J Mol Evol (1996) 43:405–412

JOURNAL OF MOLECULAR EVOLUTION

Evolution of the Larval Cuticle Proteins Coded by the Secondary Sex Chromosome Pair: *X2* and *Neo-Y* of *Drosophila miranda:* I. Comparison at the DNA Sequence Level

Manfred Steinemann,[1] Sigrid Steinemann,[1] Wilhelm Pinsker[2]

[1] Institut für Genetik, Heinrich Heine Universität Düsseldorf, Universitätsstr. 1, D-40225 Düsseldorf, FR Germany
[2] Institut für Medizinische Biologie, Abt. Genetik, Universität Wien, Währingerstr. 17, A-1090 Wien, Austria

Received: 27 December 1995 / Accepted: 30 April 1996

Abstract. The larval cuticle protein genes (*Lcp*s) represent a multigene family located at the right arm of the metacentric autosome 2 (2R) in *Drosophila melanogaster.* Due to a chromosome fusion the *Lcp* locus of *Drosophila miranda* is situated on a pair of secondary sex chromosomes, the *X2* and *neo-Y* chromosome. Comparing the DNA sequences from *D. miranda* and *D. melanogaster* organization and the gene arrangement of *Lcp1–Lcp4* are similar, although the intergene distances vary considerably. The greatest difference between *Lcp1* and *Lcp2* is due to the occurrence of a pseudogene in *D. melanogaster* which is not present in *D. miranda.* Thus the cluster of the four *Lcp* genes existed already before the separation of the *melanogaster* and *obscura* group. Intraspecific homogenizations of different cluster units must have occurred repeatedly between the *Lcp1/Lcp2* and *Lcp3/Lcp4* sequence types. The most obvious example is exon 2 of the *Lcp3* gene in *D. miranda,* which has been substituted by the corresponding section of the *Lcp4* gene rather recently. The homogenization must have occurred before the translocation which generated the *neo-Y* chromosome. *Lcp3* of *D. melanogaster* has therefore no orthologous partner in *D. miranda.* Rearrangements in the promoter regions of the *D. miranda Lcp* genes have generated new, potentially functional CAAT-box motifs. Since three of the *Lcp* alleles on the *neo-Y* are not expressed and *Lcp3* is expressed only at a reduced level, it is suggestive to speculate that the rearrangements might be involved as *cis*-regulatory elements in the up-regulation of the *X2*-chromosomal *Lcp* alleles, in *Drosophila* an essential process for dosage compensation. The *Lcp* genes on the *neo-Y* chromosome have accumulated more base substitutions than the corresponding alleles on the *X2.*

Key words: *Drosophila* — *Lcp* gene family — Gene homogenization — CAAT-box motifs — Gene phylogeny — Sex chromosomes

Introduction

Drosophila miranda shows an extraordinary karyotype, resulting from the fusion of an autosome to the *Y* chromosome (Dobzhansky 1935; MacKnight 1939; Steinemann 1982; Steinemann and Steinemann 1992; Lucchesi 1994). As a consequence, the diploid chromosome number is ten in females and nine in males. The fused autosome corresponds to chromosome element C of the ancestral chromosome set postulated by Muller (1940) for the progenitor of the genus *Drosophila.* In the sibling species *D. pseudoobscura* and *D. persimilis* this chromosome element is represented by the acrocentric autosome *3;* in the more distantly related *D. melanogaster* it forms the right arm (2R) of the metacentric autosome *2* (Patterson and Stone 1952; Steinemann et al. 1984; Steinemann and Steinemann 1990, 1992). Within the subgenus *Sophophora D. melanogaster* belongs to the *melanogaster* species group, whereas *D. miranda, D.*

Correspondence to: M. Steinemann

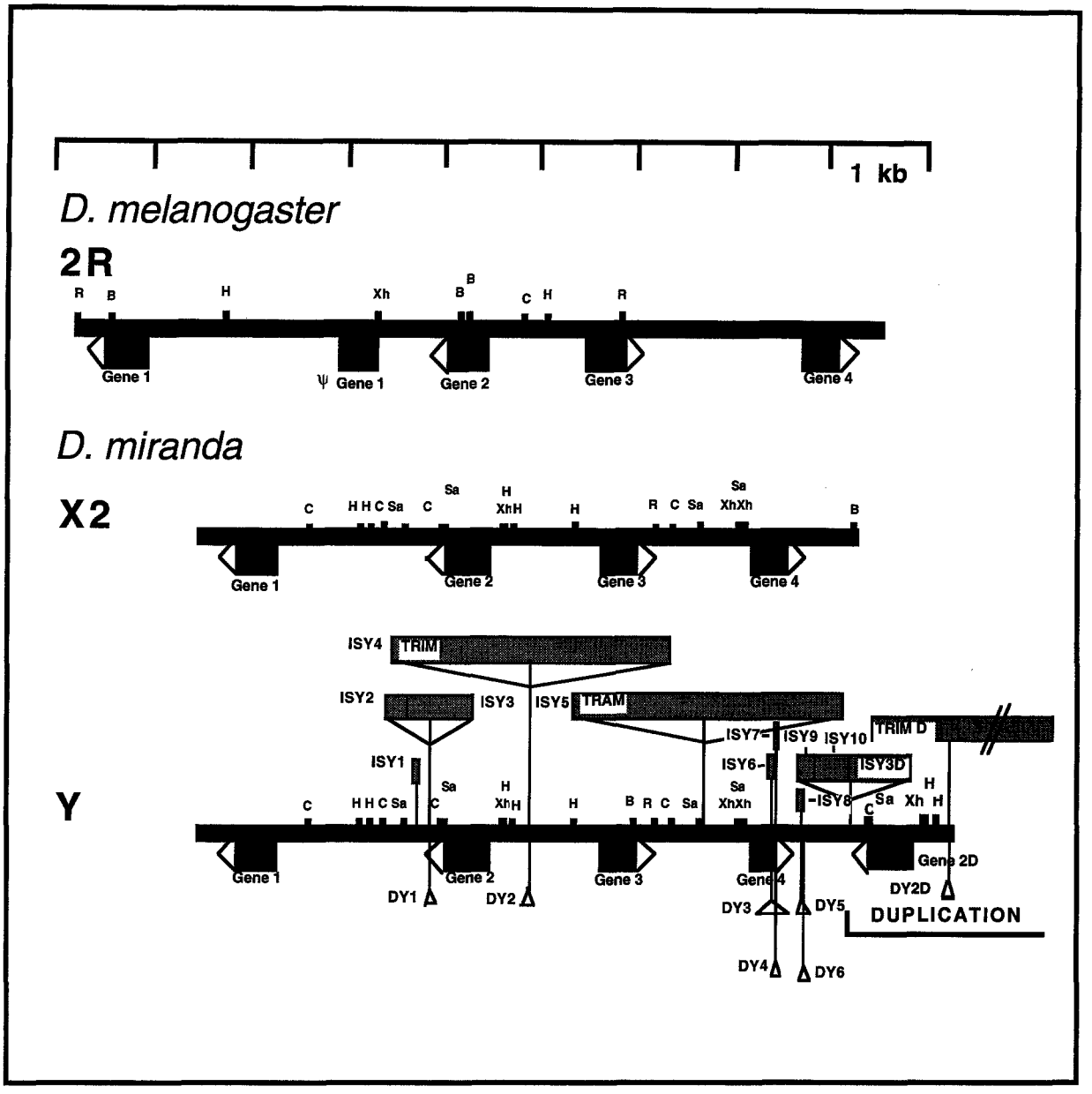

Fig. 1. Schematic view of the *Lcp* region from *D. melanogaster* and *D. miranda*. For orientation some restriction sites are included. Positions of the *Lcp1–4* genes and the orientation of transcription are indicated by *large open arrowheads*. The *Lcp* region from the right arm of chromosome 2 (*2R*) from *D. melanogaster* is redrawn from Fig. 1 in Snyder et al. (1982). The drawing of the *D. miranda Lcp* region is based on DNA sequence information from both areas on the *X2* and *neo-Y* chromosome (Steinemann and Steinemann 1992). The *Lcp* regions from the *X2* and *neo-Y* chromosome of *D. miranda* are arranged with respect to the *D. melanogaster Lcp 1–4* genes with the 5′-ends of the *Lcp2* genes aligned. Insertions (ISYs) in the *neo-Y* chromosomal *Lcp* region are indicated by *stippled boxes* and deletions (*DY*) *by triangles*. Only part of the duplicated *neo-Y* chromosomal area is shown. The duplicated *TRIM* retrotransposon (*TRIMD*) is shortened in the cartoon to be included. ψ, pseudogene in *D. melanogaster*, which is not present in the *D. miranda Lcp* region. Restriction sites: Sa, *Sac*I; H, *Hin*dIII; R, *Eco*RI; B, *Bam*HI; Xh, *Xho*I; C, *Cla*I.

pseudoobscura, and *D. persimilis* are members of the *obscura* species group. Based on sequence comparisons of the *Adh* gene the divergence time of the two groups has been estimated at 25 Mya (Russo et al. 1995). The separation of *D. miranda* from its next relatives *D. pseudoobscura* and *D. persimilis* has occurred much more recently, between 0.8 and 2.6 Mya according to mtDNA restriction analysis (Barrio et al. 1992). This date probably coincides with the chromosome fusion between autosome element C and the *Y* chromosome.

To analyze the molecular details of *Y* chromosome degeneration, we chose the larval cuticle protein genes *Lcp1–4* as test genes. These genes map to the *44D* region on the right arm of chromosome *2* in *D. melanogaster* (Snyder et al. 1981, 1982) and to chromosome *3* in *D. pseudoobscura* and *D. persimilis*. In *D. miranda* this cluster is found on the *X2* and *neo-Y* chromosomes. The complete *Lcp1–4* gene cluster from *D. miranda* has been cloned from both chromosomal locations (Steinemann and Steinemann 1990). Sequence analysis of the *X2* and

Table 1. Intergene distances (bp) in the respective *Lcp* clusters of *D. melanogaster* and *D. miranda*[a]

	D. mel. 2R	*D. mir. X2*
Lcp1–Lcpψ1	1,942	—
Lcpψ1	385	no ψ-gene
Lcpψ1–Lcp2	703	—
Lcp1–Lcp2	3,030	1,835
Lcp2–Lcp3	955	1,263
Lcp3–Lcp4	1,812	1,264

[a] Intergene distances in bp within the *Lcp1–4* gene clusters of *D. melanogaster* and *D. miranda*. Intergene distances are measured between the coding regions, ATG to the termination codons. The *D. melanogaster* data are taken from Snyder et al. (1982), including 0.8 kb of unsequenced sequences. ψ1, pseudogene 1

neo-Y chromosomal *Lcp1–4* region reveals a massive accumulation of inserted DNA sequences in the *neo-Y* chromosomal *Lcp* region including two newly identified retrotransposons *TRIM* and *TRAM*. For three of the *neo-Y Lcp* genes, *Lcp1, Lcp2,* and *Lcp4,* we could show that the *neo-Y* alleles are inactive while the *Lcp3 neo-Y* allele showed reduced activity (Steinemann and Steinemann 1992, 1993; Steinemann et al. 1993).

In *Drosophila* males crossovers are widely suppressed. Since the *neo-Y* chromosome is strictly paternally inherited, recombination between the *X2* and *neo-Y* chromosome is inhibited. Due to the lack of recombination both chromosomes are genetically isolated. Before the fusion event in the ancestor of *D. miranda,* the corresponding segments of the *X2* and *neo-Y* chromosomes were presumably homogeneous with respect to their gene content. Thus the beginning of the proposed start of the degeneration of the *neo-Y* chromosome can be dated about 0.8–2.6 Mya ago (Barrio et al. 1992). During this rather short time span the *neo-Y* has undergone radical changes. Multiple insertion events and rearrangements have disrupted the original chromosome organization. Now, the *X2* and *neo-Y* chromosomes already show a different chromosome structure.

The impact of sex chromosome differentiation at the sequence level is largely unknown. The present study examines differences in the evolutionary behavior of the *Lcp* multigene family. DNA-sequence comparisons were carried out between the genes from the two sex chromosomes of *D. miranda.* In addition, we performed interspecific comparisons between the strictly autosomal *Lcp* loci of *D. melanogaster* and the heterosomal ones in *D. miranda.*

Materials and Methods

Cloning and Sequencing of the Lcp1–4 Region. High-molecular-weight DNAs from *D. miranda* were isolated according to Steinemann (1982). Genomic EMBL4 lambda libraries from partial *Sau*3A (Boehringer Mannheim) digests were described in Steinemann and Steinemann (1990). Using a polymorphic restriction site, overlapping clones with *X2* or *neo-Y* chromosomal origin were isolated covering about 30 kb from both localizations (Steinemann and Steinemann 1990). For detailed restriction mapping the regions containing the *Lcp* genes on the *X2* and *neo-Y* chromosome were subcloned into pUC18. Cloning and standard DNA techniques were carried out according to Sambrook et al. (1989). We sequenced both strands by the dideoxy sequencing method (Sanger et al. 1977) from M13mp18/19 subclones covering the *X2* and *neo-Y* chromosomal *Lcp* genes and flanking sequences (Steinemann and Steinemann 1992) according to the protocol supplied with Sequenase (United States Biochemical, Cleveland). Including the insertions at the *neo-Y* chromosome we obtained from the *X2* and *neo-Y* more than 25 kb of sequence information (Steinemann and Steinemann 1992, 1993).

Computer Analysis. The 5′ and 3′ flanking sequences together with the coding region of the four *Lcp* genes of *D. miranda* are deposited in the EMBL gene bank, accession numbers X97809 DMLCP1X; X97810 DMLCP1Y; X97811 DMLCP2X; X97812 DMLCP2Y; X97813 DMLCP3X; X97814 DMLCP3Y; X97815 DMLCP4X; X97816 DMLCP4Y. DNA sequences were aligned using either MacMolly (Softgene, Berlin) or DNASIS (Pharmacia, Uppsala) alignment programs. DNA database screening was done using the updated EMBL and GenBank nucleotide Sequence Data Library (EMBL, Heidelberg; GenBank, NCBI, Washington). For the interspecific comparison with the *D. melanogaster Lcp1–4* sequences we used the sequence data (accession number DMCUT1/DMCUT2) from Snyder et al. (1982).

Results

Interspecific comparison of the multigene *Lcp* gene families of *D. miranda* and *D. melanogaster* showed that gene number and arrangement in the *Lcp1–4* cluster are similar in both species despite the fact that the *melanogaster* and *obscura* species group diverged 25 Mya ago (Russo et al. 1995). The tandem pairs *Lcp1–Lcp2* and *Lcp3–Lcp4* are transcribed in a head-to-head configuration (Fig. 1). The spacing between the genes varies between the two species (Table 1). In *D. melanogaster* the distance between *Lcp1* and *Lcp2* is about 1,200 bp longer than in *D. miranda.* This difference is partly due to a pseudogene (*Lcpψ 1*) in *D. melanogaster* which is not present in *D. miranda.* The correspondence of the gene arrangements suggests that the *Lcp1–4* cluster formed from a single ancestral *Lcp* locus before the separation of the *melanogaster* and *obscura* species groups.

Comparing the pair of heterosomes in *D. miranda,* dramatic structural differences are obvious between the *Lcp* regions of the *X2* and *neo-Y* chromosome (Fig. 1). As the *Lcp* DNA sequences on the *X2* and *neo-Y* chromosome have been recombinationally isolated since the divergence of the *D. miranda* and *D. pseudoobscura* species, we were interested in the evolutionary changes between the *X2* and *neo-Y* alleles within *D. miranda.* In order to trace the phylogenetic history of the cluster, the interspecies changes between *D. miranda* and *D. melanogaster* were also examined. Alignment of the *Lcp1–4* cluster was smoothly achieved in the coding regions and the adjacent noncoding sections, but reasonable alignments of the intergenic spacer regions were not possible. The alignments in the respective 5′- and 3′-flanking regions of the four *Lcp* genes were extended as far as

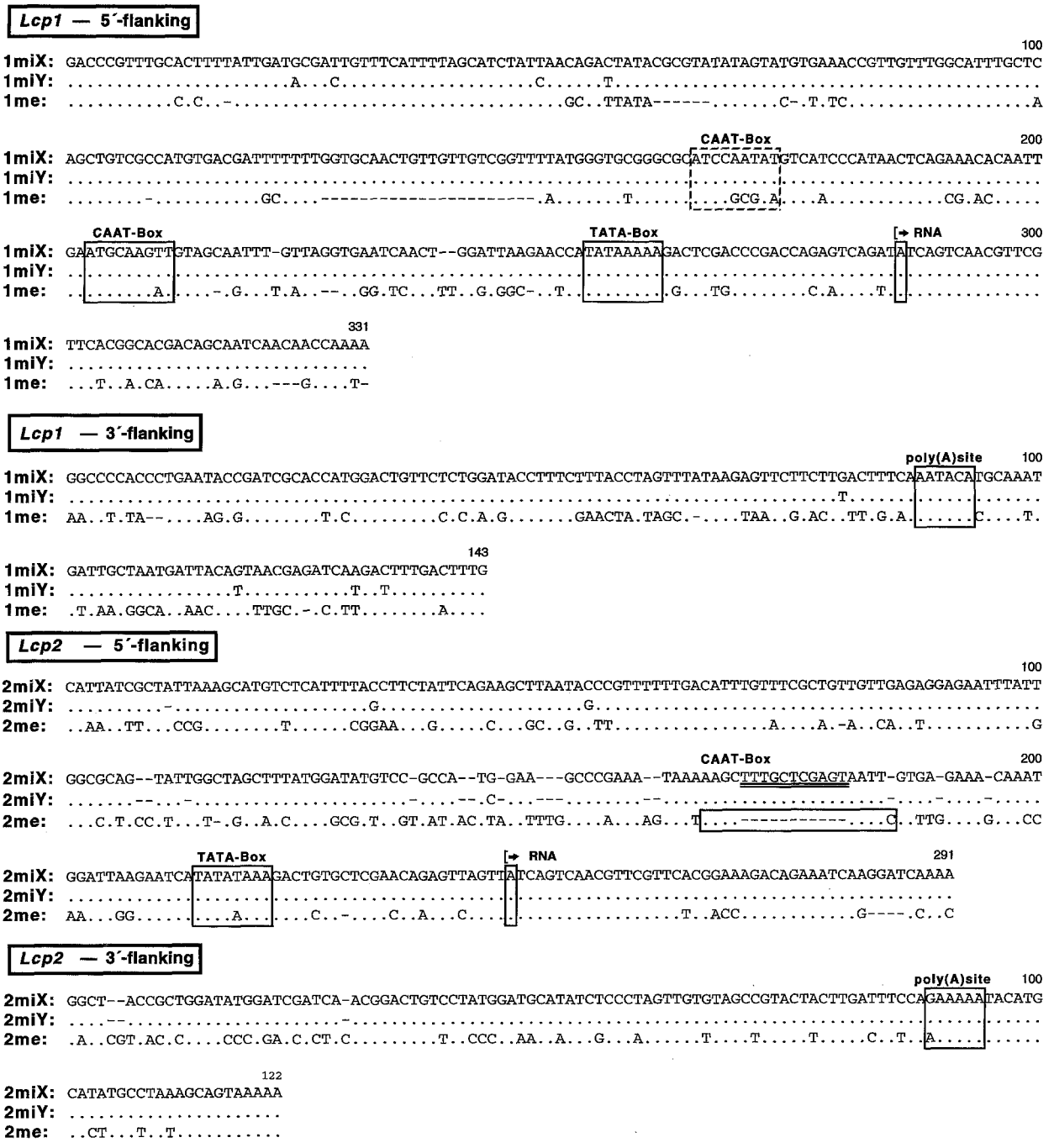

Fig. 2. Sequence alignment of the flanking 5′ and 3′ noncoding regions of the *Lcp1–4* genes from *D. miranda, X2* and *neo-Y* chromosome location, and *D. melanogaster,* right arm of autosome 2 (*2R*). The sequences are aligned to the *D. miranda X2* loci. The extension of the 5′ and 3′ sequence alignment is restricted for the stretches giving reasonable similarity. Base substitutions are shown. Sequence designations, 1miX: *Lcp1* gene from *X2,* 1miY: *Lcp1* gene from *neo-Y* and 1me: *Lcp1* gene from *D. melanogaster;* correspondingly, designations of *Lcp2, Lcp3,* and *Lcp4* sequences. Sequence identities are indicated with *points* and gaps with *dashes.* The lengths of the aligned sequences are indicated with *numbers* irrespective of functional aspects as transcription start. Start of transcription is indicated with an *arrow.* Regulatory sequence motifs are *boxed.* CAAT-box motifs are identified according to similarities to the consensus sequences from *D. melanogaster* (O'Connell and Rosbash 1984); the others, using the *D. melanogaster* sequences (Snyder et al. 1982). Additionally potential motifs are indicated with *dashed frames.* In *Lcp2* the 11-bp insert within the CAAT-box motif and in *Lcp4* the sequence of the target site duplication are *underlined.* The complete *Lcp1–4* sequences are deposited in the EMBL gene bank.

unequivocal identities could be recognized (Fig. 2). The alignment shows that most of the functionally important sequence motifs (CAAT-box, TATA-box, transcription start, poly[A] site) are conserved between *D. melanogaster* and the *X2* and *neo-Y* of *D. miranda.* In the *Lcp1* gene of *D. miranda,* a second potential CAAT-box motif (−110 to −118) occurs beside the CAAT-box (−72 to −80) at the position homologous to the *D. melanogaster* motif (Fig. 2). With respect to the consensus sequence (O'Connell and Rosbash 1984) this CAAT-box differs

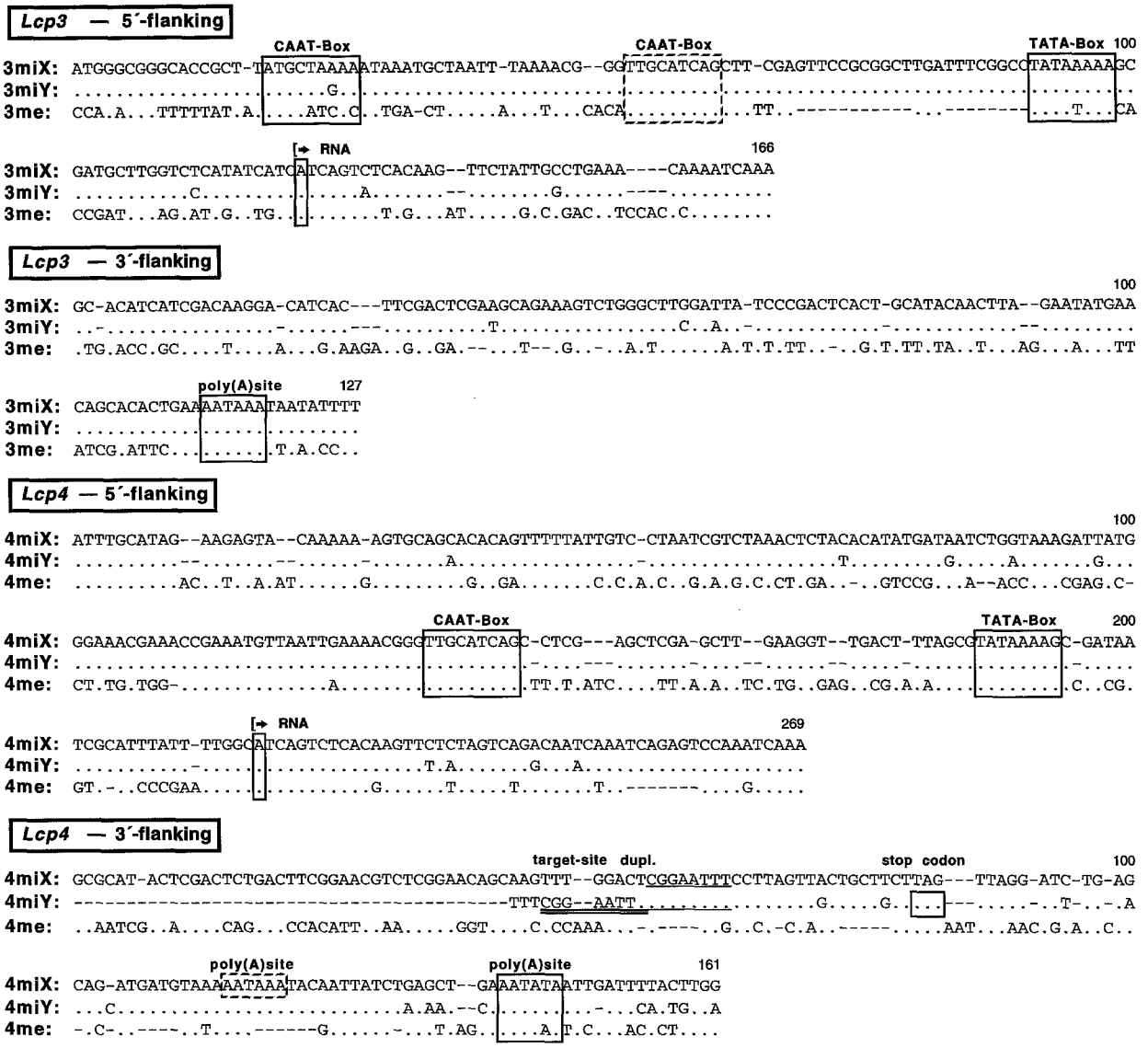

Fig. 2. Continued.

by 1-bp substitution in *D. miranda,* while the corresponding *D. melanogaster* motif shows three substitutions (Fig. 3). Thus the second motif in *D. miranda* might be closer to a functional sequence than the one of *D. melanogaster.* In the *Lcp2* gene of *D. miranda* the CAAT-box motif is disrupted into two halfs of 4 bp and 5 bp by a 11-bp insertion. Nevertheless, this gene is still expressed (Steinemann et al. 1993). Concerning *Lcp3,* the CAAT-box described for *D. melanogaster* (–75 to –83) has an equivalent motif at –91 to –99 in *D. miranda.* Compared with the consensus sequence it reveals a 1-bp substitution. In addition, in *D. miranda* another CAAT-box motif occurs further downstream (–60 to –68), which shows a perfect sequence identity to the CAAT-box motif of *Lcp4.* Due to two insertions of 11 bp and 8 bp, respectively, the position of the second CAAT-box motif is within the conventional distance from the TATA-box. In *D. melanogaster* this sequence motif is located at position –42 to –50 (Fig. 2). This position is much closer to the TATA-box and therefore probably outside of the optimal distance. At the *neo-Y,* functionally important motifs are perfectly conserved with respect to the *X2.* Comparing the *X2-* and *neo-Y-*chromosomal TATA- and CAAT-box sequences, only one base substitution could be detected. However, functional analysis had revealed that only the *neo-Y* allele of *Lcp3* is active, although at a greatly reduced level. The *neo-Y*-chromosomal *Lcp1, Lcp2,* and *Lcp4* alleles are not expressed at all (Steinemann et al. 1993).

To analyze the relationships among the cluster units, paralogous alignments were also attempted. In both species unequivocal alignments were only possible between the pairs *Lcp1/Lcp2* and *Lcp3/Lcp4,* respectively. Surprisingly, the exons of the *Lcp3/Lcp4* of *D. miranda* turned out to be almost identical (Fig. 4). In pairwise comparisons between the *Lcp3/Lcp4* genes of *D. miranda* and *D. melanogaster* (Table 2) the 5′- and 3′-noncoding sequences show a rather low similarity (35–

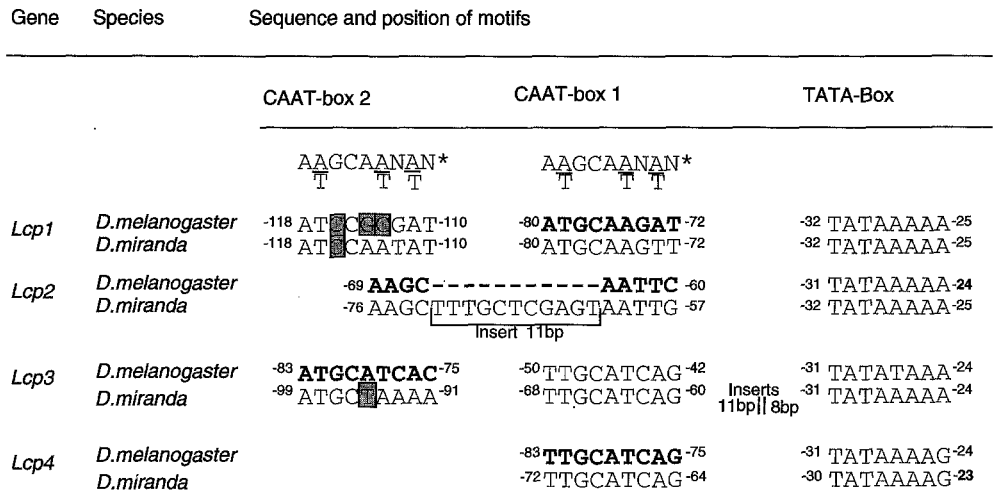

Fig. 3. Sequences and positions of the promoter motifs. Positions are indicated with reference to the transcription start. Exceptional distances of the TATA-box motifs are in *boldface.* The *D. melanogaster* CAAT-box sequences (Snyder et al. 1982, reported in O'Connell and Rosbash 1984) are in *boldface.* *For alignment of the CAAT-box motifs the consensus sequence described in O'Connell and Rosbash (1984) was used. Insertions in the *D. miranda* regions are indicated. Base substitutions with respect to the consensus sequence are *boxed and shaded.*

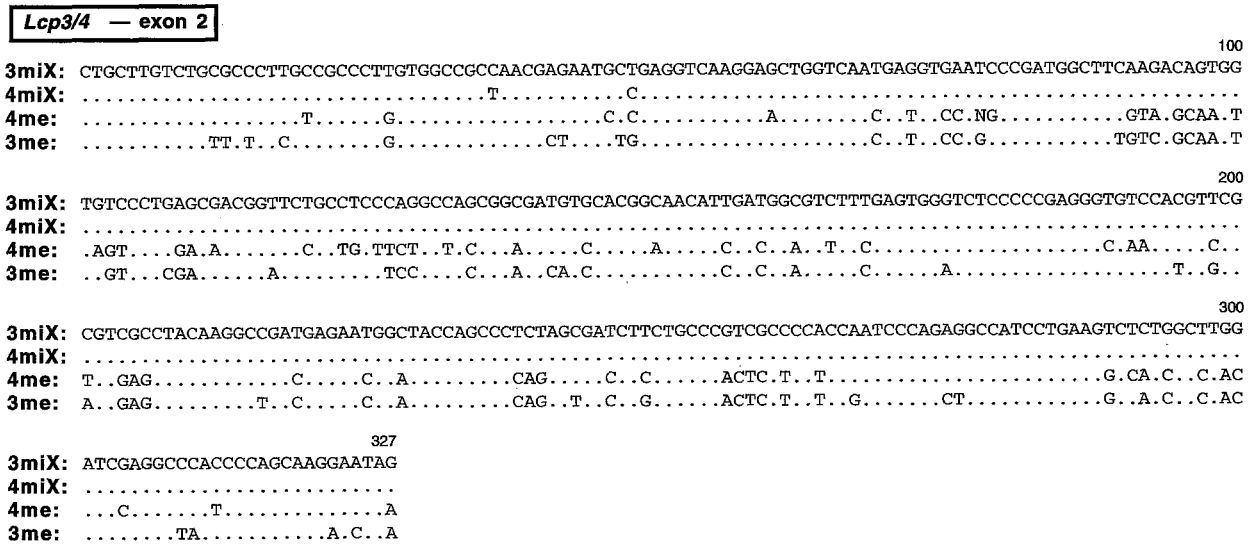

Fig. 4. Sequence alignment of exon 2 of the *Lcp3* and *Lcp4* genes from *D. miranda, X2* chromosome location, and *D. melanogaster,* right arm of autosome 2 (*2R*). The sequences are aligned to the *D. miranda X2* loci. For details see Fig. 2. *N* represents a not-determined base in Snyder et al. (1982).

39% divergence). In contrast, among the exon 2 sequences the divergence in both paralogous comparisons was significantly lower (0.6 and 16.4%) than in the orthologous comparisons (26.9–28.8%). In *D. miranda* the region of high similarity between the *Lcp3/Lcp4* genes covers the entire exon 2 (Table 3), whereas the intron and the 3′-noncoding regions are clearly distinct (not shown). In *D. melanogaster,* stretches of high similarities (positions, 49–101, 202–267) alternate with more diverged sections.

Due to the lack of recombination the *X2* and *neo-Y* chromosomes are genetically isolated. The inactivation of the *neo-Y* chromosome is expected to accelerate the rate of substitutions. To test this hypothesis we used the

Table 2. Substitutional divergence of *Lcp3* and *Lcp4* in *D. miranda* and *D. melanogaster*[a]

	Noncoding sequences			
	3me	4me	3miX	4miX
3me	—	38.0	37.5	37.2
4me	359	—	35.8	39.3
3miX	356	386	—	35.0
4miX	376	418	404	—

[a] Substitutional divergence of noncoding sections of *Lcp3* and *Lcp4* in *D. miranda* and *D. melanogaster.* The differences in % (above diagonal) and the lengths of the sections compared (below diagonal) are shown. Gaps were excluded and corrections were made for multiple hits (Jukes and Cantor 1969)

Table 3. Substitutional divergence of *Lcp3* and *Lcp4* in *D. miranda* and *D. melanogaster*[a]

	Exon 2			
	3me	4me	3miX	4miX
3me	—	16.4	28.8	29.2
4me	326	—	26.9	26.9
3miX	327	326	—	0.6
4mix	327	326	327	—

[a] Substitutional divergence of exon 2 of *Lcp3* and *Lcp4* in *D. miranda* and *D. melanogaster*. For arrangement see Table 2. In contrast to the noncoding sections the comparison of exon 2 reveals high sequence similarity among the paralogous genes *Lcp3* and *Lcp4* in both *D. melanogaster* and *D. miranda*

Table 4. Differences in the aligned sections of the *Lcp* gene cluster[a]

	Number of bp	%			
		Ident.	Gaps	Subst.	Corr.
me – miX	3,054	67.8	7.5	24.7	33.0
me – miY	3,059	66.6	7.7	25.7	34.7
miX – miY	2,992	97.4	0.5	2.1	2.1

[a] Differences in the aligned sections of the *Lcp* gene cluster. The *neo-Y* locus has diverged slightly more from the *D. melanogaster* sequence than the corresponding section of the *X2*. The difference is statistically not significant. Ident. = Identical, Gaps = unmatched positions, Subst. = substitutions, Corr. = substitutions corrected for multiple hits

Table 5. Substitutions on chromosomes *X2* and *neo-Y* of *D. miranda*[a]

	miX	miY	Total
Lcp1	2	11	13
Lcp2	3	10	13
Lcp3	2	9	11
Lcp4	1	9	10
Lcp1–4	8	39	47

[a] Substitutions on chromosome *X2* and *neo-Y* of *D. miranda*. In all four *Lcp* genes of *D. miranda* the majority of the substitutions have occurred on the *neo-Y*. On average the number of substitutions on the *neo-Y* is five times higher than on the *X2*

Lcp sequences of *D. melanogaster* as an outgroup. We compared about 3 kb from the *Lcp* cluster (Table 4). The sequence comprises the coding regions of *Lcp1–4*, the introns, and the 5′- and 3′-flanking regions as far as a reasonable alignment could be achieved. The section covered by the large deletion on the *Lcp4* gene of the *neo-Y* was excluded from the analysis. Although the percentage of identical positions with respect to the *D. melanogaster* sequences is slightly higher in the *X2* than in the *neo-Y* (67.8 vs 66.6%), this difference did not prove statistically significant in a relative-rate test. This may be due to the fact that compared to the rather high divergence between *D. melanogaster* and *D. miranda* the differentiation between *X2* and *neo-Y* is negligible. Since no data from closer relatives of *D. miranda* (e.g., *D. pseudoobscura*) are available, we used a different approach. We analyzed only those positions which differ between *X2* and *neo-Y* (Table 5). Comparing these nucleotides with the respective sites in the *D. melanogaster* sequence, we were able to decide in most cases on which of the two sex chromosomes the substitution had occurred. At the 47 nucleotide positions evaluated for this comparison 83% of the substitution could be ascribed to the *neo-Y* and only 17% to the *X2* ($\chi^2 = 20.4$, $df = 1$, $P < 0.001$).

Discussion

Although the *Lcp* genes of *D. miranda* are located on a pair of secondary sex chromosomes, their basic organization is essentially the same as in *D. melanogaster*, indicating orthologous relationship of the *Lcp1–4* loci between the two species. Evolutionarily conserved basic structures and orientation of genes within multigene families are reported as well from other *Drosophila* gene clusters, e.g., chorion genes (Martinez-Cruzado et al. 1988). In *D. miranda* extensive rearrangements have occurred on the *neo-Y* accompanied by a higher number of nucleotide substitutions which exceeds that on the *X2* by a factor of 5 (Table 5). The higher rate is not only confined to the completely inactive *Lcp1*, *Lcp2*, and *Lcp4* genes but is also true for the still weakly expressed *Lcp3* gene. It can be assumed that dosage compensation at the *X2* has probably relieved the selective constraint on the entire *Lcp* cluster of the *neo-Y*. The rather uniform increase in the number of substitutions among the four *Lcp* genes suggests that the different cluster units were not successively inactivated. It appears more likely that the *X2* took over the functions of the *neo-Y Lcp* cluster in a single step. This argument is supported by the finding that the *Lcp* cluster is positioned within the 90% of *X2*-chromosome length which stains positively with the H4.Ac16 antibody (Steinemann et al. 1996). Due to the resulting lack of selective constraint the four genes might have started to degenerate simultaneously. Another factor, which may also have contributed to the higher number of substitutions, is the smaller effective population size of the *neo-Y* (about one-third compared to the *X2*). As a consequence, random genetic drift will cause less selection against mildly deleterious genes and thus lead to faster divergence of the *Y* chromosome.

The sequence comparisons among the *Lcp* genes of *D. miranda* and *D. melanogaster* revealed some inconsistencies with the hypothesis of an ancient *Lcp* cluster that predates the separation of the two lineages. Although the general arrangement of the cluster units has remained more or less the same, there are several cases where paralogous genes showed a higher similarity than the respective orthologous genes from the other species. The most striking example is given by the *Lcp3/Lcp4* genes

of *D. miranda,* which proved to be almost identical in the region of exon 2. The similarities to the *Lcp3* gene of *D. melanogaster* in the noncoding flanking regions are low, at the 5′- as well as at the 3′-end. In contrast, the alignment between *Lcp3* and *Lcp4* extends quite far into the noncoding sections (not shown). Based on these arguments, we conclude that the *Lcp3* gene of *D. melanogaster* has no orthologous counterpart in *D. miranda.* Instead, the original *Lcp3* gene of *D. miranda* has been assimilated by the *Lcp4* gene through a process of repeated homogenization, generated by unequal crossover or, more likely, by gene conversion. The homogenization of exon 2 between *Lcp3/Lcp4* of *D. miranda* must be considered as a more recent event preceded by an earlier homogenization of the flanking regions. In *D. melanogaster* the comparisons provide evidence for homogenization of exon 2 between the *Lcp3/Lcp4* genes (Table 3) and the *Lcp1/Lcp2* genes. In the latter case (data not shown) the similarity between *Lcp1* and *Lcp2* of *D. melanogaster* is 85.6%, a value significantly higher than those from the rest of the pairwise comparisons (69.1–72.9%). Thus homogenization has apparently occurred several times within the subclusters *Lcp1/Lcp2* and *Lcp3/Lcp4,* respectively, but not between the subclusters. This result underlines the assumption that the two subclusters are functionally differentiated, whereas the tandem genes within the subclusters might be still exchangable without severe consequences (cf. Steinemann et al. this issue).

The higher degree of similarity within the subclusters is reflected as well by the precisely conserved distances of the TATA-box motifs from the transcription start. In the *Lcp1/Lcp2* subcluster the distance is −25 to −32 and in the *Lcp3/Lcp4* −24 to −31 with one exception in each subcluster, respectively. Beside *Lcp4* the promoter regions of *Lcp1, Lcp2,* and *Lcp3* show interesting rearrangements with regard to the CAAT-box motifs, cf. Fig. 3. Curiously, in the *Lcp2* promoter region an 11-bp insertion splits the CAAT-box sequence into two halfs of 4 bp and 5 bp. As the gene is expressed (Steinemann et al. 1993), flanking sequences might support the function of either half or both rest motifs. Thus, in *D. miranda* we might look at the evolution of additional CAAT-box motifs. These sequence motifs are found as well as in the *neo-Y* alleles. However, the *neo-Y Lcp1, Lcp2,* and *Lcp4* alleles are not expressed and *Lcp3* only at a greatly reduced level (Steinemann et al. 1993). Therefore the *X2* alleles must compensate for the inactive *neo-Y*-chromosomal alleles, as indicated by preliminary results (Steinemann and Steinemann unpublished results). The additional sequence motifs might in fact contribute to the up-regulation of the *X2* alleles and thus be involved in dosage compensation.

Acknowledgments. Our work was supported by a grant from the Deutsche Forschungsgemeinschaft (Ste 266/4-1).

References

Barrio E, Latorre A, Moya A, Ayala FJ (1992) Phylogenetic reconstruction of the *Drosophila obscura* group, on the basis of mitochondrial DNA. Mol Biol Evol 9:621–635

Dobzhansky TH (1935) *Drosophila miranda,* a new species. Genetics 20:377–391

Jukes TH, Cantor CR (1969) Evolution of protein molecules. In: Munro HN (ed) Mammalian protein metabolism. Academic Press, New York, pp 21–132

Lucchesi JC (1994) The evolution of heteromorphic sex chromosomes. Bioessays 16:81–83

MacKnight RH (1939) The sex-determining mechanism of *Drosophila miranda.* Genetics 24:180–201

Martinez-Cruzado JC, Swimmer C, Fenerjian MG, Kafatos FC (1988) Evolution of autosomal chorion locus in *Drosophila.* I. General organization of the locus and sequence comparisons of genes s15 and s19 in evolutionarily distant species. Genetics 119:663–677

Muller HJ (1940) Bearings of the *Drosophila* work on systematics. In: Huxley J (ed) The new systematics. Oxford University Press, Oxford, pp 185–268

O'Connell P, Rosbash M (1984) Sequence, structure, and codon preference of the *Drosophila* ribosomal protein 49 gene. Nucleic Acids Res 12:5495–5513

Patterson JT, Stones WS (1952) Evolution of the genus *Drosophila.* MacMillan, New York

Russo CAM, Takezaki N, Nei M (1995) Molecular phylogeny and divergence times of Drosophilid species. Mol Biol Evol 12:391–404

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY

Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain terminating inhibitors. Proc Natl Acad Sci USA 74:5463–5467

Snyder M, Hirsh J, Davidson N (1981) The cuticle genes of *Drosophila:* a developmental regulated gene cluster. Cell 25:165–177

Snyder M, Hunkapillar M, Yuen D, Silvert D, Fristrom J, Davidson N (1982) Cuticle protein genes of *Drosophila:* structure, organization and evolution of four clustered genes. Cell 29:1027–1040

Steinemann M (1982) Multiple sex chromosomes in *Drosophila miranda:* a system to study the degeneration of a chromosome. Chromosoma 86:59–76

Steinemann M, Steinemann S (1990) Evolutionary changes in the organization of the major LCP gene cluster during sex chromosomal differentiation in the sibling species *Drosophila persimilis, D. pseudoobscura* and *D. miranda.* Chromosoma 99:424–431

Steinemann M, Steinemann S (1992) Degenerating *Y* chromosome of *Drosophila miranda:* a trap for retrotransposons. Proc Natl Acad Sci USA 89:7591–7595

Steinemann M, Steinemann S (1993) A duplication including the *Y* allele of *Lcp2* and the *TRIM* retrotransposon at the *Lcp* locus on the degenerating *neo-Y* chromosome of *Drosophila miranda:* molecular structure and mechanisms by which it may have arisen. Genetics 134:497–505

Steinemann M, Pinsker W, Sperlich D (1984) Chromosome homologies within the *Drosophila obscura* group probed by in situ hybridization. Chromosoma 91:46–53

Steinemann M, Steinemann S, Lottspeich F (1993) How *Y* chromosomes become genetically inert. Proc Natl Acad Sci USA 90:5737–5741

Steinemann M, Steinemann S, Turner BM (1996) Evolution of dosage compensation. Chrom Res (1996) 4:185–190