J Mol Evol (2003) 57:355–362
DOI: 10.1007/s00239-003-2488-4

JOURNAL OF MOLECULAR EVOLUTION

A Nested Alpha-Amylase Gene in *Drosophila ananassae*

Jean-Luc Da Lage, Claude Maisonhaute, Frédérique Maczkowiak, Marie-Louise Cariou

Centre National de la Recherche Scientifique (CNRS), Populations, Génétique et Évolution (UPR 9034), 91198 Gif sur Yvette Cedex, France

Received: 8 October 2002 / Accepted: 28 March 2003

Abstract. The amylase gene family of *Drosophila ananassae* consists in seven copies, scattered on several chromosomal arms. We have evidenced that a member of the family, *Amy35*, lies within an intron of a gene homologous to the CG14696 gene of *D. melanogaster*. This nested arrangement seems restricted to the *D. ananassae* subgroup. The nested and the nest genes are encoded on opposite strands. Both are actively transcribed in the midgut at the same time, raising the possibility of interference between their mRNAs. Our data also help to elucidate the history of the *Amy* family, suggesting that *Amy35* arose by duplication and translocation from another ancestral locus, into a formerly short intron, in an ancestor of the subgroup.

Key words: Nested gene — Intron — *Drosophila ananassae* — Amylase — Gene duplication — CG14696

Introduction

The complete deciphering of the *Drosophila melanogaster* genome (Adams et al. 2000) is yielding a continuous stream of discoveries about molecular evolution, particularly in the field of gene and genome organization, and strengthens the interest of a comparative approach in genomics. For example, assigning exon and intron boundaries can be greatly facilitated by simply aligning a sequence of interest with the homologous *D. melanogaster* region. Discrepancies in alignment are expected to correspond to noncoding or intronic regions, whereas coding regions are better conserved. Also, if a large genomic stretch containing several genes is available in a given species, the gene order can be compared to that in *D. melanogaster*. Changes in the gene order reflect and pinpoint fine chromosomal rearrangements and may help to establish correspondences between the chromosomal arms in the species of interest and in the paradigm species.

Drosophila ananassae belongs to the *melanogaster* group, and is a member of the *ananassae* subgroup which consists of 21 species. Various aspects of its biology and genetics have been investigated (for review, see Tobari 1993). The alpha-amylase (*Amy*) gene family is a good model for studies of gene duplications in Drosophilids, and *D. ananassae* is a particularly attractive species for this purpose, because it possesses at least seven *Amy* or *Amy*-related genes, which are located on different chromosomes and are differentially expressed (Da Lage et al. 1996a, 2000). In contrast, *D. melanogaster* has only two contiguous, divergently transcribed *Amy* genes, and they are usually expressed simultaneously. Reconstituting the origin and evolution of this complex family can be facilitated by the knowledge of the situation in more or less related species.

In a previous paper (Da Lage et al. 2000), we described the organization and molecular evolution of the *Amy* genes of *D. ananassae*. Two chromosomal loci were found to harbor active genes: *Amy35* encoding AMY1 and *Amy58* (AMY2) at one locus on chromosome 2; *Amyi5* (AMY3) and *Amy4N* (AMY4) at another locus on chromosome 3. The activities of

Correspondence to: J.-L. Da Lage; *email:* jldl@pge.cnrs-gif.fr

the four isozymes are easily detected on polyacrylamide gels. *Amy35* and *Amy58* are tandemly arranged, and the region has been sequenced. There is a very high degree of sequence similarity between the two genes, extending 500 bp upstream but only 40 bp downstream. The genes in the other locus, *Amy4N* and *Amyi5*, are also very similar to each other, with almost identical surrounding regions, but the organization of this cluster (orientation, intergenic size) is unknown. As will be discussed here, the data available until now did not allow us to clarify which locus was ancestral (i.e., orthologous to that of other *Drosophila* species), and which one appeared from a duplication/translocation event.

For the purpose of the present study, the sequence of *Amy35*/*Amy58* region was extended and compared to the *D. melanogaster* genome (http://www.fruitfly.org/). This comparison shows that one copy has been inserted as a nested gene within another gene, which provides some insight into the origin of the *Amy* gene family in *D. ananassae*. We report here the organization of this region and evolutionary involvements.

Materials and Methods

Cloning and Sequencing

The clone i3a had previously been isolated from a genomic library of the African strain Taï 13-1610 of *D. ananassae*, inserted in the phage vector λ-gem11 (Promega) and partially sequenced manually over 10.5 kbp (Da Lage et al. 2000) (GenBank accession number U53698). It contained two *Amy* genes, *Amy58* and *Amy35*. The full insert length was 13.5 kbp and the remaining 3 kbp were sequenced for the present study using an ABI373 automat (Applied Biosystems). A clone of the homologous region (spanning 4938 bp) was isolated in the related species *Drosophila varians*, a remote member of the *D. ananassae* subgroup, using a minilibrary, as described earlier (Da Lage et al. 2000). The updated sequence of *D. ananassae* has been deposited to GenBank with accession number U53698/AF489266; the sequence of *D. varians* has been deposited with accession number AF486854.

Reverse-Transcriptase PCR (RT-PCR)

Rough nucleic acid extracts were obtained from one to four individuals at various stages of development (Gloor and Engels 1992). When single individuals or organs were extracted separately, several samples of the same stage/organ were assayed in subsequent RT-PCR. The RT-PCR protocol for detecting the expression of the new gene CG14696 was modified from (Huet et al. 1993). In order to avoid amplification of the genomic DNA, samples were DNase treated prior to reverse-transcription. In addition, primers were chosen on both sides of the 800-bp intron 1 of the new gene in *D. melanogaster* (CGMEL1 = GGCGGAACTGATTACGGAG; CGMEL2 = GGGCAACTCCAGCTGCGAG), or *D. ananassae* (CGANA1 = CGCCGAGCTCATTACTGAC; CGANA2 = GGCAGCTCC AGCTGGGAGC) or on both sides of intron 2 in *D. ananassae* (CGAMY351 = GCACGAGGTCACCCTAACC; CGAMY352 = TGGCCAGGCACGTAGGCAG). For each pair of primers, the expected size of the cDNA was approximately 220 bp. Fifteen microliters of RT-PCR products were visualized on 2% agarose-ethidium bromide gels. The remaining (35 μL) were purified and directly sequenced to confirm the identification and to check the intron splicing. A pair of primers with conserved sequences surrounding intron 2 were designed to check the absence of an *Amy* gene within intron 2 in more remote species (CGDIRBIS = CACACCACCACCGAGGTCTACTA; CGINTR2REV = ATRTTYTGGCCRGGCACRTAGGC). The sequences were analyzed using CLUSTALW (Thompson et al. 1994) and SEQAPP by Don Gilbert (ftp.bio.indiana.edu/molbio).

In situ Hybridization

The DNA fragments used as probes are shown in Fig. 1. Digestive tracts of third instar larvae were dissected under a binocular microscope in PBS and then fixed in 4% paraformaldehyde/PBS for 1 h at 4°C, then stored before use in 70% ethanol through 25% and 50% steps. After rinsing the tissues, hybridization was done as described by Tautz and Pfeifle (1989) and DIG-labeled probes revealed with anti-DIG antibody conjugated to alkaline phosphatase following the supplier's protocol. The digestive tracts were mounted on microscope slides in Aquatex medium (Merck, Germany). DIG-labelled DNA probes were prepared using a DIG High-Prime labeling kit (Roche) with 1 μg of DNA. The level of DIG labeling was assessed by comparison with a standard labeled DNA deposited on nitrocellulose strips as indicated by the manufacturer. Similar quantities of labeled DNA were then used for each sample tested. Photographs were taken using a standard Leitz microscope (objective 4×) and scanned before being used to reconstruct the entire tissues.

Results

Gene Organization in D. ananassae

After the *D. melanogaster* complete genome had been released on the Internet, a BLAST search (Altschul et al. 1997) was done against this genome on the intergenic region between *Amy58* and *Amy35* of *D. ananassae*. Close similarity was found between this region and the putative CG14696 gene of *D. melanogaster*, which is a part of accession number AE003689. Surprisingly, similarity was still found with this gene downstream to *Amy35*. The relevant sequence of *D. melanogaster* was then aligned to that of *D. ananassae*. The alignment allowed us to confirm the intron/exon structure previously inferred by computer (http://www.fruitfly.org/): three introns interrupt the CG14696 gene in both species, at identical positions. The first intron is of similar size in both species (805 bp in *D. melanogaster*, 817 bp in *D. ananassae*). The third intron spans 63 bp and 59 bp, respectively. In contrast, whereas in *D. melanogaster* the second intron is 59 bp long, in *D. ananassae* it encompasses the entire *Amy35* gene and its 5′ and 3′ sequences, including regulatory elements (ca. 2 kb). This implies that the *Amy35* gene is nested in the second intron of the CG14696 homolog. The two genes are encoded on opposite strands, as is often observed in nested gene arrangements (Ashburner et al. 1999). An additional neighboring gene has been

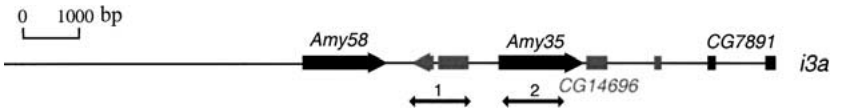

Fig. 1. Diagram of the genomic clone i3a of *D. ananassae*, with the gene arrangement inferred from sequence analysis. *Black arrows and boxes* represent coding areas in the left-to-right sense; *grey arrows and boxes* represent coding areas on the opposite strand. *Amy35* and *Amy58* are located from (Da Lage et al. 2000). *In situ* probes are indicated by two-head arrows numbered 1 (CG14696) and 2 (*Amy35*).

partially cloned at the 3′ end of the genomic clone i3a. It is encoded on the same strand as the *Amy* genes and exhibits marked homology with another putative *D. melanogaster* gene, CG7891. In both species it has an intron at the same position, which is 1009 bp in length in *D. ananassae* and 1198 bp in *D. melanogaster*. Figure 1 shows the gene organization in *D. ananassae*.

Gene Organization in D. varians

In the related species *D. varians*, we have cloned a fragment, 4938 bp in length, which is homologous to the *Amy35*/*Amy58* region. This fragment contains the 3′ part of the *Amy58* homolog and the 5′ part of the *Amy35* homolog, separated by a 4.3-kb intergenic region. The latter seems to be a pseudogene in *D. varians*, due to large deletions in the coding region. In *D. ananassae*, *Amy35* is intronless. Interestingly, there is a 60-bp insertion in the *D. varians Amy35* pseudogene, at the position where an intron is usually found in *Drosophila Amy* genes (Da Lage et al. 1996b). This insertion shows the canonical splicing sites of introns. This strongly suggests that there was an intron in *Amy35* in the ancestor of the *D. ananassae* subgroup, which was lost in *D. ananassae*. The intergenic region contains the exons 3 and 4, and the intron 3 of the CG14696 homolog, like in *D. ananassae*. Therefore, the nested gene organization is conserved between *D. ananassae* and *D. varians*. To check whether this organization is limited to the *D. ananassae* subgroup, we checked the absence of sequences related to CG14696 in the sequenced amylase regions of other *Drosophila* species, available in databases with sufficient flanking sequence, e.g., *D. kikkawai*, *D. pseudoobscura*, and *D. virilis*. We also used a pair of primers surrounding the second intron of CG14696 (CGDIRBIS and CGINTR2REV or alternatively CGAMY351 and CGAMY352) in *D. lutescens* (not shown) and *D. kikkawai* (AY239588), which are closer to *D. melanogaster* than *D. ananassae*, and in *D. subobscura* (AY239590) and *D. virilis* (AY239589), which are outgroups. In all these species, the intron was short. This strongly suggests that the ancestral CG14696 is devoid of an inserted *Amy* gene and that the nested organization is limited to the *D. ananassae* subgroup.

Chromosomal Location

The four genes *Amy58*, *Amy35*, CG14696, and CG7891 are contiguous on the 2L arm in *D. ananassae* (Fig. 1) but not in *D. melanogaster*. In the latter species, the two amylase genes, *Amy-p* and *Amy-d*, are on the 2R arm, at position 54A, whereas CG14696 is on 3R at position 86C7, and CG7891 is on 3R too, but distant, at position 84F1 (http://www.fruitfly.org/).

Analysis of CG14696

As far as we are aware, no experimental data on CG14696 have been published and no function has been proposed yet for the gene product by computer analysis (http://www.fruitfly.org). The gene is single-copy in *D. melanogaster*. We have checked by Southern blotting (not shown) that it is also true in *D. ananassae*. We have compared the protein sequences in *D. melanogaster* and *D. ananassae* (Fig. 2). The overall divergence is about 20% in amino acids and there are several deletions in *D. melanogaster*, one of which encompasses 13 amino acids. Despite this noticeable divergence, some long nucleotide stretches within the first intron (800 bp long) can be aligned, scattered along the intron. In addition, two stretches are conserved in the 5′ noncoding region: AGCTGC TCGCGTTCGG (−155/−140) and ACCAACCGAC ACCGTCGCCAAGGAG (−40/−15). In both species, the codon bias is low, as estimated by the effective number of codons (ENC) (Wright 1990): 56.1 and 56.6 in *D. melanogaster* and *D. ananassae*, respectively, which contrasts with the high bias of the nested *Amy35* gene in *D. ananassae*: ENC = 38.2. This low codon bias is generally considered to be correlated to a low level of expression of genes (Moriyama and Hartl 1993). However, our RT-PCR assays, which were performed on rough nucleic acid extracts, showed a clear amplification of CG14696 mRNA. Although this method is not quantitative, it suggested that the gene was actively transcribed in both species, at virtually all postembryonic stages (Fig. 3A and B, and not shown). Interestingly, in *D. ananassae*, an assay was done using primers that surround the large intron 2, which contains the

Fig. 2. Alignment of protein sequences of CG14696 in *D. melanogaster* (mel) and *D. ananassae* (ana).

Amy35 region (Fig. 3B). The sequence of the RT-PCR product confirmed that this intron had been correctly spliced. Detection of CG14696 transcripts in *D. melanogaster* tissues also suggested that various types of tissues produce the mRNA, with a strong signal for the midgut (Fig. 3C). In *D. ananassae*, we observed similar clear expression in the midgut, but also in midgut-ectomized larvae (not shown).

To check whether both genes were expressed in the same parts of the gut, we performed *in situ* hybridizations on *D. ananassae* larval midguts (Fig. 4). Figure 4A shows that in the absence of any probe, light pink-violet staining was evenly distributed and the hindgut showed nonspecific staining. The *Amy35* DNA probe may strongly cross-hybridize with *Amy58* transcripts. However, we have previously evidenced at the protein level through enzyme electrophoresis, that both genes have identical tissue-specificities (Da Lage et al. 1996a). *In situ* hybridization with *Amy35* probe (Fig. 4B) clearly showed that *Amy* transcripts are distributed throughout several regions of the midgut: the caeca at the very anterior region were heavily stained dark blue, like two other regions, particularly the central part, whereas the hindgut showed no specific blue staining. This is in agreement with the earlier results (Da Lage et al. 1996a). The CG14696 probe (Fig. 4C) also stained the digestive tract, in a manner similar to the *Amy* probe in the anterior midgut whereas the central part of the midgut was stained more lightly, and the posterior midgut was almost unstained. No staining was observed in the Malpighian tubules. These results suggest that both genes may be transcribed in the same tissues, although this could not be confirmed at cell level.

Discussion

Evolutionary History of the Amy *Family*

To be able to reconstitute the history of the *Amy* multigene family in *Drosophila ananassae*, we need interspecific comparisons with more or less closely related *Drosophila* species. The ancestry or novelty of the various copies cannot be ascertained from nucleic sequences alone, nor by cytogenetic localizations. It must be also reminded that duplications of *Amy* genes probably occurred independently several times in various *Drosophila* lineages. In a previous paper (Da Lage et al. 2000), we had two alternative hypotheses for the ancestral gene(s) in *D. ananassae*. The alternative hypotheses were: (1) the *Amy4N*/*Amyi5* cluster is ancestral, (2) the *Amy35*/*Amy58* cluster is ancestral. The data available until now supported weekly the first hypothesis. The present work sheds some light on the chronology of some duplications.

The nested arrangement of an *Amy* gene is not ancestral in *Drosophila*. As mentioned in the Introduction, in *D. ananassae*, there are two main amylase gene clusters: *Amy35*/*Amy58* and *Amy4N*/*Amyi5*, which are located on different chromosomes. *Amy58* and *Amy35* are organized as a tandem. We have shown in this study that *Amy35* lies within the second intron of CG14696 in *D. ananassae* and *D. varians*, thus probably in all the *D. ananassae* subgroup. This is not the case in *D. melanogaster*, in which the two genes are not even on the same chromosome. Clearly, at some time one *Amy* copy was inserted into this intron in an ancestor of the *D. ananassae* subgroup. One could imagine that the ancestral copy was translocated, but more likely, this suggests that *Amy35* is not the ancestral copy. Another *Amy* cluster of *D. ananassae*, *Amy4N*/*Amyi5*, may assume the status of ancestral locus. This ancestral status is in agreement with an argument already advanced (Da Lage et al. 2000). This argument was the synteny of *Amy4N*/*Amyi5* with the paralog *Amyrel* in several species. The cluster *Amy4N*/*Amyi5* and *Amyrel* are both on the 3L arm in *D. ananassae*. Similarly, the two *Amy* genes and *Amyrel* are on the 2R arm in *D. melanogaster* which is considered homologous to the 3L of *D. ananassae* (Kikkawa 1938; Sturtevant and Novitski 1941). In the *D. obscura* group too, the *Amy* cluster and *Amyrel* are syntenic (Da Lage et al. 1998, 2000). This suggested conservation of an ancestral state. However, the insertion of *Amy35* inside an intron is a much more powerful argument regarding the identification of the ancestral genes. We conclude that the *Amy35*/*Amy58* cluster is more recent than *Amy4N*/*Amyi5*. Since in the *D. montium* subgroup, there are also two *Amy* gene clusters, it has been proposed that the four copies two clusters

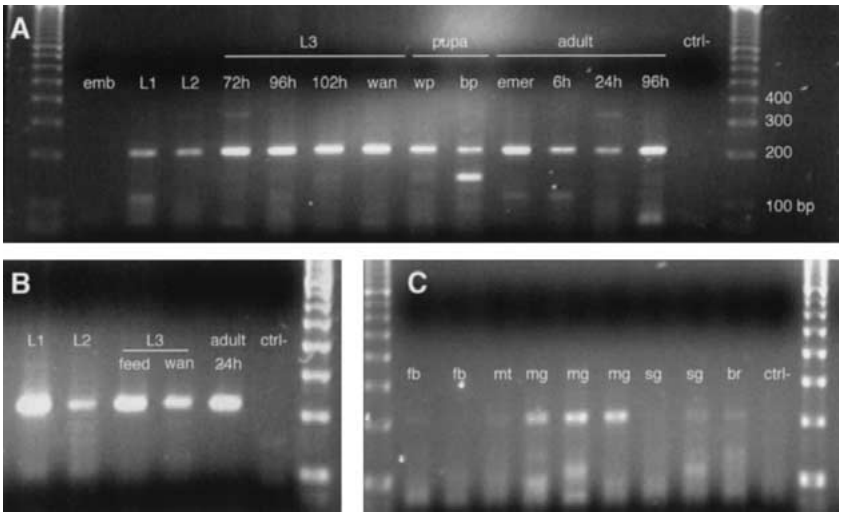

Fig. 3. Agarose gels of RT-PCR for the detection of CG14696 mRNA. **A** Temporal pattern in *D. melanogaster* Canton-S individuals at sequential developmental stages (primers CGMEL1/CGMEL2). **B** Temporal pattern in *D. ananassae*: Taï 13-1610 individuals (primers CGAMY351/CGAMY352). **C** Tissue-specificity in *D. melanogaster* individual tissues (primers CGMEL1/CGMEL2). Abbreviations: emb, embryo; L1, larval instar 1; L2, larval instar 2; L3, larval instar 3; 72 h, moult L2/L3; wan, wandering; wp, white pupa; bp, black pupa; emer, emerging adult; ctrl–, negative control; fb, fat body; mt, malpighian tubules; mg, midgut; sg, salivary gland; br, brain.

organization was ancestral to the *D. melanogaster* group, followed by the loss of one cluster in *D. melanogaster* and its subgroup (Z. Zhang, personal communication), but this hypothesis seems inconsistent with the nested arrangement found in the *D. ananassae* subgroup only. We may investigate whether *Amy35* was inserted at first into the previously short intron 2 of CG14696, or was a secondary duplication from a neighboring insertion of *Amy58*, close but outside CG14696. The analysis of the sequence shows that there are a few nucleotides similar to the coding sequence of CG14696 in the *Amy58* upstream region. This could suggest that *Amy58* arose as a duplication of *Amy35*, along with a small part of CG14696. But it could also have resulted from concerted evolution, which tends to erase traces of earlier evolutionary history. Indeed, we have shown earlier that concerted evolution occurred between *Amy35* and *Amy58* 5′ regions (Da Lage et al. 2000). Figure 5 summarizes the proposed history of the amylase family.

In a gene tree made with all three codon positions, we had previously observed that *Amy4N* (or *Amyi5*) seems less closely related to the sequence of *D. melanogaster* than *Amy35* (Da Lage et al. 2000). It is possible that the *Amy35*/*Amy58* cluster has evolved more slowly than its progenitor *Amy4N*/*Amyi5*. A similar evolutionary pattern also occurred in *D. kikkawai* (Inomata and Yamazaki 2000): the *Amy3*/*Amy4* cluster evolved faster than the *Amy1*/*Amy2* cluster. However, in this species, the two gene clusters are on the same arm, but the ancestral locus is not established. However, and more probably, it is suggested (Z. Zhang, personal communication) that the base composition biases the comparison, given the high GC3 content in *Amy35* and in *Amy* of *D. melanogaster*. A neighbor-joining tree made with the first two codon positions of all the *Amy* copies of *D. ananassae*, the two gene types of *D. kikkawai* and *D. melanogaster*, shows the two clusters of *D. ananassae* branched together, whereas *Amy1* of *D. kikkawai* remains closest to *D. melanogaster*, followed by *Amy3* of *D. kikkawai*. This topology is more evocative of independent duplication and evolution in the *D. ananassae* subgroup.

Concerning the gene organization on the chromosomes, it has been proposed from linkage group studies (Kikkawa 1938; Sturtevant and Novitzki 1941) that the 3L arm of *D. ananassae* is homologous to 2R of *D. melanogaster*. Similarly, the 2L arm of *D. ananassae* is homologous to the 3R arm of *D. melanogaster*. Our data support further this hypothesis, since the gene region studied here, located on the 2L arm, harbors CG14696 and CG7891 homologs, which are both located on the 3R arm in *D. melanogaster* (http://www.fruitfly.org/). However, as we have already mentioned, these two genes are not contiguous in *D. melanogaster*, and there is no *Amy* gene on the 3R arm. This example shows the difficulty to establish a close correspondence of chromosome arms at a detailed scale between these two species, which are both known to harbor numerous inversions (Lemeunier et al. 1986; Tobari et al. 1993). Despite this, the general correspondence between arms is still valid.

Nested Insertion and Gene Expression

Nested genes, or genes within genes, are not new objects in genome science. First findings were reported in the *Gart* locus of *Drosophila*, in which the pupal cuticle protein (*pcp*) gene lies within an intron of the purine pathway gene *ade3* (Henikoff et al. 1986; O'Hare 1986). Since then, a number of examples have been reported in *Drosophila*, most of which are encoded on opposite strands (Ashburner et al. 1999): the *dunce* gene, for instance, contains six genes, which are dispatched into two introns, and in both

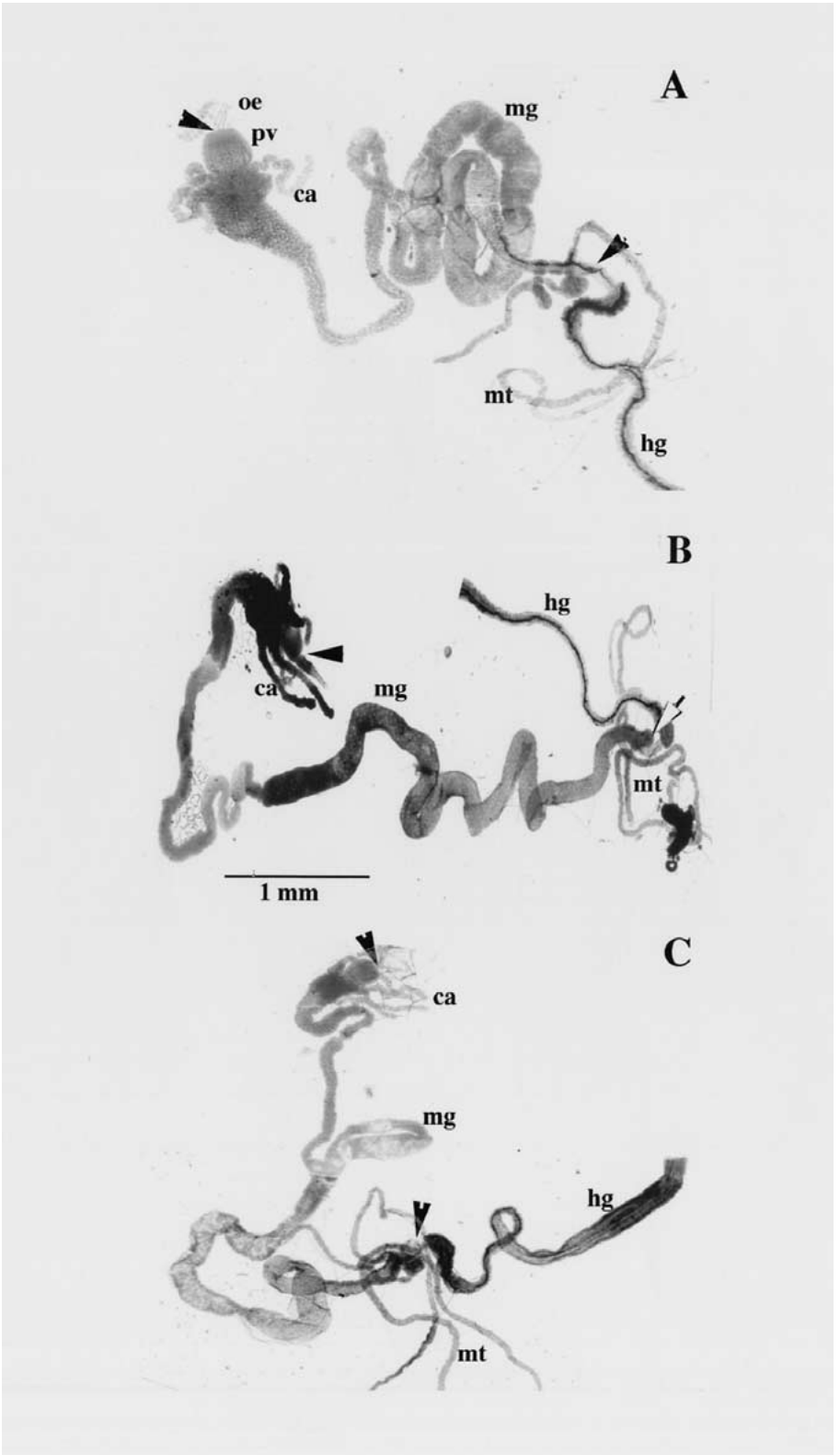

Fig. 4. *In situ* hybridizations of specific probes for either CG14696 or *Amy35* on *D. ananassae* digestive tracts. **A** Negative control (staining without DNA probe); **B** probe for *Amy35*; **C** probe for CG14696. oe, oesophagus; pv, proventricule; ca, caeca; mg, midgut; hg, hindgut; mt, Malpighian tubules; arrowheads, midgut boundaries.

senses (Furia et al. 1993); *sina* and *Rh4*, which are functionally related, this nested arrangement being not conserved in *D. virilis* (Neufeld et al. 1991). In their study of the 2.9-Mbp *Adh* region, Ashburner et al. (1999) listed 17 cases of nested genes, among which *Adh* itself, along with its relative *Adhr*, was embedded in a large intron of the *outspread* (*osp*) gene (Ashburner et al. 1999; McNabb et al. 1996). However, as stressed by Wang et al. (2000), these cases were computer-predicted and should be experimen-

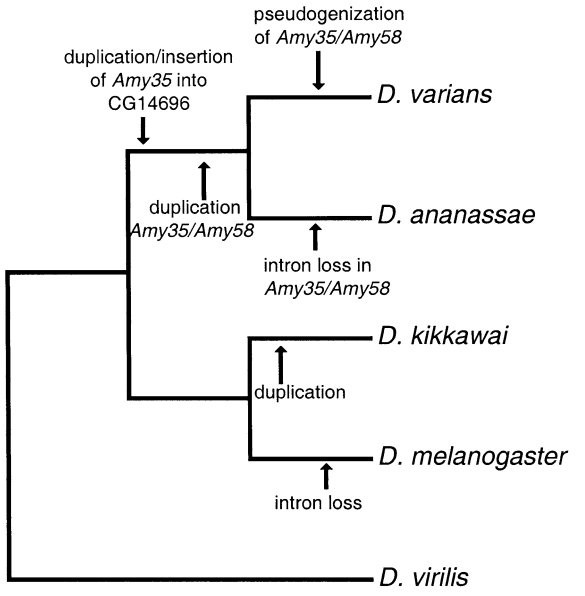

Fig. 5. Schematic consensus species tree with a tentative reconstitution of the history of molecular events in the amylase family of *D. ananassae* and its subgroup. A variant is possible, in which *Amy58* was inserted first, in the vicinity of CG14696, and then a duplication gave rise to the nested *Amy35*. A duplication event is proposed within the *D. montium* subgroup (Inomata and Yamazaki 2000).

tally evidenced. Since then, most of the nest and nested genes listed in Ashburner et al. (1999) have been detected as cDNA or EST, and therefore may be active ORFs (http://www.fruitfly.org/). If we extrapolate to the 120 Mbp of the euchromatic *D. melanogaster* genome, there may be 700–1000 cases of nested genes in *Drosophila*. The most impressive example of nested gene in *Drosophila* is *msi*, a 7.6-kb-long developmental gene, which lies in the third intron of *ymp* in *D. melanogaster*, *D. yakuba*, and *D. teissieri* (Wang et al. 2000). But the antiquity of this organization is not known at present. Indeed, comparative data between *Drosophila* species are still rare for nested genes. In a few cases, it has been shown that the arrangement is ancient: the *Gart* locus is conserved in *D. pseudoobscura*, a member of the *D. obscura* group (Henikoff and Eghtedarzadeh 1987); the nested organization of *l(2)tid*/*l(2)not* is conserved in *D. virilis*, although the intriguing antisense *l(2)rot*, transcribed from the strand opposite of *l(2)tid* but not translated, is absent in this species (Kaymer et al. 1997; Kurkiz-Dumke et al. 1997). In the published cases, however, the history of the nesting has not been reconstructed, nor dated. For example, in the *Adh* region, no comparative study is available yet, which would enable one to assess when the various nested arrangements occurred. In contrast, the insertion of an amylase gene within CG14696 may be dated as relatively recent, since it is restricted to the *D. ananassae* subgroup. It may have occurred less than 20 million years ago, after the split from *D. melanogaster–D. montium* ancestors.

Nested arrangements are puzzling, in that transcripts may interfere with each other. If the mRNA from the nested gene is encoded on the opposite strand, it could hybridize with the pre-mRNA or with spliced intron particles from the nest gene. It has been suggested and indeed observed that there should be no overlap in the peak expression patterns of the two genes. It is the case for the *ymp*/*msi* pair: *ymp* is expressed in the testis, *msi* in the sensilla (Wang et al. 2000). Another example is *Adh* and *osp* in embryo (McNabb et al. 1996), but some cases have been found with apparently simultaneous expression (e.g., the same two genes in the adult female ovary). In our study, we observed that CG14696 was expressed at every post-embryonic developmental stage and in several organs, especially the midgut. *Amy35* (and its twin neighbor *Amy58*) is actively transcribed in third instar larval midgut (Da Lage et al. 1996a). *In situ* hybridizations confirm that both *Amy35* and CG14696 are active at the same time in the same parts of the gut. Theoretically, in this case, interference could occur between their mRNAs before processing, if the primary transcripts came into contact, but this is difficult to show. Further studies should be done to find out whether the two genes are active in different cell types within the midgut.

It has been suggested that functional coupling between the nest and nested genes could explain the conservation of a nested organization between remote species (Henikoff and Eghtedarzadeh 1987; Kaymer et al. 1997). However, it is reasonable to assume that in *D. ananassae*, the arrangement we studied is not constrained by functional requirements, but is a contingent event, since the nested organization is recent. On the other hand, this arrangement has not been eliminated by natural selection, and therefore, despite the potential interference phenomenon, seems not to be detrimental to metabolism.

Acknowledgments. We are grateful to two anonymous referees for their helpful comments. This work was funded by the CNRS PICS nr. 607.

References

Adams MD, Celniker SE, Holt RA, et al. (2000) The genome sequence of *Drosophila melanogaster*. Science 287:2185–2195

Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. Nucleic Acids Res 25:3389–3402

Ashburner M, Misra S, Roote J, Lewis SE, Blazej R, Davis T, et al. (1999) An exploration of the sequence of a 2.9-Mb region of the genome of *Drosophila melanogaster*: The *Adh* region. Genetics 153:179–219

Da Lage J-L, Klarenberg A, Cariou M-L (1996a) Variation in sex-, stage- and tissue-specific expression of the amylase genes in *Drosophila ananassae*. Heredity 76:9–18

Da Lage J-L, Maczkowiak F, Cariou M-L (2000) Molecular characterization and evolution of the amylase multigene family of *Drosophila ananassae*. J Mol Evol 51:391–403

Da Lage J-L, Renard E, Chartois F, Lemeunier F, Cariou ML (1998) *Amyrel*, a paralogous gene of the amylase gene family in *Drosophila melanogaster* and the *Sophophora* subgenus. Proc Nat Acad Sci USA 95:6848–6853

Da Lage J-L, Wegnez M, Cariou M-L (1996b) Distribution and evolution of introns in Drosophila amylase genes. J Mol Evol 43:334–347

Furia M, D'Avino PP, Crispi S, Artiaco D, Polito LC (1993) Dense cluster of genes is located at the ecdysone-regulated 3C puff of *Drosophila melanogaster*. J Mol Biol 231:531–538

Gloor G, Engels W (1992) Single fly DNA preps for PCR. DIS 71:148–149

Henikoff S, Eghtedarzadeh MK (1987) Conserved arrangement of nested genes at the Drosophila *Gart* locus. Genetics 117:711–725

Henikoff S, Keene MA, Fechtel K, Fristrom JW (1986) Gene within a gene: Nested Drosophila genes encode unrelated proteins on opposite DNA strands. Cell 44:33–42

Huet F, Ruiz C, Richards G (1993) Puffs and PCR: the in-vivo dynamics of early gene expression during ecdysone responses in *Drosophila*. Development 118:613–627

Inomata N, Yamazaki T (2000) Evolution of nucleotide substitutions and gene regulation in the amylase multigenes in *Drosophila kikkawai* and its sibling species. Mol Biol Evol 17:601–615

Kaymer M, Debes A, Kress H, Kurkiz-Dumke U (1997) Sequence, molecular organization and products of the *Drosophila virilis* homologs of the *D. melanogaster* nested gene *lethal*(*2*) *tumorous imaginal discs* [*l*(*2*)*tid*] and *lethal*(*2*) *neighbour of tid* (*l*(*2*)*not*]. Gene 204:91–103

Kikkawa H (1938) Studies on the genetics and cytology of *Drosophila ananassae*. Genetica 20:458–516

Kurkiz-Dumke U, Kaymer M, Gundacker D, Debes A, Labitzke K (1997) Gene within gene configuration and expression of the *Drosophila melanogaster* genes *lethal*(*2*)*neighbour of tid* [*l*(*2*)*not*] and *lethal* (*2*) *relative of tid* [*l*(*2*)*rot*]. Gene 200:45–58

McNabb S, Greig S, Davis T (1996) The alcohol dehydrogenase gene is nested in the *outspread* locus of *Drosophila melanogaster*. Genetics 143:897–911

Moriyama EN, Hartl DL (1993) Codon usage bias and base composition of nuclear genes in Drosophila. Genetics 134:847–858

Neufeld TP, Carthew RW, Rubin GM (1991) Evolution of gene position: Chromosomal arrangement and sequence comparison of the *Drosophila melanogaster* and *Drosophila virilis sina* and *Rh4* genes. Proc Nat Acad Sci USA 88:10203–10207

O'Hare K (1986) Genes within genes. TIG 2:33

Sturtevant AH, Novitski E (1941) The homologies of the chromosome elements in the genus *Drosophila*. Genetics 26:517–541

Tautz D, Pfeifle C (1989) A nonradioactive in situ hybridization method for the localization of specific RNAs in Drosophila embryos reveals translational control of the segmentation gene *hunchback*. Chromosoma 98:81–85

Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. Nucleic Acids Res 22:4673–4680

Tobari YN (1993) *Drosophila ananassae*, genetical and biological aspects. Japan Scientific Societies Press, Tokyo

Tobari YN, Goñi B, Tomimura Y, Matsuda M (1993) Chromosomes. In: Tobari YN (ed) *Drosophila ananassae*, genetical and biological aspects. Japan Scientific Societies Press, Tokyo, pp 23–48

Wang W, Zhang J, Alvarez C, Llopart A, Long M (2000) The origin of the *Jingwei* gene and the complex modular structure of its parental gene, *Yellow Emperor*, in *Drosophila melanogaster*. Mol Biol Evol 17:1294–1301

Wright F (1990) The "effective number of codons" used in a gene. Gene 87:23–29