Two Medfly Promoters That Have Originated by Recent Gene Duplication Drive Distinct Sex, Tissue and Temporal Expression Patterns

George K. Christophides,* Ioannis Livadaras,† Charalambos Savakis† and Katia Komitopoulou*

*Department of Genetics and Biotechnology, School of Biological Sciences, University of Athens, Panepistimiopolis, Athens 157 01, Greece and †Institute of Molecular Biology and Biotechnology, Foundation for Research and Technology, Heraklion 711 10, Crete, Greece

ABSTRACT

Genes encoding predominantly male-specific serum polypeptides (MSSPs) in the medfly *Ceratitis capitata* are members of a multigene family that are structurally similar to the genes encoding odorant binding proteins of insects. To study the transcriptional regulation of the genes *MSSP-α2* and *MSSP-β2*, overlapping fragments of their promoters, containing the 5′ UTRs and 5′ flanking regions, were fused to the *lacZ* reporter gene and introduced into the medfly genome via *Minos*-mediated germline transformation. Transgenic flies were functionally assayed for β-galactosidase activity. Despite their extensive sequence similarity, the two gene promoters show distinct expression patterns of the reporter gene, consistent with previously reported evidence for analogous transcriptional activity of the corresponding endogenous genes. The *MSSP-α2* promoter drives gene expression specifically in the fat body of the adult males, whereas the *MSSP-β2* promoter directs gene expression in the midgut of both sexes. In contrast, similar transformation experiments in *Drosophila melanogaster* showed that both promoters drive the expression of the reporter gene in the midgut of adult flies of both sexes. Thus, the very same *MSSP-α2* promoter fragment directs expression in the adult male fat body in Ceratitis, but in the midgut of both sexes in Drosophila. Our data suggest that through the evolution of the *MSSP* gene family a limited number of mutations that occurred within certain *cis*-acting elements, in combination with new medfly-specific *trans*-acting factors, endowed these recently duplicated genes with distinct sex-, tissue-, and temporal-specific expression patterns.

THE Mediterranean fruit fly *Ceratitis capitata* is a major agricultural pest throughout the tropics and subtropics. It attacks soft fruit crops and vegetables, thereby causing immense devastation with grave economic consequences. One of the most effective methods of medfly control in current use is the sterile insect technique (SIT), whereby insects are irradiated in mass rearing facilities and then released for nonproductive mating in the wild. However, released females diminish the efficiency of the method because they compete with wild flies in mating to sterile males. Thus a major advance in SIT would be to generate genetic sexing strains producing only male flies. For this reason, molecular mechanisms regulating the expression of sex-specific genes received great attention during the last decade. The promoter sequences of such genes could be useful tools for the expression of genes that may serve in genetic sexing, when introduced in the medfly genome using transgenic technologies.

Toward this aim, genes involved in sex determination and development have been extensively studied in the medfly. Homologues of two members of the sex determination pathway, the *sex-lethal* and *double-sex*, have been isolated; however, *sex lethal* does not produce sex-specific transcripts as it does in Drosophila (SACCONE *et al.* 1998a,b). A number of other female-specific genes have also been characterized, such as genes encoding the major egg yolk polypeptides (RINA and SAVAKIS 1991), chorion genes (KONSOLAKI *et al.* 1990; TOLIAS *et al.* 1990; VLACHOU *et al.* 1997), and genes coding for the antibacterial peptide ceratotoxins (MARCHINI *et al.* 1993; ROSETTO *et al.* 1997). Recently, we have reported the isolation, evolution, and expression of a multigene family encoding the male-specific serum polypeptides (MSSPs; THYMIANOU *et al.* 1998; CHRISTOPHIDES *et al.* 2000).

The *MSSP* gene family consists of seven members classified in three subgroups according to the degree of the deduced polypeptide similarity: two *MSSP-α*, three *MSSP-β*, and two *MSSP-γ*. All *MSSP-α* and *MSSP-β* genes are tandemly arranged in a compact cluster spanning a 35-kb genomic region. It was demonstrated that all *MSSP-α* and *MSSP-β* genes and their 5′ and 3′ flanking regions have originated by recent multiple duplication events and show a remarkably high degree of similarity. The MSSP polypeptides are synthesized mainly in the fat body of adult males and secreted into the hemolymph

Corresponding author: Katia Komitopoulou, Department of Genetics and Biotechnology, School of Biological Sciences, University of Athens, Panepistimiopolis, Athens 15701, Greece.
E-mail: akomitop@biology.db.uoa.gr

where they are detected as homo- or heterodimers (Katsoris *et al.* 1990; Thymianou *et al.* 1995). Small amounts of *MSSP* transcripts and polypeptides have also been identified in the fat body of adult females and the midgut in both sexes. Their structural similarity with members of the odorant binding protein (OBP) family of insects predicts a potential function in binding and transporting volatile substances or other hydrophobic molecules (Thymianou *et al.* 1998; Christophides *et al.* 2000).

The ability to conduct *in vivo* functional studies and genome manipulation in nondrosophilid insects was first provided by *Minos* element-mediated germline transformation of the medfly (Loukeris *et al.* 1995a). Since then a number of transposable elements such as *piggyBac* in the medfly (Handler *et al.* 1998) and *mariner* (Coates *et al.* 1998) and *Hermes* (Jasinskiene *et al.* 1998) in the yellow fever mosquito *Aedes aegypti* have been used successfully as germline transformation vectors. The *Hermes* transformation system has already been used for the study of the promoter function of genes expressed specifically in the salivary glands of *A. aegypti* (Coates *et al.* 1999). Such studies will inevitably increase our understanding of the molecular mechanisms regulating the expression of genes in these species, the outcomes of which could be instrumental for the development of novel genetic strategies for the control of insects of economic and medical importance.

We report here the functional analysis of overlapping promoter fragments of the genes *MSSP-α2* and *MSSP-β2* in transgenic lines of *C. capitata* and *Drosophila melanogaster*. We demonstrate that the *MSSP-α2* promoter drives strong male-specific gene expression in the medfly and could be used for the construction of genetic sexing strains. A 320-bp fragment of this promoter was shown to direct the basic fat body-specific expression of the *lacZ* reporter gene in male adults. To our surprise, the analogous fragment of the *MSSP-β2* gene showed midgut-specific expression of *lacZ* in both sexes. Both fragments are flanked by regions that confer enhancement of *lacZ* expression in most of the lines. Interestingly, the two promoters share 94.5% similarity. In Drosophila, both promoters appear to function in a midgut-specific manner, similar to the *MSSP-β2* in the medfly. These data allow us to propose that the midgut-specific transcriptional regulation of the *MSSP-β2* gene has been conserved among the two species. They suggest that a complex regulatory mechanism, involving transcriptional activation of *MSSP-α2* in the male fat body and suppression in the midgut, has been established during the evolution of the *MSSPs* in the medfly. Taking note of the high degree of identity of the two promoters, it appears that a limited number of mutations within their *cis*-acting elements, combined with the establishment of a complex regulatory network, proved adequate to give distinct promoter functions to the corresponding genes.

MATERIALS AND METHODS

Plasmid constructions: Two overlapping fragments of the 5′ flanking regions and the 5′ untranslated regions (UTRs) of the *MSSP-α2* and *MSSP-β2* genes were fused to the *lacZ* reporter gene and fusions were introduced into the medfly genome via *Minos* element-mediated germline transformation. The *MSSP* fragments were subcloned and prepared in several steps using a series of plasmids pBS KS II (Stratagene, La Jolla, CA) and pHSS6 (Seifert *et al.* 1986). Complete details are available (Christophides 2000). An overview follows.

pMiα2PS-lacZ and pMiα2PL-lacZ transposon plasmids: The two overlapping *MSSP-α2* promoter fragments named α2PS (α2 promoter short) and α2PL (α2 promoter long), respectively, were generated from an *MSSP-α2* genomic subclone, carrying sequences −2430 to +534 by PCR and restriction enzyme digests. The α2PS fragment (−283 to +37) was PCR amplified from the genomic DNA template using two primers each comprising a 5′ end flanking sequence *Cla*I site (ms5FC1: −465 **CCATCGATGG**TAAGAGACAGCAGCTAC and ms5RC2: +37 **CCATCGATGG**TGAAGTACGTTTGGGTC; *Cla*I site and flanking sequences are shown in boldface type), digested with *Eco*RI/*Cla*I, and cloned into the corresponding sites of the pHSS6 vector. This plasmid was named pHα2PS. The α2PL fragment (−522 to +37) was amplified from the genomic DNA template using the primers ms5FC2 (−522 **CCATCGAT GG**GCCAAACATGATGGCG) and ms5RC2, digested with *Cla*I, and cloned into the respective site of pHSS6. The resulting plasmid was named pHα2PL. The fusion gene *Adh/lacZ*/SV40 was derived from the vector pDM79 (Mismer and Rubin 1987) as an *Eco*RI fragment and ligated to the unique *Eco*RI site of pBS KS II, producing the pBS-*lacZ* plasmid. Subsequently, it was digested with *Hin*dIII and *Bam*HI and inserted into *Hin*dIII/*Bam*HI sites of plasmids pHα2PS and pHα2PL. The resulting plasmids were digested with *Not*I and the fusions were inserted as *Not*I cassettes into pTZMiCcwNotI vectors modified from the original transformation vector pMihsCcw (Loukeris *et al.* 1995b) by the authors.

pMiβ2PS-lacZ and pMiβ2PL-lacZ transposon plasmids: The two *MSSP-β2* promoter fragments named β2PS (β2 promoter short) and β2PL (β2 promoter long) derived from an initial *MSSP-β2* genomic subclone carrying sequences from −502 to +461 by PCR and restriction enzyme digests. The β2PS promoter was PCR amplified using the primers ms5FR3 (−287 **CGAATTCC**GGTTCGTGAAATCAGT; *Eco*RI site and flanking nucleotides are shown in boldface type) and ms5RC2. The ms5FR3 generates a 5′ end flanking *Eco*RI site, since the endogenous one, compared to *MSSP-α2*, has been eliminated because of an A/T transversion at position −277 (Figure 1). The PCR product was digested with *Eco*RI/*Cla*I and the derived fragment was ligated to corresponding sites of the pHSS6 vector. This plasmid was called pHβ2PS. The *Hin*dIII/*Bam*HI restriction fragment from the pBS-*lacZ* plasmid containing the *Adh/lacZ*/SV40 gene fusion was inserted into the *Hin*dIII/*Bam*HI sites of the pHβ2PS vector. The resulting plasmids were digested with *Not*I and the fusions were inserted as *Not*I cassettes into the pTZMiCcwNotI vector. The β2PL promoter was amplified by PCR using the oligonucleotide ms5RC-2 and the M13 reverse primer (Stratagene). The PCR product was digested with *Cla*I and *Eco*RI and the 542-bp *Cla*I/*Eco*RI fragment was ligated to pBS KS II and subsequently excised as a *Hin*dIII/*Bam*HI fragment (−485 to +37). This fragment was coligated with the 4.38-kb *Hin*dIII/*Bam*HI fragment from the pBS-*lacZ* plasmid to the *Bam*HI site of pHSS6 vector. The promoter-reporter gene fusion was then inserted as a *Not*I cassette in the pTZMiCcwNotI vector. At critical steps, plasmid clones were sequenced to confirm the expected sequence and define precisely the junction areas.

DNA preparations and sequence analysis: Plasmid DNA used for microinjections was prepared using the QIAGEN (Chatsworth, CA) Plasmid Midi kit and plasmid DNA used in subcloning procedure with standard protocols (Sambrook *et al.* 1989). For nucleotide sequence determination, DNA was prepared using either the QIAGEN Plasmid Midi kit or CirclePrep Spin Midi kit (BIO101) and sequences were determined by the dideoxynucleotide chain termination method (Sanger *et al.* 1977). Sequence alignment was obtained using CLUSTALW (Thompson *et al.* 1994). Sequences reported in this article have been deposited into the EMBL data bank with accession nos. under Y19145 and Y19147 for *MSSP-α2* and *MSSP-β2*, respectively.

Germline transformation: The transformation procedure was based on the *Minos* element-mediated germline transformation technique, described by Loukeris *et al.* (1995a,b), with minor modifications. Each of the four fusion constructs (600 μg/ml) and *Minos*-helper pHSS6hsMi2 (300 μg/ml) was injected into preblastoderm medfly *w* embryos. Emerged G0 adults were backcrossed to *w* adults in groups. Each male was crossed to 5 *w* females and each female to 3 *w* males. To induce expression of the *w* gene (Zwiebel *et al.* 1995) from the hsp70 promoter, G1 pupae were exposed daily to a 39° heat shock for 1 hr. Transformed lines were established by individual crosses of G2 transformed flies and massive crosses of the G3 progeny that were homozygous for the insertion as determined by the eye coloration.

Staining for β-galactosidase activity: Adult flies of transformed lines were initially injected with 4% paraformaldehyde in PEM buffer (0.1 M Pipes pH 6.9, 1 mM MgSO$_4$, 2 mM EGTA) and fixed for 10 min at room temperature. Gross sections of whole flies were stained with 0.2% X-gal as previously described (Simon and Lis 1987; Hama *et al.* 1990).

Western blot analysis: Fat body tissue (15 flies) and midguts (20 flies) were dissected from adult transgenic flies and the control *w* strain in 1× PBS buffer, homogenized in 200 μl 0.25 M Tris-Cl at pH 6.8, and sonicated in temperate conditions. Equivalent samples were electrophoresed in 15% SDS-polyacrylamide gel, electroblotted (BIO-RAD semidry blotter) onto nitrocellulose membrane (Hybond C, Amersham), and probed with anti-β-galactosidase antibody (Boehringer Mannheim, Indianapolis). The IgGs were localized with peroxidase-labeled second antibody (RaM/PO from Nordic, Tilburg, The Netherlands) and detected by chemiluminescense (ECL Western Blotting Detection Reagents, Amersham, Buckinghamshire, UK).

β-Galactosidase assay: β-Galactosidase activity was determined spectrophotometrically by following the hydrolysis of o-nitrophenyl β-D-galactopyranoside (ONPG). Assay conditions were as follows: individual transgenic flies were homogenized in 300 μl 0.25 M Tris-Cl at pH 6.8, sonicated for 15 min in mild conditions, and centrifuged (10 krpm) for 10 min at 4°. Total protein extracts from five synchronized 5-day-old males were quantified by the Bradford reaction using BSA as standard control and assayed for β-galactosidase activity as described in Sambrook *et al.* (1989) for mammalian cell extracts. Incubation period with ONPG chromogen was set at 30 min at 37°. The absorbance of reaction mixtures was measured at 420 nm. Standard dilutions of commercial pure β-galactosidase (Sigma, St. Louis) were used as a positive control and protein extracts of the medfly *w* strain were used as a negative control.

RESULTS

Experimental strategy and medfly transformation: To investigate the sex, tissue, and temporal specificity of regulation of the *MSSP* multigene family, promoter functional analysis of the genes *MSSP-α2* and *MSSP-β2* was performed using the *Minos* transformation system (Loukeris *et al.* 1995b). Comparison of the 5′ flanking regions and 5′ UTRs of the two genes showed a limited number of differences in nucleotide sequences (Figure 1A). A copy of the *mariner* transposable element (Gomulski *et al.* 1997) is detected at position −468 in the *MSSP-α2* gene promoter, resulting in interruption of the homology of the two promoters (Christophides *et al.* 2000). In a 504-bp DNA fragment from −1 up to the *mariner* insertion the two promoters share 94.5% overall identity.

A number of potential regulatory modules are present within these regions. A putative transcription initiation site determined by similarity to the consensus sequence of the arthropod initiator TCAGT (Cherbas and Cherbas 1993) is located 39 bp upstream of the ATG translational initiation codon. A presumptive TATA box (GTA TAAAT) is located at position −31 in both genes. A putative steroid hormone response element (GGTCAT CTAATGACC) is present at nucleotides −213/−199 in *MSSP-α2* and −209/−195 in *MSSP-β2*. The GGTCA inverted repeats constituting the palindromic sequence are separated by five intervening nucleotides (reviewed by Tsai and O'Malley 1994). Two direct repeats of the GGTCA motif are located ∼135 and 180 bp upstream of the palindromic sequence in both genes.

For the construction of the final *Minos*-based transposon plasmids presented in Figure 1B, two overlapping promoter fragments of each gene containing the 5′ UTRs and 5′ flanking regions were fused to the recombinant AUGβ-gal reporter gene (Mismer and Rubin 1987; Thummel *et al.* 1988). The AUGβ-gal gene hybrid includes a 127-bp fragment containing the translational start codon of the Drosophila *Adh* gene (Benyajati *et al.* 1983). This results in the synthesis of a fusion protein containing 30 amino acids of ADH at its N terminus. The SV40 terminator sequence was also included in the fusion to allow appropriate post-transcriptional RNA processing (Subramani and Southern 1983). The two short *MSSP* promoter fragments were named PS (promoter short) whereas the two long fragments were named PL (promoter long). PS fragments contained the region −283/+37 of the *MSSP-α2* gene (α2PS) and the region −287/+37 of the *MSSP-β2* gene (β2PS), respectively. In the PL fragments the 3′ ends remained the same as in PS, whereas the 5′ ends were extended up to −522 of the *MSSP-α2* gene (α2PL) and −485 of the *MSSP-β2* gene (β2PL). The 5′ UTRs were included in all constructs since it has been reported that MSSP gene expression is regulated not only at the transcriptional but also at the translational level (Thymianou *et al.* 1995). The 3′ termini of all fragments ended 2 bp upstream of the methionine initiation codon of the *MSSP* genes, thus interrupting the Kozak sequence

(CAAA; Kozak 1986); the 5′ untranslated region of the Drosophila *Adh* gene provides an identical motif.

The medfly transformation was carried out using the Minos transformation system as described elsewhere (Loukeris *et al.* 1995a,b). The results of all four transformation experiments are documented in Table 1. Verification of transformation was determined by Southern analyses of transformant DNA cut with *Pst*I and *Eco*RI

A

B

TABLE 1

Results obtained from the four independent germline transformations of the medfly

Experimental round	Plasmid fusion	G0 adults	Transgenic lines	Efficiency %
A	α2PS-*lacZ*	404	4	1
	β2PL-*lacZ*	299	3	0.7
B	α2PL-*lacZ*	368	13	3.5
	β2PS-*lacZ*	486	9	2.5

and probed with the *Minos* inverted repeats as well as by *in situ* hybridization to polytene chromosomes (data not shown).

The *MSSP-α2* gene promoter is capable of promoting the expression of *lacZ* in the fat body of adult males: Four independent transgenic lines (26, 27, 29, and 36) were established after the injections of medfly *w* embryos with the pMiα2PS-*lacZ* construct and *lacZ* expression was examined by chromogen X-gal staining. To eliminate background caused by β-galactosidase that is expressed by bacteria residing mostly in the midgut, transgenic flies were treated with tetracycline added in larval food (0.001%) for at least two generations before staining. In all lines, the reporter gene was expressed exclusively in the fat body of adult males (Figure 2A, top left), although variability of expression levels was observed. The strongest expression was detected in line 36, while in line 26 β-galactosidase activity was faint. Transgene expression started ∼72 hr after eclosion and both peripheral and deep fat bodies were stained. The maximum β-galactosidase activity was observed 5 to 6 days after eclosion. Thereafter the amount of the enzyme decreased gradually but remained detectable even until the 14th day. Relative to the endogenous MSSP protein accumulation pattern in the fat body starting 24 hr after eclosion (KATSORIS *et al.* 1990; THYMIANOU *et al.* 1995), the *lacZ* expression appears to be ∼50 hr delayed.

Thirteen transgenic lines obtained by transformation with the α2PL-*lacZ* fusion showed the same sex- and tissue-specific expression pattern compared to α2PS-*lacZ* (Figure 2A, top right and bottom left); however, β-galactosidase was detected in the fat body of adult

males within the first 30 hr after eclosion. Furthermore, throughout the first 13 days of adult male life in ∼50% of the lines, β-galactosidase output levels were estimated to be higher than in α2PS lines. In medfly, endogenous β-galactosidase activity was detected, similar to Drosophila (SCHNETZER and TYLER 1996). This activity was detected in pericardial cells in transgenic and nontransgenic *w* flies (Figure 2A, bottom, left and right).

The levels of β-galactosidase activity were quantified spectrophotometrically in five synchronized, 5-day-old male flies of all α2PL lines and the two stronger α2PS lines (36 and 29). As shown in Table 2, more than half of the α2PL lines showed stronger activity than the α2PS lines, in agreement with the results obtained by the X-gal staining, indicating that the region −522/−284 may act as transcriptional enhancer. The variability in β-galactosidase expression is attributed to chromosomal position effects and correlated with analogous variability in eye color (Table 2). However, in some of the lines the levels of expression of the *w* gene were inconsistent with the expression of the reporter gene, suggesting promoter interference phenomena. This was more evident among lines α2PL8A and 8B and α2PS29 and 36.

To study the transcriptional function of the α2PL promoter fragment more thoroughly, β-galactosidase activity was measured in line α2PL8A throughout adult development (Figure 3). Five males and five females were examined for 13 days after eclosion. β-Galactosidase activity started to be detectable within the first 24 hr of adult male life and increased drastically, reaching maximum levels during the fifth and sixth day. In the next 24 hr β-galactosidase activity underwent a more than twofold decrease and thereafter decreased gradu-

FIGURE 1.—Sequences used for functional analysis of *MSSP-α2* and *MSSP-β2* gene promoters. (A) Comparison of the 5′ flanking and the 5′ UTR regions of genes *MSSP-α2* and *MSSP-β2*. Numbers on the right indicate nucleotide positions. A putative steroid hormone response element (GGTCATCTAATGACC), direct GGTCA repeats, and TATA box are shaded. Transcription initiation sites are indicated by an arrow. The two overlapping fragments used for the promoter functional analysis of *MSSP-α2* are α2PL (−522/+37) and α2PS (−283/+37). Similarly, the analogous fragments used for *MSSP-β2* promoter analysis are β2PL (−485/+37) and β2PS (−287/+37). The first nucleotide of the putative modules and the 5′ end of the promoter fragments described above are shown by numbers and dots. The position of the *mar2* element is indicated by an arrow. (B) Schematic presentation of the four constructs used for the promoter functional analysis. The 5′ flanking and 5′ UTR fragments were fused to the *lacZ* reporter gene (shown in dark gray) followed by the SV40 terminator (T; unshaded) and introduced into the *Minos* transformation vector pTZMiCcwNotI, which is marked with the medfly white gene (light gray). ML and MR indicate the left- and right-end parts of *Minos*, respectively. The *Hsp*70 promoter (P) and terminator (T) are unshaded.
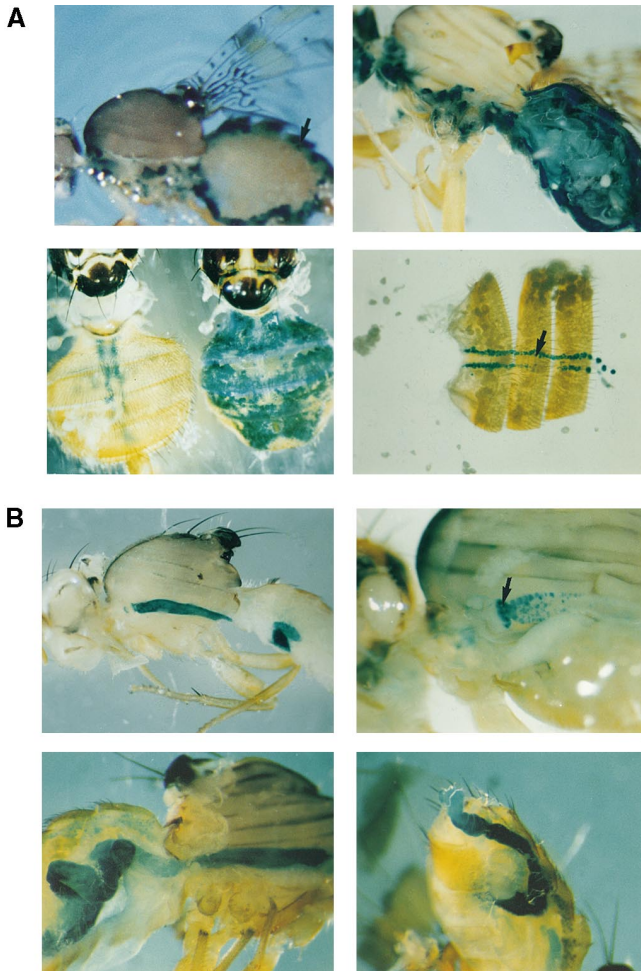
FIGURE 2.—(A) *MSSP-α2* promoter-directed expression of the reporter gene *lacZ* in medfly adults, visualized by X-gal staining in 5-day-old adult flies. The transgene expression driven by both α2PS and α2PL fragments is restricted to the fat body of males. (A, top left) Expression of β-galactosidase in the fat body (arrow) that is found at the periphery of the abdomen of an α2PS29 male. (A, top right) β-Galactosidase activity in the fat body of an α2PL11B male. The levels of expression driven by the α2PL fragment are much higher than the expression levels of the α2PS fragment. (A, bottom left) Comparison of β-galactosidase activity of a female (left) and a male (right) adult of the transgenic line α2PL8A. The female adults show only endogenous activity in the pericardial cells similarly to control *w* flies. (A, bottom right) Endogenous activity is detected in *w* female medflies (as well as in males) in pericardial cells found as a row on either side of the midline. (B) β-Galactosidase activity in β2PS and β2PL transgenic lines. The X-gal staining (blue) was performed in 2-day-old flies. The two overlapping promoter fragments of the *MSSP-β2* gene direct the expression in the midgut in both sexes. (B, top left) β2PS5 male, (B, top right) high magnification showing the *lacZ* staining of the pro-ventriculus (indicated by arrow) and midgut epithelial cells in a β2PL16 female, (B, bottom left) β2PL16 female, and (B, bottom right) β2PL28 male.

ally. Activity in transgenic female flies perfectly matched the base line and thus is not illustrated in the diagram.

The *MSSP-β2* gene promoter directs the expression of *lacZ* in the midgut in both sexes: When β2PS and

TABLE 2

β-Galactosidase activity in α2PS and α2PL transgenic males

Line	β-Gal activity[a]	*white* expression[b]
w	27 ± 7	−
α2PS29	56 ± 14	+++
α2PS36	90 ± 21	+
α2PL2A	68 ± 8	+
α2PL2B	74 ± 11	++
α2PL3A	59 ± 5	+/−[c]
α2PL3B	369 ± 75	++
α2PL4	79 ± 25	+
α2PL5	384 ± 111	++
α2PL7	156 ± 14	+
α2PL8A	1359 ± 162	+
α2PL8B	523 ± 62	+++
α2PL11A	119 ± 19	+
α2PL11B	1971 ± 420	++++
α2PL12	58 ± 8	+/−[c]
α2PL13	75 ± 9	+

[a] Values represent means ±SE about the mean that was determined from five independent experiments and given as OD (420 nm) × 10^{-3}.

[b] Eye coloration as determined by visual inspection varying from very weak yellow (+) to dark red (++++).

[c] Eye coloration was hardly detected (+/−).

β2PL transgenic lines were tested for β-galactosidase staining, we observed that only the midgut was stained in both sexes (Figure 2B). This expression pattern was developmentally regulated; it started in the pupal stage, a few hours before adult eclosion, remained at maximum levels until the third day, and then declined. Seven days after eclosion the enzyme was not significantly detected. Furthermore, as determined by visual inspection, lines carrying the longer *MSSP-β2* promoter fragment (16 and 28) clearly presented higher *lacZ* expression than the β2PS lines, although we have not determined this increase precisely.

The synthesis of β-galactosidase was also tested by Western blot analysis in fat bodies of 5-day-old adults of several α2PS and α2PL lines and in midguts of 2-day-old flies of β2PS and β2PL lines, where the maximum transgene expression was observed. In Figure 4, the synthesis of β-galactosidase in lines α2PL8A and β2PL16 is shown. The *w* strain was used as a negative control and the endogenous MSSP polypeptides as internal references. β-Galactosidase was detected only in the male fat body of line α2PL8A and the midgut of both sexes of line β2PL16, confirming the results presented above.

Both *MSSP-α2* and *MSSP-β2* gene promoters act in a midgut-specific manner in Drosophila: To investigate whether the distinct promoter function of *MSSP-α2* and *MSSP-β2* genes remained conserved through evolution, the constructs containing the long promoter-*lacZ* fusions (α2PL-*lacZ* and β2PL-*lacZ*) were introduced into
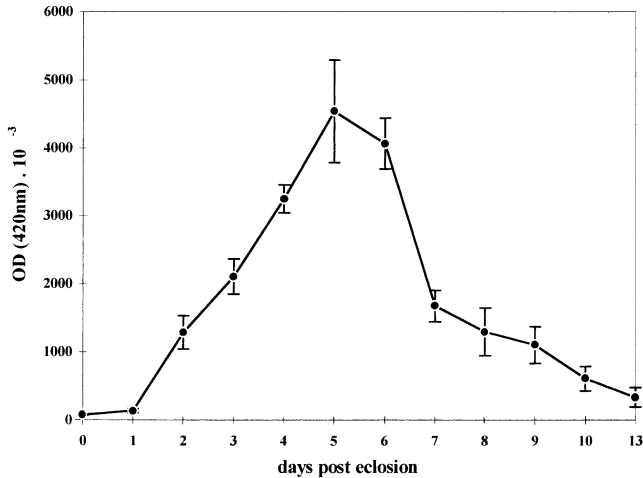
FIGURE 3.—Rate of β-galactosidase synthesis throughout the adult male development in line α2PL8A. β-Galactosidase activity was determined spectrophotometrically in protein extracts prepared from five synchronized adult flies each day after eclosion and is given as OD (420 nm) $\times 10^{-3}$. Transgene expression starts from the first day after eclosion and reaches the maximum level 4 to 5 days later. Within the next days the expression decreases gradually. Points (●) represent the average values of five separate experiments and bars indicate the standard errors.

the Drosophila genome according to LOUKERIS *et al.* (1995b). Interestingly, when transgenic flies were tested for β-galactosidase activity, both promoters were found to express *lacZ* in the midgut of both sexes (Figure 5). The pattern of expression was identical to that obtained with the β2PL-*lacZ* transgene in the medfly.

DISCUSSION

Individual members of the MSSP multigene family are expressed in distinct sex, tissue, and temporal patterns: The genomes of most organisms contain multiple copies of genes that are closely related in structure and
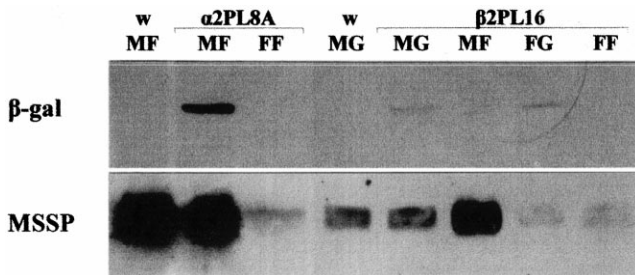

FIGURE 4.—Immunodetection of β-galactosidase in the fat body and midgut of transgenic flies. Western blotting was performed in protein extracts prepared from the fat body of 5-day-old α2PL8A and the midgut of 2-day-old β2PL16 flies, using antibody against the transgenic β-galactosidase (MF, male fat body; MG, male gut; FF, female fat body; FG, female gut). The medfly *w* strain was used as control. Immunodetection of MSSP polypeptides served as internal reference.
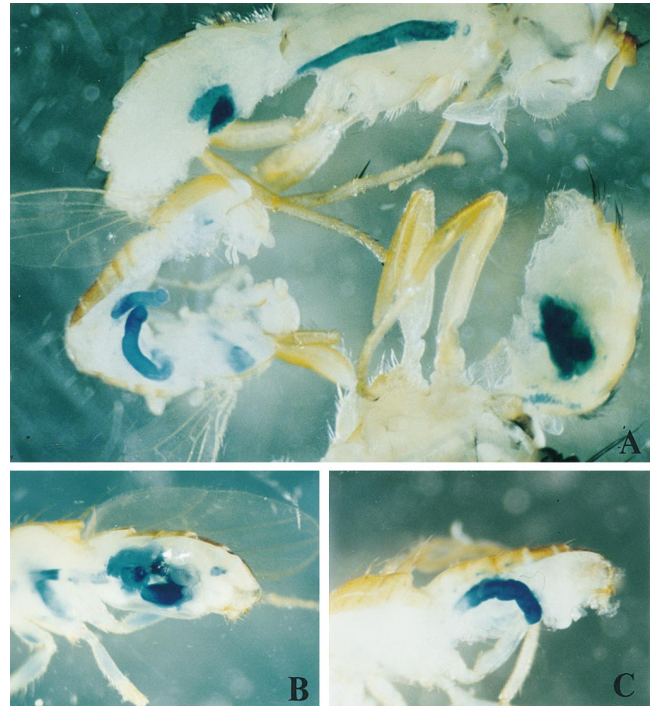

FIGURE 5.—*MSSP-α2* and *MSSP-β2* promoters drive the expression of β-galactosidase in Drosophila midgut in both sexes. (A) Expression of the *lacZ* driven by the β2PL promoter fragment in a male Drosophila (bottom left) and male medfly (top and bottom right). B and C show the expression of β-galactosidase directed by the α2PL promoter fragment in a midgut-specific manner in Drosophila, similar to β2PL.

function. Such multigene families consist of genes originating by gene duplication events, which retain a certain degree of sequence similarity. Different members of these families are frequently arranged in compact clusters, a feature that often results in concerted evolution (ARNHEIM 1983). Although duplicated genes initially have fully overlapping, redundant functions, if gene dosage is not critical the selective constraint becomes less for the extra copy, and it can evolve to have a slightly different function, while the original function of the gene is kept in the other copy. Recently, FORCE *et al.* (1999) suggested that complementary degenerative mutations in different regulatory elements of duplicated genes can facilitate the preservation of both duplicates, thereby increasing long-term opportunities for the evolution of new gene functions.

The *MSSP* genes belong to a multigene family that has originated by gene duplications of one ancestral gene (CHRISTOPHIDES *et al.* 2000). The most closely related members of the family, the *MSSP-α* and *MSSP-β* genes, are tandemly organized in a compact cluster. The very high degree of identity, both in their coding and surrounding regions, predicts that they have arisen by very recent gene duplications. Alternatively, they might have evolved under high selective constraint, possibly undergoing concerted evolution.

Initially, MSSP-α and MSSP-β polypeptides were characterized as male specific and restricted to the fat body (Katsoris *et al.* 1990; Thymianou *et al.* 1995), features that led us to the assumption that all five genes have redundant function. However, small amounts of both polypeptides were also detected in the female fat body as well as in the midgut of both sexes; quantification of the mRNA levels showed that they are about 500 times higher in the male fat body than in the remaining tissues. As determined by RT-PCR experiments, transcripts of the *MSSP-α1* gene are present in the fat body as well as in the midgut of both sexes, whereas *MSSP-α2* transcripts are restricted to the male fat body, suggesting that these genes may be expressed in a distinct manner (Christophides *et al.* 2000). Analogous discrimination of transcripts produced by individual *MSSP-β* genes was not achieved because of their extensive sequence similarity. The results obtained by functional analysis of two *MSSP* gene promoters and presented herein confirm this hypothesis and further suggest that this expression pattern is due to the distinct sex-, tissue-, and temporal-specific transcriptional activity of individual members of the family: the *MSSP-α2* promoter drives expression exclusively in the fat body of adult males, whereas *MSSP-β2* directs the expression only in the midgut of both sexes in a different temporal profile. Since both types of MSSP-α and -β polypeptides are predominantly synthesized in the male fat body, at least one of the remaining *MSSP-β* genes (-β*1* or -β*3*) must be expressed in a male fat body-specific manner, analogous to the *MSSP-α2* gene. The high sequence similarity in the regulatory regions of all *MSSP-α/β* genes, despite their different expression patterns, suggests that primary DNA sequences are under strong constraint to remain similar in sequence, while acquiring new abilities to bind different *trans*-acting factors. Recent studies on the *even-skipped* stripe 2 expression in Drosophila showed that binding-site differences in stripe 2 enhancer have functional consequences, but they are masked by other coevolved differences. This stabilizing selection has maintained phenotypic conservation of *eve* expression in various Drosophila species but has allowed mutational turnover of functionally important sites (Ludwig *et al.* 1998, 2000).

Why the members of the MSSP family are expressed in this distinct sex-, tissue-, and temporal-specific manner is an interesting question. The MSSP proteins are homo- or heterodimers of closely related polypeptides with sequence similarity to OBPs, predicting a potential function in binding and transporting pheromones or other hydrophobic molecules (Christophides *et al.* 2000). Thus, the distinct expression patterns of the *MSSP* genes may result in the formation of various dimers with slightly different specificities, which are distributed differentially and in different quantities in sexes and tissues in a temporally regulated profile.

Medfly transformation analysis of the regulatory ele-

ments of genes *MSSP-α2* and *MSSP-β2* was performed using two nested fragments of each gene promoter fused to the *lacZ* reporter gene. Sequences of the *MSSP-α2* gene located between +37 and −283 are responsible for a basic male fat body expression pattern, though a temporal delay was observed, compared to the endogenous synthesis of the MSSP-α polypeptides (Katsoris *et al.* 1990; Thymianou *et al.* 1995). The region from −284 to −522 corrects the temporal profile, as shown by comparison of β-galactosidase expression in two almost equally expressed α2PS and α2PL lines (data not shown), possibly due to the existence of early activation element(s). The very same region causes transcriptional enhancement of transgene expression in ∼50% of the α2PL lines. However, more α2PS lines should be examined to confirm the presence of general enhancer elements in this region. On the other hand, the region of *MSSP-β2* mapped within +37 and −287 drives the basic midgut expression pattern although any difference in the temporal profile was not precisely determined. The region from −288 to −485 does not affect the sex and tissue specificity of the promoter but it may confer transcriptional enhancement, similar to the analogous region of the *MSSP-α2*.

Comparison of the short regulatory fragments used in this analysis showed that the 5′ UTRs are identical in both genes, suggesting that they do not participate directly in their distinct expression. Thus *cis*-regulatory elements responsible for the basic sex, tissue, and temporally distinct expression pattern of the two genes are included from −1 to −283 in *MSSP-α2* and from −1 to −287 in *MSSP-β2*. These regions have 12 nucleotide variations, two single nucleotide deletions, and one deletion of five nucleotides, all dispersed along the sequence. It is known that DNA-protein interactions leading to regulated gene expression depend on the use of either weak but multiple binding sites or a few contact sites with stronger binding affinities; in the *MSSP* promoters the latter is most probable. Therefore within the *cis*-acting elements of these regions, a few nucleotides may be absolutely required while the rest may be noncritical. Mutations of these critical nucleotides may be able to modulate the promoter function of the two genes. The limited sequence differences could be exploited to identify the *trans* factors responsible, beginning with gel-shift experiments.

The palindromic motif GGTCATCTAATGACC, which is mapped at −213 of *MSSP-α2* and −209 of *MSSP-β2*, presumably is not involved in these different sex and developmental regulation processes, since it is present in both promoters. Thus, if a biological role were to be assigned to this sequence, it could not be related directly to the sex and developmental specificity of the promoters but rather to a general hormone response. Alternatively, ecdysteroid receptor isoforms or ligands could be expressed in a differential sex- and tissue-specific manner. Two single GGTCA motifs that are present in

both long promoter fragments at positions −395 and −348 in *MSSP-α2* and −394 and −347 in *MSSP-β2* may act as enhancing elements cooperatively with the palindromic sequence (ANTONIEWSKI *et al.* 1996).

Comparison of the *MSSP* promoter functions in *C. capitata* and *D. melanogaster*: To investigate whether the promoter function of *MSSP-α2* and *MSSP-β2* genes has been maintained through evolution, we transformed *D. melanogaster* with the constructs containing the long promoter fragments of both genes. Surprisingly, in transgenic Drosophila both promoters act in the same manner, expressing the reporter gene in the midgut of both sexes in a similar temporal profile. These results suggest that the male fat body transcriptional activity of the *MSSP-α2* gene may have been established in the medfly after the phylogenetic separation from the common ancestor with Drosophila. Therefore, *cis*-regulatory sequences are not recognized or *trans*-acting factors controlling this expression pattern are not found in Drosophila. As a result of that, a "default" midgut expression is elicited. Clearly, the *MSSP-α2* gene promoter bears all the *cis* information required to direct the expression in the midgut, but this function is suppressed in the medfly. This rather complex regulatory network, involving both activation in the male fat body and suppression in the midgut of the *MSSP-α2* promoter, is an interesting system for future studies.

Such evolutionary comparisons of specific developmental systems studied at the molecular level are necessary for the understanding of forces postulated to be at work in the fixation of different regulatory patterns observed in the species analyzed. Moreover, new data are providing evidence for the idea that physiological and developmental differences between species are predominantly due to changes in regulatory networks rather than in structural genes. In the medfly, mature adult males sexually excite and call their mates for courtship by secreting a mixture of pheromones and other volatile substances through their erect anal ampoules (BAKER *et al.* 1985). As described above, the MSSP proteins are putative carriers of such hydrophobic molecules. Therefore, their presence in high amounts in the male hemolymph during the first days of the adult male development may indicate a potential role in binding and transporting these male-specific pheromones or chemical cues. According to this hypothesis, the male-specific transcription of *MSSPs* may be absent in Drosophila because of differences in their sexual behaviors. In contrast, the fact that midgut-specific expression has been conserved may predict a general function necessary in both species. Our data suggest that the establishment of the *MSSP* multigene family is consistent with the classic theory that gene duplications have played an important role during evolution in the development of novel morphologies, physiologies, and behaviors (HALDANE 1933a,b; OHNO 1970; HOOD *et al.* 1975; TARTOF 1975; FRYXELL 1996; SHARMAN and HOLLAND 1996).

The medfly male-specific promoter could be used in genetic sexing: Beyond the interest in the complex regulatory networks that control the expression of the *MSSP* genes and evolutionary forces that determined the establishment of this gene family, a major goal of this study was the isolation of a medfly strong male-specific promoter. The *MSSP-α2* promoter may prove to be a useful tool for the expression of selectable genes in transgenic strains, such that only males will be mass-produced under certain conditions (LOUIS *et al.* 1987). Furthermore, the availability of two nested promoter fragments that differ substantially in their expression level provides the ability for controlled expression levels of selectable genes, a feature useful in both applied and basic studies. We have already used the short MSSP-α2 promoter fragment for the expression of the Drosophila alcohol dehydrogenase (ADH; BENYAJATI *et al.* 1983) in the medfly. Preliminary results show that the transgenic enzyme is active, leading to an approximately twofold increase of total ADH activity in male compared to female transgenic adults.

LITERATURE CITED

ANTONIEWSKI, C., B. MUGAT, F. DELBAC and J. A. LEPESANT, 1996 Direct repeats bind the EcR/USP receptor and mediate ecdysteroid responses in *Drosophila melanogaster*. Mol. Cell. Biol. **16:** 2977–2986.

ARNHEIM, N., 1983 Concerted evolution of multigene families, pp. 38–61 in *Evolution of Genes and Proteins*, edited by M. NEI and K. KOEHN. Sinauer Associates, Sunderland, MA.

BAKER, R., R. H. HERBERT and G. G. GRANT, 1985 Isolation and identification of the sex pheromone of the Mediterranean fruit fly *Ceratitis capitata*. J. Chem. Soc., Chem. Commun., 824–825.

BENYAJATI, C., N. SPOEREL, H. HAYMERLE and M. ASHBURNER, 1983 The messenger RNA for alcohol dehydrogenase in *Drosophila melanogaster* differs in its 5′ end in different developmental stages. Cell **33:** 125–133.

CHERBAS, L., and P. CHERBAS, 1993 The arthropod initiator: the capsite consensus plays an important role in transcription. Insect Biochem. Mol. Biol. **23:** 81–90.

CHRISTOPHIDES, G. K., 2000 Characterization of sex-specific gene promoters and their usage in the construction of transgenic strains of the Mediterranean fruit fly *Ceratitis capitata*. Ph.D. Thesis, School of Biological Sciences, University of Athens.

CHRISTOPHIDES, G. K., A. C. MINTZAS and K. KOMITOPOULOU, 2000 Organization, evolution and expression of a multigene family encoding putative members of the Odorant Binding Protein family in the medfly *Ceratitis capitata*. Insect Mol. Biol. **9:** 185–195.

COATES, C. J., N. JASINSKIENE, L. MIYASHIRO and A. A. JAMES, 1998 *Mariner* transposition and transformation of the yellow fever mosquito, *Aedes aegypti*. Proc. Natl. Acad. Sci. USA **95:** 3748–3751.

COATES, C. J., N. JASINSKIENE, G. B. POTT and A. A. JAMES, 1999 Promoter-directed expression of recombinant fire-fly luciferase in the salivary glands of *Hermes*-transformed *Aedes aegypti*. Gene **226:** 317–325.

FORCE, A., M. LYNCH, F. B. PICKETT, A. AMORES, Y. L. YAN *et al.*, 1999 Preservation of duplicate genes by complementary, degenerative mutations. Genetics **151:** 1531–1545.

FRYXELL, K. J., 1996 The coevolution of gene family trees. Trends Genet. **12:** 364–369.

GOMULSKI, L. M., C. TORTI, A. R. MALACRIDA and G. GASPERI, 1997 *Ccmar1*, a full-length *mariner* element from the Mediterranean fruit fly, *Ceratitis capitata*. Insect Mol. Biol. **6:** 241–253.

HALDANE, J. B. S., 1933a The part played by recurrent mutation in evolution. Am. Nat. **69:** 446–455.

HALDANE, J. B. S., 1933b *The Causes of Evolution.* Longwood Green, London.

HAMA, C., Z. ALI and T. B. KORNBERG, 1990 Region-specific recombination and expression are directed by portions of the Drosophila engrailed promoter. Genes Dev. **4:** 1079–1093.

HANDLER, A. M., S. D. MCCOMBS, M. J. FRASER and S. H. SAUL, 1998 The lepidopteran transposon vector, *piggyBac*, mediates germline transformation in the Mediterranean fruit fly. Proc. Natl. Acad. Sci. USA **95:** 7520–7525.

HOOD, L., J. H. CAMPBELL and S. C. R. ELGIN, 1975 The organization, expression, and evolution of antibody genes and other multigene families. Annu. Rev. Genet. **9:** 305–353.

JASINSKIENE, N., C. J. COATES, M. Q. BENEDICT, A. J. CORNEL, C. S. RAFFERTY *et al.*, 1998 Stable transformation of the yellow fever mosquito, *Aedes aegypti*, with the *Hermes* element from the housefly. Proc. Natl. Acad. Sci. USA **95:** 3743–3747.

KATSORIS, P. G., M. MAVROIDIS and A. C. MINTZAS, 1990 Identification and characterization of male specific serum proteins in the Mediterranean fruit fly *Ceratitis capitata*. Insect Biochem. **20:** 653–657.

KONSOLAKI, M., K. KOMITOPOULOU, P. P. TOLIAS, D. L. KING, C. SWIMMER *et al.*, 1990 The chorion genes of the medfly, *Ceratitis capitata*, I: structural and regulatory conservation of the s36 gene relative to two Drosophila species. Nucleic Acids Res. **18:** 1731–1737.

KOZAK, M., 1986 Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eucariotic ribosomes. Cell **44:** 283–292.

LOUIS, C., C. SAVAKIS and F. C. KAFATOS, 1987 Possibilities for genetic engineering in insects of economic interest, pp. 4–57 in *Fruit Flies: Proceedings of the Second International Symposium*, edited by A. P. ECONOMOPOULOS. Elsevier, Amsterdam.

LOUKERIS, T. G., B. ARCA, I. LIVADARAS, G. DIALEKTAKI and C. SAVAKIS, 1995a Introduction of the transposable element *Minos* into the germ line of *Drosophila melanogaster*. Proc. Natl. Acad. Sci. USA **92:** 9485–9489.

LOUKERIS, T. G., I. LIVADARAS, B. ARCA, S. ZABALOU and C. SAVAKIS, 1995b Gene transfer into the medfly, *Ceratitis capitata*, with a *Drosophila hydei* transposable element. Science **270:** 2002–2005.

LUDWIG, M. Z., N. H. PATEL and M. KREITMAN, 1998 Functional analysis of eve stripe 2 enhancer evolution in Drosophila: rules governing conservation and change. Development **125:** 949–958.

LUDWIG, M. Z., C. BERGMAN, N. H. PATEL and M. KREITMAN, 2000 Evidence for stabilizing selection in a eukaryotic enhancer element. Nature **403:** 564–567.

MARCHINI, D., P. C. GIORDANO, R. AMONS, L. F. BERNINI and R. DALLAI, 1993 Purification and primary structure of ceratotoxin A and B, two antibacterial peptides from the female reproductive accessory glands of the medfly *Ceratitis capitata* (Insecta: Diptera). Insect Biochem. Mol. Biol. **23:** 591–598.

MISMER, D., and G. M. RUBIN, 1987 Analysis of the promoter of the ninaE opsin gene in *Drosophila melanogaster*. Genetics **116:** 565–578.

OHNO, S., 1970 *Evolution by Gene Duplication.* Springer Verlag, New York.

RINA, M., and C. SAVAKIS, 1991 A cluster of vitellogenin genes in the Mediterranean fruit fly *Ceratitis capitata*: sequence and structural conservation in dipteran yolk proteins and their genes. Genetics **127:** 769–780.

ROSETTO, M., T. DE FILIPPIS, A. G. MANETTI, D. MARCHINI, C. T. BALDARI *et al.*, 1997 The genes encoding the antibacterial sex-specific peptides ceratotoxins are clustered in the genome of the medfly *Ceratitis capitata*. Insect Biochem. Mol. Biol. **27:** 1039–1046.

SACCONE, G., F. DI PAOLA, G. TESTA, A. PANE, G. DE MARTINO *et al.*, 1998a Finding sex-specifically expressed genes to develop a medfly transgenic sexing strain: Drosophila as a tool box for preliminary studies. Second research co-ordination meeting organized by FAO/IAEA Division of Nuclear Techniques in Food and Agriculture, Penang, Malaysia.

SACCONE, G., I. PELUSO, D. ARTIACO, E. GIORDANO, D. BOPP *et al.*, 1998b The *Ceratitis* capitata homologue of the Drosophila sex-determining gene sex-lethal is structurally conserved, but not sex-specifically regulated. Development **125:** 1495–1500.

SAMBROOK, J., E. F. FRITSCH and T. MANIATIS, 1989 *Molecular Cloning: A Laboratory Manual.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

SANGER, F., S. NICKLEN and A. R. COULSON, 1977 DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **12:** 5463–5467.

SCHNETZER, J. W., and M. S. TYLER, 1996 Endogenous beta-galactosidase activity in the larval, pupal, and adult stages of the fruit fly, *Drosophila melanogaster*, indicates need for caution in *lacZ* fusion-gene studies. Biol. Bull. **190:** 173–187.

SEIFERT, H. S., E. Y. CHEN, M. SO and F. HEFFRON, 1986 Shuttle mutagenesis: a method of transposon mutagenesis for *Saccharomyces cerevisiae*. Proc. Natl. Acad. Sci. USA **83:** 735–739.

SHARMAN, A. C., and P. W. H. HOLLAND, 1996 Conservation, duplication, and divergence of developmental genes during chordate evolution. Netherl. J. Zool. **46:** 47–67.

SIMON, J. A., and J. T. LIS, 1987 A germline transformation analysis reveals flexibility in the organization of heat shock consensus elements. Nucleic Acids Res. **15:** 2971–2988.

SUBRAMANI, S., and P. J. SOUTHERN, 1983 Analysis of gene expression using simian virus 40 vectors. Anal. Biochem. **135:** 1–15.

TARTOF, K. D., 1975 Redundant genes. Annu. Rev. Genet. **9:** 355–385.

THOMPSON, J. D., D. G. HIGGINS and T. J. GIBSON, 1994 Clustal W improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice. Nucleic Acids Res. **22:** 4673–4680.

THUMMEL, C. S., A. M. BOULET and H. D. LIPSHITZ, 1988 Vectors for Drosophila P-element-mediated transformation and tissue culture transfection. Gene **74:** 445–456.

THYMIANOU, S. P., G. CHRYSANTHIS, K. A. PETROPOULOU and A. C. MINTZAS, 1995 Developmentally regulated biosynthesis of two male specific serum polypeptides in the fat body of the medfly *Ceratitis capitata*. Insect Biochem. Mol. Biol. **25:** 915–920.

THYMIANOU, S., M. MAVROIDIS, G. KOKOLAKIS, K. KOMITOPOULOU, A. ZACHAROPOULOU *et al.*, 1998 Cloning and characterization of a cDNA clone encoding a male-specific serum protein of the Mediterranean fruit fly, *Ceratitis capitata*, with sequence similarity to odourant-binding proteins. Insect Mol. Biol. **7:** 345–353.

TOLIAS, P. P., M. KONSOLAKI, K. KOMITOPOULOU and F. C. KAFATOS, 1990 The chorion genes of the medfly, *Ceratitis capitata*. II. Characterization of three novel cDNA clones obtained by differential screening of an ovarian library. Dev. Biol. **140:** 105–112.

TSAI, M., and B. W. O'MALLEY, 1994 Molecular mechanisms of action of steroid/thyroid receptor superfamily members. Annu. Rev. Biochem. **63:** 451–486.

VLACHOU, D., M. KONSOLAKI, P. P. TOLIAS, F. C. KAFATOS and K. KOMITOPOULOU, 1997 The autosomal chorion locus of the medfly *Ceratitis capitata*. I. Conserved synteny, amplification and tissue specificity but sequence divergence and altered temporal regulation. Genetics **147:** 1829–1842.

ZWIEBEL, L. J., G. SACCONE, A. ZACHAROPOULOU, N. J. BESANSKY, G. FAVIA *et al.*, 1995 The white gene of *Ceratitis capitata*: a phenotypic marker for germline transformation. Science **270:** 2005–2008.

Communicating editor: A. J. LOPEZ