The Autosomal Chorion Locus of the Medfly *Ceratitis capitata*. I. Conserved Synteny, Amplification and Tissue Specificity but Sequence Divergence and Altered Temporal Regulation

Dina Vlachou,* Mary Konsolaki,† Peter P. Tolias,‡ Fotis C. Kafatos§ and Katia Komitopoulou*

**Department of Biology, Division of Genetics and Biotechnology, University of Athens, Panepistimiopolis, Kouponia, Athens 15701, Greece, †Department of Molecular Biology, Princeton University, Princeton, New Jersey 08544, ‡Public Health Research Institute, New York, New York 10016 and §European Molecular Biology Laboratory, 69117 Heidelberg, Germany*

Manuscript received October 28, 1996
Accepted for publication September 17, 1997

ABSTRACT

We report the isolation, full sequence characterization, amplification and expression properties of medfly chorion genes corresponding to the autosomal chorion locus of Drosophila. These genes are found adjacent to the paramyosin gene and are organized in the same order and tandem orientation as their Drosophila homologues, although they are spaced further apart. They show substantial sequence divergence from their Drosophila homologues, including novel peptide repeats and a new spacing of the tyrosines, which are known to be cross-linked in Dipteran chorion. The genes are amplified and expressed during oogenesis, as in Drosophila. Three of them are expressed in the same relative temporal order as in Drosophila but the fourth gene, the homologue of *s15*, shows a clear shift to an earlier expression period. This is the first known instance of changed temporal regulation in dipteran chorion genes.

THE autosomal Drosophila gene cluster encoding the *s18*, *s15*, *s19* and *s16* chorion proteins is a favorite system for comparative chromosomal analysis and molecular evolution. This locus has been fully sequenced in four distant species of the genus Drosophila widely diverged during a period of 60–80 million years (LEVINE and SPRADLING 1985; WONG *et al.* 1985; MARTINEZ-CRUZADO *et al.* 1988; FENERJIAN *et al.* 1989; SWIMMER *et al.* 1990). Despite considerable sequence diversification, this locus has remained highly conserved in organization throughout the genus: the same four chorion genes are found in all four species in the same order, in a compact tandem organization spanning approximately the same length of DNA. Moreover, all four species possess an unrelated highly conserved gene just downstream of the chorion cluster; this gene, originally described as *NC-ORF* (MARTINEZ-CRUZADO *et al.* 1988), has been identified recently as the paramyosin gene (VINOS *et al.* 1991, 1992; BECKER *et al.* 1992; MAROTO *et al.* 1995, 1996). The highly conserved gene organization of the locus is paralleled by conserved developmental regulatory mechanisms. Throughout the genus, these chorion genes amplify differentially as a single unit in the follicular epithelial cells during late oogenesis (SPRADLING and MAHOWALD 1980; SPRADLING 1981; GRIFFIN-SHEA *et al.* 1982; DELIDAKIS and KAFATOS 1987; SWIMMER *et al.* 1990). Moreover, these genes are expressed at characteristic, brief periods during late oogenesis, which are gene-specific and completely conserved within Drosophila (FENERJIAN *et al.* 1989). Possibly associated with these conserved regulatory features is the presence of short conserved sequence motifs in the proximal 5′ flanking DNA, embedded within otherwise highly diverged DNA sequences (MARTINEZ-CRUZADO *et al.* 1988; FENERJIAN *et al.* 1989; SWIMMER *et al.* 1990). The autosomal chorion cluster also has been studied in more closely related taxa of Hawaiian Drosophila, confirming the above features and illuminating the processes of its sequence evolution, including an unusually rapid evolutionary rate of divergence of the encoded proteins (MARTINEZ-CRUZADO 1990).

Recently, the Mediterranean fruitfly *Ceratitis capitata* (Wiedermann), frequently called medfly, has become a favorite experimental system for molecular studies in Diptera, both fundamental and applied in orientation. The applied interest results from its nature as one of the most destructive and costly agricultural pests, causing immense devastation to coffee and >200 fruit crops (CHRISTENSON and FOOTE 1960). The medfly originated in eastern Africa but subsequently invaded most of the tropical and subtropical countries of the word, such as South and Central America, Australia (BACK and PEMBERTON 1918; SAUL 1986) and recently California (CAREY 1991). An effective method for its eradication or control is sterile insect technique (SIT), which involves sterilization by radiation (ANWAR *et al.* 1971). The prospects for improving the SIT by construction of a stable genetic sexing medfly strain, permitting se-

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

lection and release of only male sterile flies (LOUIS *et al.* 1987), have been enhanced by the recent development of a successful transformation system for *C. capitata* (LOUKERIS *et al.* 1995). This breakthrough also makes the medfly especially attractive for fundamental comparative studies of insect gene regulation and molecular evolution. Such studies are facilitated by the ease of rearing the medfly (wherever this is not imprudent or illegal for reasons of agricultural quarantine), the ease of collecting and injecting its embryos, and its phylogenetic position, which is close enough to Drosophila to facilitate gene isolation by cross-hybridization or PCR but far enough to be informative. The medfly belongs to Tephritidae, a family of Diptera related to but distinct from the drosophilids, and is thought to have shared a last common ancestor with Drosophila ca. 120 mya (BEVERLEY and WILSON 1984) or approximately double the time since the separation of the ancestors of *Drosophila melanogaster* and *D. virilis*.

In previous studies, we have isolated four medfly chorion cDNAs. Two of them, *C2* and *C5*, were found to cross-hybridize with and correspond to the sex-linked *s36* and *s38* chorion genes of Drosophila (KONSOLAKI *et al.* 1990; TOLIAS *et al.* 1990); they were renamed *Ccs36* and *Ccs38*. These genes are clustered at position 70B of the 5L chromosome of *C. capitata*, which has been homologized with the *X* chromosome of *D. melanogaster* (ZACHAROPOULOU 1990; ZACHAROPOULOU *et al.* 1992). *Ccs36* and *Ccs38* also show sequence conservation and the same regulatory properties as their homologues in Drosophila: they are expressed with strict developmental specificity in early stages of choriogenesis and are differentially amplified in the ovarian follicle cells (KONSOLAKI *et al.* 1990; TOLIAS *et al.* 1990; MOUZAKI and MARGARITIS 1991).

In contrast, we had been unable to isolate the autosomal chorion genes or their cDNAs by cross-hybridization. The *C1* and *C4* chorion cDNAs are encoded by genes that are clustered at 88B of the 6L chromosome (ZACHAROPOULOU *et al.* 1992) and are amplified and expressed in late choriogenesis, as are the products of the Drosophila autosomal chorion cluster. However, their size and expression patterns correlate with the synthesis of polypeptides of substantially greater apparent size, and preliminary partial sequencing of *C1* and *C4* did not reveal obvious similarities to the Drosophila autosomal chorion cluster. Therefore, we resorted to an indirect cloning strategy, which is explained below and which proved successful. This permitted complete characterization of the structural and regulatory properties of the autosomal chorion genes of the medfly, which show both conserved and novel features relative to the Drosophila homologues.

MATERIALS AND METHODS

Screening of the genomic library: The medfly genomic library (RINA and SAVAKIS 1991) was a kind gift of the authors. Approximately 300,000 plaques were screened in two batches by hybridization at 55° (CHURCH and GILBERT 1984) to an 1.7-kb *Eco*RI-*Bam*HI genomic fragment, containing the two last exons of *D. melanogaster* paramyosin (described as *NC-ORF*) (FENERJIAN *et al.* 1989). Preparation of phage and plasmid DNA, agarose gel electrophoresis of DNA and labeling of DNA probes by nick-translation were carried out using standard procedures (SAMBROOK *et al.* 1989). Southern blot analysis was performed on GeneScreen nylon membrane (DuPont), using the hybridization conditions described by CHURCH and GILBERT (1984).

Construction of cDNA clones: The *Ccs18* and *Ccs19* cDNA clones were previously identified as *C4* and *C1* (TOLIAS *et al.* 1990) and had been selected from a Ceratitis ovarian cDNA library. cDNA clones of the *Ccs16* and *Ccs15* chorion genes were prepared by the rapid amplification of cDNA ends (RACE) method as described (FROHMAN 1990). Briefly, 1 μg total RNA extracted from Ceratitis ovaries was reverse transcribed using a primer consisting of oligo(dT) (17 residues) linked to a unique 17 nucleotide adapter. Amplification of *Ccs16* and *Ccs15* cDNAs was subsequently performed using the universal dT_{17}-adapter primer, and a gene-specific primer corresponding to the ATG region of each gene. The sequences of the specific primers, defined from genomic sequencing, were 5′ ATGCTCCGCTTCACGGTT 3′ for *Ccs16*, and 5′ ATGCAACGATTCATTTGC 3′ for Ccs15. The PCR products were directly subcloned into Pcr II vector, using the TA cloning kit (Invitrogen) and amplified after transformation as described by the manufacturer.

Sequence analysis and alignments: Genomic and cDNA clones were mapped with various restriction enzymes and subcloned into M13mp18 and M13mp19 or pBluescript II KS (Stratagene). Further subcloning was performed after deletions of clones and subclones using Exonuclease III (HENIKOFF 1987). Clones and subclones were amplified after transformation with $CaCl_2$/RbCl, and DNA was extracted from miniplasmid preparations by alkaline lysis (SAMBROOK *et al.* 1989). Overlapping restriction fragments were sequenced in both directions using the chain termination procedure (SANGER *et al.* 1977) with [^{35}S-dATP] (NEN) (BIGGEN *et al.* 1983) and Sequenase V2.0 (U.S. Biochemicals) (TABOR and RICHARDSON 1987). Samples were resolved in 4 or 5% polyacrylamide, 7.5 M urea gels, and autoradiography was performed using Kodak X-Omat film. Each strand was sequenced at least three times. Sequences were identified by comparison with data from the SWISSPROT database using the BLAST program (ALTSCHUL *et al.* 1990). Other computer programs used were CLUSTAL W (THOMPSON *et al.* 1994), DNA analysis programs developed by PUSTELL and KAFATOS (1982, 1984) and the GCG secondary structure predictions programs (version 8, Genetics Computer Group, Inc).

Amplification analysis: Genomic DNA was prepared from male flies or hand-dissected ovaries from 3–4-day-old female flies as described (HOLMES and BONNER 1973). The DNA was restricted with *Eco*RI or *Hin*dIII, electrophoresed in Tris-borate agarose gels and transferred to nylon membranes (GeneScreen). The filters were hybridized (CHURCH and GILBERT 1984) with ^{32}P-labeled *Eco*RI restriction fragments containing chorion genes by nick translation (RIGBY *et al.* 1977) or random hexanucleotide priming (FEINBERG and VOGELSTEIN 1983). Hybridizations were carried out at 72°.

RNA blot and developmental Northern analysis: RNAs were extracted from male flies, whole ovaries or stage-specific follicles of the medfly, according to the JACOBS-LORENA protocol (1980) as modified by H. BOUHIN *et al.* (1992). For RNA dot blots, performed according to KAFATOS *et al.* (1979), 20 male flies or 10 hand-dissected ovaries were homogenized in 0.5 ml of lysis buffer (2% SDS, 0.1 M Tris-HCl, pH 9.5, 0.2 M

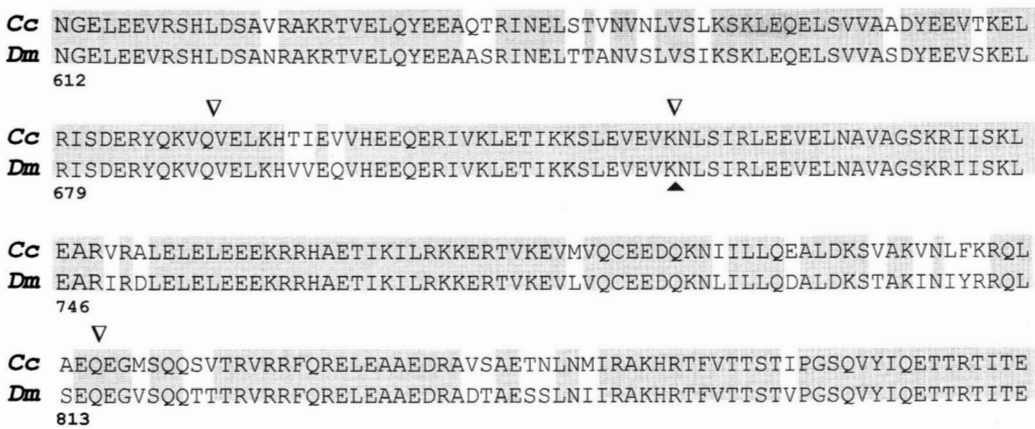

FIGURE 1.—Sequence alignment of the C-terminal regions of the deduced *D. melanogaster* and *C. capitata* paramyosin proteins. Residues identical in the two species are shaded. Amino acids are numbered relative to the Drosophila sequence. Intron positions of the paramyosin gene in *C. capitata* (▽) and *D. melanogaster* (▲) are indicated.

EDTA, 5% ethanol containing 2 mg of heparin and 1.5 mg of proteinase K) and extracted twice with 1 vol phenol (saturated with 1 M Tris, pH 9.5)/0.5 vol chloroform, once with 0.5 vol phenol/1 vol chloroform and once with 1 vol chloroform. After ethanol precipitation, total nucleic acids were dissolved in double distilled water, incubated for 1 hr at 37° with RNAase-free pancreatic DNAase I to a final concentration of 2 μg/ml and reprecipitated with ethanol. For developmental Northern analysis, RNAs were extracted from pools of 50 developing follicles staged according to their size, proportion of oocyte *vs.* nurse cells, and formation of the chorion layers (MOUZAKI and MARGARITIS 1991). Samples were resolved by electrophoresis on 1.5% agarose gels containing formaldehyde according to SAMBROOK *et al.* (1989). RNA blot analysis was performed using rapid alkaline transfer to a neutral nylon membrane (GeneScreen) as described by LOW and RAUSCH (1994). The membranes were baked at 80° for 2 hr and hybridized as previously using as probes cDNA chorion clones labeled by random hexanucleotide priming (FEINBERG and VOGELSTEIN 1983). Used blots were stripped of probes by boiling for 10–30 min in a solution of 0.010 M Tris-HCl, pH 7.5–8.0, 0.001 M EDTA, 1% SDS, before reusing.

RESULTS

Isolation of the autosomal chorion gene cluster in *C. capitata*: Initial attempts to clone the medfly homologues of the Drosophila autosomal chorion locus by low stringency hybridization, using Drosophila chorion genes as probes, were unsuccessful. This was not unexpected considering the rapid rates of divergence within the genus Drosophila (MARTINEZ-CRUZADO 1990). On the assumption that the paramyosin gene is well conserved in the medfly as in other organisms (BECKER *et al.* 1992; VINOS *et al.* 1992) and that this gene is closely linked to a chorion gene cluster as in all Drosophila species examined (FENERJIAN *et al.* 1989), we screened for the paramyosin gene to gain an indirect entry into the chorion locus.

The medfly genomic library was screened under permissive criteria with a genomic fragment containing the last two exons of the *D. melanogaster* paramyosin gene. Two positive lambda clones were isolated and subjected to Southern hybridization with the same probe, and the cross-hybridizing restriction fragments were mapped, subcloned and partially sequenced. Indeed they encompassed the 3′ terminal region of the paramyosin gene, as shown by alignment of the encoded protein with the corresponding *D. melanogaster* sequence (89% identity and 97% similarity over 264 amino acid residues, aligned without gaps; Figure 1). In the sequenced region, one intron was found at the same location as in *D. melanogaster*, while two others were unique to the medfly (Figure 1). Once the position and the orientation of the Ceratitis paramyosin homologue had been determined by sequencing, all *Eco*RI fragments downstream of its last intron were tested for putative chorion genes (which should only be expressed in the female ovary) (TOLIAS *et al.* 1990; MOUZAKI and MARGARITIS 1991) by being used as probes in dot hybridization of RNAs from Ceratitis males and ovaries (see below). Four successive *Eco*RI fragments hybridized exclusively with ovarian RNA. The fragment most distant from the paramyosin gene was used to isolate a third chromosomal clone bearing one additional fragment that showed ovarian expression. Plasmid subclones were constructed, restriction mapped and partially sequenced, resulting in the detection of four putative chorion genes and their positioning on the 26-kb map of this chromosomal walk (Figure 2).

In situ hybridization to polytene chromosomes, with a biotinylated genomic fragment containing the 3′ exons of the cloned *C. capitata* paramyosin gene, revealed a single band at position 88B of the *6*L chromosome (A. ZACHAROPOULOU, personal communication), where the genes corresponding to the previously isolated *C1* and *C4* chorion cDNAs (TOLIAS *et al.* 1990) had already been localized (ZACHAROPOULOU *et al.* 1992). Indeed, hybridization and sequence analysis confirmed that two of the genes in the chromosomal walk of Figure 2 correspond to the *C1* and *C4* chorion cDNAs. For the remaining two genes, cDNAs were constructed by the RACE-PCR method, using as primers a dT_{17} adapter and synthetic oligonucleotides, corresponding to the respective start of the ORF as determined from the partial genomic sequence (see MATERIALS AND METHODS). The cDNA and predicted protein sequences of the four chorion genes are presented in Figures 3–6,

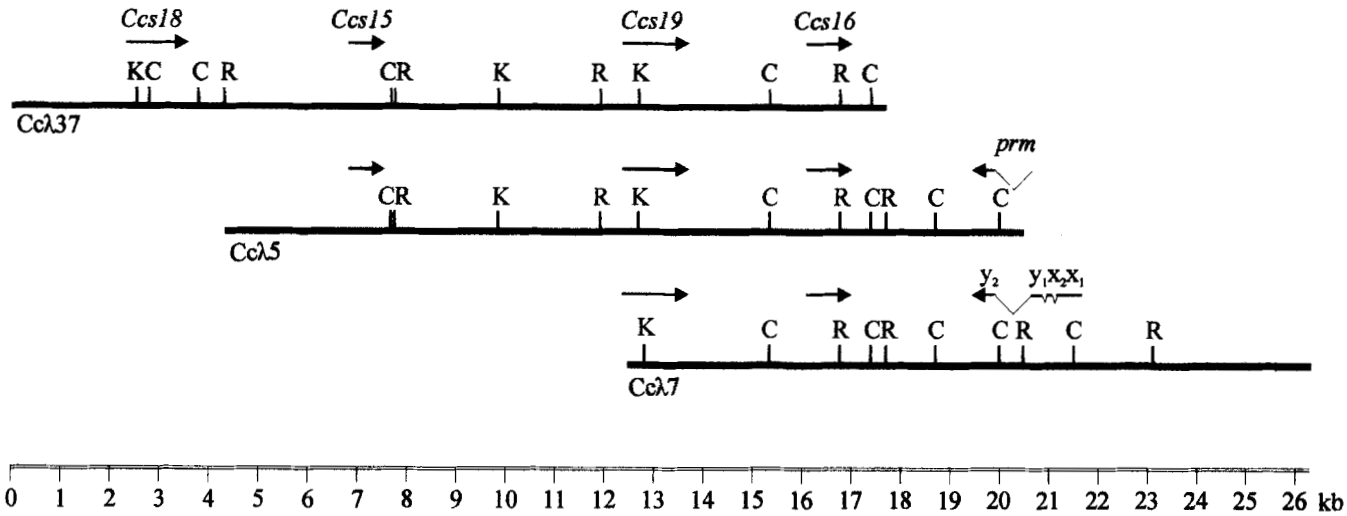

FIGURE 2.—A chromosomal walk consisting of three lambda clones and spanning five genes. Partial restriction maps of cloned segments were generated by standard prosedures (E, *Eco*RI; C, *Cla*I; K, *Kpn*I). The location and direction of transcription of the paramyosin gene (*prm*) were defined by cross-hybridization to the Drosophila homologue and sequence analysis of the appropriate restriction fragments. Note that the sequenced segment of this gene contains four exons (Named x1, x2 and y1, y2 relative to the previously described x, y exons of Drosophila; FENERJIAN *et al.* 1989), which are shown as an arrow interrupted by three introns. Chorion genes were identified downstream of *prm*, by dot blot hybridization to ovarian RNA, and then fully characterized by subcloning and sequencing. Their transcriptional orientations are indicated by solid arrows.

```
+1    ATTCAATAATATTATTAAACTAGAATTACGCAAA  ATG AAT AGA TTC CTT TGT ACC TTT GCC GCT ATT GTC GCC
                                          M   N   R   F   L   C   T   F   A   A   I   V   A

+74   GTT GCC AAT GGT TAC GCA GTT GGT GGA GGT GGT GGT TAT GGT GGT CGC GGT GGT AGC GGT ACC GTA
        V   A   N   G   Y   A   V   G   G   G   G   G   Y   G   G   R   G   G   S   G   T   V

+140  ATC GGT GGA CAA GCC TAC CAA ATT TTG CCT GCT TTG CAA GTG CAA ACT ATT GCT GCC GCT GGT GGC
       I   G   G   Q   A   Y   Q   I   L   P   A   L   Q   V   Q   T   I   A   A   A   G   G

+206  TCC TCA GCT GGT TAT GGT GGC TCC TCA GCT GGT TAT GGT GCA TCG TCG GGT AGC TAT GGT GCA TCG
       S   S   A   G   Y   G   G   S   S   A   G   Y   G   A   S   S   G   S   Y   G   A   S

+272  TCG GGT GGT TAC GGT GGC TCA TCG AAT GGA TAT GGT GCG TCC AGT GCA CCA TCG ATT GAT ATT GGA
       S   G   G   Y   G   G   S   S   N   G   Y   G   A   S   S   A   P   S   I   D   I   G

+338  CAA TTA TTG GCC GCC GTT GGT GGT GAC CTT ACC GCA CAA GAA GCT GCC CAA TTA GTA AAT TCT TTG
       Q   L   L   A   A   V   G   G   D   L   T   A   Q   E   A   A   Q   L   V   N   S   L

+404  CCT TCA GCT GGC GGT CCA ATC ATT GAC ACC AGT GGT TCA TCG GCT GGT AGC AGC CAT CAA GGT TCA
       P   S   A   G   G   P   I   I   D   T   S   G   S   S   A   G   S   S   H   Q   G   S

+470  TAC CCT TCG GGC GGT AAT CTA GCT TAT GTT ATT CAA TCT GGT GGT AGC TCA TAC TCT GCA CCA GCA
       Y   P   S   G   G   N   L   A   Y   V   I   Q   S   G   G   S   S   Y   S   A   P   A

+535  CCT GCT GCT TCA TAC AGT GCC CCA GCA CCA GCG CCA GCT GCT TCG TAC AGT GCT CCA GCT CCT TCT
       P   A   A   S   Y   S   A   P   A   P   A   P   A   A   S   Y   S   A   P   A   P   S

+602  TAC AGT GCT CCA GCA CCA GCG CCA GCT CCT TCT TAC AGT GCT CCA GCT CCT TCT TAC AGT GCT CCA
       Y   S   A   P   A   P   A   P   A   P   S   Y   S   A   P   A   P   S   Y   S   A   P

+668  GCT CCT TCT TAC AGT GCT CCA GCA CCA GCG CCC GCA CCA GCT GCA TAT AGT GCT CCT GCA CCT GCT
       A   P   S   Y   S   A   P   A   P   A   P   A   P   A   A   Y   S   A   P   A   P   A

+734  GTA TAC AGT GCT CCT GCA CCA GCT GCA TAT AGT GCT CCT GCA CCT GCT GTA TAC AGT GCT CCA GCA
       V   Y   S   A   P   A   P   A   A   Y   S   A   P   A   P   A   V   Y   S   A   P   A

+800  CCA GCG CCC GCA CCA GCT GCA TAT AGT GCT CCA GCA CCT GCT GCA TAC AGT GCT CCG GCT CCT GCT
       P   A   P   A   P   A   A   Y   S   A   P   A   P   A   A   Y   S   A   P   A   P   A

+866  GCA TAC AGT GCT CCT GCA TCA TCT GGC TAT GGT GCC TCA GCT CCT GCC GCC GCT GCT CCA GCT GCA
       A   Y   S   A   P   A   S   S   G   Y   G   A   S   A   P   A   A   A   A   P   A   A

+932  GCA CAT CAG CCA TCA GCA GCC GCT GCT CGC AGC TAT ATC TCA GGC AGT TAC GGT GCT GCC TAC GCA
       A   H   Q   P   S   A   A   A   A   R   S   Y   I   S   G   S   Y   G   A   A   Y   A

+998  CCC GCA CCC GCC CCT GCT GCT GGT GGT GCT TAT TAA ATCTGTGATATTAAATATTTTCACTAAATTCAGTGTTCG
       P   A   P   A   P   A   A   G   G   A   Y  end

+1073 ATCAAAAGATAGAATAAAACGCGATATTTCGAACAAAG
```

FIGURE 3.—Complete cDNA sequence of medfly *Ccs18* (*C4* cDNA clone) and the encoded protein. Numbers refer to the nucleic acid sequence. The putative polyadenylation signal AATAAA is bold and underlined. The best conserved peptide sequence detected after alignment with all Drosophila *s18* proteins is boxed, and 14 imperfect tandem repeats of the heptapeptide SYSAPAP are bracketed; deviations from this consensus are dotted. The repeats are spaced by pure alanine and proline segments, which are also found elsewhere in the protein (dashed underlining). Tyrosines showing periodicities of six, seven, eight or nine residues are circled. Possible cleavage sites of the signal peptide are shown with vertical arrows; the most probable is that preceding a tyrosine, which is also seen in Drosophila chorion proteins.

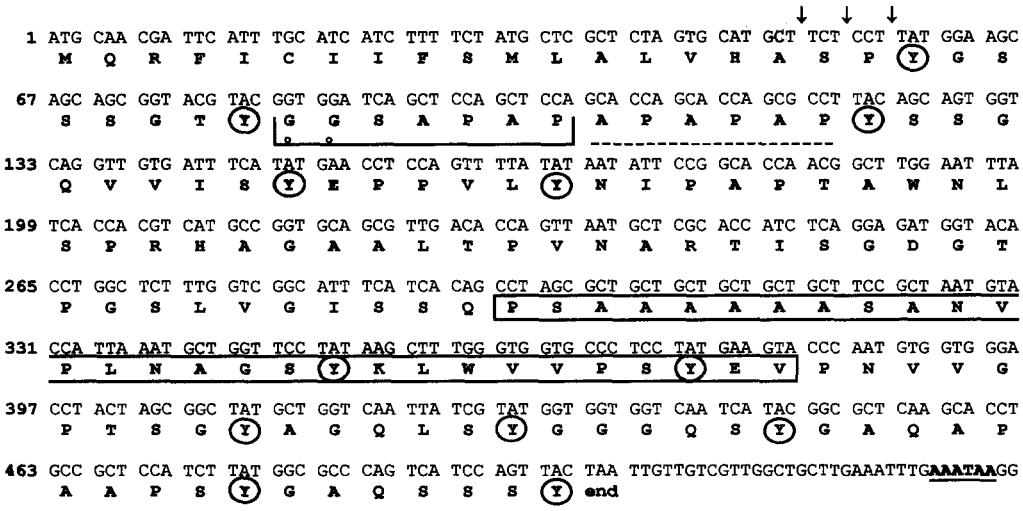

FIGURE 4.—cDNA and deduced animo acid sequence of medfly *Ccs15*. This cDNA was produced by the RACE-PCR method (MATERIALS AND METHODS), using as gene-specific primer the oligonucleotide ATGCAACGATTCATTTGC, defined by preliminary genomic sequencing of the beginning of the ORF. Annotations as in Figure 3.

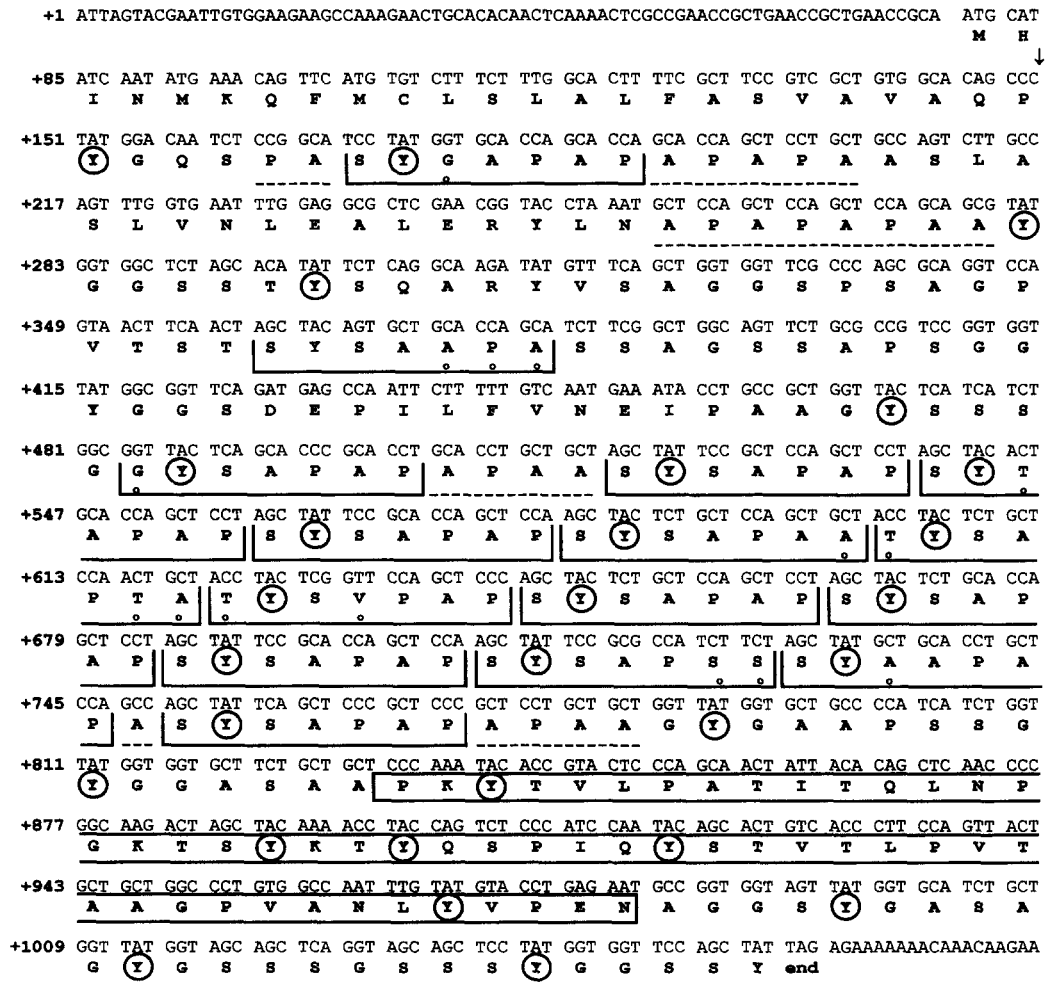

FIGURE 5.—CompletecDNA sequence of medfly *Ccs19* (*C1* cDNA clone). Annotations as in Figure 3.

```
                                                                  ↓  ↓        ↓  ↓
  1 ATG CTC CGC TTC ACG GTT TGC TTG GTG GCC ACG GCC GTG TGT ATA TGC ACA TCC GCG TTG GCT TCT
    M   L   R   F   T   V   C   L   V   A   T   A   V   C   I   C   T   S   A   L   A   S

 67 CCA CAA TAT GGC AGC GCT CCA ACT CCA GCA GCA TCC TCG TAT GGC GGG AGT GGT GGT TAT GCC GGT
    P   Q   Y   G   S   A   P   T   P   A   A   S   S   Y   G   G   S   G   G   Y   A   G

133 GCC GGT GCG AGT GGC GCA CCA GCA GCT CAG GTA ATT GAA TTG CCC GCC AAT GTT GAA GCT CAA GCG
    A   G   A   S   G   A   P   A   A   Q   V   I   E   L   P   A   N   V   E   A   Q   A

199 GCT GCT TTG ACA AAT GCT GCC GGC GCT AAG GCG ACA GCT GCC AAA ATC AAT GCA CTC AAA TGG GAT
    A   A   L   T   N   A   A   G   A   K   A   T   A   A   K   I   N   A   L   K   W   D

265 TTG CTC AAT TTG TAC GGT TAT GAA ATC GGT TAT CCA TTG TTG GTG AAG AAA TAC GGA CCA TTG ACC
    L   L   N   L   Y   G   Y   E   I   G   Y   P   L   L   V   K   K   Y   G   P   L   T

331 TCA TTG TTC TCC GCC TCG TTG CCA CCA CGC AGT TTT GTT GGA AAG ATC GAT CCT GCT TTC TTG AAA
    S   L   F   S   A   S   L   P   P   R   S   F   V   G   K   I   D   P   A   F   L   K

397 GAC TCT TAC GGC AAG ATC AAA TAT GTC GGC GAA AGT TCG ATC GGT GTT GCG GCT ATT TAA AATCGTG
    D   S   Y   G   K   I   K   Y   V   G   E   S   S   I   G   V   A   A   I  end

464 CGCCTAAAACGCGACTTGTAAATTATTGTTAAAGTAGCTAAGAATAAAACTATACTTTTTTAAGTGG
```

FIGURE 6.—cDNA sequence of medfly *Ccs16*. This cDNA was produced by the RACE-PCR method, using as gene-specific primer, the oligonucleotide ATGCTCCG CTTCACGGTT, defined by preliminary genomic sequencing of the beginning of the ORF. Annotations as in Figure 3.

according to their order in the locus relative to the paramyosin gene.

Characterization of the medfly proteins: The predicted polypeptides were used to search the SWISSPROT database through the BLASTX program (ALTSCHUL *et al.* 1990). The closest matches were detected with the Drosophila *s18, s15, s19, s16* proteins, permitting the *C. capitata* genes to be renamed *Ccs18*

s18

```
Dms18    32...VGGYAYQVQPALTVKAI...48
Dss18    32...IGAYAYQVQPALTVQAI...48
Dvs18    30...VGGYAYQVQPALTVKAI...46
Dgs18    30...VGGLAYQVQPALTVSSI...46
Ccs18    36...IGGQAYQILPALQVQTI...52

consensus     VGGYAYQVQPALTV-AI
```

s15

```
Dms15      70...GSAAAAAAASAAAVNPGTYKQYAIPSYEL...98
Dss15      71...SRASAAAASAAAAIAPGSYSQYAIPRYEI...99
Dvs15      71...PSASAAAAAASAGIRPGRYEQAAVIGYDL...99
Dgs15      57...PSAAAAAAAASSGVNSGLYNQRGVIGYEL...85
Ccs15      99...PSAAAAAASANVPLNAGSYKLWVVPSYEV..127

consensus       PSAAAAAAAA-A--NPG-Y-Q-AVP-YEL
```

s19

```
Dms19    106...PRWTVQPAGATLLYPGQNNYKAYVSPPEYSKVILPIRPAAPVAKLFVPEN...155
Dss19    107...PRWTVQPAGATLLYPGQNNYRAYVSPPEYTKVVLPVRPAEPVAKLYIPEN...156
Dvs19    134...PRWSVQPAGTTLLYPGQNSYRRYASPPEYTKVVLPVRAAPPVAKLYLPEN...183
Dgs19    139...PRFTVQPAGATLLYPGQNSYRRISSPVEYSKVILPVRAAAPVAKLYIPQN...188
Ccs19    252...PKYTVLPATITQLNPGKTSYKTYQSPIQYSTVTLPVTAAGPVANLYVPEN...301

consensus       PRWTVQPAGATLLYPGQNSYR-Y-SPPEYSKV-LPVRAA-PVAKLY-PEN
```

s16

```
Dms16    45...EAQASALTNAAGAAASAAK...63
Dss16    46...EAQAAALTNAAGAAASAAK...64
Dvs16    45...EAQAAALTNAAGAAASSAK...63
Dgs16    48...EAQAAALTNAAGAAAASAK...66
Ccs16    63...EAQAAALTNAAGAKATAAK...81

consensus       EAQAAALTNAAGAAASAAK
```

FIGURE 7.—Identification of the *C. capitata* (*Cc*) chorion genes, by comparison with known *D. melanogaster* (*Dm*), *D. subobscura* (*Ds*), *D. virilis* (*Dv*) and *D. grimshawi* (*Dg*) sequences. These conserved, uninterrupted segments were identified in BLASTX alignments and are shown as boxes in Figures 3–6. Shading indicates residues that do not vary in more than one sequence, and consensus is defined by three or more sequences.

TABLE 1

Comparison of amino acid residues in chorion proteins of *C. capitata* and *D. melanogaster*

	s18		s15		s19		s16	
AA type	Cc	Dm	Cc	Dm	Cc	Dm	Cc	Dm
Nonpolar								
Ala A	96[a]	28	27[b]	18	73[a]	23	29	29
Val V	10[b]	16	2[b]	7	12	10	9	9
Leu L	9	6	9[c]	5	13[b]	6	15	15
Ile I	10[b]	4	7	4	5	8	8[b]	1
Pro P	45[a]	15	20[a]	7	48[a]	13	10	8
Met M	1	3	2	1	3	1	1	2
Phe F	2	1	2	1	3	3	4	2
Trp W	0	1	2	0	0	1	1	2
Polar								
Gly G	46[a]	28	19	21	30[b]	23	16	14
Ser S	54[a]	18	25[a]	8	63[a]	16	14[b]	8
Thr T	5	2	6[c]	2	15[a]	3	7	4
Cys C	1	3	1	1	1	3	3	3
Tyr Y	29[a]	17	12[c]	16	33[a]	19	9	10
Asn N	5	4	6	5	7	8	4	4
Gln Q	10[c]	6	7[c]	3	7[b]	12	3	3
Acidic								
Asp D	3	1	1	1	1	1	3	6
Glu E	1	4	2	5	5	7	4	5
Basic								
Lys K	0[b]	9	1	2	4	4	9	7
Arg R	3	4	3[c]	7	2[b]	12	2	5
His H	2	2	2	1	1	0	0	1
Total AA	332	172	166	115	326	173	151	138
kDa	30.4	17.3	16.5	12.0	31.5	18.5	15.2	14.3
pI	6.5	11.1	10.2	11.2	7.0	12.0	10.2	9.3

a, *b*, and *c* indicate differences in numbers of specific residues between *C. capitata* and *D. melanogaster* chorion proteins in either direction. *a*, >10 residues; *b*, >4 residues; and *c*, >3 residues.

(*C4*), *Ccs15*, *Ccs19* (*C1*), and *Ccs16*, respectively, in increasing proximity to the paramyosin gene. Figure 7 presents the most conserved peptide sequences that show diagnostic similarities to the aligned sequences of the corresponding proteins in four diverse Drosophila species (LEVINE and SPRADLING 1985; WONG *et al.* 1985; MARTINEZ-CRUZADO *et al.* 1988; FENERJIAN *et al.* 1989; SWIMMER *et al.* 1990), unequivocally establishing the identity of the medfly genes. These diagnostic sequences are boxed in Figures 3–6. Additional similarities exist, but the fruitfly and medfly genes are highly diverged for much of their length, consistent with our inability to identify them by cross-hybridization. For convenience, we will use the term "autosomal chorion locus" for this medfly gene cluster, as in Drosophila.

The most striking difference is that the medfly chorion proteins are invariably longer than their Drosophila homologues: by 9% in the case of *Ccs16*, 44% for *Ccs15*, 88% for *Ccs19* and 93% for *Ccs18* (Table 1). This disparity would be even greater if the putative signal peptides, which are usually of similar length, were omitted from the comparison (Figures 3–6). It should be noted that all *C. capitata* chorion components migrate on SDS gels more slowly than their expected molecular weights, but *Ccs19* (*C1*) and *Ccs18* (*C4*) show an unusually large retardation: they behave as proteins of 68 and 59 kD, respectively (according to TOLIAS *et al.* 1990) or 66 and 61 kD (according to MOUZAKI and MARGARITIS 1991), although they are 31.5 and 30.4 kD, including the signal peptide, as determined by sequencing (Table 1). Slower migration on SDS gels has been noted in some gel systems for *D. melanogaster* chorion proteins (WARING and MAHOWALD 1979) and may largely reflect unusual features of sequence and structure. However, early attempts at detecting chorion glycosylation showed the presence of 6% sugars (PETRI *et al.* 1976), and KOCH and SPITZER (1982) reported that chorion was heavily labeled after *in vivo* incubation with radioactive sugars.

The increased length of the medfly sequences as compared with Drosophila is correlated with substantial enrichment in a few amino acids (Table 1). More than half the increased length of *Ccs15* is due to proline and serine. *Ccs19* and *Ccs18* are strongly enriched in these same two amino acids plus alanine and tyrosine, while they also show strong enrichment in one additional

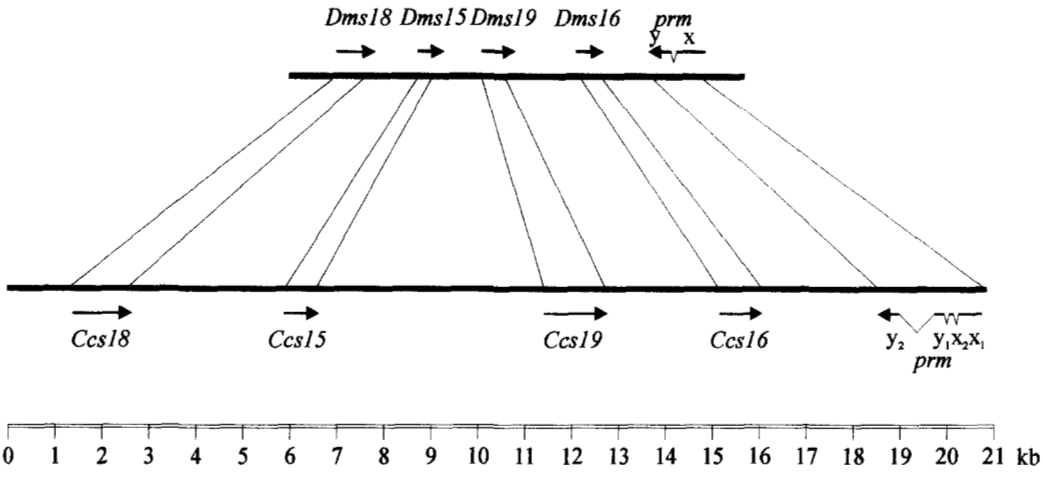

FIGURE 8.—Comparison of the autosomal chorion locus in *D. melanogaster* and *C. capitata* encompassing the paramyosin gene (*prm*) and four similarly organized chorion genes. In Ceratitis, the locus shows size expansion in both coding (especially *s18* and *s19*) and intergenic regions. Note that the sequenced portion of the paramyosin gene shows three introns as compared with one in the Drosophila homologue (see Figure 1).

amino acid each: threonine in the case of *Ccs19*, and glycine in the case of *Ccs18*. In both *Ccs19* and *Ccs18*, a substantial part of the increased length is due to a novel segment consisting of imperfect tandem repeats of the same heptapeptide consensus sequence, SYSAPAP (bracketed in Figures 3 and 5). There are 13 such repeats in the case of *Ccs19* and 14 in *Ccs18*, either directly abutting each other or separated by one to five residues, which are invariably alanine or proline (dashed underlining in Figures 3 and 5). Only one imperfect repeat of this consensus is encountered in *Ccs15* and is followed by three repeats of AP (Figure 4). The SYSAPAP consensus is not encountered in *Ccs16* (Figure 6) nor in any of these four chorion proteins in any Drosophila species that has been studied to date.

In addition to the interrelated differences in length, amino acid content and occurrence of SYSAPAP-like repeats, the medfly chorion proteins differ from their *D. melanogaster* homologues in isoelectric point. *Ccs18* and *Ccs19* are substantially less basic than their *D. melanogaster* homologues, and *Ccs15* is also somewhat less basic. Again, *Ccs16* is deviant in that it is actually more basic than *Dms16*.

The locus as a whole differs substantially in length, because of substantially longer intergenic DNA in *C. capitata* as compared with *D. melanogaster* or other Drosophila species (MARTINEZ-CRUZADO *et al.* 1988). Thus the distance from the 5′ end of s18 to the 3′ end of the paramyosin gene is 6.9 kb in *D. melanogaster* and 17.2 kb in *C. capitata* (Figure 8). Nevertheless, the order and orientation of the genes, with respect to each other as well as to the neighboring paramyosin gene, are perfectly conserved (Figure 8). Each medfly gene has a single intron, at the same position as the Drosophila homologue (data not shown).

The autosomal chorion genes are amplified in *C. capitata*: In all Drosophila species examined, the autosomal chorion genes amplify differentially in the ovarian follicle cells during oogenesis (SPRADLING and MAHOWALD 1980; SPRADLING 1981; ORR *et al.* 1984; MARTINEZ-CRUZADO *et al.* 1988). This is also true in *C. capitata* (Figure 9). Genomic DNA was prepared from either males or total choriogenic ovaries. Equal amounts of each DNA were digested with *Eco*RI and blot hybridized with three cloned fragments encompassing chorion genes. In each case, autoradiography revealed amplification in ovarian relative to male DNA (Figure 9A). This was confirmed (Figure 9B) by comparison of blots of *Hin*dIII—digested DNA hybridized with chorion probes and a single copy control probe (the *Adh2* cDNA isolated by BOURTZIS and SAVAKIS, unpublished results). Surprisingly, several amplified bands were seen with each of the chorion probes (Figure 9A), suggesting the existence of restriction fragment length polymorphism in the autosomal chorion locus of *C. capitata*; the different intensities of the bands presumably correspond to the frequencies of the alleles in the medfly population. The polymorphism was confirmed by analysis of additional *C. capitata* genomic clones and by Southern analysis from ovaries of individual females (D. VLACHOU and K. KOMITOPOULOU, unpublished results) and is interesting in view of the commonly held view that medfly populations have a low degree of polymorphism.

The chorion genes are expressed specifically in the ovary but show some changes in temporal expression: To test whether expression of autosomal chorion genes is ovarian-specific in the medfly, as it is in Drosophila (GRIFFIN-SHEA *et al.* 1982; FENERJIAN *et al.* 1989), equal amounts of RNAs from choriogenic ovaries and from adult males were subjected to dot blot analysis. We used as probes the same three cloned fragments encompassing chorion genes as in the previous experiment. Indeed, expression was abundant in the ovaries but not detectable in males (Figure 10).

Detailed comparisons of the temporal specificities of the four chorion genes in *D. melanogaster* and *C. capitata* showed clear evidence of regulatory evolution (Figure 11). Staging the choriogenic follicles (egg chambers) in a comparable manner was not straightforward because their morphology differs in the two species (MARGARITIS 1985). We used the staging criteria established by MOUZAKI and MARGARITIS (1991), based on overall size,

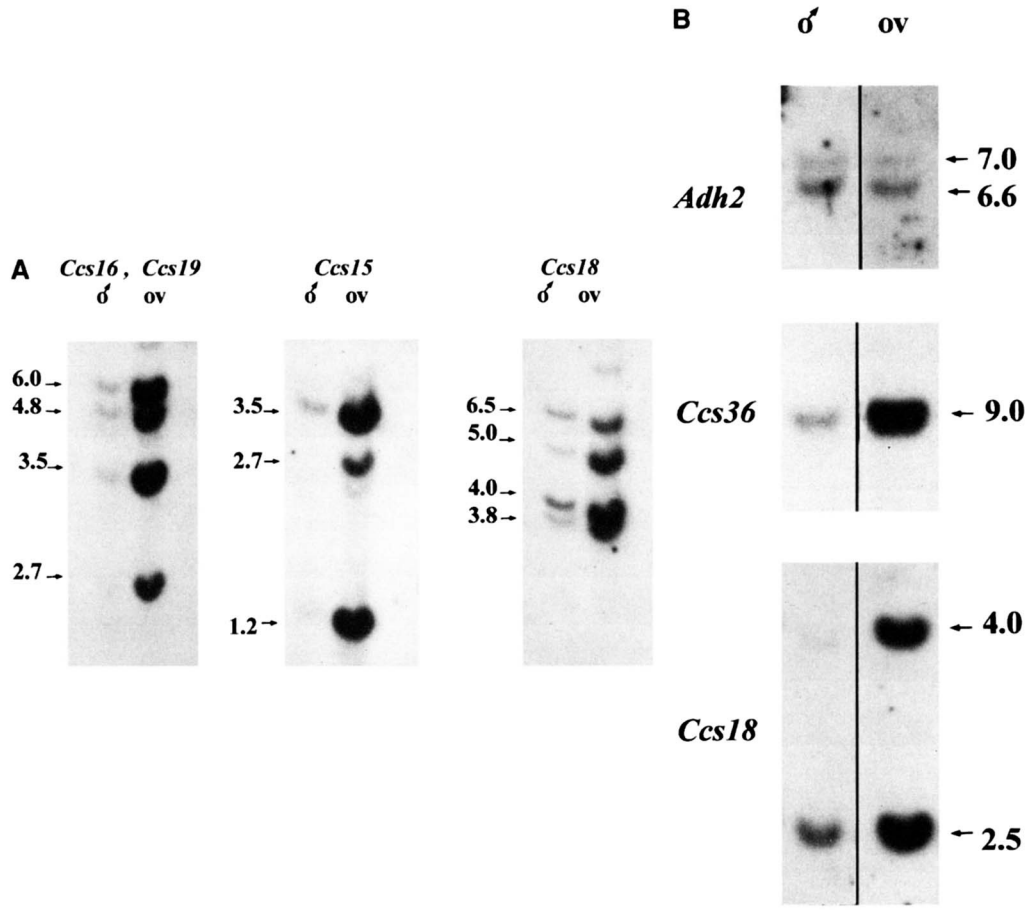

FIGURE 9.—Amplification of the medfly chorion genes in choriogenic ovaries. (A) High molecular weight genomic DNAs were isolated from males and choriogenic ovaries (ov) from *C. capitata* and digested with *Eco*RI; equal aliquots were blot hybridized as described in the text. ^{32}P-labeled *Eco*RI fragments, encompassing *Ccs16* plus *Ccs19*, *Ccs15* and *Ccs18*, respectively, as defined by ovarian RNA hybridization (Figure 10) were used as probes. Note that each probe hybridizes to more than one *Eco*RI band in both males and female ovaries and that these bands are all specifically amplified in ovaries. The weakest bands seen in the ovarian lanes were also seen in the male lanes upon long exposure (data not shown). Additional studies showed that this band multiplicity is due to restriction fragment length polymorphism in the autosomal locus (data not shown). The varying intensities of hydridizing bands indicate different representation of the alleles in the *C. capitata* population. (B) An additional Southern blot hybridization experiment using *Hin*dIII digested genomic DNA. Note that both sex-linked (*Ccs36*) and autosomal (*Ccs18*) chorion genes are amplified in the ovary, unlike the *Adh2* single copy control.

proportion of the oocyte *vs.* nurse cells, and formation of the chorion layers. The clearly stronger stage 12 expression of *s19* in medfly as compared with Drosophila should be viewed with caution because of this uncertainty in *absolute* equivalence of the stages. However, we were able to compare with complete confidence the *relative* temporal specificities of the four genes relative to each other by the simple expedient of using the same RNA blot for all of them. Following the procedure described by FENERJIAN *et al.* (1989), we isolated RNAs from pools of large numbers (50) of staged follicles of *C. capitata,* fractionated the RNAs by electrophoresis, transferred them to nylon filters (MATERIALS AND METHODS), and probed the filters sequentially with the four chorion genes. Thus the temporal RNA pattern of each gene served as an internal reference for the others, permitting consistent analysis of the relative timing of gene expression. In *D. melanogaster* (Figure 11, right), the genes are expressed in the order *s19, s16, s18* and *s15*, from earliest to latest expressed gene. In particular, the mRNA abundance ratios at stages 13 and 14 demonstrate that *s15* is expressed in an overlapping but clearly later manner relative to *s18.* The same temporal order is preserved in all Drosophila species (FENERJIAN *et al.* 1989). In contrast, repeated experiments in *C. capitata* (Figure 11, left and data not shown) revealed a clear alteration of the gene expression order. In particular, *Ccs15* has the same RNA expression pattern as *Ccs16* and is clearly earlier than *Ccs18.*

DISCUSSION

Extensive conservation of chromosomal linkage groups has been observed in Diptera, including both Drosophila and Ceratitis (FOSTER *et al.* 1981; LOUKAS and KAFATOS 1986; MALACRIDA *et al.* 1986; WHITING *et*

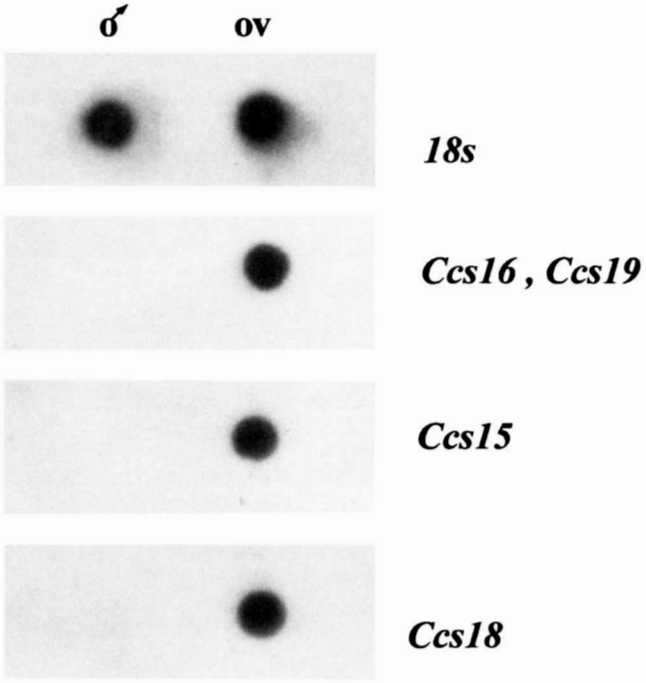

FIGURE 10.—Dot blot analysis of RNAs isolated from males and female ovaries (MATERIALS AND METHODS). All *Eco*RI restriction fragments mapping downstream of the paramyosin gene (see Figure 2) were used as probes and showed hybridization with ovarian but not male RNA, demonstrating the ovarian specificity expected of chorion genes. Ribosomal *18s* hybridization served as a control demonstrating comparable RNA loading.

al. 1989; ZACHAROPOULOU 1990). This synteny occurs despite the occurrence of a number of inversions and other chromosomal rearrangements; for example, the haploid female genome in the medfly includes three metacentric, two submetacentric and one acrocentric chromosome, while in Drosophila it includes two metacentric, one acrocentric and one dot chromosome. The possibility of retention of close linkage with the paramyosin gene was the basis of our successful strategy for cloning the chorion locus. First we cloned the paramyosin gene by homology and showed that it is extensively conserved relative to the single-copy paramyosin gene of Drosophila (BECKER *et al.* 1992; VINOS *et al.* 1992), which maps close to the autosomal chorion genes (FENERJIAN *et al.* 1989). As predicted, a short chromosomal walk and characterization of the transcription units downstream of the paramyosin gene revealed the existence of homologues of the four Drosophila autosomal genes, two of which were identified subsequently with the *C1* and *C4* sequences.

The study of this gene cluster complements earlier studies of the evolution of the autosomal chorion locus in diverse species of the genus Drosophila (MARTINEZ-CRUZADO *et al.* 1988; FENERJIAN *et al.* 1989; MARTINEZ-CRUZADO 1990; SWIMMER *et al.* 1990). By comparison of the organization and sequence of the locus in a fly belonging to a non-Drosophila family, we have confirmed several highly conserved features while also detecting novel events in the evolution of the locus.

The most impressively conserved feature is the presence of the same genes in both Drosophila and Ceratitis, in the same order and orientation. This invariance for ~120 million years (BEVERLEY and WILSON 1984) is in striking contrast to the radical differences in chorion gene sequences and their chromosomal organization in flies and silkmoths, which have been separated twice as long in evolution (KAFATOS *et al.* 1987). The conserved organization within the Diptera may be related to the apparently ancient invention of chorion gene amplification, a developmental mechanism that does not occur in silkmoths and permits rapid eggshell formation during oogenesis with a small number of structural genes (SPRADLING and MAHOWALD 1980; SPRADLING 1981; OSHEIM and MILLER 1983; OSHEIM *et al.* 1988). Amplification has now been documented for both the autosomal (this report) and the sex-linked chorion genes of the medfly (KONSOLAKI *et al.* 1990; TOLIAS *et al.* 1990) as in Drosophila. Quantification of the level of amplification (as in SPRADLING 1981) indicates values of 16× for the sex-linked locus and 14× for the autosomal locus, whereas in Drosophila the autosomal locus is amplified more extensively. This putative difference will need to be confirmed by quantification in staged egg chambers rather than whole medfly ovaries. In Drosophila species, an essential amplification control element (ACE) has been mapped upstream of gene *s18* (ORR-WEAVER and SPRADLING 1986; ORR-WEAVER *et al.* 1989; SWIMMER *et al.* 1990) and accessory amplification enhancing elements (AEEs) appear to map in the DNA flanking the other genes, permitting high-level amplification of the locus as a single amplicon (DELIDAKIS and KAFATOS 1987, 1989). Detailed sequence alignments of the intergenic sequences in the autosomal chorion locus of *C. capitata* and *D. melanogaster* reveal well-conserved elements that may be candidate *cis*-regulatory elements for amplification as well as expression in late oogenesis (D. VLACHOU and K. KOMITOPOULOU, unpublished results).

Despite the conserved organization of the chorion locus, its length varies substantially, by 149%. All intergenic regions are longer in the medfly, by factors ranging from 2× for the *s19/s16* intergenic region to 5× for *s15/s19*. Sequence analysis of these regions will be reported elsewhere (D. VLACHOU and K. KOMITOPOULOU, unpublished results). The medfly paramyosin gene is also substantially longer in the medfly, showing three introns within a region where the Drosophila homologue has only one. If the compact nature of the Drosophila genome is a secondary character, this well-studied multigene locus manifests a process of compaction in Drosophila, eliminating presumably dispensable sequences in the intergenic regions as well as whole introns.

The proteins encoded by the chorion genes have re-

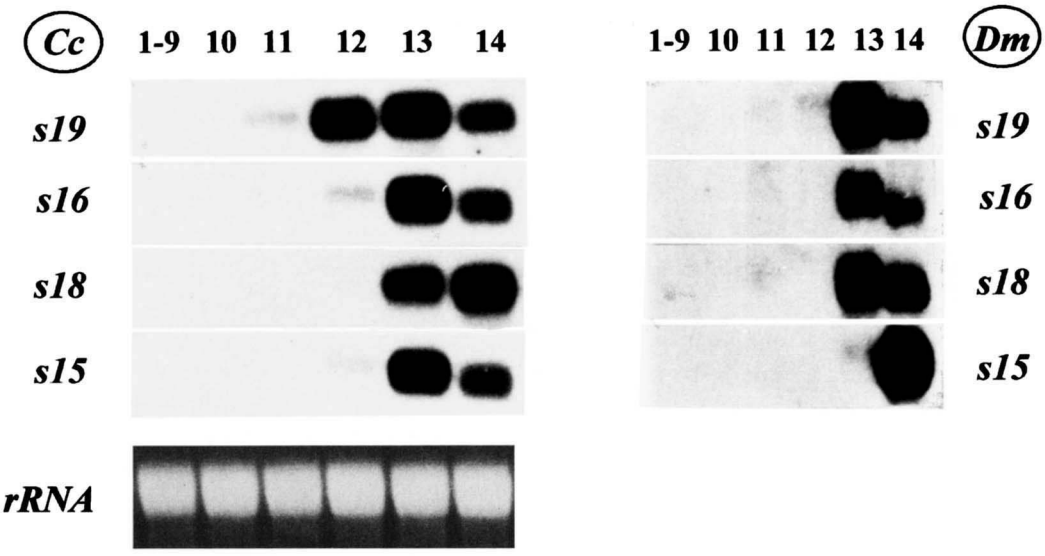

FIGURE 11.—Northern analysis of autosomal chorion gene expression in *C. capitata* and *D. melanogaster*. The relative temporal expression profiles of the four chorion genes during oogenesis are compared with confidence because the same RNA blot of the appropriate species was used for all four genes (in four cycles of hybridization, stripping of label and reusage with the next gene). Expression patterns are arranged according to the order of temporal specificity of the genes in *D. melanogaster*. Follicles were staged using criteria comparable in the two species (MATERIALS AND METHODS), and 5 μg of total RNA was prepared from each pool, electrophoresed, blotted and probed with homospecific probes. Note the clear shift of temporal specificity of *s15*, which is later than *s18* in *D. melanogaster* but earlier in *C. capitata*, as shown by the intensity ratios of the signals at stages 13 and 14; weak expression of *s15* but not *s18* at stage 12 is also seen in the medfly experiments, making the expression pattern of *s15* as early as that of *s16*. Total RNA loading was equal in all *C. capitata* samples, as shown by the ethidium bromide stained *rRNA* control. *D. melanogaster* patterns are from FENERJIAN *et al.* (1989).

tained sufficient similarities at the sequence level to be identified with their Drosophila homologues but have otherwise diverged drastically, confirming the generalization that chorion proteins evolve very rapidly (MARTINEZ-CRUZADO 1990). Nevertheless, secondary structure predictions (computer program GCG) suggest that the proteins are highly structured, with β sheets alternating with β turns (data not shown), as is true for chorion proteins of other insects (HAMODRAKAS *et al.* 1985; HAMODRAKAS 1992). In line with the larger size of this locus as a whole in Ceratitis, the proteins are invariably longer than their Drosophila homologues. This difference is minimal in the case of *Ccs16*, which is the least rapidly evolving autosomal chorion protein, both between the genera and among different Drosophila species (FENERJIAN *et al.* 1989). This protein is also unusual in that the alteration of putative β sheets and β turns is interrupted centrally by a predicted α helix (data not shown); this domain includes the segment that is highly conserved in all Dipteran species studied (Figure 7).

The genes of the second chorion locus, located in a separate chromosome (*Ccs36* and *Ccs38*), are conservative in both sequence and length (KONSOLAKI *et al.* 1990; TOLIAS *et al.* 1990). At the other extreme, the *Ccs18* and *Ccs19* proteins are nearly double the length of their Drosophila homologues. The difference is largely due to a novel tandem array of consensus SYSAPAP repeats, not encountered in Drosophila. This array is found in the central region of *Ccs19* and near the C terminus in *Ccs18*; individual repeats that conform to the same consensus are also seen in the aminoterminal region of *Ccs19* as well as *Ccs15*. The repeats are often associated with extensions consisting purely of alanine and proline; similar AP-segments are also found in isolation elsewhere in the sequences (dashed underlining in Figures 3–5). Moreover, a related octapeptide repeat, Y(G or S)AAPAAS, is found in the C-terminal region of the *Ccs36* sequence (AGGELI *et al.* 1991). Because of the total absence of these features from Drosophila, we favor the hypothesis that the SYSAPAP repeats and AP segments evolved *de novo* during the period separating the two families of flies, although we cannot eliminate the possibility that they were lost selectively from Drosophila. In any case, it is notable that these features appear or disappear in a concerted manner from *s18*, *s19* and to a lesser extent *s15*, proteins which do not otherwise show extensive sequence homologies.

Another notable feature of the medfly chorion proteins is the semiregular distribution of tyrosines. In all four sequences, the vast majority of tyrosine residues are found at a distance of six to nine residues from each other, both in the SYSAPAP repeat array and elsewhere (Figures 3–6). The mature Drosophila chorion shows extensive tyrosine cross-links (PETRI *et al.* 1976). We suspect that the average eight-residue tyrosine periodicity in these medfly proteins reflects a species-specific dityrosine bonding pattern. Different tyrosine periodicities, related to GYGG repeats, have been noted in other

chorion proteins (HAMODRAKAS *et al.* 1989; AGGELI *et al.* 1991).

The temporal specificities of the autosomal medfly chorion genes are of special interest. Their Drosophila homologues are all expressed late in oogenesis in a completely conserved sequential and overlapping pattern, which is the same in four different species, ranging ≤60–80 million years since their last common ancestor (FENERJIAN *et al.* 1989). This expression pattern, from earliest to latest, is *s19* < *s16* < *s18* < *s15*, as exemplified in Figure 11 for *D. melanogaster*. Irregular low-level expression at prechoriogenic or very early choriogenic stages is sometimes seen in Drosophila (THIREOS *et al.* 1980; FENERJIAN *et al.* 1989), but it has no bearing on the strict sequential pattern during choriogenesis. This pattern is retained in *C. capitata* as far as *s19*, *s16* and *s18* are conserved. However, *Ccs15* shows a markedly changed temporal specificity, which is as early as that of *Ccs16* and clearly earlier than that of *Ccs18*. This is the first temporal regulatory change that has been detected for any Dipteran chorion gene. Like *s19*, *s16* and *s18*, the early chorion genes *s36* and *s38* show similar expression patterns in Ceratitis and Drosophila (TOLIAS *et al.* 1990). In Drosophila, *s15* is expressed in a very short period of ca. 2 hr at the end of choriogenesis (PETRI *et al.* 1976; WARING and MAHOWALD 1979). Thus, we expect that its substantially earlier expression in the medfly is due to evolutionary change in the transcriptional *cis*-regulatory elements, or the corresponding *trans* factors, although change at a posttranscriptional level of regulation cannot be rigorously excluded.

Early transformation experiments established that the *Dms15* gene can be expressed autonomously when transferred to multiple chromosomal sites, together with its regulatory elements (WAKIMOTO *et al.* 1986). Transformation with smaller constructs established that short sequences of DNA upstream of the *Dms15* transcriptional start site are enough for normal expression (MARIANI *et al.* 1988; ROMANO *et al.* 1988). Several *cis* elements were identified as necessary for developmentally correct expression. One of them, TCACGT, is typical of all chorion genes examined to date, being found in all Drosophila species, in *C. capitata* (KONSOLAKI *et al.* 1990; D. VLACHOU and K. KOMITOPOULOU, unpublished results) and almost all silkmoths (SPOEREL *et al.* 1986; MITSIALIS *et al.* 1989). Deletion mutations between −189 and −39 of the *Dms15* promoter led to quantitatively reduced or temporally disrupted specificity, suggesting that this region encompasses many positive and negative, partially degenerate *cis*-regulatory elements that specify the highly precise expression pattern of the gene (MARIANI *et al.* 1996). Similar transformation experiments are necessary to reveal the molecular basis of the unusual evolutionary change in temporal regulation in the medfly. Importantly, such analysis can now be performed reciprocally, because of the recent establishment of a robust germ-line transformation method in the medfly (LOUKERIS *et al.* 1995).

We thank G. CHRISTOPHIDIS for excellent technical assistance, A. ZACHAROPOULOU for *in situ* hybridization information and S. HAMODRAKAS for discussions of protein secondary structure predictions. This work was supported by the University of Athens and by the General Secretariat of Research and Technology of the Greek Ministry of Development.

LITERATURE CITED

AGGELI, A., S. J. HAMODRAKAS, K. KOMITOPOULOU and M. KONSOLAKI, 1991 Tandemly repeating peptide motifs and secondary structure in *Ceratitis capitata* eggshell proteins *Ccs36* and *Ccs38*. Int. J. Biol. Macromol. **13:** 307–315.

ALTSCHUL, S. F., W. GISH, W. MILLER, E. W. MYERS and D. J. LIPMAN, 1990 Basic local alignment search tool. J. Mol. Biol. **215:** 400–410.

ANWAR, M., D. L. CHAMBERS, I. OHINATA and R. M. KOBAYASHI, 1971 Radiation sterilization of the mediterranean fruit fly (Diptera: Tephritidae): comparison of spermatogenesis in flies treated as pupae or adults. Ann. Entomol. Soc. Am. **64:** 627–633.

BACK, E. A., and C. E. PEMBERTON, 1918 The mediterranean fruit fly in Hawaii. In *United States Department of Agriculture Bulletin 536*. Government Printing Office, Washington, DC.

BECKER, D. K., P. T. O'DONNELL, J. M. HEITZ, M. VITO and S. I. BERNSTEIN, 1992 Analysis of Drosophila paramyosin: identification of a novel isoform which is restricted to a subset of adult muscles. J. Cell Biol. **116:** 669–681.

BEVERLEY, S. M., and A. C. WILSON, 1984 Molecular evolution in Drosophila and the higher Diptera. II. A time scale for fly evolution. J. Mol. Evol. **21:** 1–13.

BIGGEN, M. D., T. J. GIBSON and G. F. HONG, 1983 Buffer gradient gels and ^{35}S label as an aid to rapid DNA sequence determination. Proc. Natl. Acad. Sci. USA **80:** 3963–3965.

BOUHIN, H., J. P., C. B. QUENNEDEY and J. DELACHAMBRE, 1992 Developmental profiles of epidermal mRNAs during the pupal-adult molt of *Tenebrio molitor* and isolation of a cDNA clone encoding an adult cuticular protein: effects of a juvenile hormone analogue. Dev. Biol. **149:** 112–122.

CAREY, R. J., 1991 Establishment of the mediterranean fruit fly in California. Science **253:** 1369–1373.

CHRISTENSON, L. D., and R. H. FOOTE, 1960 Biology of the fruit flies. Annu. Rev. Entomol. **5:** 171–192.

CHURCH, G. M., and W. GILBERT, 1984 Genomic sequencing. Proc. Natl. Acad. Sci. USA **81:** 1991–1995.

DELIDAKIS, C., and F. C. KAFATOS, 1987 Amplification of a chorion gene cluster in Drosophila is subject to multiple *cis*-regulatory elements and to long-range position effects. J. Mol. Biol. **197:** 11–26.

DELIDAKIS, C., and F. C. KAFATOS, 1989 Amplification enhancers and replication origins in the autosomal chorion gene cluster of Drosophila. EMBO J. **8:** 891–901.

FEINBERG, A. P., and B. VOGELSTEIN, 1983 A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. **132:** 6–13.

FENERJIAN M. G, J. C. MARTINEZ-CRUZADO, C. SWIMMER, D. L. KING and F. C. KAFATOS, 1989 Evolution of the autosomal chorion cluster in Drosophila. II. Chorion gene expression and sequence comparisons of the *s16* and *s19* genes in evolutionarily distant species. J. Mol. Evol. **29:** 108–125.

FOSTER, G. G., M. J. WHITTEN, C. KONOVALOV, S. T. A. ARNOLD and G. MAFFI, 1981 Autosomal genetic maps of the Australian Sheep Blowfly, *Lucillia cuprina dorsalis* R-D (Dipteral, Calliphoridae) and possible correlations with the linkage maps of *Musca domestica* L, and *Drosophila melanogaster* (Mg.). Genet. Res. **37:** 55–69.

FROMAN, M. A., 1990 RACE: Rapid amplification of cDNA ends, pp. 28–38 in *PCR Protocols: A Guide to Methods and Applications*. Academic Press, New York.

GRIFFIN-SHEA, R., G. C. THIREOS and F. C. KAFATOS, 1982 Organization of a cluster of four chorion genes in Drosophila and its relationship to developmental expression and amplification. Dev. Biol. **91:** 325–336.

HAMODRAKAS, S. J., 1992 Molecular architecture of helicoidal proteinaceous eggshells. Results Probl. Cell Differ. **19:** 115–186.

HAMODRAKAS, S. J., T. ETMEKTZOGLOU and F. C. KAFATOS, 1985 Amino acid periodicities and their structural implications for the evolutionarily conservative central domain of some siklmoth chorion proteins. J. Mol. Biol. **186:** 583–589.

HAMODRAKAS S. J., A. BATRINOU and T. CHRISTOPHORATOU, 1989 Structural and functional features of Drosophila chorion protein *s36* and *s38* from analysis of primary structure and infrared spectroscopy. Int. J. Biol. Macromol. **11:** 307–313.

HENIKOFF, S., 1987 Exonuclease III generated detetions for DNA sequence analysis. Promega Notes No. 8.

HOLMES, D. S., and J. BONNER, 1973 Preparation, molecular weight, base composition, and secondary structure of giant nuclear ribonucleic acid. Biochemistry **12:** 2330–2338.

JACOBS-LORENA, M., 1980 Dosage of *5S* and ribosomal genes during oogenesis of *Drosophila melanogaster.* Dev. Biol. **80:** 134–145.

KAFATOS, F. C., C. W. JONES and A. EFSTRATIADIS, 1979 Determination of nucleic acid sequence homologies and relative concentrations by a dot hybridization procedure. Nucleic Acids Res. **7:** 1541–1552.

KAFATOS, F. C., S. A. MITSIALIS, H. T. NGUYEN, N. SPOEREL, S.G TSITILOU *et al.*, 1987 Evolution of structural genes and regulatory elements for the insect chorion, pp. 161–178 in *Development as an Evolutionary Process,* Vol. 8, edited by R. RAFF and E. C. RAFF. A. R. Liss, New York.

KOCH, E. A., and R. H. SPITZER, 1982 Autoradiographic studies of protein and polysaccharide synthesis during vitellogenesis in Drosophila. Cell Tissue Res. **224:** 315–333.

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.

LEVINE, J. F, and A. C. SPRADLING, 1985 DNA sequence of a 3.8 kilobase pair region controlling Drosophila chorion gene amplification. Chromosoma **92:** 136–142.

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

LOUKAS, M., and F. C. KAFATOS, 1986 The actin loci in the genus Drosophila: establishment of chromosomal homologies among distantly related species by *in situ* hybridization. Chromosoma **94:** 297–308.

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

LOW, R., and T. RAUSCH, 1994 Sensitive, nonradioactive Northern blots using alkaline transfer of total RNA and PCR-amplified biotinylated probes. BioTechniques **17:** 1026–1030.

MALACRIDA A., G. GASPERI, G. F. BISCALDI and R. MILANI, 1986 Persistence of linkage relationships among enzyme loci in some Dipteran species. Atti. Assoc. Genet. Ital. **31:** 121–122.

MARGARITIS, L. H., 1985 Comparative study of the eggshell of the fruit flies *Dacus oleae* and *Ceratitis capitata* (Diptera: Trypetidae). Can. J. Zool. **63:** 2194–2206.

MARIANI, B. D., J. R. LINGAPPA and F. C. KAFATOS, 1988 Temporal regulation in development: negative and positive *cis*-regulators dictate the precise timing of exression of a Drosophila chorion gene. Proc. Natl. Acad. Sci. USA **85:** 3029–3033.

MARIANI, B. D., M. J. SHEA, M. J. CONBOY, I. CONBOY, D. L. KING *et al.*, 1996 Analysis of regulatory elements of the developmentally controlled chorion *s15* promoter in trangenic Drosophila. Dev. Biol. **174:** 115–124.

MAROTO M., J. J. ARREDONDO, M. SAN ROMAN, R. MARCO and M. CERVERA, 1995 Analysis of the paramyosin/miniparamyosin gene. Miniparamyosin is an independently transcribed, distinct paramyosin isoform, widely distributed in invertebrates. J. Biol. Chem. **270:** 4375–4382.

MAROTO, M., J. J ARREDONDO, D. GOULDING, R. MARCO, B. BULLARD *et al.* 1996 Drosophila paramyosin/minparamyosin gene products show a large diversity in quantity. localization, and isoform pattern: a possible role in muscle maturation and function. J. Cell Biol. **134:** 81–92.

MARTINEZ-CRUZADO, J. C., 1990 Evolution of the autosomal chorion cluster in Drosophila. IV. The Hawaiian Drosophila: rapid protein evolution and constancy in the rate of DNA divergence. J. Mol. Evol. **31:** 402–423.

MARTINEZ-CRUZADO, J. C., C. SWIMMER, M. G. FENERJIAN and F. C. KAFATOS, 1988 Evolution of the 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:** 663–667.

MITSIALIS, S. A., S. VELETZA and F. C. KAFATOS, 1989 Transgenic regulation of moth chorion gene promoters in Drosophila: tissue, temporal and quantitative control of four bidirectional promoters. J. Mol. Evol. **29:** 486–95.

MOUZAKI, D. G., and L. H. MARGARITIS, 1991 Choriogenesis in the medfly *Ceratitis capitata* (Wiedermann) (Diptera: Tephretidae) Int. J. Insect Morphol. Embryol. **29:** 51–68.

ORR, W., K. KOMITOPOULOU and F. C. KAFATOS, 1984 Mutants suppressing in *trans* chorion gene amplification in Drosophila. Proc. Natl. Acad. Sci. USA **81:** 3773–3777.

ORR-WEAVER, T. L., and A. C. SPRADLING, 1986 Drosophila chorion gene amplification requires an upstream region regulating *s18* transcription. Mol. Cell Biol. **6:** 4624–4633.

ORR-WEAVER, T. L., C. G. JOHNSTON and A. C. SPRADLING, 1989 The role of ACE3 in Drosophila chorion gene amplification. EMBO J. **8:** 4153–4162.

OSHEIM, Y. N., and O. L. MILLER, JR., 1983 Novel amplification and transcriptional activity of chorion c4 genes in *Drosophila melanogaster* follicle cells. Cell **33:** 543–553.

OSHEIM, Y. N., O. L. MILLER, JR. and A. L. BEYER, 1988 Visualization of *Drosophila melanogaster* chorion genes undergoing amplification. Mol. Cell. Biol. **8:** 2811–2821.

PETRI, W. H., A. R. WYMAN and F. C. KAFATOS, 1976 Specific protein synthesis in cellular differentiation. III. The eggshell proteins of *Drosophila melanogaster* and their program of synthesis. Dev. Biol. **49:** 185–199.

PUSTELL, J., and F. C. KAFATOS, 1982 A high speed, high capacity homology matrix: zooming through SV40 and polyoma. Nucleic Acids Res. **10:** 4765–4782.

PUSTELL, J., and F. C. KAFATOS, 1984 A convenient and adaptable package of computer programs for DNA and protein sequence management, analysis and homology determination. Nucleic Acids Res. **12:** 643–655.

RIGBY, P. W., M. DIECKMANN, C. RHODES and P. BERG, 1977 Labeling deoxyribonucleic acid to high specific activity *in vitro* by nick translation with DNA polymerase I. J. Mol. Biol. **113:** 237–248.

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.

ROMANO, C. P., B. BIENZ-TADMOR, B. D. MARIANI and F. C. KAFATOS, 1988 Both early and late Drosophila chorion gene promoters confer correct temporal, tissue and sex specificity on a Reporter *Adh* gene. EMBO J. **7:** 783–790.

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

SAUL, S. H., 1986 Genetics of the mediterranean fruit fly (*C. capitata:* Wiedmann). Agric. Zool. Rev. **1:** 73–108.

SPOEREL, N., H. T. NGUYEN and F. C. KAFATOS, 1986 Gene regulation and evolution in the chorion locus of *B. mori*: structural and developmental characterization of four eggshell genes and their flanking DNA regions. J. Mol. Biol. **190:** 23–35.

SPRADLING, A. C., 1981 The organization and amplification of two chromosomal domains containing Drosophila chorion genes. Cell **27:** 193–201.

SPRADLING, A. C., and A. P. MAHOWALD, 1980 Amplification of genes for chorion proteins during oogenesis in *Drosophila melanogaster.* Proc. Natl. Acad. Sci. USA **77:** 1096–1100.

SWIMMER C., M. G. FENERJIAN, J. C. MARTINEZ-CRUZADO and F. C. KAFATOS, 1990 Evolution of the autosomal chorion cluster in Drosophila. III. Comparison of the *s18* gene in evolutionarily distant species and heterospecific control of chorion gene amplification. J. Mol. Biol. **215:** 225–235.

TABOR, S., and C. C. RICHARDSON, 1987 DNA sequence analysis with

a modified bacteriophage T7 DNA polymerase. Proc. Natl. Acad. Sci. USA **84:** 4767–4771.

THIREOS, G., R. GRIFFIN-SHEA and F. C. KAFATOS, 1980 Untranslated mRNA for a chorion protein of *Drosophila melanogaster* accumulates transiently at the onset of specific gene amplification. Proc. Natl. Acad. Sci. USA **77:** 5789–5793.

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.

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.

VINOS, J., A. DOMINGO, R. MARKO and M. CERVERA, 1991 Identification and characterization of *Drosophila melanogaster* paramyosin. J. Mol. Biol. **220:** 687–700.

VINOS, J., M. MAROTO, R. GARESSE, R. MARCO and M. CERVERA, 1992 *Drosophila melanogaster* paramyosin: developmental pattern, mapping and properties deduced from its complete coding sequence. Mol. Gen. Genet. **231:** 385–394.

WAKIMOTO, B. T., L. J. KALFAYAN and A. C. SPRADLING, 1986 Developmentally regulated expression of Drosophila chorion genes introduced at diverse chromosomal positions. J. Mol. Biol. **187:** 33–45.

WARING, G. L., and A. P. MAHOWALD, 1979 Identification and time of synthesis of chorion proteins in *Drosophila melanogaster.* Cell **16:**599–607.

WHITING, J. H., JR., M. D. PLILEY, J. L. FARMER and D. E. JEFFERY, 1989 *In situ* hybridization analysis of chromosomal homologies in *Drosophila melanogaster* and *Drosophila virilis.* Genetics **122:** 99–109.

WONG, Y-C., J. PUSTELL, N. SPOEREL and F. C. KAFATOS, 1985 Coding and potential regulatory sequences of a cluster of chorion genes in *Drosophila melanogaster.* Chromosoma **92:** 124–135.

ZACHAROPOULOU, A., 1990 Polytene chromosome maps in the medfly *Ceratitis capitata.* Genome **33:** 184–197.

ZACHAROPOULOU, A., M. FRISARDI, C. SAVAKIS, A. S. ROBINSON, P. TOLIAS *et al.*, 1992 The genome of the Mediterranean fruitfly *Ceratitis capitata.* Localization of molecular markers by *in situ* hybridization to salivary gland polytene chromosomes. Chromosoma **101:** 448–455.

Communicating editor: V. G. FINNERTY