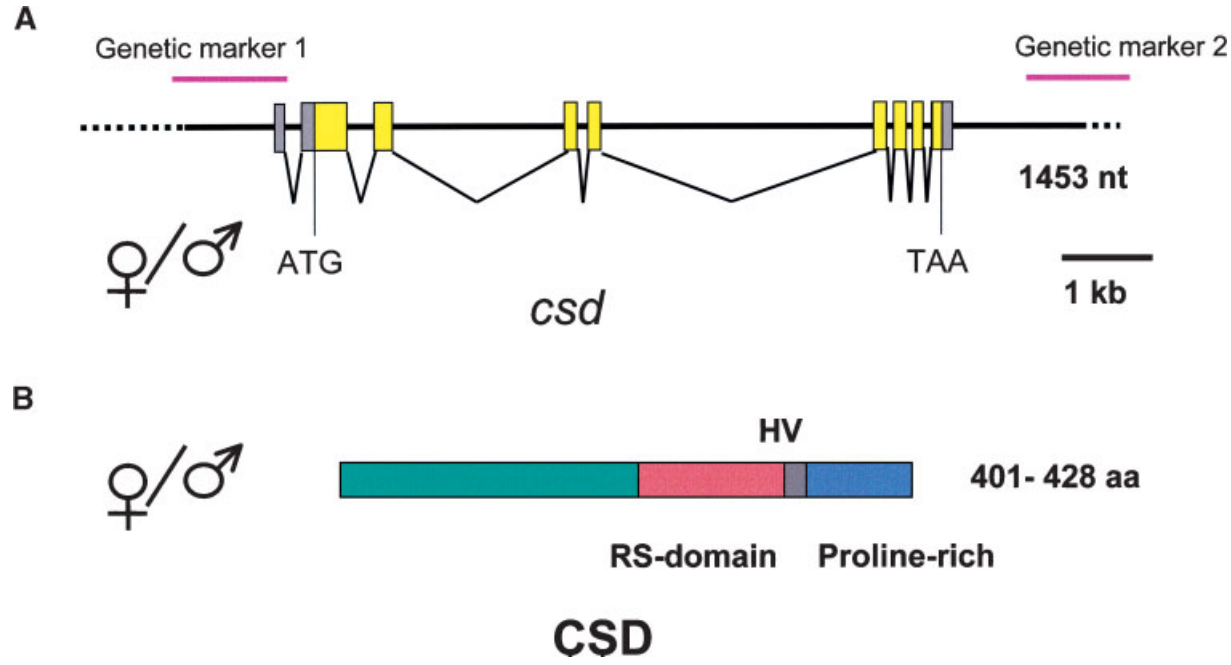

Figure 3. Schematic presentation of the genomic organization and proposed protein structure of *csd*. **A:** Genomic region encompassing *csd* and the two closely linked genetic markers (marker 1 and 2). The genetic markers 1 and 2 are always heterozygous in females; this identifies the sex-determining locus. Exons are indicated by boxes and the deduced ORF is marked in yellow. The estimate of the transcript is indicated in nucleotides (nt), which represents just one allelic variant. No sex-specific differences were found among the transcripts. Predicted translational start and stop sites are indicated. **B:** The schematic presentation of the predicted domain structure of *csd*. The region rich in arginine/serine (RS domain) is marked in red, the hypervariable region that has different number of repeats in various alleles is shown in grey (HV) and the proline-rich region is marked in blue. The number of amino acids of the CSD protein varies between alleles as indicated.[55]

highly between alleles and has variable number of asparagines/tyrosine repeats (HV in Fig. 3B). Repression of *csd* by the RNA interference technique[31] in females that were heterozygous for *csd* resulted in a full developmental switch into male gonad development. Repression of *csd* in males, however, had no phenotypic effect. These results indicate that *csd*, when derived from different alleles, is functionally active and initiates female development. When *csd* is derived from one allele, it is not functionally active and the default program of male development ensues.

Sex-determining alleles and proposed molecular function

The fascinating question is how the diversity of allelic specificities results in a stable regulatory signal that initiates sexual development. Given that about 19 sex-determining alleles segregate in populations,[22] this would require that the 171 possible heterozygous combinations of allelic *csd* proteins are active, while 19 combinations derived from the same allele are non-active. In the subsequent paragraphs, I will explore the potential source of allelic specificities, possible concepts underlying the molecular decision making, and plausible signals that initiate female development.

Single amino acid differences found among the sequences are a potential source of allelic specificity. Most single differences occur in the RS and proline-rich domain and are more often found in the interspersed amino acid residues in between the arrays of arginine/serine or proline residues. It is now widely accepted that RS domains[30,32] have protein-binding function, some with important specificity.[33] Protein binding has been also attributed to proline-rich domains[34,35] that interact with its cognate domains. In addition to single amino acid differences, deletions and insertions are possible sources of allelic specificity. The most apparent of these differences are found in the hypervariable region (HV in Fig. 3B) with repeats that appear a different number of times in different alleles. This repeating unit consists of one to four asparagines (N), followed by one tyrosine (Y). These length differences may alter the functioning of neighboring domains, or the repeat may itself be functional. Differences in phosphorylation sites are also possible sources of specificity. Predicted phosphorylation sites vary between allelic CSD polypeptides (data not shown). Several studies have shown that SR-protein function is regulated through phosphorylation of their RS domains by multiple kinases.[36,37] Finally, the various sources of specificity are not mutually exclusive, and allelic specificity may be a combination of any or all of the sequence attributes discussed above.

Based on the previous conclusions, we present three models to explore how the different specificities can combine and form an active or a non-active molecule (Fig. 4). The conceptual distinction between these models is attributed to

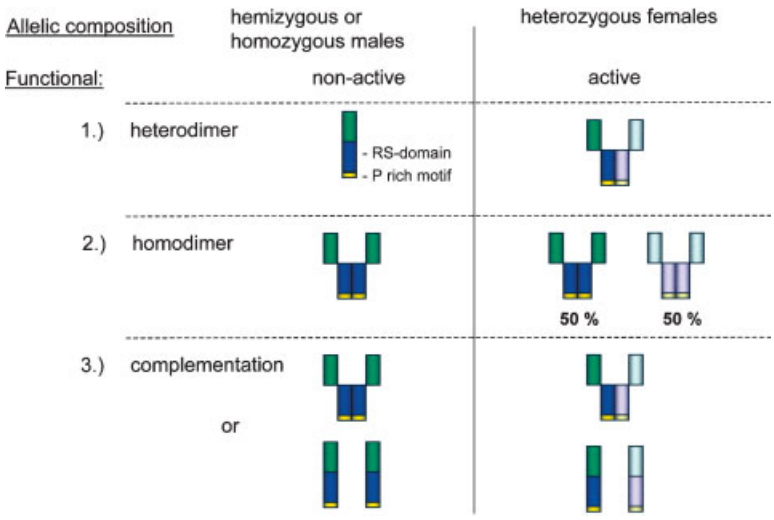

Figure 4. Models for protein association of two *csd* allelic polypeptides and their functional consequences. Color intensities represent different specificities of *csd* alleles. The most-variable parts of the allelic sequences (RS domain, proline-rich domain, hypervariable part) are marked in blue and yellow while the more-conserved N-terminal part is shown in green. **Model 1:** polypeptides that are derived from different alleles associate and form heteromers that are functionally active. The non-functional species is the monomer that is present in haploid and diploid males. **Model 2:** polypeptides that are derived from the same allele associate and form homomers. Consequently, two homomer species exist in females, while there is only one homomeric species in males. **Model 3:** no binding but activity differences exist depending on whether polypeptides derive from the same or different alleles. Different complementation groups in different alleles complement each other resulting in a functional protein in heterozygous females. I predict that no differences in binding properties of polypeptides are found whether they are derived from the same or different alleles.

differences in binding abilities of the allelic polypeptides that modulate protein activity. In the first two models, binding differences exist depending on whether polypeptides combine from the same or different alleles while, in the third model, activity differences occur without the distinction in binding. In model 1, only polypeptides derived from different alleles are capable of association. Hence, heteromers are formed and are functionally active only in females. In males, polypeptides that have the same specificity whether they are derived from the same or two identical alleles cannot associate and they are not active. This model is analogous to the self-incompatibility mechanisms in the fungus *Ustilago*.[38] In model 2, only polypeptides with the same specificity can combine to form homomers; polypeptides that are derived form different alleles cannot associate. Homomeric molecules exist in both sexes, the difference being whether one or two species of homomer exist (in males and females, respectively). This mechanism would require distinguishing the dosage discrepancies between all homozygous and heterozygous allelic combinations. Model 3 is based on the assumption of interallelic (intragenic) functional complementation (complementation model). Single allelic proteins are non-functional, but the combination of different alleles in the heterozygous state results in functional complementation and yields an active protein. Under the complementation model, no binding differences occur whether the polypeptides derive from the same or different alleles. For the complementation model to hold, 19 sex-determining alleles would necessitate 19 distinct non-functional versions that complement into active protein in all various heterozygous combinations.

Protein interactions among SR-type proteins can be functionally altered in the splicing process.[29,39] For example, TRA protein influences the binding properties of TRA-2 through cooperative interaction, and thus directs the regulated splicing of dsx.[40,41] Although most SR-type function is attributed to coactivation, it is also possible that *csd* might function as a repressor similar to the ASF/SF2 proteins.[42] CSD may thus be an activator of female- or a repressor of male-specific splicing. Such a mechanism for *csd* would indicate a remarkable relation to the sex-determining cascade in the fruitfly (see below): the splicing of downstream genes can be shifted to the female pattern if CSD is active while a default male-specific pattern is chosen if CSD is non-active.

Implications for the evolution of sex-determining cascades in insects

In the next paragraph, I will compare the sex-determining cascade of *D. melanogaster* (Fig. 5A, for review see Refs. 1,26) and some available molecular data from other dipteran species, such as *Musca domestica* and *Ceratitis capitata* (Fig. 5B shows the *C. capitata* case),[43–45] and how this relates to the sex-determining system of the honey bee,[9] which is estimated to have separated from the dipteran insects about 270 million years ago (Fig. 5C). This will provide some insights into the differences and some common principles of sexual regulation among insects, that emerge from the available data.

(i) In dipteran species, the primary decision is based on sex-specific factors that are present differently in the two sexes (marked in blue, Fig. 5A,B). These initial signals vary considerably between species: in *D. melanogaster*, the X:A ratio regulates *Sex-lithal (Sxl)* transcription while, in *M. domestica* and *C. capitata*, the presence or absence of a dominant male factor (*M*) on the Y chromosome is the initial signal (Fig. 5B), although this signal has not yet been identified.[46,47] In the honey bee, the initial signal depends on whether the allelic composition is identical or different. All alleles in the honey bee have the same potential to initiate male and female development. *csd* thus defines a novel class of initiating primary signals in which the genetic composition of the gamete has no predictive value for sex tendency prior to fertilization. Taken together, it appears that initial signals in insects can differ considerably in both molecular mechanism of activation and relative position in the regulative pathway (Fig. 5).

(ii) The initial signal is transferred to switch genes that have two states of activity (yellow box in Fig. 5). This is true, for example, for *Sxl* or *transformer (tra)* in the fruitfly, or the *tra* orthologue in *Ceratitis*. In these systems, the male state is non-active and occurs by default. For instance, in fruitfly females, SXL causes female-specific splicing of the *tra* transcript, which encodes an active TRA protein. In males, the absence of SXL results in a default splicing pattern of *tra* transcript which is non-functional. *csd* also acts as a switch gene and has similarly to *Drosophila* and *Ceratitis* in that it is in the active state in females and in the non-active state in males. In contrast, however, CSD is itself both the initiating signal and the switch, and regulation of its switch activity occurs at the post-translational level. Despite the difference, CSD belongs, like TRA, to the family of SR-type proteins that have sequences rich in arginine (R) and serine (S). Although they differ in sequence, both proteins have an RS and a proline-rich domain in common. *csd*, however, lacks the 11 amino acid long conserved motif found among the available TRA sequences of the Drosophilidae and *Ceratits*.[9,45]

(iii) On one level of the cascade, the sex-specific signal is maintained in the male or female mode throughout development (marked in pink in Fig. 5). In the fruitfly, SXL protein regulates its own splicing via a positive feed-back loop that maintains the sex-specific signal. *Sxl* has no plausible role in the maintenance of the signal in the non-drosophilid dipterans because both sexes have identical transcripts.[26] There is some evidence that, in *Ceratitis*, the sexual fate throughout development is maintained at the level of the *tra* gene by a positive feedback loop, which is similar to the one found in the fruitfly at the level of *Sxl*.[45] In the honey bee, we have no evidence so far for a mechanism that maintains the sexual fate throughout development. One possibility is that the permanent

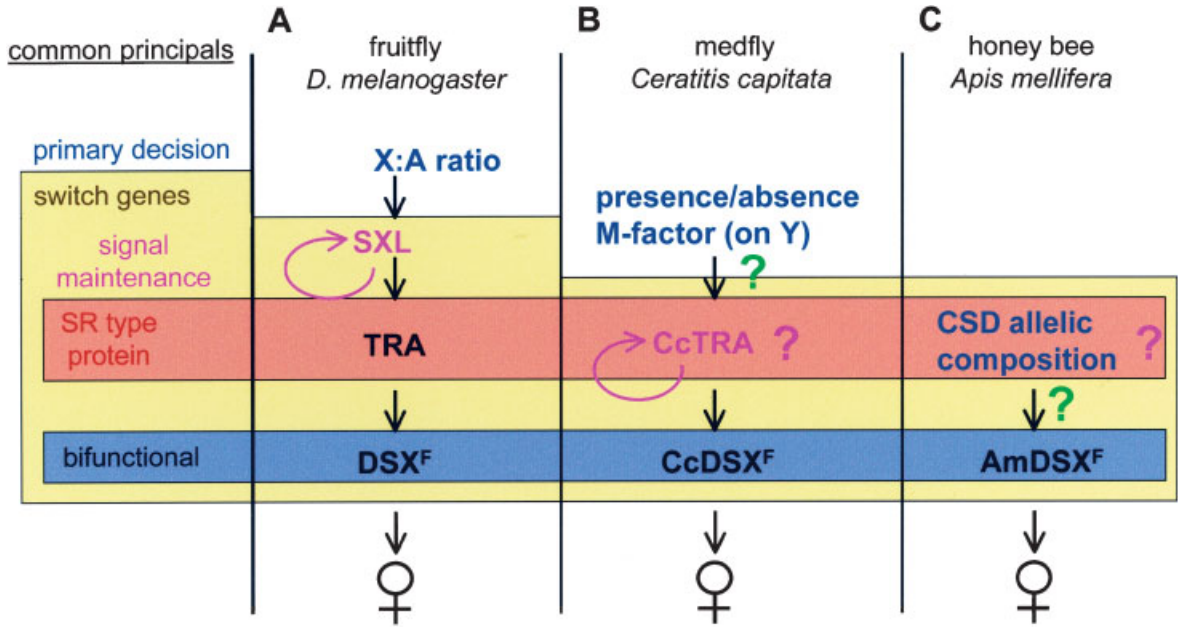

Figure 5. Sex-determining pathways in insects in which molecular elements upstream of the common *dsx* gene have been identified: the dipteran insects *Drosophila* **(A)** and medfly *Ceratitis* **(B)** as well as the hymenopteran insect, *Apis*, the honey bee **(C)**. The available data on molecular elements are presented. The left column indicates some common principles that emerge from these regulative hierarchies. The diverse initial signals correspond to the different sex-determining mechanisms and are marked in blue. Note that the factor *M* in *Ceratitis* has not been identified. Switch genes that have different active states and are present at different levels of the cascade are shown in a yellow box, the level of regulation in which SR-type proteins are involved are indicated with a red box. The conserved *dsx* gene, which is sex-specific spliced and has a proposed function in male and in female development, is marked with a blue box. The level of regulation at which the signal is maintained over development is marked in pink, while the pink question marks indicate that this principle is not full understood in *Ceratitis* and *Apis*. Functional relationships of elements that have not been proven so far are indicated by green question marks. Note that for *Drosophila* only the pathway of somatic sexual regulation is presented.

transcription of *csd*, which starts after cell formation, is a constant source of sexual identity. In summary, the available data suggest that the maintenance of the sex-specific signal throughout development is attributed to different molecular elements on different levels of the hierarchy. Whether the molecular mechanism is the same or is different has yet to be proven.

(iv) *doublesex (dsx)*, the gene at the end of the primary cascade of the fruitfly is a bifunctional switch gene in that it can encode both an active male and an active female protein, which are both involved in somatic differentiation (blue box in Fig. 5). The switch between male and female *dsx* transcript in the fruitfly is regulated via the TRA protein and alternative splicing. The *dsx* gene is well conserved not only in its structure but also in its transcription pattern in a variety of dipteran insects (*Musca*, *Ceratitis*, *Megaselia*, *Bactrocera*, Ref. 26) and in the moth *Bombyx*.[48] Although the gene is apparently under sexual regulation, its major sexual differentiation function has only been proven outside of *Drosophila* in the housefly *Musca*.[44] An orthologous *dsx* transcript is present in the honey bee and encodes a putative male- and female-specific protein by alternative splicing (R. Becher and M. Beye, unpublished results).

Comparison of the identified elements among insects suggests that, at some downstream level, the regulative hierarchy elements are conserved (Fig. 5). This finding supports the "bottom-up" hypothesis of Wilkins,[49] which suggests that the most-ancient genes of sexual regulation operate at the bottom of the cascade and that, during the course of evolution, new regulatory elements have been recruited upstream. These changes, however, do not occur at a constant fashion across the different levels of the hierarchies. Regulative changes of genes in the hierarchy upstream of *dsx* occur at varies phylogenetic distances. This is illustrated by the fact that possibly the *dsx* gene have been involved in sexual differentiation for more than 270 million years, while at the upstream level of SR-type proteins, the hierarchies diversify and different modes of SR-type protein activations are found among *Drosophila*, *Ceratitis* and *Apis*. At least SXL and the X:A signal functions appear to be conserved inside the Drosophilidae that represent about 60 million years of evolution,[50] but SXL has

no obvious sex-determining role in non-drosophilid species that have been separated more than 100 million years ago.[51,52] In the honey bee, no further upstream element above the SR-type protein has been recruited over the last 270 million years.[9] The comparative findings suggest that regulative changes upstream of *dsx* appear to be much easier to accommodate than at the downstream level, which is consistent with some theoretical considerations on the evolution of hierarchies.[50] The changes are thought to occur more upstream especially under conditions when the more downstream genes have pleiotropic effects. Indeed, there is accumulating evidence that *dsx* (the most-conserved gene in the sex-determining hierarchy of insects) has a strong pleiotropic effects in the fruitfly.[53,54]

Plasticity of regulation among SR-type proteins, a key to understand the diversity of sex-determining systems?

How can the diverse initial signals of sex determination activate the common pathway of sexual regulation? A possible answer comes from the astonishing differences in regulation methods that are found among the identified SR-type proteins. In a complementary system, CSD is active if the polypeptides are derived from different alleles while, in a genic balance (X:A) system, TRA function depends on the correct splicing of *tra* transcripts via SXL protein. In the *M*-factor and Y-chromosome system of *Ceratitis*, SXL protein has no obvious sex-determining function but functional TRA is most likely obtained via a positive feed-back loop in which TRA correctly splices its own transcript.[45] These findings suggest that different upstream or initial signals have been recruited over the last 270 million years resulting in various mechanisms of SR-type protein activation. Thus, the astonishing plasticity of regulation among SR-related proteins, which resulted in various sex-specific signals and the proposed conserved downstream function, can help to explain the diversity of sex-determining systems found today.

Future directions

The regulation of sex using allelic specificities in different combinations is a fascinating problem. We are now beginning to understand how the decision is made at the level of polypeptides. The isolation of different alleles that are functionally active if two specificities are combined offers the opportunity to dissect the source of specificity and to understand the nature of molecular decision. This will allow us to understand how the enormous number of 171 possible heterozygous combinations can result in an active protein while 19 possible homozygous combinations can result in non-active protein. The first molecular population genetics analyses of 14 allelic *csd* nucleotide sequences have found signatures of positive selection that are candidate regions of the protein to dissect their allelic specificity.[55]

The structural and possible functional similarity of the initial signal *csd* in the honey bee and the downstream regulator *tra* in the fruitfly offers the opportunity to elucidate how different sex-determining systems have evolved over the last 270 million years. Further studies will show whether *csd* is involved in alternative splicing and whether *csd* directly targets *dsx* primary transcripts, and thus whether *csd* functionally corresponds to *tra*.

The finding that exceptions to the complementary sex determination mode occur within the Hymenoptera (habitually inbreeding species),[23] and that haplodiploidy occurs in perhaps 20% of all known animal species offers the opportunity to study the evolution of haplodiploid sex determination mechanisms in a wide range of phylogenetically diverse species.

Acknowledgments

The author thanks very much C. Whitfield for critical reviewing the manuscript. The author is grateful to M. Hasselmann, T. Gempe, A. Wilkins and R. Nöthiger for helpful discussion.

References

1. Cline TW, Meyer BJ. 1996. Vive la Difference: Males vs females in flies vs worms. Annu Rev Genet 30:637–702.
2. Goodfellow PN, Lovell-Badge R. 1993. *Sry* and sex determination in mammals. Annu Rev Genet 27:71–92.
3. Whiting PW. 1943. Multiple alleles in complementary sex determination of *Habrobracon*. Genetics 28:365–382.
4. Beye M, Moritz RFA, Epplen C. 1994. Sex linkage in the honeybee *Apis mellifera* detected by multilocus DNA fingerprinting. Naturwissenschaften 81:460–462.
5. Beye M, Crozier RH, Crozier YC, Moritz RFA. 1996. Mapping the sex locus of the honeybee (*Apis mellifera*). Naturwissenschaften 83: 424–426.
6. Hunt GJ, Page RE. 1994. Linkage analysis of sex determination in the honey bee (*Apis mellifera*). Mol Gen Genet 244:512–518.
7. Hunt GJ, Page RE Jr. 1995. Linkage map of the honey bee, *Apis mellifera*, based on RAPD markers. Genetics 139:1371–1382.
8. Beye M, Hunt GJ, Page RE, Fondrk MK, Grohmann L, et al. 1999. Unusually high recombination rate detected in the sex locus region of the honey bee (*Apis mellifera*). Genetics 153:1701–1708.
9. Beye M, Hasselmann M, Fondrk MK, Page RE, Omholt SW. 2003. The gene *csd* is the primary signal for sexual development in the honeybee and encodes an SR-type protein. Cell 114:419–429.
10. Page RE Jr., Gadau J, Beye M. 2002. The emergence of hymenopteran genetics. Genetics 160:375–379.
11. Dzierzon J. 1845. Gutachten über die von Herrn Direktor Stöhr im ersten und zweiten Kapitel des General-Gutachtens aufgestellten Fragen. Eichstädter Bienenzeitung 1:109–113; 119–121.
12. Siebold CTh. 1856. Die Drohneneier sind nicht befruchtet. Bienenzeitung 12:181–184.
13. Nachtsheim H. 1916. Zytologische Studien über die Geschlechtsbestimmung bei der Honigbiene (*Apis mellifera*). Arch Zellforsch 11:169–241.
14. Mable BK, Otto SP. 1998. The evolution of life cycles with haploid and diploid phases. Bioessays 20:453–462.
15. Crozier RH, Pamilo P. 1996. Evolution of social insect colonies. Oxford: Oxford University Press.
16. Bull JJ. 1983. Evolution of Sex Determining Mechanisms. Menlo Park, Calif.: Benjamin/Cummings Publishing Company.
17. McClung CE. 1902. The accessory chromosome—sex determinant? Biol Bull Mar Biol Lab, Woods Hole 3:43–84.
18. Wilson EB. 1905. The chromosomes in relation to determination of sex in insects. Science 22:500–502.

19. Bridges CB. 1916. Non-disjunction as proof of the chromosome theory of heredity. Genetics 1:1–52.
20. Bridges CB. 1925. Sex in relation to genes and chromosomes. Am Nat 59:134.
21. Whiting PW. 1933. Selective fertilization and sex determination in Hymenoptera. Science 78:537–538.
22. Adams J, Rothman ED, Kerr WE, Paulino ZL. 1977. Estimation of the number of sex alleles and queen maitings from diploid male frequencies in a population of *Apis mellifera*. Genetics 86:583–596.
23. Cook JM. 1993. Sex determination in the hymenoptera: A review of models and evidence. Heredity 71:421–435.
24. Dobson SL, Tanouye MA. 1998. Evidence for a genomic imprinting sex determination mechanism in *Nasonia vitripennis* (Hymenoptera; Chalcidoidea). Genetics 149:233–242.
25. Poirié M, Périquet G, Beukeboom LW. 1992. The hymenopteran way of determining sex. Semin Dev Biol 3:357–361.
26. Schutt C, Nothiger R. 2000. Structure, function and evolution of sex-determining systems in Dipteran insects. Development 127:667–677.
27. Hasselmann M, Fondrk MK, Page RE, Beye M. 2001. Fine scale mapping in the sex locus region of the honey bee (*Apis mellifera*). Insect Mol Biol 10:605–608.
28. Beye M, Poch A, Burgtorf C, Moritz RFA, Lehrach H. 1998. A gridded genomic library of the honeybee (*Apis mellifera*): A reference library system for basic and comparative genetic studies of a Hymenopteran Genome. Genomics 49:317–320.
29. Hastings ML, Krainer AR. 2001. Pre-mRNA splicing in the new millennium. Curr Opin Cell Biol 13:302–309.
30. Graveley BR. 2000. Sorting out the complexity of SR protein functions. RNA 6:1197–1211.
31. Beye M, Hartel S, Hagen A, Hasselmann M, Omholt SW. 2002. Specific developmental gene silencing in the honey bee using a homeobox motif. Insect Mol Biol 11:527–532.
32. Tacke R, Manley JL. 1999. Determinants of SR protein specificity. Curr Opin Cell Biol 11:358–362.
33. Dauwalder B, Mattox W. 1998. Analysis of the functional specificity of RS domains in vivo. EMBO J 17:6049–6060.
34. Kay BK, Williamson MP, Sudol M. 2000. The importance of being proline: the interaction of proline-rich motifs in signaling proteins with their cognate domains. FASEB J 14:231–241.
35. Sudol M, Sliwa K, Russo T. 2001. Functions of WW domains in the nucleus. FEBS Lett 490:190–195.
36. Wang HY, Lin W, Dyck JA, Yeakley JM, Songyang Z, et al. 1998. SRPK2: a differentially expressed SR protein-specific kinase involved in mediating the interaction and localization of pre-mRNA splicing factors in mammalian cells. J Cell Biol 140:737–750.
37. Yeakley JM, Tronchere H, Olesen J, Dyck JA, Wang HY, et al. 1999. Phosphorylation regulates in vivo interaction and molecular targeting of serine/arginine-rich pre-mRNA splicing factors. J Cell Biol 145:447–455.
38. Kamper J, Reichmann M, Romeis T, Bolker M, Kahmann R. 1995. Multiallelic recognition: nonself-dependent dimerization of the bE and bW homeodomain proteins in *Ustilago maydis*. Cell 81:73–83.
39. Black DL. 2003. Mechanisms of alternative pre-messenger RNA splicing. Annu Rev Biochem 72:291–336.
40. Hoshijima K, Inoue K, Higuchi I, Sakamoto H, Shimura Y. 1991. Control of *doublesex* alternative splicing by *transformer* and *transformer-2* in *Drosophila*. Science 252:833–836.
41. Tian M, Maniatis T. 1993. A splicing enhancer complex controls alternative splicing of *doublesex* pre-mRNA. Cell 74:105–114.
42. Wang J, Xiao SH, Manley JL. 1998. Genetic analysis of the SR protein ASF/SF2: interchangeability of RS domains and negative control of splicing. Genes Dev 12:2222–2233.
43. Dubendorfer A, Hediger M, Burghardt G, Bopp D. 2002. *Musca domestica*, a window on the evolution of sex-determining mechanisms in insects. Int J Dev Biol 46:75–79.
44. Hediger M, Burghardt G, Siegenthaler C, Buser N, Hilfiker-Kleiner D, et al. 2004. Sex determination in *Drosophila melanogaster* and *Musca domestica* converges at the level of the terminal regulator *doublesex*. Dev Genes Evol 214:29–42.
45. Pane A, Salvemini M, Bovi PD, Polito C, Saccone G. 2002. The *transformer* gene in *Ceratitis capitata* provides a genetic basis for selecting and remembering the sexual fate. Development 129:3715–3725.
46. Hilfiker-Kleiner D, Dubendorfer A, Hilfiker A, Nothiger R. 1993. Developmental analysis of two sex-determining genes, M and F, in the housefly, *Musca domestica*. Genetics 134:1187–1194.
47. Willhoeft U, Franz G. 1996. Identification of the sex-determining region of the *Ceratitis capitata* Y chromosome by deletion mapping. Genetics 144:737–745.
48. Ohbayashi F, Suzuki MG, Mita K, Okano K, Shimada T. 2001. A homologue of the *Drosophila doublesex* gene is transcribed into sex-specific mRNA isoforms in the silkworm, *Bombyx mori*. Mol Genet Metab 128:145–158.
49. Wilkins AS. 1995. Moving up the hierarchy: A hypothesis on the evolution of a genetic sex determination pathway. Bioessays 17:71–77.
50. Marin I, Baker BS. 1998. The evolutionary dynamics of sex determination. Science 281:1990–1994.
51. Saccone G, Peluso I, Artiaco D, Giordano E, Bopp D, et al. 1998. The *Ceratitis capitata* homologue of the *Drosophila* sex-determining gene *Sex-lethal* is structurally conserved, but not sex-specifically regulated. Development 125:1495–1500.
52. Meise M, Hilfiker-Kleiner D, Dübendorfer A, Brunner C, Nöthiger R, et al. 1998. *Sex-lethal*, the master sex-determining gene in *Drosophila*, is not sex-specifically regulated in *Musca domestica*. Development 125:1487–1494.
53. Garrett-Engele CM, Siegal ML, Manoli DS, Williams BC, Li H, et al. 2002. *intersex*, a gene required for female sexual development in *Drosophila*, is expressed in both sexes and functions together with *doublesex* to regulate terminal differentiation. Development 129:4661–4675.
54. Christiansen A, Keisman E, Ahmad S, Baker B. 2002. Sex comes in from the cold: the integration of sex and pattern. Trends Genet 18:510.
55. Hasselmann M, Beye M. 2004. Signatures of selection among sex-determining alleles of the honey bee. Proc Natl Acad Sci USA 101:4888–4893.
56. http://www.rug.nl/biologie/onderzoek/onderzoekGroepen/evolutionaryGenetics/evogen/