DOI: 10.1534/genetics.107.084244

Positive Selection at the Binding Sites of the Male-Specific Lethal Complex Involved in Dosage Compensation in Drosophila

Doris Bachtrog[1]

Department of Integrative Biology, University of California, Berkeley, California 94720

Manuscript received November 6, 2007
Accepted for publication August 9, 2008

ABSTRACT

In many taxa, males and females differ with respect to their sex chromosomes, and dosage compensation mechanisms have evolved to equalize X-linked gene transcription. In Drosophila, the male-specific lethal (MSL) complex binds to hundreds of sites along the male X chromosome and mediates twofold hypertranscription of the single male X. Two recent studies found evidence for lineage-specific adaptive evolution in all five core protein-coding genes of the MSL complex in *Drosophila melanogaster*. In particular, dramatic positive selection was detected in domains shown to be responsible for their specific targeting to the X chromosome. Here I use population genetics to show that three previously characterized MSL-binding DNA segments on the X themselves underwent adaptive evolution in *D. melanogaster*, but not in its close relatives *D. simulans* and *D. yakuba*. MSL components have been shown to not correctly target the *D. melanogaster* X chromosome in hybrids between *D. melanogaster* and *D. simulans*. My finding supports the idea of selection-driven coevolution among DNA-protein interactions of the dosage compensation machinery and suggests that misregulated dosage compensation could contribute to male hybrid inviability in Drosophila.

In many eukaryotic organisms, the two sexes have different chromosomal configurations. Typically, one sex (called heterogametic) has a pair of morphologically different chromosomes, whereas the other sex (called homogametic) has two identical members of each chromosomal pair. Morphologically and genetically distinct X and Y chromosomes involved in sex determination have evolved independently many times in both animals and plants (Bull 1983; Charlesworth 1996; Rice 1996). It is generally believed that the X and Y chromosomes originated from an initially identical chromosomal pair, with morphological differentiation being a byproduct of the degeneration of the chromosome that is present only in the heterogametic sex and thus is completely sheltered from genetic recombination (*i.e.*, the Y chromosome in species with XY males and XX females). Indeed, a striking common feature of many taxa is the almost complete erosion of genes from the Y chromosome.

The degeneration of the Y chromosome creates the problem of reduced gene dosage of X-linked genes in males. While females have two doses of each X-linked gene, males—after the Y-linked copy degenerates—only have one. Many genes on the X chromosome, however, will frequently have the same optimal level of expression in both sexes. This implies strong selective pressure to evolve compensatory mechanisms to equalize the level of products of X-linked genes in both sexes. Indeed, in several cases, the possession of genetically eroded Y chromosomes is known to be associated with dosage compensation, such that the activity of most X-linked genes is effectively the same in males and females (Lucchesi 1978). This phenomenon was first discovered in *Drosophila melanogaster* (Bridges 1922) and termed "dosage compensation" (Muller *et al.* 1931).

In *D. melanogaster*, the mechanism of dosage compensation and its components have been extensively studied. Dosage compensation in flies occurs by doubling the transcription rate of X-linked genes in males (Baker *et al.* 1994). This is mediated by the protein product of five known genes: *maleless* (*mle*), the *male-specific lethal* (*msl*) genes *msl-1*, *msl-2*, and *msl-3*, and *males absent on the first* (*mof*), together with the noncoding RNA products of two additional genes, *roX1* and *roX2* (RNA on the X) (Baker *et al.* 1994; Marín *et al.* 2000; Meller 2000). The products of all these genes function together in a ribonucleoprotein complex, termed the male-specific lethal (MSL) complex, to mediate dosage compensation by altering the chromatin structure of the X chromosome in males (Baker *et al.* 1994; Marín *et al.* 2000; Meller 2000). The MSL complex is reproducibly associated with specific sets of hundreds of positions along the length of the polytenized salivary gland X chromosome in males (Baker *et al.* 1994; Marín *et al.* 2000; Meller 2000).

The processes that generate the observed distribution of the MSL complex along the male X chromosome and the molecular nature of the MSL-binding sites remain mysterious. Recent ChIP-chip analyses have revealed sev-

[1]*Address for correspondence:* Department of Integrative Biology, University of California, 3060 Valley Life Sciences Bldg., Berkeley, CA 94720. E-mail: dbachtrog@berkeley.edu

eral hundred regions where the MSL complex is bound along the male X chromosome (Alekseyenko *et al.* 2006; Dahlsveen *et al.* 2006; Gilfillan *et al.* 2006). To date, only three such binding sites have been characterized on the X of *D. melanogaster* in detail. Two of these sites correspond to parts of the *roX* genes themselves on the *D. melanogaster* X chromosome (Kageyama *et al.* 2001; Park *et al.* 2003). The third site is an intergenic region from the *18D10* region of the X chromosome (Oh *et al.* 2004). When a *roX* or *18D10* transgene is inserted on an autosome, the MSL complex is recruited to this site, and neighboring sites sometimes show MSL binding as well (Kelley *et al.* 1999; Oh *et al.* 2003).

Interestingly, two recent studies found evidence for adaptive evolution in all five (or four of five) protein-coding genes of the MSL complex in *D. melanogaster* (Levine *et al.* 2007; Rodriguez *et al.* 2007). In particular, dramatic positive selection was detected in MSL1 and MSL2 protein domains shown to be responsible for their specific targeting to the X chromosome (Levine *et al.* 2007; Rodriguez *et al.* 2007). This signature of positive selection at an essential protein-DNA interface of the complex was surprising and may suggest that X chromosomal MSL-binding DNA segments themselves are changing rapidly. Here, I study patterns of molecular evolution and population variation of the *roX* genes and the putative binding site *18D10* in *D. melanogaster* populations to test whether the MSL-binding sites are indeed under positive selection in the *D. melanogaster* lineage. For comparisons, the same regions were sequenced in *D. simulans*, where no evidence of adaptive protein evolution was detected (Levine *et al.* 2007; Rodriguez *et al.* 2007), and *D. yakuba* populations, a species where we have no *a priori* expectation of positive selection operating on the MSL-binding sites. Indeed, I find various signatures of positive selection acting on the MSL-binding sites in *D. melanogaster*, but little evidence for adaptive evolution in *D. simulans* or *D. yakuba*. This finding supports the idea of selection-driven coevolution among DNA-protein interactions of the dosage compensation machinery in the *D. melanogaster* lineage.

MATERIALS AND METHODS

Data collection: For the sequence variability study, 24 lines of a *D. melanogaster* population from Zimbabwe (collected by B. Ballard in 2002), 24 lines of a *D. simulans* population from Madagascar (Dean and Ballard 2004), and 24 lines of a *D. yakuba* population from Cameroon (collected by the author and P. Andolfatto in 2002) were used. Genomic DNA was extracted from a single male of each line using the Puregene DNA extraction kit (Puregene). PCR products were amplified as ~1400-bp fragments from genomic DNA using primers and conditions available on request. PCR products were used as sequencing templates after treatment with the SAP/EXO reagent. Gene-specific internal primers and the original amplification primers were used for sequencing with the BigDye 3.0 cycle sequencing kit (Applied Biosystems, Foster City, CA) following the manufacturer's protocol. Sequences were run on an ABI 3730 automated sequencer, and each fragment was sequenced at least once on both strands. I obtained polymorphism data for three regions located on the X chromosome that contain the *roX1* binding site (Kageyama *et al.* 2001; Park *et al.* 2003), the *roX2* binding site (Kageyama *et al.* 2001; Park *et al.* 2003), and the *18D10* binding site (Oh *et al.* 2004) and adjacent flanking regions. For the two *roX* genes, the sequenced flanking region is part of the transcribed noncoding *roX* RNA, while it consists of intergenic sequence for *18D10*. Binding and nonbinding regions were defined by experimental data derived from DNase protection assays and genetic assays defining minimal fragments necessary for location of the MSL complex *in vivo* (Kageyama *et al.* 2001; Park *et al.* 2003; Oh *et al.* 2004). For locus- and species-specific estimates of variability see Table 1. For comparison, polymorphism data of six previously studied regions from the same populations of *D. melanogaster*, *D. simulans*, and *D. yakuba* were included (Andolfatto 2005; Bachtrog *et al.* 2006; Haddrill *et al.* 2008), consisting of a total of 4.3 kb of protein-coding sequence.

DNA polymorphism, divergence, and evolutionary analysis: Nucleotide sequences were edited using Sequencher 4.1 (Gene Codes, Ann Arbor, MI) and aligned by eye. Multiple hit sites were included in the analysis, but polymorphisms overlapping indel variation was excluded. Population-genetic parameters were estimated on the basis of the number of silent polymorphic sites and the average silent pairwise diversity. All summary statistics were calculated using Perl scripts written by the author and P. Andolfatto. A maximum likelihood version of the HKA test was used to test for departures from neutrality (Wright and Charlesworth 2004). Synonymous sites, while probably themselves under selection in some Drosophila lineages, are used as a neutral standard to detect selective constraints at noncoding DNA since they appear to evolve under less constraint than synonymous sites (Andolfatto 2005). For lineage-specific estimates of divergence, I reconstructed a *D. melanogaster–D. simulans* ancestor sequence (ANC) by maximum likelihood, using *D. yakuba* as an outgroup, as implemented in the *baseml* (for noncoding regions) or *codeml* (for coding regions) programs of PAML (Yang 1997). To calculate lineage-specific estimates of adaptive evolution at the MSL-binding sites, the software package *DoFE* (kindly provided by A. Eyre-Walker) was used, using divergence estimates to the reconstructed ancestral sequence.

RESULTS

I sequenced ~1.4 kb of each of the three binding sites, and its surrounding genomic regions, from multiple strains of a *D. melanogaster*, *D. simulans*, and *D. yakuba* population. The *roX* genes and the *18D10* region have been extensively studied experimentally (Kageyama *et al.* 2001; Park *et al.* 2003; Oh *et al.* 2004), which allows me to divide the fragments into binding and nonbinding regions. Summary statistics for polymorphisms and divergence at these regions are given in Table 1. For comparison, synonymous diversity at six protein-coding X-linked loci studied previously in the same lines of *D. melanogaster*, *D. simulans*, and *D. yakuba* are given (Andolfatto 2005; Bachtrog *et al.* 2006; Haddrill *et al.* 2008). Diversity is reduced at the three noncoding fragments studied here compared to synonymous site polymorphism (Table 1). In particular, levels of polymorphisms are 2- to 10-fold lower in the three noncoding regions studied, compared to polymorphism levels at synonymous sites (Table 1). Likewise, levels of divergence between species are lower at the noncoding

TABLE 1

Polymorphism (π) and divergence (D_{xy}) of MSL-binding regions in *D. melanogaster*, *D. simulans*, and *D. yakuba* (%)

		D. melanogaster			*D. simulans*			*D. yakuba*		
		π	D_{xy}[a]	D_{xy}ANC[b]	π	D_{xy}[a]	D_{xy}ANC[b]	π	D_{xy}[a]	D_{xy}ANC[b]
18D10	Sequenced region (1.2 kb)	0.89	8.18	2.65	1.49	7.35	3.43	0.57	17.10	10.55
	Binding site (0.6 kb)	0.22	6.44	1.78	1.21	4.59	2.28	0.50	12.25	7.56
	Nonbinding region (0.6 kb)	1.53	9.85	3.51	1.77	10.26	4.65	0.63	21.40	13.22
roX1	Sequenced region (1.3 kb)	1.15	6.57	4.27	1.14	6.97	2.74	0.15	12.19	9.63
	Binding site (0.2 kb)	0.26	5.09	3.26	0.11	4.96	1.78	0.04	5.36	3.53
	Nonbinding region (1.1 kb)	1.35	6.89	4.49	1.36	7.41	2.95	0.17	13.75	11.02
roX2	Sequenced region (1.3 kb)	0.24	5.60	2.92	0.96	5.44	1.83	0.41	13.64	10.49
	Binding site (0.1 kb)	0.06	8.59	5.34	0.91	4.87	3.02	0.06	13.82	10.34
	Nonbinding region (1.2 kb)	0.26	5.26	2.63	1.02	6.02	1.69	0.45	13.62	10.51
X-linked	(6 loci, 4 kb)	2.97	14.68	8.71	2.51	14.57	5.36	1.50	28.57	21.76

For comparison, synonymous site diversity of 6 X-linked protein-coding genes sequenced in the same populations is given. Binding and nonbinding regions were defined by experimental data derived from DNase protection assays and genetic assays defining minimal fragments necessary for location of the MSL complex *in vivo* (Kageyama *et al.* 2001; Park *et al.* 2003; Oh *et al.* 2004). The sequenced flanking region (nonbinding region) is part of the transcribed noncoding *roX* RNA for *roX1* and *roX2*, while it consists of intergenic sequence for *18D10*.

[a] D_{xy} is the weighted average pairwise divergence per site (%) to an outgroup species (*D. simulans* for the *D. melanogaster* and *D. yakuba* populations and *D. melanogaster* for the *D. simulans* population), corrected for multiple hits (Jukes–Cantor).

[b] D_{xy}ANC is the weighted average pairwise divergence per site (%) to a reconstructed ancestral species, corrected for multiple hits (Jukes–Cantor).

segments than at synonymous sites (Table 1). Reduced levels of polymorphism and divergence emerge as a general property of noncoding DNA in Drosophila (Andolfatto 2005; Haddrill *et al.* 2005, 2008; Bachtrog and Andolfatto 2006; Halligan and Keightley 2006), reflecting functional constraints on noncoding DNA. However, levels of polymorphism at the binding sites themselves are even lower compared to their flanking genomic regions, while divergence appears reduced to a lesser extent (Table 1).

Reduced diversity at MSL-binding sites relative to divergence could result from adaptive evolution at these regions. To investigate this possibility, the Hudson–Kreitman–Aguade (HKA) test (Hudson *et al.* 1987) can be used to test for heterogeneity in levels of polymorphism and divergence among unlinked genomic regions. A maximum-likelihood approach of the HKA test (Wright and Charlesworth 2004) was used to compare rates of interspecies divergence and intraspecies polymorphism at the six protein-coding loci and the three MSL-binding sites. I find that variation is significantly reduced at the MSL-binding sites in *D. melanogaster* compared to synonymous sites, given levels of divergence (Table 2). In contrast, a model of neutral evolution at the MSL-binding sites cannot be rejected for *D. simulans* and *D. yakuba* (Table 2). Thus, the results from the HKA test are compatible with recent positive selection at the MSL-binding sites in *D. melanogaster*, but not its close relatives *D. simulans* and *D. yakuba*.

I also tested for positive selection using the McDonald–Kreitman (MK) test (McDonald and Kreitman 1991). This test was originally designed to test for selection at

TABLE 2

HKA tests on synonymous polymorphism and divergence at six X-linked markers *vs.* MSL-binding sites, using maximum likelihood

Model	ln L	*k* (18D10)	*k* (roX1)	*k* (roX2)	Likelihood-ratio statistic (d.f.)	*P*-value
D. melanogaster neutral ($k = 1$)	−56.7	1	1	1	13.4 (3)	0.004*
Selection on MSL-binding sites	−50.0	0.35	0.15	0.16		
D. simulans neutral ($k = 1$)	−57.1	1	1	1	4.6 (3)	0.204
Selection on MSL-binding sites	−54.9	1.78	0.26	1.53		
D. yakuba neutral ($k = 1$)	−53.9	1	1	1	6.9 (3)	0.075
Selection on MSL-binding sites	−50.4	0.71	0.22	0.09		

D. simulans was used as the outgroup for the *D. melanogaster* or *D. yakuba* populations, and *D. melanogaster* was used as the outgroup for the *D. simulans* population. The selection parameter *k* indicates how much diversity is elevated at the locus over neutral expectation given levels of divergence. *$P < 0.05$.

TABLE 3

McDonald–Kreitman tests at MSL-binding sites *vs.* flanking regions

	Pairwise divergence					Lineage-specific divergence				
	Binding site		Nonbinding site			Binding site		Nonbinding site		
	P	*D*	*P*	*D*	*P*-value	*P*	*D*_ANC	*P*	*D*_ANC	*P*-value
D. melanogaster										
18D10	10	35	39	49	0.009*	10	9	33	13	0.117
roX1	2	11	71	58	0.006*	2	7	67	35	0.014*
roX2	1	10	25	56	0.122	1	6	25	27	0.097
Combined	13	56	135	163	<0.001*	13	22	125	75	0.005*
D. simulans										
18D10	39	19	70	45	0.838	32	7	58	16	0.757
roX1	3	10	79	61	0.021*	3	3	78	20	0.12
roX2	35	24	39	33	0.779	11	3	60	12	0.459
Combined	77	53	188	139	0.671	46	13	196	48	0.403
D. yakuba										
18D10	15	59	24	114	0.760	15	37	23	71	0.782
roX1	1	11	11	132	0.766	1	7	11	107	0.831
roX2	1	17	41	137	0.068	1	13	40	104	0.078
Combined	17	87	76	383	0.545	17	57	74	282	0.722

P, number of polymorphic sites; *D*, number of divergent sites from an outgroup species (*D. simulans* for the *D. melanogaster* and *D. yakuba* populations and *D. melanogaster* for the *D. simulans* population); *D*_ANC is the number of divergent sites from a reconstructed outgroup species. *P*-values are from Fisher's exact test. $*P < 0.05$.

protein-coding regions, and examines whether the ratio of synonymous polymorphism and fixed differences equals that for replacement changes, as expected under neutral evolution. This test can be modified to test for selection in noncoding DNA (Ludwig and Kreitman 1995; Kohn *et al.* 2004; Andolfatto 2005). Here, I test whether the MSL-binding sites show excess fixed differences relative to polymorphisms, compared to their flanking genomic region (Table 3). Note that the MK test assumes no recombination between selected and neutral sites and is also valid for completely unlinked sites. Neither assumption is likely in the case of binding sites and their flanking regions, and the *P*-values for MK tests are expected to be slightly larger if there is partial linkage within loci (Andolfatto 2005). Nevertheless, I find strong support for adaptive evolution at the MSL-binding sites in *D. melanogaster*, both for individual loci and the combined data set (2/3 loci show evidence of positive selection and only the *roX2* binding site, the smallest binding region analyzed, is marginally not significant), but little evidence for adaptive evolution in *D. simulans* (1/3 loci shows evidence of positive selection) and *D. yakuba* (0/3 loci show evidence of positive selection; Table 3). Thus, the MK test also suggests adaptive evolution acting on the MSL-binding sites in *D. melanogaster*, but little in *D. simulans* or *D. yakuba*. A similar pattern emerges if lineage-specific estimates of divergence are used (Table 3).

I also use an extension of the MK test to estimate α, the fraction of nucleotides in the MSL-binding sites that have been fixed by positive selection. I use flanking regions as a neutral standard and employ three slightly different approaches to obtain lineage-specific estimates of α (using divergence to the reconstructed ancestral sequence). I find that between 59 and 73% of substitutions in MSL-binding sites along the *D. melanogaster* lineage were driven by adaptive evolution (95% confidence intervals 38–88%, Table 4). In contrast, there is little evidence for positive selection driving fixations at the MSL-binding sites in either *D. simulans* (α ranging from 6 to 17% with all 95% C.I. overlapping zero) or *D. yakuba* (see Table 4). Similar estimates of α are obtained following the method of Andolfatto (2005), using synonymous sites as a neutral standard (data not shown).

DISCUSSION

Two recent studies have shown that the proteins involved in dosage compensation have undergone adaptive protein evolution in *D. melanogaster*, but not in its close relatives (Levine *et al.* 2007; Rodriguez *et al.* 2007). In particular, the protein domains of MSL1 and MSL2 that are responsible for their specific targeting to the X chromosome harbor strong signatures of positive selection (Levine *et al.* 2007; Rodriguez *et al.* 2007). This rapid divergence at the DNA binding domains should create selective pressure for the corresponding MSL-binding sites to coevolve on the X chromosome. Here, I show that three previously characterized MSL-binding sites indeed have undergone adaptive evolution in the

TABLE 4

Lineage-specific estimates of the proportion of nucleotide substitutions driven by positive selection (α) in the MSL-binding regions (and 95% confidence intervals), using flanking regions as a neutral standard

Method	*D. melanogaster* α (95% C.I.)	*D. simulans* α (95% C.I.)	*D. yakuba* α (95% C.I.)
Fay *et al.* (2002)	0.65 (0.49, 0.85)	0.13 (−0.26, 0.74)	−0.14 (−1.11, 0.80)
Smith and Eyre-Walker (2002)	0.59 (0.50, 0.85)	0.06 (−0.24, 0.75)	−0.21 (−0.47, 0.80)
Bierne and Eyre-Walker (2004)	0.73 (0.38, 0.88)	0.17 (−0.83, 0.60)	0.09 (−0.71, 0.54)

D. melanogaster lineage, but not in *D. simulans* or *D. yakuba*. This finding supports the idea of selection-driven coevolution among DNA-protein interactions of the dosage compensation machinery. Note that the *roX* RNAs themselves are part of the dosage compensation complex; thus the adaptive evolution seen at the *roX* genes could also be driven by coevolution between the components of the MSL protein complex and their interacting RNAs within the MSL complex. However, the adaptive signals observed at the *roX* genes are localized to the MSL-binding sites, suggesting that this evolution is indeed driven by changed binding preferences of the MSL complex.

The selective pressure that drives this coevolution between the proteins of the MSL complex and its binding sites is not known, and it appears likely that such coevolution might come at a substantial selective cost, as the changing binding specificity of MSL proteins requires several hundred MSL-binding sites to change in concert. It was speculated that some conflict scenarios, either involving male-killing bacteria that specifically detect components of the MSL complex or defense against transposable elements could serve as the selective agents for this rapid evolution of the MSL components in *D. melanogaster* (Levine *et al.* 2007; Rodriguez *et al.* 2007). While it is unclear what is driving this rapid evolution, my finding of positive selection at the MSL-binding sites demonstrates that the protein-coding changes observed at the MSL genes indeed alter the binding specificity of the MSL complex, and that there is selective pressure for the binding sites to respond to this changing binding specificity. This supports the idea of substantial selective costs involved in this coevolution between the MSL complex and its binding sites on the X chromosome of *D. melanogaster*.

One consequence of rapid evolution between the components of the MSL machinery is that they might result in incompatibilities between species. Indeed, several studies have shown that the X chromosome plays a disproportionally large role in postzygotic isolation in animals (Coyne and Orr 2004; Masly and Presgraves 2007), and that it is preferentially the heterogametic sex that is inviable or sterile in species hybrids (Coyne and Orr 2004). One possibility is that both the large X effect and Haldane's rule are caused by a breakdown in dosage compensation in hybrids (Orr 1989). In support of this idea, the MSL components do not correctly target the *D. melanogaster* X chromosome in hybrids between *D. melanogaster* and *D. simulans* (Pal Bhadra *et al.* 2006). Note however, that hybrid lethality is not sex specific in this species pair, since hybrid females homozygous for the *D. melanogaster* X chromosome also die at the same stage of development, so dosage compensation defects are unlikely to be the sole explanation of hybrid lethality (Bolkan *et al.* 2007). Selection-driven coevolution among molecular interactors within species—as seen for the dosage compensation complex in *D. melanogaster*—might commonly cause the incidental evolution of incompatible interactions seen in hybrids between species. For instance, *Nup96*, a gene involved in a lethal hybrid incompatibility between *D. melanogaster* and *D. simulans* encodes part of a protein complex, and all five members studied in this complex also show evidence for recurrent adaptive protein evolution (Presgraves and Stephan 2007).

Little is known about the molecular nature of the vast majority of MSL-binding sites in Drosophila. Only the three binding sites studied here have been characterized in detail on the Drosophila X (Kageyama *et al.* 2001; Park *et al.* 2003; Oh *et al.* 2004). ChIP-chip experiments were used in several recent studies to identify additional genomic regions of the Drosophila X chromosome that are bound by the MSL complex (Alekseyenko *et al.* 2006; Dahlsveen *et al.* 2006; Gilfillan *et al.* 2006) and have offered a number of interesting insights. First, >700 binding regions for the MSL complex were identified, encompassing more than half the genes on the Drosophila X chromosome (Alekseyenko *et al.* 2006; Dahlsveen *et al.* 2006; Gilfillan *et al.* 2006). Interestingly, MSL binding clearly favors genes over intergenic regions and binds most strongly to the 3′ end of transcription units (Alekseyenko *et al.* 2006; Dahlsveen *et al.* 2006; Gilfillan *et al.* 2006). The three regions studied here might thus be somewhat unusual, since all three represent noncoding regions. If MSL binding normally occurs in coding regions, there might be additional costs associated with changing the binding specificity of the MSL complex, due to constraints imposed by the amino acid sequence of a protein. However, a recent study suggested that the MSL complex uses high-affinity sites (those studied here) to initially recognize the X chromosome and then associates with many

of its targets through sequence-independent features of transcribed genes (Larschan *et al.* 2007). This mechanism of targeting the X chromosome would substantially reduce the selective cost of evolving new binding-site preferences in *D. melanogaster*, since only high-affinity sites would be required to coevolve. It will be of great interest to compare patterns of polymorphisms and divergence along regions of the *D. melanogaster* X chromosome that recruit the MSL complex *vs.* segments that show no binding.

Growing evidence suggests that positive selection operates on noncoding DNA genomewide in Drosophila (Andolfatto 2005; Haddrill *et al.* 2008). Up to 20% of substitutions fixed between *D. melanogaster* and *D. simulans* at intergenic regions are estimated to be driven by adaptive evolution (Andolfatto 2005), using synonymous sites as a neutral standard. I estimate the fraction of adaptive divergence at the MSL-binding sites along the *D. melanogaster* branch to range between 60 and 70%, while little adaptive evolution is inferred in the *D. simulans* or *D. yakuba* lineage. Thus, while adaptive evolution at noncoding DNA may occur genomewide in Drosophila, the magnitude of positive selection observed at the MSL-binding sites in *D. melanogaster* appears well above background levels of adaptation observed in Drosophila. This supports the notion of selection-driven coevolution among DNA-protein interactions of the dosage compensation machinery in the *D. melanogaster* lineage.

I am grateful to Peter Andolfatto for technical assistance and Mia Levine for comments on the manuscript. This work was funded by a National Institutes of Health grant (GM-076007) to D.B. and an A. P. Sloan fellowship in molecular and computational biology.

LITERATURE CITED

Alekseyenko, A., E. Larschan, W. Lai, P. Park and M. Kuroda, 2006 High-resolution ChIP-chip analysis reveals that the Drosophila MSL complex selectively identifies active genes on the male X chromosome. Genes Dev. **20:** 848–857.

Andolfatto, P., 2005 Adaptive evolution of non-coding DNA in Drosophila. Nature **437:** 1149–1152.

Bachtrog, D., and P. Andolfatto, 2006 Selection, recombination, and demographic history in *Drosophila miranda*. Genetics **174:** 2045–2059.

Bachtrog, D., K. Thornton, A. Clark and P. Andolfatto, 2006 Extensive introgression of mitochondrial DNA relative to nuclear gene flow in the *Drosophila yakuba* species group. Evol. Int. J. Org. Evol. **60:** 292–302.

Baker, B., M. Gorman and I. Marín, 1994 Dosage compensation in Drosophila. Annu. Rev. Genet. **28:** 491–521.

Bierne, N., and A. Eyre-Walker, 2004 The genomic rate of adaptive amino acid substitution in Drosophila. Mol. Biol. Evol. **21:** 1350–1360.

Bolkan, B. J., R. Booker, M. L. Goldberg and D. A. Barbash, 2007 Developmental and cell cycle progression defects in Drosophila hybrid males. Genetics **177:** 2233–2241.

Bridges, C. B., 1922 The origin of variations in sexual and sex-limited characters. Am. Nat. **56:** 51–63.

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.

Coyne, J. A., and H. A. Orr, 2004 *Speciation.* Sinauer Associates, Sunderland, MA.

Dahlsveen, I., G. Gilfillan, V. Shelest, R. Lamm and P. Becker, 2006 Targeting determinants of dosage compensation in Drosophila. PLoS Genet. **2:** e5.

Dean, M. D., and J. W. Ballard, 2004 Linking phylogenetics with population genetics to reconstruct the geographic origin of a species. Mol. Phylogenet. Evol. **32:** 998–1009.

Fay, J. C., G. J. Wyckoff and C. I. Wu, 2002 Testing the neutral theory of molecular evolution with genomic data from Drosophila. Nature **415:** 1024–1026.

Gilfillan, G., T. Straub, E. de Wit, F. Greil, R. Lamm *et al.*, 2006 Chromosome-wide gene-specific targeting of the Drosophila dosage compensation complex. Genes Dev. **20:** 858–870.

Haddrill, P., D. Bachtrog and P. Andolfatto, 2008 Positive and negative selection on noncoding DNA in *Drosophila simulans*. Mol. Biol. Evol. **25:** 1825–1834.

Haddrill, P., B. Charlesworth, D. Halligan and P. Andolfatto, 2005 Patterns of intron sequence evolution in Drosophila are dependent upon length and GC content. Genome Biol. **6:** R67.

Halligan, D., and P. Keightley, 2006 Ubiquitous selective constraints in the Drosophila genome revealed by a genome-wide interspecies comparison. Genome Res. **16:** 875–884.

Hudson, R. R., M. Kreitman and M. Aguade, 1987 A test of neutral molecular evolution based on nucleotide data. Genetics **116:** 153–159.

Kageyama, Y., G. Mengus, G. Gilfillan, H. Kennedy, C. Stuckenholz *et al.*, 2001 Association and spreading of the Drosophila dosage compensation complex from a discrete roX1 chromatin entry site. EMBO J. **20:** 2236–2245.

Kelley, R., V. Meller, P. Gordadze, G. Roman, R. Davis *et al.*, 1999 Epigenetic spreading of the Drosophila dosage compensation complex from roX RNA genes into flanking chromatin. Cell **98:** 513–522.

Kohn, M., S. Fang and C. Wu, 2004 Inference of positive and negative selection on the 5′ regulatory regions of Drosophila genes. Mol. Biol. Evol. **21:** 374–383.

Larschan, E., A. A. Alekseyenko, A. A. Gortchakov, S. Peng, B. Li *et al.*, 2007 MSL complex is attracted to genes marked by H3K36 trimethylation using a sequence-independent mechanism. Mol. Cell **28:** 121–133.

Levine, M., A. Holloway, U. Arshad and D. Begun, 2007 Pervasive and largeley lineage-specific adaptive protein evolution in the dosage compensation complex of *Drosophila melanogaster*. Genetics **177:** 1959–1962.

Lucchesi, J. C., 1978 Gene dosage compensation and the evolution of sex chromosomes. Science **202:** 711–716.

Ludwig, M. Z., and M. Kreitman, 1995 Evolutionary dynamics of the enhancer region of even-skipped in Drosophila. Mol. Biol. Evol. **12:** 1002–1011.

Marín, I., M. Siegal and B. Baker, 2000 The evolution of dosage-compensation mechanisms. BioEssays **22:** 1106–1114.

Masly, J., and D. Presgraves, 2007 High-resolution genome-wide dissection of the two rules of speciation in Drosophila. PLoS Biol. **5:** e243.

McDonald, J. H., and M. Kreitman, 1991 Adaptive protein evolution at the Adh locus in Drosophila. Nature **351:** 652–654.

Meller, V., 2000 Dosage compensation: making 1X equal 2X. Trends Cell Biol. **10:** 54–59.

Muller, H. J., B. B. League and C. A. Offermann, 1931 Effects of dosage changes of sex-linked genes, and the compensatory effect of other gene differences between male and female. Anat. Rec. **51:** 110.

Oh, H., Y. Park and M. Kuroda, 2003 Local spreading of MSL complexes from roX genes on the Drosophila X chromosome. Genes Dev. **17:** 1334–1339.

Oh, H., J. Bone and M. Kuroda, 2004 Multiple classes of MSL binding sites target dosage compensation to the X chromosome of Drosophila. Curr. Biol. **14:** 481–487.

Orr, H., 1989 Does postzygotic isolation result from improper dosage compensation? Genetics **122:** 891–894.

Pal Bhadra, M., U. Bhadra and J. Birchler, 2006 Misregulation of sex-lethal and disruption of male-specific lethal complex localization in Drosophila species hybrids. Genetics **174:** 1151–1159.

Park, Y., G. Mengus, X. Bai, Y. Kageyama, V. Meller *et al.*, 2003 Sequence-specific targeting of Drosophila roX genes

by the MSL dosage compensation complex. Mol. Cell **11:** 977–986.

Presgraves, D., and W. Stephan, 2007 Pervasive adaptive evolution among interactors of the Drosophila hybrid inviability gene, Nup96. Mol. Biol. Evol. **24:** 306–314.

Rice, W. R., 1996 Evolution of the Y sex chromosome in animals. BioScience **46:** 331–343.

Rodriguez, M., D. Vermaak, J. Bayes and H. Malik, 2007 Species-specific positive selection of the male-specific lethal complex that participates in dosage compensation in Drosophila. Proc. Natl. Acad. Sci. USA **104:** 15412–15417.

Smith, N. G., and A. Eyre-Walker, 2002 Adaptive protein evolution in Drosophila. Nature **415:** 1022–1024.

Wright, S., and B. Charlesworth, 2004 The HKA test revisited: a maximum-likelihood-ratio test of the standard neutral model. Genetics **168:** 1071–1076.

Yang, Z., 1997 PAML: a program package for phylogenetic analysis by maximum likelihood. Comput. Appl. Biosci. **13:** 555–556.

Communicating editor: D. M. Rand