Copyright © 2004 by the Genetics Society of America

The Evolution of the Drosophila Sex-Determination Pathway

Andrew Pomiankowski,*[,1] Rolf Nöthiger[†] and Adam Wilkins[‡]

**Department of Biology, University College, London NW1 2HE, United Kingdom, [†]Zoological Institute, University of Zürich, CH-8057 Zürich, Switzerland and [‡]BioEssays, Cambridge CB1 1HN, United Kingdom*

Manuscript received September 22, 2003
Accepted for publication December 10, 2003

ABSTRACT

The molecular complexity of the Drosophila somatic sex-determination pathway poses formidable intellectual challenges for attempts to explain its evolutionary origins. Here we present a reconstruction of how this regulatory cascade might have evolved in a step-by-step fashion. We illustrate how mutations in genes, which were already part of the pathway or were recruited as new regulators of the pathway, were favored by sexual selection acting on the discriminatory sex-determining signal. This allows us to explain the major features of the pathway, including multiple promoter sites, alternative splicing patterns, autoregulation, and stop codons. Our hypothesis is built on the available data from Drosophila and other insect species, and we point out where it is amenable to further experimental and comparative tests.

CONSIDERING the seemingly simple task that the Drosophila sex-determination pathway performs, namely the production of males and females, it exhibits a remarkable degree of complexity (reviewed in Schutt and Nöthiger 2000). A primary signal, which in itself consists of several interacting gene products, initiates an intricately integrated multi-genic cascade whose components show a variety of regulatory mechanisms. These include multiple promoter sites, alternative splicing patterns, autoregulation, and the presence of stop codons. Comparative analyses of Dipteran relatives of Drosophila and more distantly related species reveal further structural and regulatory complexity, as the sex-determination mechanisms employ some of the same genes but frequently in different ways along with a suite of other control elements.

Although there is a voluminous literature on the diversity of sex-determination systems in general and even within particular phylogenetic groups such as the Diptera (see, for instance, Bull 1983; Marin and Baker 1998), comparatively little effort has been given to considering how specific sex-determination systems may have evolved. The early theoretical literature on the evolution of sex determination is extensive (see Bull 1983 for a review) but has not proved very illuminating, principally because it is too abstract, concentrating on conditions for the spread of hypothetical "modifiers." The classical literature, in the virtual absence of genetic and molecular information, was unable to address how real genetic networks were constructed. Today, however, our knowledge about the genetic and molecular basis of several sex-determination systems (Cline and Meyer 1996; Schutt and Nöthiger 2000) renders an evolutionary approach both possible and desirable to make sense of the evident variety and complexity of these systems. A few analyses have made a start in this direction, focusing on the known sex-determination pathways in insects and nematodes and their possible evolutionary origins (Nöthiger and Steinmann-Zwicky 1985; Hodgkin 1992; Wilkins 1995; Raymond *et al.* 1998). In particular, Nöthiger and Steinmann-Zwicky (1985) showed how the multiplicity of insect sex-determination systems might, in principle, reflect diversity only in the most upstream switches of the pathway, while Wilkins (1995) argued that the long *Caenorhabditis elegans* pathway might have grown by successive addition of upstream control elements to an ancient conserved downstream module.

None of the previous discussions, however, have explained the complexity seen in sex-determination pathways in terms of evolutionary dynamics and selectional forces. Our goal is to redress this gap through a hypothetical reconstruction of the main evolutionary steps that led to the Drosophila sex-determination system. We have chosen to concentrate on Drosophila as this is the best characterized of all the sex-determination pathways and, by virtue of its complexity, provides a challenging test of our general approach. In addition, there is now considerable knowledge about sex determination in other Dipteran insects, which permits informative comparisons and the inference of ancestral states. Our focus is on the underlying genetic events, rather than on morphological or developmental change. By specifying (a) the order in which genes were added to the pathway and (b) the selective reasons for their recruitment, our reconstruction can be broken down into a series of hypotheses, many of which can be tested via comparative studies with other species.

Our wider aim is to develop a framework to study the

[1]*Corresponding author:* Department of Biology, University College, 4 Stephenson Way, London NW1 2HE, United Kingdom.
E-mail: ucbhpom@ucl.ac.uk

evolution of the apparently quite different mechanisms of sex determination seen among Diptera (*e.g.*, Drosophila, Ceratitis, Musca, and Sciara), other insects (*e.g.*, Lepidoptera and Hymenoptera), and beyond (*e.g.*, nematodes and mammals). The burgeoning wealth of data on the sex-determination pathways of animal species should serve to facilitate both the formulation of new hypotheses and the testing of ideas, including the ones we propose here.

A major point in the scheme to be described is that it relies on sexual selection as a principal motor for evolutionary change in sex-determining systems. Sexual selection is known to be a strong and temporally variable selective force that has contributed to the exaggeration and diversity of secondary sexual characters involved in courtship display and mating success (Darwin 1871; Andersson 1994). It is not commonly appreciated, however, that sexual selection can also act on the primary mechanism that determines sex. The differences in behavior, physiology, morphology, etc., that affect sex-specific fitness arise from differential gene expression, which is set up and maintained by the sex-determining gene network. Hence, genetic variation in the sex-determining mechanism, even in the primary sexual signal (*e.g.*, in the strength or timing of this signal), will sometimes have consequences for sexual fitness. We show how this can occasionally lead to major transitions in the sex-determining mechanism, such as the recruitment of new elements, changes in heterogamety, new promoter regions, or alternative splicing.

We begin with a hypothetical ancestral state, from which we derive, through a series of mutational changes, the current system of *Drosophila melanogaster*. The reconstruction that we outline shows how a sequence of individually conventional mutational changes could have generated the pathway that determines somatic sex in Drosophila today. We will begin with a short review of the contemporary system and then proceed to our conceptual reconstruction of its possible evolution from a much simpler ancestral state.

SEX DETERMINATION IN DROSOPHILA

The basic features of the somatic sex-determination cascade in *D. melanogaster* are outlined in Figure 1 (for a complete description, see Schutt and Nöthiger 2000). Many other genes are essential to the pathway, but these are equally expressed in both sexes and thus have no discriminatory role. The primary genetic signal is provided by the ratio of X-linked numerator genes [three *sisterless* genes (*sisA, sisB*, and *sisC*) and *runt* (*run*)] to one major autosomal denominator gene, *deadpan* (*dpn*). In females, with two X chromosomes, this X:A ratio is 2:2 while in males, which carry only one X, it is 1:2. The products of these genes are transcription factors that regulate the expression of *Sex-lethal* (*Sxl*). This gene is unusual in having a stop codon (UAG) embedded in exon 3. Exon 3, therefore, has to be removed during RNA processing for transcripts to produce functional SXL protein. In females, the products of two doses of the X-linked numerators activate the early promoter of *Sxl* (*Pe*) shortly after fertilization at the cellular blastoderm stage (Estes *et al.* 1995). The *Pe* promoter produces RNA transcripts from which exons 2 and 3 are constitutively spliced out, resulting in an early burst of active SXL protein. *Pe*, however, is only transiently active between embryonic cleavage cell cycles 12 and 14 and is quickly replaced by the maintenance promoter *Pm*, which is active in both sexes and is not regulated by the numerator and denominator transcription factors. Transcripts from this promoter do not undergo constitutive excision of exons 2 and 3. Nevertheless, in females, exon 3 is spliced out of *Sxl* primary transcripts because the SXL protein that was initially produced from the *Pe* promoter can bind to its own pre-mRNA and, as a splice enhancer, enforces the elimination of exon 3. This establishes an autoregulatory loop, which maintains itself throughout development. Females, therefore, continue to produce SXL protein. In contrast, there is no alternative splicing in males because they do not produce the initial burst of SXL protein from *Pe*, which cannot be activated by only a single dose of the numerators. Hence in males, the autoregulatory loop is never established. Male *Sxl* transcripts produced from the late *Pm* promoter retain exon 3, and this results in premature termination of translation and absence of functional SXL protein.

Sxl codes for an RNA-binding protein that regulates production of not only its own transcripts but also those of *transformer* (*tra*), the next gene in the sex-determination pathway (Figure 1). Like *Sxl, tra* produces transcripts that contain several stop codons at the beginning of exon 2. In females, SXL protein blocks the canonical splice site and forces use of a cryptic splice site just downstream of the stop codons. This creates an open reading frame, which now allows the production of active TRA protein. In males, however, the absence of SXL results in mRNAs that retain the stop codons in exon 2, which leads to premature termination of translation and absence of any functional TRA protein. *tra* codes for another RNA-binding protein that causes alternative splicing of *doublesex* (*dsx*), the next downstream element in the pathway. In males, the absence of TRA protein results in the default splice of *dsx* transcripts and the loss of exon 4. Hence male *dsx* mRNA contains exons 1–3 and 5–6. This produces the male-specific DSXM isoform. In females, in contrast, the presence of TRA protein, together with the cofactor TRA2, initiates an alternative splicing pattern, which includes and terminates with exon 4. Thus female *dsx* mRNA contains exons 1–4 and produces the female-specific DSXF isoform. Most somatic sexual characters are differentially determined by the two *dsx* proteins. These act as transcription factors that sex-specifically enhance or repress

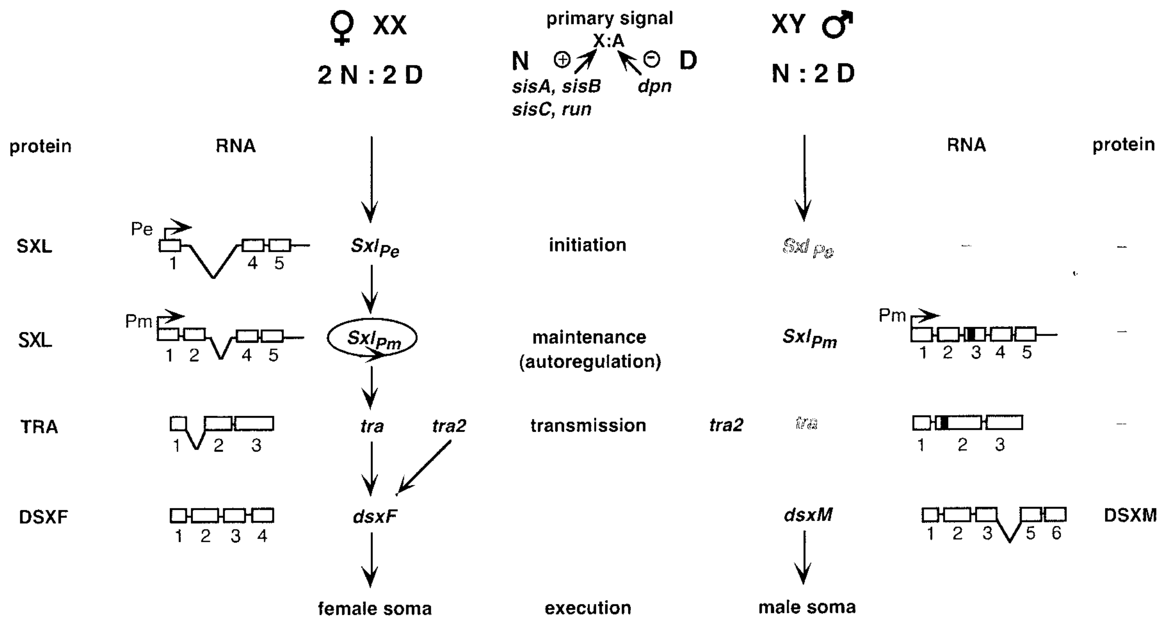

Figure 1.—The genetic cascade regulating somatic sexual development in Drosophila today. The female pathway is on the left; the male pathway is on the right. Open gene symbols indicate functional inactivity. The gene *tra2* is active in both sexes, but is required only in the female soma. Vertical arrows show the flow of information. The genes forming the primary signal (N, numerators; D, denominators) and *dsx* encode transcription factors; *Sxl*, *tra*, and *tra2* encode proteins involved in RNA processing. Boxes represent exons; for *Sxl*, only the first 5 exons (of 10) are drawn. Bars in exon 3 of *Sxl* and in exon 2 of *tra* symbolize stop codons.

a number of downstream male- and female-specific genes, which implement the two different routes of sexual differentiation (Christiansen *et al.* 2002).

Less is understood about the genetic basis of sex determination in related insects, but it is clear that there is both evolutionary conservation and divergence. A significant conserved element is the last regulatory gene in the cascade, *dsx*. The transcripts of *dsx* are alternatively spliced in males and females in several insects, including other Diptera such as *Bactrocera tryoni* (Shearman and Frommer 1998), *Ceratitis capitata* (Pane *et al.* 2002), *Megaselia scalaris* (Sievert *et al.* 1997), and *Musca domestica* (Dubendorfer *et al.* 2002) and, more distantly, the Lepidopteran *Bombyx mori* (Ohbayashi *et al.* 2001). Sex-determination genes with strong sequence similarity to *dsx* in the regions encoding their DNA-binding domains have also been found in a diverse range of still more distantly related animals from nematode worms to mammals (Raymond *et al.* 1998, 2000). This phylogenetic conservation suggests an ancient origin and role for *dsx* in sex determination.

Other genes in the Drosophila pathway appear to be more recent additions. *tra* is thought to be a common factor in Dipteran insects as *dsx* undergoes sex-specific alternative splicing in several species of this group (Sievert *et al.* 1997; Shearman and Frommer 1998; Dubendorfer *et al.* 2002; Pane *et al.* 2002; Hediger *et al.* 2004). This was confirmed for the fruit fly *C. capitata*, in which TRA protein causes sex-specific alternative splicing of *dsx* (Pane *et al.* 2002). Recently, a gene with molecular similarity to *tra* has been identified outside the Diptera in the honeybee *Apis mellifera* (Beye *et al.* 2003), where again it appears to be involved in *dsx* splicing. However, *tra* is unlikely to be involved in the Lepidoptera, as *dsx* in the silkworm *B. mori* lacks TRA-binding sites and has a reversed pattern of alternative splicing with the default splice in females (Ohbayashi *et al.* 2001; Suzuki *et al.* 2001). Outside the insects, there is no evidence for the involvement of *tra* in sex determination in any species. For instance, the *dsx* homolog *mab-3* of the nematode *C. elegans* is under completely different genetic control (Shen and Hodgkin 1988). This suggests that the role of *tra* in sex determination is limited to certain insect groups. The involvement of *Sxl* in sex determination appears to be even more restricted than that of *tra*, being limited to the genus Drosophila (Bopp *et al.* 1996; Penalva *et al.* 1996). In *Chrysomya rufifacies* (Muller-Holtkamp 1995), *M. scalaris* (Sievert *et al.* 1997), *M. domestica* (Meise *et al.* 1998), and *C. capitata* (Saccone *et al.* 1998), *Sxl* is equally expressed in both sexes and thus cannot act as a discriminator of sex (the function of *Sxl* in these Dipteran species is not known). A similar limitation to the genus Drosophila is likely to pertain to the numerator and denominator genes involved in sex determination (Erickson and Cline 1998).

The general picture is diversity among upstream and conservation among downstream control genes, as proposed by Wilkins (1995). The Drosophila cascade has apparently seen the repeated recruitment of upstream

elements as regulators of the sex-determination pathway. To envisage how this could have happened, we have reconstructed the evolution of the current genetic system of sex determination in Drosophila from an ancestral state that had *dsx* as the discriminatory signal of the pathway. This seems a reasonable assumption, given that *dsx* is involved in sex determination in all insect species examined, whereas *tra*, *Sxl*, and the numerator/denominator genes are more phylogenetically restricted. In our reconstruction, we have concentrated on explaining the evolution of the pathway from the X:A ratio to *dsx* for somatic sex determination. The evolution of germline sex determination and dosage compensation has been left for the future (see DISCUSSION). Our efforts have been guided by the extensive experimental knowledge of gene expression in Drosophila and related insect species. Many parts of the evolutionary sequence that we are proposing can, in principle, be tested by future comparative studies.

For the special case of Drosophila, we begin with a putative ancestral sex-determination pathway and postulate a degree of weakness in it that could not be "solved" by simple changes in gene expression of the component elements. We then show how a single genetic change could have rectified the initial weakness while simultaneously laying the ground for further change. The complete sequence of changes that we propose to account for the present-day Drosophila sex-determination system involves both the addition of new control elements to the gene network and major changes to the component genes.

SEX-DETERMINATION PATHWAY EVOLUTION

The ancestral state

We assume that control of sex determination in the ancestor of Drosophila was through heterogamety at the *dsx* locus. In the proposed ancestral state, males were heterogametic dsx^M/dsx^+, and females were homogametic dsx^+/dsx^+. A tra^+ allele was at fixation and produced TRA protein equally in both sexes. TRA acted as a female splice enhancer of dsx^+ transcripts, as occurs today in conjunction with its cofactor TRA2 (SCHUTT and NÖTHIGER 2000), but had no effect on dsx^M transcripts. Hence, dsx^+/dsx^+ individuals generated transcripts that followed the female splice to give DSXF and a female phenotype. dsx^M/dsx^+ individuals produced a greater amount of DSXM and a male phenotype. In this scheme, the dsx^M allele behaves as a dominant, generating transcripts that exclusively follow the male splice. dsx^M mutants with these characteristics have been reported in *D. melanogaster* (BAKER and WOLFNER 1988; NAGOSHI and BAKER 1990). It has also been shown that Drosophila with a surplus of DSXM over DSXF (XX flies with two copies of dsx^M plus one of dsx^+) are phenotypically male (NÖTHIGER and LEUTHOLD *et al.* 1987).

In addition to DSXM, the dsx^M/dsx^+ genotype produced DSXF from its dsx^+ allele, thereby potentially reducing the strength of the male-determining signal in this genotype. We hypothesize that this ambiguity in *dsx* expression was the main selective force that led to the first expansion of the pathway with the evolution of *tra* as a differential upstream regulator of *dsx*. Mutant forms of dsx^+ with less expression and hence less DSXF production would have been favored in males. But such mutants would have reduced DSXF production in females and would have been disadvantageous in this sex. The sharing of gene expression across the sexes limited the possible improvement of sex-specific adaptation by mutational change in the *dsx* gene (RICE 1984, 1998).

Conversion of *tra* to an upstream regulator of *dsx*

Two features of the contemporary *tra* gene need to be accounted for: first, *tra* carries a stop codon in exon 2, and second, this part of the exon is skipped in females. In our model, these features evolved in two steps, each resulting in an increase in the distinctiveness of the two sexes. In the first stage, a mutation occurred in the tra^+ allele, creating a stop codon (UAG) in exon 2. This mutation, which we call tra^S, caused premature termination of translation of *tra* transcripts and production of a truncated and inactive form of the TRA protein, similar to that seen in *D. melanogaster* today. The tra^S mutation was beneficial in males (tra^S/tra^+; dsx^M/dsx^+) because less TRA protein lowered the efficiency of the female splice of dsx^+ transcripts and hence increased the production of DSXM relative to DSXF (Table 1). Conversely, the reduction in DSXF in mutant females (tra^S/tra^+; dsx^+/dsx^+) was disadvantageous as the female splice was less efficient and these females suffered a reduction in DSXF, possibly even accompanied by production of some DSXM. Note that this reconstruction is consistent with the observation that flies heterozygous for a null allele of *tra* are not sex reversed but female in *D. melanogaster* (STURTEVANT 1945).

The balance of benefit to males and harm to females determined the fate of the tra^S mutant. This can be followed by assigning fitness values to genotypes (Table 1) and allowing evolution using a standard population genetic simulation (Figure 2; see APPENDIX for details). When rare, the tra^S mutant invades if the gain in male heterozygote (tra^S/tra^+; dsx^+/dsx^M) fitness is greater than the loss in female heterozygote (tra^S/tra^+; dsx^+/dsx^+) fitness (*i.e.*, $m > f$ in Table 1). The mating of male and female mutant heterozygotes generates tra^S homozygous males, which produce no TRA protein and lack any female splice of dsx^+ transcripts and so are male irrespective of their *dsx* genotype.

Our simulations show that if the fitness of tra^S homozygous males was greater than or equal to the fitness of tra^S heterozygous males (*i.e.*, $m2 \geq m$), then the tra^S allele would rise to a frequency of 0.75 (Figure 2). This

TABLE 1

Fitness of tra^S stop mutants

Sex	Genotype		Fitness		Comments
	tra	*dsx*	Male	Female	
Female	tra^+/tra^+	dsx^+/dsx^+		1	Original female
Male		dsx^+/dsx^M	1		Original male
Female	tra^+/tra^S	dsx^+/dsx^+		$1-f$	Less TRA, worse female
Male		dsx^+/dsx^M	$1+m$		Less TRA, better male
Male	tra^S/tra^S	dsx^+/dsx^+	$1+m2$		No TRA, even better male
Male		dsx^+/dsx^M	$1+m2$		No TRA, even better male

causes elimination of the dsx^M allele and fixation of the dsx^+ allele (Figure 2). As a result, the *tra* gene becomes the upstream regulator of *dsx*, with females being heterozygous (tra^+/tra^S) and males being homozygous (tra^S/tra^S). Thus, the discriminating signal moved one gene upward, from *dsx* to *tra*, with a simultaneous reversal in heterogamety from male to female. Our hypothesis shows how a stop mutation in *tra* ameliorated the problem that arose because the main sex-determining products, DSXM and DSXF, initially were both present in males.

In *D. melanogaster*, *dsx* and *tra* are linked, and recombination is absent from males. Our simulations show that the same evolutionary transition (*i.e.*, *tra* as the new upstream regulator with the loss of dsx^M) occurs if the tra^S mutation arises in linkage with dsx^+, although the conditions for spread are more restrictive than those with free recombination (see APPENDIX for details). In contrast, if the tra^S mutation occurs in linkage with dsx^M, it is limited to males, as dsx^M is a dominant masculinizer. The tra^S dsx^M mutant spreads as it reduces the amount of TRA produced by males. But due to the lack of male recombination, the tra^S mutant cannot cause the elimination of the dsx^M allele or the recruitment of *tra* as an upstream discriminatory regulator of sex determination (at equilibrium, females remain tra^+ dsx^+ homozygotes and males become tra^S dsx^M/tra^+ dsx^+ heterozygotes).

Recruitment of *Sxl*

The evolution of tra^S, although of overall benefit, itself caused a problem. Females now produced TRA from a single tra^+ allele rather than from two copies and were less efficient in splicing dsx^+ transcripts in the female mode and may even have produced some DSXM. To some extent, these disadvantages could have been counteracted by selection for higher expression of the single tra^+ allele. However, we know that evolution took a different path, which led to the recruitment of *Sxl*.

Sxl is a general RNA-binding protein that has multiple roles in RNA processing and translation suppression (KELLY *et al.* 1997; GEBAUER *et al.* 2003). In insects other

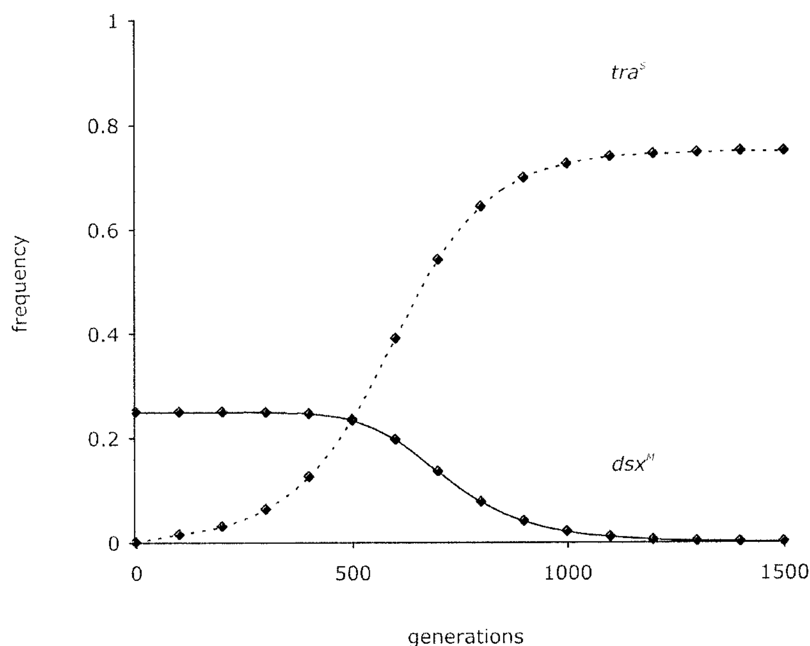

FIGURE 2.—The spread of the tra^S mutant is shown (broken line) from introduction at low frequency (0.01) to its equilibrium frequency (0.75). This leads to the loss of the dsx^M allele (solid line). The tra^S allele is set to have a selective advantage in males of $m = m2 = 0.03$, and a selective disadvantage in females of $f = 0.015$. The initial frequency and exact values of the selection coefficients do not alter the outcome as long as $m = m2 > f$. In the simulation shown, the *dsx* and *tra* loci were assumed to be unlinked and to undergo free recombination in both sexes (see the APPENDIX for the case of linkage).

TABLE 2

Fitness of *Sxl*^{F} mutants capable of splicing *tra*^{S}

	Genotype		Fitness		
Sex	*tra*	*dsx*	Male	Female	Comments
Female	Sxl^+/Sxl^+	tra^+/tra^S		1	Original female
Male		tra^S/tra^S	1		Original male
Female	Sxl^+/Sxl^F	tra^+/tra^S		$1 + g$	More TRA, better female
Female		tra^S/tra^S		$1 + g2$	More TRA, better female

than Drosophila, however, *Sxl* is not involved in somatic sex determination, as discussed previously. Hence, we make the parsimonious assumption that the ancestral version Sxl^+ was likewise not involved and had no, or at least no significant, binding affinity for tra^+ or tra^S transcripts. In the lineage leading to Drosophila, we assume that the tra^S allele had two important preexisting features: a poly(U)-binding site in the intron upstream of exon 2 and a cryptic splice acceptor site downstream of the stop codon in exon 2. Both of these are present in the *tra* allele of *D. melanogaster* (Sosnowski *et al.* 1989) and *C. capitata* (Pane *et al.* 2002). Neither of them would have had any consequences in the ancestral sex-determination system.

We propose that the first step in the recruitment of *Sxl* was the occurrence of a mutant allele, which we designate Sxl^F, with affinity for the poly(U)-rich binding site in tra^S. The binding of SXLF protein to tra^S transcripts blocked this canonical splice acceptor site at the start of exon 2 and forced acceptance of the downstream cryptic splice site by the splicing machinery. The resulting removal of the stop codon from tra^S transcripts converted this null allele into one that produced active TRA protein. We assume that this TRA protein, although somewhat shorter at its N terminus, retained normal activity, as the product of the contemporary Drosophila *tra* gene does.

The Sxl^F mutant in heterozygous form (Sxl^F/Sxl^+) had the following consequences for sexual phenotype and fitness (Table 2). In tra^+/tra^S female heterozygotes, the splicing-mediated conversion of the tra^S allele from a null to a functional form allowed both *tra* alleles to produce TRA protein. In tra^S/tra^S homozygotes, both alleles also produced TRA protein, causing feminization, given that the amount of TRA was similar to that found in Sxl^F/Sxl^+; tra^+/tra^S heterozygotes. This is a reasonable assumption as heterozygotes for null alleles of *Sxl* in *D. melanogaster*, where both *tra* alleles carry stop codons, are female.

The Sxl^F allele spread if selection favored an increase in TRA production in females (*i.e.*, g, $g2 > 0$, Table 2). But owing to the pattern of inheritance, this caused a large increase only of tra^S homozygous females. The reason for this is that Sxl^F/Sxl^+; tra^S/tra^+ females generate equal numbers of Sxl^F/Sxl^+; tra^S/tra^+ and Sxl^F/Sxl^+; tra^S/tra^S daughters, whereas Sxl^F/Sxl^+; tra^S/tra^S females produce only Sxl^F/Sxl^+; tra^S/tra^S female offspring. The spread of the Sxl^F allele thus caused the elimination of the tra^+ allele and fixation of tra^S, resulting in a population of Sxl^+/Sxl^+; tra^S/tra^S males producing no TRA and of Sxl^F/Sxl^+; tra^S/tra^S females producing TRA from both tra^S alleles. Assuming that *Sxl* and *tra* were unlinked, as they are in *D. melanogaster*, our simulations show that the frequency of Sxl^F rises to 0.25.

The addition of Sxl^F as an upstream regulator of *tra* boosted female fitness at no cost to males. Females remained the heterogametic sex. We note, however, that *Sxl* is a poor splice enhancer of *tra* in *D. melanogaster*, with over half of *tra* mRNAs in females retaining the stop codon in exon 2 (Belote *et al.* 1989). Thus, females would have benefited from two doses of *Sxl*. But this genotype could not arise in this system, because Sxl^F heterozygotes were females, as the equivalent genotype is in *D. melanogaster* today (Cline and Meyer 1996). Thus homozygous Sxl^F individuals could not have been generated. This imbalance led to the next evolutionary step.

Addition of a stop codon in *Sxl*

In *D. melanogaster*, *Sxl*, like *tra*, has an exonic stop codon (in exon 3), which is skipped in females, but prevents production of SXL protein in males (see Figure 1). Sex-specific alternative splicing and stop codons are absent from copies of *Sxl* in other insects in which *Sxl* has no role in sex determination (Muller-Holtkamp 1995; Sievert *et al.* 1997; Meise *et al.* 1998; Saccone *et al.* 1998). This implies that a stop codon mutation was added to *Sxl* in the lineage leading to Drosophila. We assume this occurred after Sxl^F was recruited to the sex-determination pathway. Our hypothesis is that the stop mutation, which we call Sxl^{FS}, initially promoted female fitness. This effect is paradoxical at first glance, as a stop mutation is expected to reduce protein production. But, as we will see, the ability of *Sxl* to autoregulate by self-splicing allowed the evolution of an allele with stop codons, which resulted in SXL protein being produced

TABLE 3

Fitness of Sxl^{FS} stop mutants

Genotype	Dose of active *Sxl*	Sex	Fitness		Comments
			Male	Female	
Sxl^{+}/Sxl^{+}	0	Male	1		Original male
Sxl^{+}/Sxl^{F}	1	Female		1	Original female
Sxl^{+}/Sxl^{FS}	0	Male	1		Stop mutant male
Sxl^{F}/Sxl^{FS}	2	Female		$1 + h$	More SXL, better female
Sxl^{FS}/Sxl^{FS}	0	Male	1		*Sxl* stop homozygote male

from both copies of *Sxl* in females and by neither copy in males.

We make the key assumption that *Sxl* autoregulation arose prior to the origin of the stop mutation. Specifically, the mutation in Sxl^{F} that allowed recognition of *tra* poly(U)-rich sequences also permitted the recognition of poly(U)-rich sequences within Sxl^{F} transcripts. Such poly(U)-rich tracts occur in *D. melanogaster* in both introns that neighbor the exon with the stop codons. These attract SXL binding and the splicing out of this exon in *Sxl* transcripts (Horabin and Schedl 1993). We propose that the ancestral Sxl^{+} allele lacked these poly(U)-rich sequences as autoregulation of *Sxl* is not known outside the Drosophilids. In the Drosophila lineage, these sequences were introduced and amplified, perhaps by replication slippage or unequal recombination, after the Sxl^{F} allele had evolved. This hypothesis can be tested by examination of the genomic sequence of *Sxl* in non-Drosophilids. While the exact timing of the proposed event is not crucial, we must postulate that some poly(U) sequences existed in Sxl^{F} and allowed self-splicing of exon 3 from Sxl^{F} transcripts prior to the origin of the *Sxl* stop mutation. The introduction of autoregulation presumably was advantageous to females carrying the Sxl^{F} allele, but made no distinct change to the sex-determining system.

Consider the effect of a stop mutation Sxl^{FS}. The mutation arose in the Sxl^{F} allele in a female, which then generated Sxl^{+}/Sxl^{FS} heterozygotes (Table 3). These genotypes produced no SXLF protein as Sxl^{FS} transcripts carry the stop codon and so were converted into males. But when mated to Sxl^{+}/Sxl^{F} females, some of the progeny were Sxl^{F}/Sxl^{FS} heterozygotes. These female individuals were favored by selection as they produced SXLF protein from both alleles. SXLF was constitutively produced from the Sxl^{F} allele and this protein spliced out the stop-containing exon from Sxl^{FS} transcripts. In turn, these females could have mated with Sxl^{+}/Sxl^{FS} males to generate Sxl^{FS} homozygotes, which produced only transcripts carrying the stop codon. These individuals were male. Selection in favor of Sxl^{FS} in females (Table 3) led to its spread, replacing Sxl^{+} and resulting in an equilibrium frequency of 0.75 with Sxl^{F}/Sxl^{FS} females and Sxl^{FS}/Sxl^{FS} males.

It is worth emphasizing the counterintuitive nature of the Sxl^{FS} stop mutation. Its spread boosts SXLF production in females, now from both *Sxl* alleles, without any cost to males, who continue to produce only nonfunctional *Sxl* transcripts. The spread of Sxl^{FS} again causes no change in heterogamety, with females remaining the heterogametic sex. Note that we assume that the original version of Sxl^{+} did not have the ability to autoregulate through self-splicing. If it did, a stop mutation would not have conferred any benefit, as Sxl^{+}/Sxl^{FS} mutants would have produced SXLF protein and thus have been female. In this situation, homozygous Sxl^{FS} males could not have been generated.

The early promoter of *Sxl* and the recruitment of *sis*

In *D. melanogaster* today, all copies of *Sxl* carry a stop codon in exon 3 (Figure 1). In females, the stop codon is removed, allowing the production of SXL protein. This process is initiated by the turning on of the early promoter (*Pe*) of *Sxl* in females due to a twofold higher dose of the X-linked numerator genes *sisA*, *sisB*, *sisC*, and *run*. We envisage that the early promoter evolved first, followed by the recruitment of a numerator gene as an upstream control element. Once this system of genetic control was established, further numerators were added.

We can reconstruct these evolutionary steps by considering a mutation arising in the first intron of Sxl^{FS}, which generated a new early promoter site, *Pe*. We designate this allele Sxl^{FSPe}. For simplicity, we consider a single numerator gene, *sis*; the fact that several *sis* genes are required today does not demand a qualitatively different argument. Since the numerator genes of *D. melanogaster* have developmental roles other than sex determination (Cline and Meyer 1996; Walker *et al.* 2000), we assume that the sis^{+} allele was originally at fixation and acted as a general transcription factor with expression early in development. We further assume that the SIS protein also bound to the *Pe* promoter, leading to the early production of Sxl^{FSPe} transcripts. As today, mRNAs deriv-

TABLE 4

Fitness of the early promoter *Pe* of *Sxl*

Sex	Genotype		Fitness		Comments
	sis	*Sxl*	Male	Female	
Female	sis^+/sis^+	Sxl^F/Sxl^{FS}		1	Original female
Male		Sxl^{FS}/Sxl^{FS}	1		Original male
Female		Sxl^F/Sxl^{FSPe}		$1 + j$	*Pe* active, better female
Female		Sxl^{FS}/Sxl^{FSPe}		$1 + j$	*Pe* active, better female

ing from the *Pe* promoter lacked exons 2 and 3, which led to an early burst of SXLF protein. This was sufficient to activate the autoregulatory loop and to maintain the production of SXLF protein when the maintenance promoter, *Pm*, took over.

The Sxl^{FSPe} allele was initially a heterozygote with either Sxl^F or Sxl^{FS}. These genotypes received an early burst of SXLF protein from the Sxl^{FSPe} allele and later produced SXLF from both alleles when *Pm* took over (Table 4). Both genotypes were female. Selection favored them as they produced higher titers of SXLF protein. As males were Sxl^{FS} homozygotes, the Sxl^{FSPe} allele was thereafter found only in Sxl^{FSPe}/Sxl^{FS} heterozygous females. Given that selection favored these females (*i.e.*, $j > 0$, Table 4), the Sxl^{FSPe} allele spread and replaced *SxlF*, the allele that did not contain a stop codon. After this point, females were Sxl^{FSPe}/Sxl^{FS} heterozygotes and males were Sxl^{FS} homozygotes. We note that any mutation to create a *Pe* promoter in the Sxl^F allele would have been selectively neutral, as the Sxl^F allele already produced SXLF protein from the *Pm* promoter. For selection to have favored *Pe*, it must have occurred in an allele that already contained a stop codon.

We can now see how *sis* alleles were recruited as regulators of *Sxl* at the top of the sex-determination pathway. We assume that the *sis* locus was linked to the same chromosome as *Sxl* and that there was no recombination in males, as seen today in *D. melanogaster*. The effect of *sis* alleles is dose dependent in *D. melanogaster* (Cline 1993), so we further assume that sis^- null mutants failed to activate the early promoter *Pe* and so acted as dominant masculinizers (Table 5). The original sis^- null mutant could have arisen in linkage with either the Sxl^{FSPe} or the Sxl^{FS} allele. For simplicity, we consider a mutation in linkage with Sxl^{FSPe}, which occurred in the germline of a female (see appendix for the alternative pattern of linkage). This mutant chromosome produced a sis^- Sxl^{FSPe}/sis^+ Sxl^{FS} male because a single dose of sis^+ does not activate *Pe*. Matings of this male with standard sis^+ Sxl^{FSPe}/sis^+ Sxl^{FS} females gave rise to sis^- Sxl^{FSPe}/sis^+ Sxl^{FSPe} offspring, which were males homozygous for Sxl^{FSPe}. Further "backcrosses" of these males with nonmutant females in the F_3 and in later generations produced females homozygous for sis^+ Sxl^{FSPe}. Assuming that two active early promoters were advantageous to females, because they established the autoregulatory loop more quickly or with greater reliability ($k > 0$, Table 5), the Sxl^{FSPe} allele could spread. In turn, this selected for the sis^- allele because of its male-determining activity when combined with Sxl^{FSPe} homozygosity.

The final system consisted of females that were sis^+ Sxl^{FSPe}/sis^+ Sxl^{FSPe} and males that were sis^- Sxl^{FSPe}/sis^+ Sxl^{FSPe}. The Sxl^{FS} allele was lost, the Sxl^{FSPe} allele went to fixation, and the *sis* locus took over as the upstream regulator of sex determination. This set the stage for the last step, the degeneration of the chromosome carrying sis^- to become the Y (see discussion). The sis^- allele constitutes a second paradoxical mutation in our scheme; sis^- initially promoted the development of one sex (male), but ultimately contributed to improving the fitness of the other sex (female).

To sum up, our model proposes that the major features of *Sxl* as seen in Drosophila sex determination evolved in the following order: the ability of SXL protein to bind poly(U) in *tra*, *Sxl* autoregulation, the addition of a stop codon, followed by that of the early promoter, and then the recruitment of a *sis* null as an upstream regulator of *Sxl*. This final step led to a change from female heterogamety to homogametic sis^+/sis^+ females and heterogametic sis^-/sis^+ males. All these changes served to limit and strengthen the autoregulatory loop of *Sxl* in females. The reliability of this signal was augmented by the recruitment of more numerator genes with mutant null alleles linked to the original *sis* null and of an autosomal denominator gene that acted as an antagonist to the numerators. The introduction of a denominator converted the analog system (more or less SIS protein) into a digital system (active or inactive SIS protein), thereby improving the reliability of *Sxl* regulation, *i.e.*, "on" in females and "off" in males (Cline 1993).

DISCUSSION

Our reconstruction sets out a hypothesis for the evolution of the gene network that determines sex in Drosophila. We assume that the system evolved from an ancestral state in which *dsx* provided the discriminatory

TABLE 5

Fitness of *sis* null mutants

	Genotype		Fitness		
Sex	*sis*	*Sxl*	Male	Female	Comments
Male	sis^+/sis^+	Sxl^{FS}/Sxl^{FS}	1		Original male
Female		Sxl^{FS}/Sxl^{FSPe}		1	Original female
Female		Sxl^{FSPe}/Sxl^{FSPe}		$1 + k$	Two doses of *Pe*, better female
Male	sis^-/sis^+	Sxl^{FS}/Sxl^{FS}	1		Same as original male
Male		Sxl^{FS}/Sxl^{FSPe}	1		*Pe* inactive, male
Male		Sxl^{FSPe}/Sxl^{FSPe}	1		*Pe* inactive, male

signal and a *tra* allele without stop codons facilitated female-specific splicing of dsx^+ transcripts. There followed a series of gene recruitments and major transitions at the top of the pathway in the following order (Figure 3): (1) a stop codon in *tra*, which created a null allele; (2) the recruitment of *Sxl* as an RNA-binding factor, which caused alternative splicing of *tra* transcripts and the removal of the exon containing the stop codon; (3) *Sxl* autoregulation through the presence of SXL-protein-binding sites in *Sxl* transcripts; (4) a stop

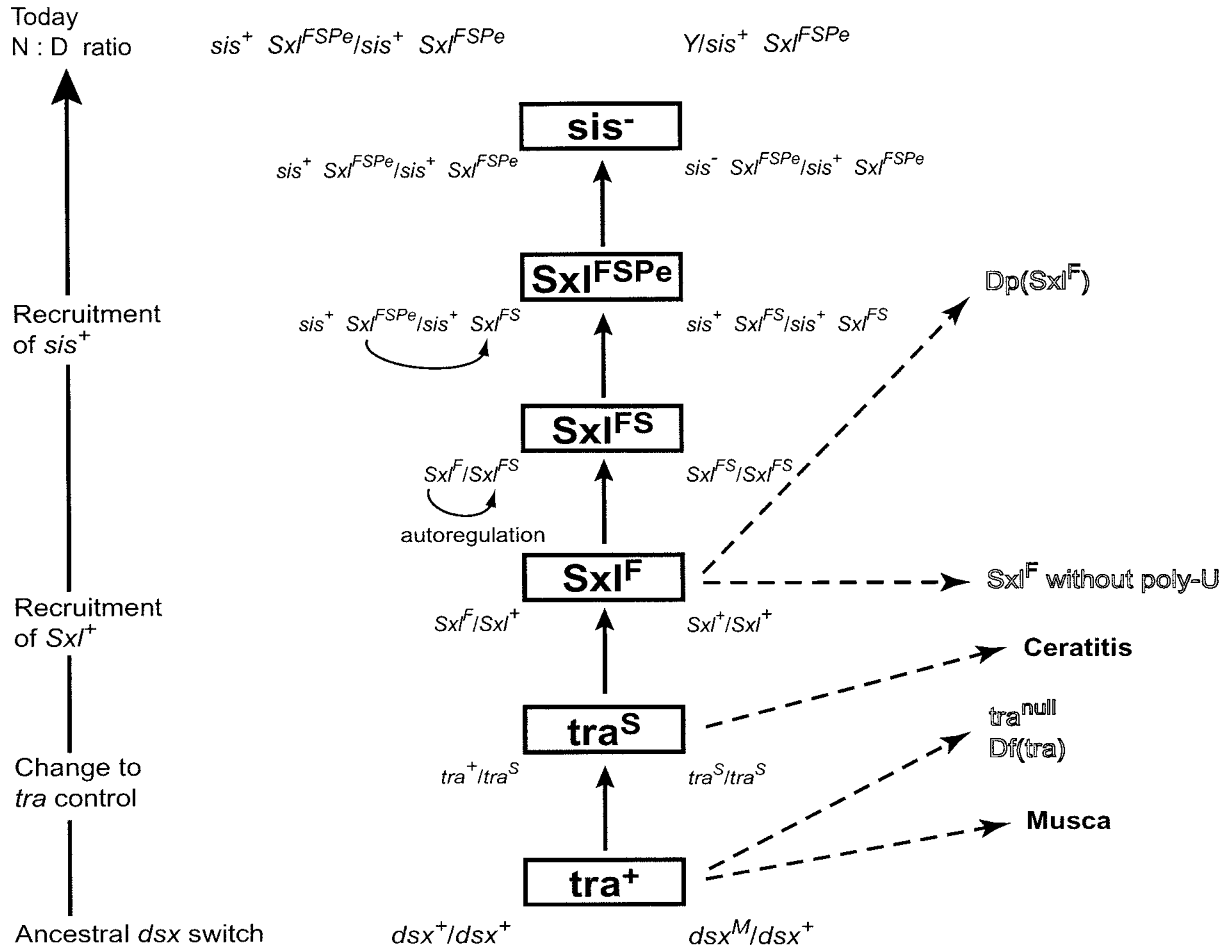

Figure 3.—Sequence of events leading to the Drosophila sex-determining pathway. Alleles in boxes are shown in the order in which they are proposed to have arisen. Below the boxes are the genotypes (in italics) that determine the sex at a given phase of evolution: females on the left side, males on the right side. Vertical solid arrows mark the direct phylogenetic route as described in the text. Broken arrows pointing to the right indicate possible side branches, that is, alternatives that were not taken or were discarded (gene symbols in outlined type) or may have led to systems operating in other Diptera. N, numerators; D, denominators; Df, deficiency; Dp, duplication

codon in *Sxl*, which created a null allele; (5) the presence of an early promoter of *Sxl* that was activated by binding of SIS transcription factor, leading to the early production of *Sxl* transcripts that lost the stop-containing exon; and (6) a null allele of *sis*, which caused dose-dependent activation of the early promoter of *Sxl*. Each of the transitions was favored because they increased the distinctiveness of gene expression in one sex or the other and hence the strength and reliability of the sex-determining signal. We postulate that sexual selection was the fundamental force driving these evolutionary changes in the sex-determining mechanism, as sexual selection leads to differential effects on the two sexes and is known to be a strong force that can yield rapid evolutionary change (Andersson 1994).

In the course of developing our scheme, we considered many other possibilities, which were rejected. The main criteria used for rejection were incompatibility with current knowledge about the sex-determining genes in *D. melanogaster* and parsimony. For example, we postulate that a null allele of the original numerator gene was unable to activate the early promoter of *Sxl*. This follows from the observation that the numerator genes in *D. melanogaster* act in such a dose-dependent manner (Cline 1993). We therefore rejected the alternative idea that there was active recruitment of a sis^+ allele to turn on the early promoter of *Sxl*.

We also rejected some hypotheses because they were implausible and required too many additional steps or failed to account for some aspects of the Drosophila pathway. For example, if the original Sxl^F mutant had been a recessive, Sxl^F/Sxl^+ heterozygotes would have been male (tra^S/tra^S) or female (tra^S/tra^+). Matings between these genotypes would have generated Sxl^F homozygotes that would have been female. Because Sxl^F homozygotes produce SXL protein from both alleles, we assume that they were favored by selection, and the spread of this allele would have resulted in the fixation of tra^S with Sxl^F/Sxl^F females and Sxl^F/Sxl^+ males. However, it is difficult to see how the *Sxl* stop codon could subsequently have been favored. This allele would have been a dominant male determiner, but in heterozygous or homozygous condition it would not have contributed any obvious benefit to male fitness. It is likewise difficult to see how autoregulation could then have been favored. Finally, the hypothesis of Sxl^F recessivity is contradicted by the fact that an active *Sxl* allele today is dominant over a null allele; so we would have to find, in addition, an explanation for a change in dominance.

In our presentation, we reconstructed a direct and short route from *dsx* to *Sxl* and the numerator/denominator system of Drosophila today (Figure 3). It is conceivable, even likely, that evolution has tried alternative routes. For example, a *tra* null or a deficiency for *tra* would have improved the original situation just as well as did a stop codon in *tra*. But with a *tra* null, there would have been no chance to recruit a correcting upstream regulator that could have turned *tra* "on" in females and "off" in males. Similarly, Sxl^F without poly(U)-binding sites would have helped females at that stage of evolution (Figure 3), but could not later have acquired autoregulation. A duplication of Sxl^F would also have increased the amount of SXL protein; but without the early promoter, it could not have come under transcriptional control by *sis*.

The linear pathway of Drosophila as it exists today is probably the stem left over from a "shrub" whose many side branches have disappeared or led to the pathways now encountered in other insects. In our view, the pathway has gone through many "trials and errors," with the outcome not being the optimal solution, but just the one that evolved through short-term advantage. Another area of uncertainty is the temporal sequence of the events underlying the evolution of the Drosophila sex-determining pathway. Were the steps evenly distributed or clumped in time? Did each step reach equilibrium before the next was initiated or did some changes occur simultaneously? For example, one can imagine that *Sxl* was recruited before its target tra^S allele had reached an equilibrium frequency of 0.75. It is beyond the scope of this article to deal with these possibilities and complications.

We have not considered the evolution of dosage compensation, which is also under the control of *Sxl*, through the repression in females of the male-essential gene *msl2* (Kelley *et al.* 1997). We believe that dosage compensation was added after the recruitment of *Sxl* and *sis* to the somatic sex-determination pathway. Dosage compensation is required once X and Y chromosomes have differentiated and the dose of X-linked genes needs to be equalized across the sexes (Charlesworth 1996). We propose that this occurred after the recruitment of *Sxl* and *sis*, both of which are X-linked in Drosophila today. The key was the functional loss of one or more of the numerator genes in males. As a dominant masculinizer, the chromosome carrying the sis^- null allele was limited to males, in which recombination is reduced or absent; thus, deleterious mutations could have accumulated on this chromosome, leading to its gradual degeneration until it eventually became the Y (Lucchesi 1978). Its homolog carrying the sis^+ allele became the X. The numerators themselves are transcription factors with important roles in processes other than sex determination; for example, *sisA* is involved in midgut formation (Walker *et al.* 2000), and *sisB* has multiple roles in bristle formation and neurogenesis (for a review see Cline and Meyer 1996). Thus, there must have been selective pressure for upregulation of the single copy of these numerators and other vital genes on the X chromosome in males. In *D. melanogaster*, this is achieved through the binding of the *msl* complex to the single X in males, which causes enhanced transcription (for review see Lucchesi 1996). In females, this complex does not form because *Sxl* inhibits *msl2*, an essential

component of the *msl* complex (Kelley *et al.* 1997). We have not discussed this step in detail, but these ideas could be tested by looking at the phylogenetic distribution of sex-specific expression in the numerators, *Sxl* and *msl* genes.

A second area that we have neglected is germline sex determination. This is much less well understood than somatic sex determination. Interestingly, *Sxl* is essential to oogenesis, but not via its downstream somatic targets, *tra* and *msl2*. In additon, *Sxl* in the germline is not regulated by the numerator/denominator ratio (for review see Steinmann-Zwicky 1994). As for dosage compensation, we suspect that these germline-specific features involving *Sxl* followed the recruitment of *Sxl* to the somatic sex-determination pathway. We should also note that some of the genes of the somatic pathway of *D. melanogaster* have other sex-determination functions; for example, *tra* controls sex-specific expression of *fru* in the central nervous system (Heinrichs *et al.* 1998). Again, we suspect that these are secondary adaptations and additions made once the basic pathway has been established. Future data from other insects may clarify this point.

In this article, we have proposed a step-by-step hypothesis for the evolution of the Drosophila sex-determination system from a hypothetical ancestral state. Comparative data were used to infer the ancestral state and the general order in which genes were added to the sex-determination pathway. In particular, comparative data support the assumption that *dsx* was ancestral, *tra* was added next, followed by *Sxl* and *sis*. In principle, we might have used an alternative approach. This would have been to deduce the evolutionary history from comparative analyses of related sex-determination systems, for example, those of other insects like *B. mori*, *C. capitata*, or *M. domestica*. The best-characterized is that of the Mediterranean fruit fly, Ceratitis (Pane *et al.* 2002), which shares many features with *D. melanogaster* (*tra* stop codons, *tra* control of *dsx*), but also differs in important aspects (no sex-determining role of *Sxl* or *sis*, dominant masculinizing factors, *tra* autoregulation, and possibly maternal initiation of *tra* autoregulation). We have not followed this approach because these other systems are far less well characterized, and it is difficult or impossible to infer the sequence of evolutionary events before we have more genetic and molecular information about sex determination in these related insect genera.

We have put forward our hypothesis and set out its assumptions so as to stimulate further research. Our evolutionary hypothesis is as simple as we could make it while being consistent with the known facts. Even though some or many of the details may well need to be refined in the light of subsequent comparative and experimental findings, we believe our general approach is a constructive one: proposing the order in which the major features of sex-determining genes were recruited and how selection favored these changes. This is an advance on an overly abstract evolutionary analysis that does not take into account the known facts about the gene networks involved. It also provides an informative perspective on the known developmental and molecular biology of Drosophila sex determination. This system is as complicated as it is not because of some intrinsic engineering constraints that require it to be so, but because it is a reflection of its evolutionary "bricolage" (Jacob 1977; Duboule and Wilkins 1998).

We thank Daniel Bopp for his fruitful discussion of our work. This interdisciplinary research was carried out through a number of meetings sponsored in part by the British Council and the Wissenschaftskolleg zu Berlin.

LITERATURE CITED

Andersson, M., 1994 *Sexual Selection*. Princeton University Press, Princeton, NJ.

Baker, B. S., and M. F. Wolfner, 1988 A molecular analysis of *double-sex*, a bifunctional gene that controls both male and female sexual differentiation in *Drosophila melanogaster*. Genes Dev. **2:** 477–489.

Belote, J. M., M. McKeown, R. T. Boggs, R. Ohkawa and B. A. Sosnowski, 1989 The molecular genetics of *transformer*, a genetic switch controlling sexual differentiation in *Drosophila*. Dev. Genet. **10:** 143–154.

Beye, M., M. Hasselmann, M. K. Fondrk, R. E. Page and S. W. Omholt, 2003 The gene *csd* is the primary signal for sexual development in the honeybee and encodes an SR-type protein. Cell **114:** 419–429.

Bopp, D., G. Calhoun, J. I. Horabin, M. Samuels and P. Schedl, 1996 Sex-specific control of *Sex-lethal* is a conserved mechanism for sex determination in the genus *Drosophila*. Development **122:** 971–982.

Bull, J. J., 1983 *Evolution of Sex Determining Mechanisms*. Benjamin/Cummings, Menlo Park, CA.

Charlesworth, B., 1996 The evolution of chromosomal sex determination and dosage compensation. Curr. Biol. **6:** 149–162.

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

Cline, T. W., 1993 The *Drosophila* sex determination signal: How do flies count to two? Trends Genet. **9:** 385–390.

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

Darwin, C. R., 1871 *The Descent of Man and Selection in Relation to Sex*. John Murray, London.

Dubendorfer, A., M. Hediger, G. Burghardt and D. Bopp, 2002 *Musca domestica*, a window on the evolution of sex-determining mechanisms in insects. Int. J. Dev. Biol. **46:** 75–79.

Duboule, D., and A. S. Wilkins, 1998 The evolution of "bricolage." Trends Genet. **14:** 54–59.

Erickson, J. W., and T. W. Cline, 1998 Key aspects of the primary sex determination mechanism are conserved across the genus *Drosophila*. Development **125:** 3259–3268.

Estes, P. A., L. N. Keyes and P. Schedl, 1995 Multiple response elements in the *Sex-lethal* early promoter ensure its female-specific expression pattern. Mol. Cell. Biol. **15:** 904–917.

Gebauer, F., M. Grskovic and M. W. Hentze, 2003 Drosophila Sex-lethal inhibits the stable association of the 40S ribosomal subunit with msl-2 mRNA. Mol. Cell **11:** 1397–1404.

Hediger, M., G. Burghardt, C. Siegenthaler, N. Buser, D. Hilfiker-Kleiner *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.

Heinrichs, V., L. C. Ryner and B. S. Baker, 1998 Regulation of sex-specific selection of *fruitless* 5′ splice sites by *transformer* and *transformer-2*. Mol. Cell. Biol. **18:** 450–458.

Hodgkin, J., 1992 Genetic sex determination mechanisms and evolution. BioEssays **14:** 253–261.

Horabin, J. I., and P. Schedl, 1993 Sex-lethal autoregulation requires multiple cis-acting elements upstream and downstream of the male exon and appears to depend largely on controlling the use of the male exon 5′ splice site. Mol. Cell. Biol. **13:** 7734–7746.
Jacob, F., 1977 Evolution and tinkering. Science **196:** 1161–1166.
Kelley, R. L., J. Wang, L. Bell and M. I. Kuroda, 1997 *Sex-lethal* controls dosage compensation in *Drosophila* by a non-splicing mechanism. Nature **387:** 195–199.
Lucchesi, J. C., 1978 Gene dosage and the evolution of sex chromosomes. Science **202:** 711–716.
Lucchesi, J. C., 1996 Dosage compensation in *Drosophila* and the 'complex' world of transcriptional regulation. BioEssays **18:** 541–547.
Marin, I., and B. S. Baker, 1998 The evolutionary dynamics of sex determination. Science **281:** 1990–1994.
Meise, M., D. Hilfiker-Kleiner, A. Dubendorfer, C. Brunner, R. Nöthiger *et al.*, 1998 *Sex-lethal*, the master sex-determining gene in *Drosophila*, is not sex-specifically regulated in *Musca domestica*. Development **125:** 1487–1494.
Muller-Holtkamp, F., 1995 The *Sex-lethal* gene homologue in *Chrysomya rufifacies* is highly conserved in sequence and exon–intron organization. J. Mol. Evol. **41:** 467–477.
Nagoshi, N. N., and B. S. Baker, 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 Dev. **4:** 89–97.
Nöthiger, R., and M. Steinmann-Zwicky, 1985 A simple principle for sex determination in insects. Cold Spring Harbor Symp. Quant. Biol. **50:** 615–629.
Nöthiger, R., M. Leuthold, N. Anderson, P. Gerschwiler, A. Gruter *et al.*, 1987 Genetic and developmental analysis of the sex-determining gene "double sex" (*dsx*) of *Drosophila melanogaster*. Genet. Res. **50:** 113–123.
Ohbayashi, F., M. G. Suzuki, K. Mita, K. Okano and T. Shimada, 2001 A homologue of the *Drosophila doublesex* gene is transcribed into sex-specific mRNA isoforms in the silkworm, *Bombyx mori*. Comp. Biochem. Physiol. **128:** 145–158.
Pane, A., M. Salvemini, P. Bovi, C. Polito and G. Saccone, 2002 The transformer gene in *Ceratitis capitata* provides a genetic basis for selecting and remembering the sexual fate. Development **129:** 3715–3725.
Penalva, L. O. F., H. Sakamoto, A. Navarro-Sabate, E. Sakashita, B. Granadino *et al.*, 1996 Regulation of the gene *Sex-lethal*: a comparative analysis of *Drosophila melanogaster* and *Drosophila subobscura*. Genetics **144:** 1653–1664.
Raymond, C. S., C. E. Shamu, M. M. Shen, K. J. Seifert, B. Hirsch *et al.*, 1998 Evidence for evolutionary conservation of sex-determining genes. Nature **391:** 691–695.
Raymond, C. S., M. W. Murphy, M. G. O'Sullivan, V. J. Bardwell and D. Zarkower, 2000 Dmrt1, a gene related to worm and fly sexual regulators, is required for mammalian testis differentiation. Genes Dev. **14:** 2587–2595.
Rice, W. R., 1984 Sex chromosomes and the evolution of sexual dimorphism. Evolution **41:** 911–914.
Rice, W. R., 1998 Male fitness increases when females are eliminated from gene pool: implications for Y chromosome. Proc. Natl. Acad. Sci. USA **95:** 6217–6221.
Saccone, G., I. Pelusio, D. Artiaco, E. Giordiano, D. Bopp *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.
Schutt, C., and R. Nöthiger, 2000 Structure, function and evolution of sex-determining systems in Dipteran insects. Development **127:** 667–677.
Shearman, D. C. A., and M. Frommer, 1998 The *Bactrocera tryoni* homologue of the *Drosophila melanogaster* sex-determination gene *doublesex*. Insect Mol. Biol. **7:** 355–366.
Shen, M. M., and J. Hodgkin, 1988 *mab-3*, a gene required for sex-specific yolk protein expression and a male-specific lineage in *C. elegans*. Cell **54:** 1019–1031.
Sievert, V., S. Kuhn and W. Traut, 1997 Expression of the sex determining cascade genes *Sex-lethal* and *doublesex* in the phorid fly *Megaselia scalaris*. Genome **40:** 211–214.
Sosnowski, B. A., J. M. Belote and M. McKeown, 1989 Sex-specific alternative splicing of RNA from the *transformer* gene results from sequence dependent splice site blockage. Cell **58:** 449–459.
Steinmann-Zwicky, M., 1994 *Sxl* in the germline of *Drosophila*: a target for somatic late induction. Dev. Genet. **15:** 265–274.
Sturtevant, A. H., 1945 A gene in *Drosophila melanogaster* that transforms females into males. Genetics **30:** 297–299.
Suzuki, M. G., F. Ohbayashi, K. Mita and T. Shimada, 2001 The mechanism of sex-specific splicing at the doublesex gene is different between *Drosophila melanogaster* and *Bombyx mori*. Insect Biochem. Mol. Biol. **31:** 1201–1211.
Walker, J. J., K. K. Lee, R. N. Desai and J. W. Erickson, 2000 The *Drosophila melanogaster* sex determination gene *sisA* is required in yolk nuclei for midgut formation. Genetics **155:** 191–202.
White, M. J. D., 1973 *Animal Cytology and Evolution*. Cambridge University Press, Cambridge, UK.
Wilkins, A. S., 1995 Moving up the hierarchy: a hypothesis on the evolution of a genetic sex determination pathway. BioEssays **17:** 71–77.

Communicating editor: T. Schüpbach

APPENDIX

Simulation technique: Simulations were carried out to follow the fate of mutant alleles, introduced at low frequency in an infinite diploid population with non-overlapping generations. The sex of the possible genotypes was defined (Tables 1–5). The next step was to apply selection and then to allow random mating between the survivors and the production of offspring for the next generation. So the flow of events iterated was zygotes → selection → reproduction. We initially assumed that different loci were not linked and so underwent free recombination in both sexes. Simulations were run until equilibrium conditions were reached (defined as decreasing changes in allele frequency of $<10^{-6}$/ generation at all loci).

Initial conditions were set up by finding the equilibrium gene frequency when there was only allelic variation at the original locus. A mutant allele was then introduced at the second locus at low frequency (usually 10^{-3}) and followed until equilibrium was reached. A range of values for the selective coefficients in Tables 1–5 (*i.e.*, m, $m2$, f, g, $g2$, h, j, and k) was examined to identify the general patterns reported in the text. Only positive values of the selection coefficients were considered. For simplicity, no conditions of heterozygote advantage were investigated (*i.e.*, $m2 \geq m$ in Table 1 and $g2 \geq g$ in Table 2).

Linkage: In *D. melanogaster*, *dsx* and *tra* are located on chromosome 3, and *Sxl* and *sis* are located on the X chromosome. Nothing is yet known about the linkage of *tra* and *dsx*, or of *Sxl* and *sis*, outside of the Drosophilids. In addition, in Drosophila and many, if not all, families of the higher Diptera, male meiosis is achiasmatic and recombination is thought to take place only in females (White 1973). In the following paragraph, we investigate how linkage and the absence of male recombination influenced the evolution of tra^S and sis^- alleles.

For *dsx-tra*, the evolutionary outcome depends on the initial linkage of the tra^S mutation. If the tra^S mutation occurs in linkage with dsx^+, then tra^S heterozygous males

(tra^S dsx^+/tra^+ dsx^M) and females (tra^S dsx^+/tra^+ dsx^+) will form, and matings between them will generate tra^S homozygous males. When rare, the conditions for the spread of the tra^S mutant linked to dsx^+ are approximately twice as restrictive as for the unlinked case. This can be explained as follows. When tra^S is linked to dsx^+, it can never occur in physical linkage with the dsx^M allele, as dsx^M is limited to males and there is no recombination in males. This means that the tra^S dsx^+ linked mutant is more likely to occur in females than in males. Specifically, when the tra^S mutant is rare, tra^S dsx^+/tra^+ dsx^M males typically mate with standard females (tra^+ dsx^+/tra^+ dsx^+). The offspring of this mating that carry the tra^S mutation are always female. The net effect is that the tra^S allele is about twice as likely to be in a female and so suffers about twice as much negative selection (in females) as positive selection (in males). In contrast, when tra^S is unlinked, it is equally likely to occur in males as in females because tra^S/tra^+; dsx^+/dsx^M males mated to standard females have offspring with the tra^S mutation that are equally likely to be males as females. Using the selective coefficients in Table 1, the condition for the spread of tra^S when linked to dsx^+ is approximately $m > 2f$. The alternative linkage pattern for tra^S is with dsx^M, which is discussed in the text.

For *Sxl-sis*, the evolutionary outcome is independent of the initial linkage of the sis^- mutation. We discuss linkage of sis^- to Sxl^{FSPe} in the text. Alternatively, this mutant may have occurred in linkage with the Sxl^{FS} allele. If we assume that this mutation arose in the germline of a female, then the offspring carrying the chromosome would have been male sis^- Sxl^{FS}/sis^+ Sxl^{FS}. Matings of this male with standard sis^+ Sxl^{FSPe}/sis^+ Sxl^{FS} females gave rise to sis^- Sxl^{FS}/sis^+ Sxl^{FSPe}, which were also male. Further "backcrosses" of these males with nonmutant females in the F_3 and in later generations produced females homozygous for sis^+ Sxl^{FSPe}, which are favored by selection (Table 5). This led to the spread of both the Sxl^{FSPe} and the sis^- alleles. At equilibrium males were sis^- Sxl^{FS}/sis^+ Sxl^{FSPe} and females were sis^+ Sxl^{FSPe}/sis^+ Sxl^{FSPe}. This differs from the case with original linkage to the Sxl^{FSPe} allele only in the retention of Sxl^{FS} in coupling with sis^- in males. When this chromosome degenerates to become the Y, Sxl^{FS} will have disappeared.