J Mol Evol (2000) 51:391–403
DOI: 10.1007/s002390010102

JOURNAL OF MOLECULAR EVOLUTION

Molecular Characterization and Evolution of the Amylase Multigene Family of *Drosophila ananassae*

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

UPR 9034 Populations, génétique et évolution, C.N.R.S. 91198 Gif sur Yvette Cedex, France

Received: 30 May 2000 / Accepted: 17 July 2000

Abstract. *Drosophila ananassae* is known to produce numerous alpha-amylase variants. We have cloned seven different *Amy* genes in an African strain homozygous for the AMY1,2,3,4 electrophoretic pattern. These genes are organized as two main clusters: the first one contains three intronless copies on the 2L chromosome arm, two of which are tandemly arranged. The other cluster, on the 3L arm, contains two intron-bearing copies. The amylase variants AMY1 and AMY2 have been assigned to the intronless cluster, and AMY3 and AMY4 to the second one. The divergence of coding sequences between clusters is moderate (6.1% in amino acids), but the flanking regions are very different, which could explain their differential regulation. Within each cluster, coding and noncoding regions are conserved. Two very divergent genes were also cloned, both on chromosome 3L, but very distant from each other and from the other genes. One is the *Amyrel* homologous (41% divergent), the second one, *Amyc1* (21.6% divergent) is unknown outside the *D. ananassae* subgroup. These two genes have unknown functions.

Key words: Amylase — Gene duplication — Multigene family — Drosophila ananassae — Intron

Introduction

The alpha-amylase gene-enzyme system has been studied in many bacteria, plants, and animals. This enzyme plays a major role in the digestive processes involving carbohydrates by hydrolyzing starch from food substrates in smaller sugars, such as maltose and glucose.

Despite a large amount of sequence variation, the tertiary structure of the enzyme and several amino acid motifs are conserved among all living organisms studied so far (Janecek 1997). The exon-intron structure of the gene is variable concerning the number of introns and their insertion points. Duplications of the coding gene *Amy* have been reported for many plants and animals. In animals, duplications were found in primates and rodents (Groot et al. 1989; Meisler and Ting 1993; Nielsen 1977) and in arthropods, mainly insects (Baker et al. 1990; Hickey et al. 1987; Laulier 1988; Van Wormhoudt and Sellos 1996). In *Drosophila*, molecular techniques have permitted a large screening of the genus after the cloning of the *D. melanogaster* copies (Boer and Hickey 1986; Gemmill et al. 1985). A number of species of the subgenus *Sophophora* have been evidenced for multigenic structure of the *Amy* locus (Tadlaoui-Ouafi 1993; Da Lage et al. 1992; Inomata et al. 1997). The structural organization of the genes has been detailed in *D. melanogaster* and its subgroup (two close copies divergently transcribed) (Payant et al. 1988; Shibata and Yamazaki 1995) and in *D. pseudoobscura* (one to three copies, with pseudogenes [Brown et al. 1990; Popadic et al. 1996]). In both cases, the copies remained grouped in a single cluster.

Recently, a very divergent amylase-related gene, *Amyrel*, has been described in several species from the *Sophophora* subgenus (Da Lage et al. 1998). This gene, which has a specific intron position, has been found physically distant from the classical genes although the

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

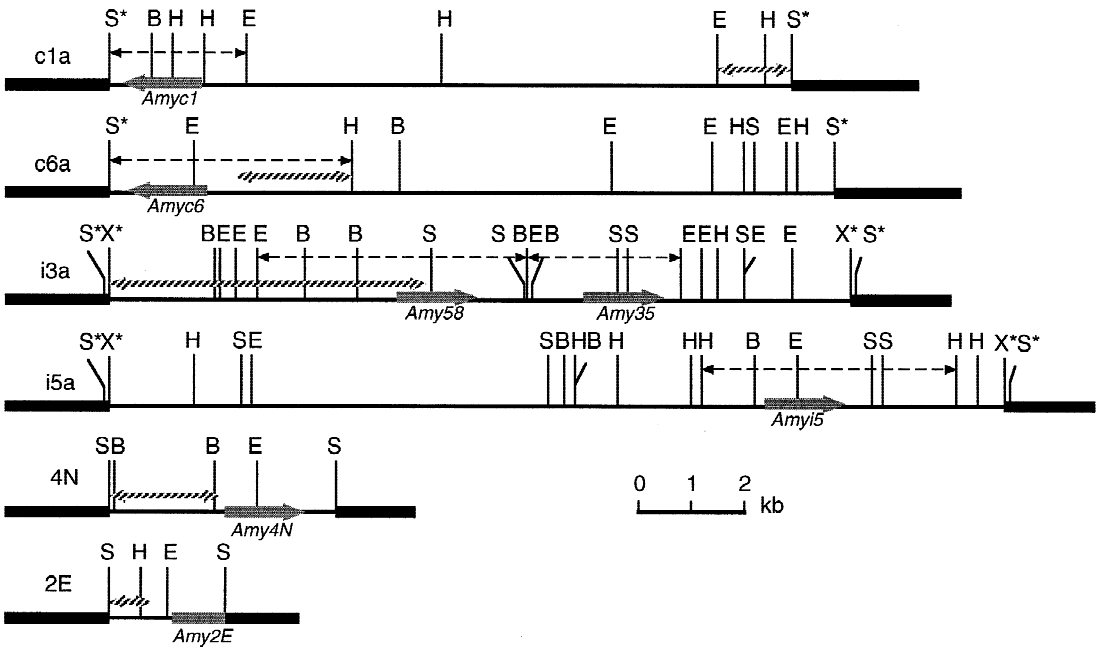

Fig. 1. Genomic clones containing *Amy* genes in *Drosophila ananassae* (strain Taï 13-1610). Thick lines are the vectors: λ-EMBL3 for c1a and c6a, λ-gem11 for i3a and i5a. The clones are oriented from left to right arms of the phage vectors. 4N and 2E are in pUC18 plasmids. Grey arrows are the *Amy* coding sequences. Dashed arrows indicate subclones used for sequencing and somatic transformations. Hatched arrows indicate the regions used as probes for in situ hybridizations. Restriction sites are: S, *Sal* I; X, *Xho* I; B, *Bam* HI; H, *Hin* dIII; E, *Eco* RI. Asterisks indicate sites that belong only to the vectors.

localization of *Amyrel* relative to *Amy* is variable between species. Further experiments suggest that *Amyrel* is present throughout the *Drosophila* genus (unpublished), but its function is presently unknown.

However, except for *Amyrel*, which is easily identified and considered to be orthologous in all *Drosophila* species, the data suggest that most often the duplication events were independent between *Drosophila* lineages. The intriguing fact that so many organisms have undergone independent *Amy* gene duplications remains an exciting problem. Functional evolutions have been suggested or evidenced in a few cases of *Amy* duplications (Da Lage et al. 1996a; Meisler et al. 1986). However, within a given taxonomic group like the genus *Drosophila*, the physiological benefits for multicopy-bearing species versus single copy–bearing species are not clear.

In the present study, we describe seven members of the amylase family in *Drosophila ananassae*, the most complex *Amy* family known to date in *Drosophila.* This widespread tropical and domestic species (Tobari 1993) commonly exhibits complex electrophoretic patterns (Da Lage et al. 1989). From an African population, we selected a strain that expressed four different proteins, encoded by four different genes (Cariou and Da Lage 1993; Da Lage et al. 1992). We have shown that these different genes are subject to tissue-specific regulation. In other *D. ananassae* strains or populations, a clear stage-specific regulation is commonly observed (Da Lage et al. 1996a). This African strain was used in the present work. In *D. ananassae*, the complexity of the *Amy* family illustrates several ways by which a multigene family may evolve.

Materials and Methods

We used the strain Taï 13-1610 of *D. ananassae* from Ivory Coast. It had been made homozygous for the *Amy* genotype *Amy1,2,3,4/ Amy1,2,3,4* (Da Lage et al. 1992). High molecular weight DNA was prepared to make a genomic library. The DNA was partially digested by *Sau* 3A and fragments of 12–18 kb were recovered and ligated into λ-EMBL3 and λ-gem11-*Bam* H1-cut phage vectors (Promega). The yield was low, and all the preparations were plated. The library was screened with the pORM7 probe from D.A. Hickey (Ottawa University, Canada), which contains a cDNA of *D. melanogaster Amy.* Positive clones were cultured and phage DNA was extracted according to Qiagen protocols. Restriction maps were done prior to subcloning *Amy* genes into pUC plasmids (Fig. 1). Mini-libraries were also performed in *Sal* I-cut pUC18 plasmid to clone additional copies (see Results). Nested deletions (kit from Pharmacia) were performed prior to sequencing. All clones were manually sequenced on both strands (Sanger et al. 1977). The sequence data were deposited in Genbank with the accession numbers U534770–U53480 and U53698–U53699. Subclones containing adjacent regions (indicated on Fig. 1) were used as probes for in situ localization on the polytene chromosomes. The protocol for in situ hybridization was as in Da Lage et al. (1992).

Genic amplifications were performed to study the upstream regions of some *Amy* genes in various *D. ananassae* strains. Specific primers were designed for each of *Amy35* and *Amy58* 5′ regions: 35AM = 5′ACGTCTTGACACTGAACATC3′; 58AM = 5′TATCCGATTCAGATATTCC3′; the reverse primer is at the beginning of the coding sequence and is common to both forward primers: DEBREV = 5′CCCAGGAAGTTCTCGCACTC3′. To detect some polymorphic sites, three reverse primers were designed for use with 35AM: TYP1 = 5′TTCTTTGCACTTGGTGG3′; TYP2 = 5′GGGGAGTTTCTTTTCGG3′; TYP3 = 5′CTTTGCACTTGATAACG3′. PCR cycles were performed with standard conditions: denaturation at 94°C for 25 s, annealing at 60°C for 1 min, elongation at 72°C for 1 min, 35 cycles. When direct sequencing was not feasible (indel polymorphism), the PCR products were cloned in the pGEM-T cloning vector (Promega)

and sequenced on a AB1373 sequencer. These data were deposited in Genbank with accession numbers AF238900–AF238977.

Transient expression of amylase genes was performed by injecting the whole subclones *Amy35, Amy58, Amy4N, Amyi5, Amyc1,* and *Amyc6* (see Fig. 1) in embryos of a *Amy*-null *D. melanogaster* strain. The plasmid DNAs were dissolved in KCl 5 m*M*, phosphate buffer pH 6.8, 0.1 m*M* at 500 ng/μl. Surviving third-instar larvae were fed on nonsugared axenic medium and electrophorized for amylase activity according to Da Lage et al. (1989).

Molecular data were analyzed with SEQAPP for Macintosh by Don Gilbert (*ftp.bio.indiana.edu/molbio)*, CLUSTAL W (Thompson et al. 1994) and MEGA (Kumar et al. 1993). Substitution rates were computed with KESTIM (Comeron 1995).

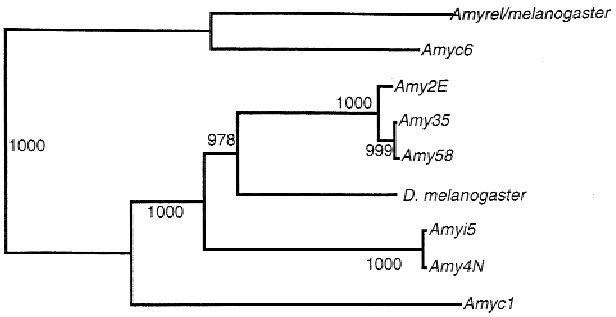

Fig. 2. Neighbor-joining tree (made with CLUSTAL W) of the *Amy* coding sequences from *D. ananassae* and *D. melanogaster* (accession number X04569 for *Amy,* AF022713 for *Amyrel*). To include *Amy2E* in the tree, only the first 909 bp were used. Bootstrap values are shown (1000 replicates).

Results

Number and Localization of Amy *Copies*

Although a total of 15,000 pfu only were screened in the phage genomic library, four positive clones were isolated at moderate stringency: *c1a, c6a, i3a,* and *i5a.* Clones *c1a* and *c6a* were in λ-EMBL3 and clones *i3a* and *i5a* were in λ-gem11. Figure 1 shows the restriction maps of the four clones and the positions of the coding sequences. A tandem of *Amy* genes was evidenced in *i3a* only. Each of the other three clones had a single *Amy* gene. The gene regions were subcloned into pUC plasmids and sequenced. Comparing the restriction maps of the genomic clones to Southern hybridizations of whole genomic DNA enabled us to identify some of the genes on Southern blots, especially in *Sal* I digests (see Fig. 4b in Da Lage et al. 1992). However, two bands (2.3 and 4.4 kb) were fragments that had not been cloned yet. Therefore we performed a minilibrary of *Sal* I fragments of the desired sizes and we obtained the two copies, named 4N (4.4–kb band) and 2E (2.3-kb band). As a whole, seven different *Amy* copies were cloned from the strain Taï 13-1610, six of which were complete. *Amy2E* (from clone 2E) was interrupted by a *Sal* I site at position 909.

Adjacent sequences were used as probes for chromosomal localization. *Amy35, Amy58,* and *Amy2E* map at 37C on chromosome 2L, *Amy4N* and *Amyi5* map at 81C on chromosome 3L. These two loci had been previously identified with the *D. melanogaster* coding sequence as a probe (Da Lage et al. 1992). *Amyc1* maps at 74A on chromosome 3L, inside the 3LA cosmopolitan terminal inversion (Tobari et al. 1993) and *Amyc6* maps at 76C, on the same arm but outside the 3LA inversion (not shown). The probe used for *Amy35* and *Amy58* also hybridized to more than 15 euchromatic sites on arms XL, 2L, 2R, and 3R, which suggested the proximity of a repeated element. Therefore all the available 5′ region of *Amy58* was sequenced. This repeated element is under study in our laboratory. A number of euchromatic and centromeric labellings were also observed with the probe for *Amy4N* (see below).

The neighbor-joining tree constructed from the seven aligned coding sequences (Fig. 2) shows two main groups of genes, which correspond to the cytogenetical groups: a first group clusters *Amy35, Amy58,* and *Amy2E;* a second group includes *Amy4N* and *Amyi5.* The divergence within clusters is very low, but is much higher between groups. *Amyc1 and Amyc6* cannot be assigned to any gene group, as they are very divergent from each group and from each other. Their chromosomal localizations are isolated from other *Amy* genes. The position of *D. melanogaster* will be studied separately.

Intragroup Comparisons

The Intronless Genes: Amy35, Amy58, Amy2E. *Amy35* and *Amy58* are tandemly arranged 2 kb apart and are transcribed in the same direction. *Amy2E* maps at the same cytological locus, but its position and distance from the former two genes is still unknown. The three genes are intronless, unlike the majority of *Drosophila Amy* genes (Da Lage et al. 1996b). The coding sequences are very similar. Ks and Ka substitution rate-per-site values are given in Table 1. Because the sequence data available for *Amy2E* are restricted to the region upstream of the *Sal* I site at position 904–909, this copy was not included in Table 1. However, in the available region the sequence is highly similar to *Amy35* and *Amy58:* 12 nucleotide substitutions over 909 bp, all specific to *Amy2E,* 6 of which are nonsynonymous: A128G, K235R, V239A, N278D (changing electric charge), S289G, F301L.

There are only nine nucleotide substitutions between *Amy35* and *Amy58* (0.6%), three of which are nonsynonymous: V63A, R387K, D394G (changing charge). The two genes encode putative proteins of 494 amino acids (like in *D. melanogaster*), and their electric charge differ by one unit, the product of *Amy35* should migrate faster. This has been confirmed by somatic transformation experiments in *Amy*-null *D. melanogaster*, which have shown that *Amy35* encodes AMY1 and *Amy58* encodes AMY2 (Fig. 3).

In this gene cluster the codon preference is mainly

Table 1. Ks and Ka substitution rate-per-site values (Ks are above the diagonal; Ka are below)

Gene	*Amy35*	*Amy58*	*Amy4N*	Amy-i5	*Amy-C1*	*Amy-c6*
Amy35	—	0.016	1.067	1.065	1.290	1.078
Amy58	0.003	—	1.096	1.094	1.286	1.068
Amy4N	0.041	0.044	—	0.011	1.708*	1.367
Amy-i5	0.042	0.045	0.001	—	1.710*	1.358
Amy-c1	0.144	0.147	0.152	0.150	—	1.530
Amy-c6	0.366	0.368	0.363	0.365	0.391	—

The correcting method was the Kimura two-parameter model, except asterisks, where it was not applicable and where the Jukes and Cantor one-parameter model was used instead. The *Amy2E* gene was not included.

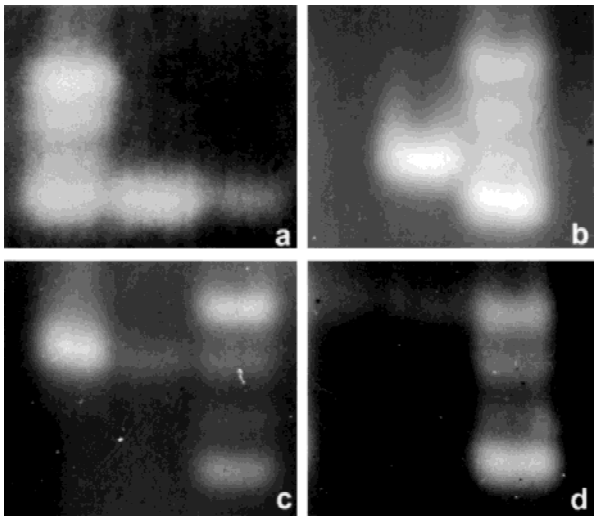

Fig. 3. Transient expression of amylase genes from *D. ananassae* injected in *D. melanogaster Amy*-null embryos. Injected third-instar larvae (pooled by two to four individuals) were electrophorized on a polyacrylamide gel with a AMY1,2,3,4 ladder, **a:** *Amy35:* AMY1; **b:** *Amy58:* AMY2; **c:** *Amyi5:* AMY3; 6: **d:** *Amy4N:* AMY4.

similar to that of *D. melanogaster,* except for aspartic acid: the ratio GAC/GAU is 32/2 in *D. melanogaster,* whereas it is 20/11 for *Amy35.* Actually, the situation in *D. melanogaster* may be unusual, since aspartic acid generally shows no codon preference in *Drosophila* (Moriyama and Powell 1997). The base usage (Table 2) is similar to *D. malanogaster* for the first and second positions in codons, due to constraints for the conservation of the protein sequence, but is significantly different at the third position. The G + C content at the third codon position (GC_3) is 74.9% for *Amy35* and 74.5% for *Amy58,* which is lower than in *D. melanogaster* (88.3%).

The noncoding regions of *Amy35* and *Amy58* are very similar (90%) over almost 500 bp 5′ to the translation start ATG, but the similarity vanishes then abruptly (Fig. 4A). *Amy2E* shares the last 350 bp of its 5′ region with the former two copies. Putative regulatory sites have been found. The putative TATA box lies at −58 and the putative CAAT box at −85. The putative transcription start is at −27, by analogy with the sequences of *D. melanogaster, D. erecta, D. pseudoobscura,* and *D. virilis* (Magoulas et al. 1993). In −166, there is a motif similar to the conserved sequence TTGTGATAAGC, involved in midgut amylase expression (Magoulas et al. 1993). Farther upstream of *Amy2E,* in position −720 (not shown), a tandem repeat was found, as a perfectly duplicated 138-bp stretch. The two duplicates are separated by 2 bp only. No homology was found in database searches.

On the other hand, the 3′ regions of *Amy35* and *Amy58* are conserved (81%) only within the first 42 bp after the stop codon and contain a putative polyadenylation site (Fig. 4B). These data suggest that the boundaries of the duplication event between these two copies are about −450 and +1530 or that extensive gene conversion occurred between these positions. In the next paragraph, the 5′ regions of these two copies are analysed in several *D. ananassae* strains to detect such events.

Population Analysis of the Upstream Regions of Amy35 *and* Amy58. The last 500 bp of the 5′ regions of *Amy35* and *Amy58* are almost perfectly duplicated (Fig. 4A), except a discrepancy between −283 and −258: the *Amy35* sequence is different from *Amy58* and *Amy2E*. The presence or absence of this discrepancy was investigated in several strains of *D. ananassae* from various origins. We amplified and sequenced about 700 bp of these regions with primers specific for either *Amy35* or *Amy58* 5′ regions (see Materials and Methods). The results (Table 3) show that the discrepancy observed in our reference strain Taï 13-1610 upstream of *Amy35* is widespread but is absent from several samples. This polymorphic region shows three alleles, which are not clearly correlated to the geographical origin of the flies. One allele is identical to the sequence in *Amy58* (normal type), the second one is the "Taï" type, the third one is specific from Bangalore (India) (Fig. 5A). In contrast, the 5′ region of *Amy58* is monomorphic at this position, all populations have the normal type only. We never found the Taï type simultaneously upstream to both *Amy35* and *Amy58,* which could have been a clue for a recent concerted evolution event. Another polymorphic pattern has been found closer to the translation start, between −137 and −124 (Fig. 5B). Three haplotypes were isolated in the *Amy35* upstream region, in homozygous or heterozygous flies. A striking result is that the Bouaké strain harbors the three types in a single indi-

Table 2. χ^2 values for comparisons of base usage between *Amy* gene copies of *D. ananassae* and *D. melanogaster*. One copy per group was used, since the intragroup differences were negligible.

Gene copy	*Amy35*	*Amy4N*	Amy-c1	Amy-c6	Position in codon
Amy4N	0.104				1
	0.151				2
	11.8**				3
Amy-c1	0.943	0.631			1
	1.472	2.353			2
	52.2***	18.8***			3
Amy-c6	2.209	2.391	3.115		1
	1.638	2.615	0.198		2
	6.331	17.4***	43.9***		3
melanogaster	0.648	0.490	1.495	1.020	1
	0.129	0.009	2.143	2.351	2
	49.8***	78.7***	151.0***	41.6***	3

** $p < 0.01$; *** $p < 0.001$.

vidual. To check for contamination, several flies from this strain were amplified with primers selective for the three types (see Materials and Methods). The other strains were also checked and compared to the results of sequencing as a control. The results confirm that most flies from Bouaké have the three haplotypes (not shown). We conclude that in this strain, there is at least one additional *Amy* copy almost identical to *Amy35* in its 5′ region. In *Amy58,* only haplotype 1 was found (Table 3), except in the Bouaké strain, in which type 1 and type 3 were amplified in the same fly. Thus, since type 3 is present only upstream of *Amy35* and not upstream of *Amy58,* except in Bouaké, it suggests a possible occurrence of a concerted evolution event, such as gene conversion, from *Amy35* toward *Amy58* in this particular strain.

The Intron-Containing Genes: Amy4N, Amyi5. *Amy4N* and *Amyi5* are almost identical. Their divergence is limited to four nucleotide substitutions in the coding sequence, one of which is an amino acid replacement that modifies the electric charge of the protein by one unit (N279D). They both have a short intron (61 and 62 bp, respectively) between positions 177 and 178, at the usual position (Da Lage et al. 1996b). The introns differ by one substitution and one single-base indel. The physical arrangement of the two genes is not known presently. The putative proteins encoded by *Amy4N* and *Amyi5* are 495 aa long. Somatic transformations have shown that *Amyi5* encodes AMY3 and *Amy4N* encodes AMY4 (Fig. 3).

The base usage is similar to that of *D. melanogaster* at the first two codon positions but is very different at the third position, as observed above for the *Amy35* group (Table 2). Accordingly the codon usage is significantly different from *D. melanogaster,* and generally less biased: indeed the GC_3 is much lower (65%). Most remarkable discrepancies are for Asp, like in the *Amy35* group, and also for Asn: AAT/AAC is 11/27 for *Amy4N* and 3/32 in *D. malanogaster;* and for Glu: GAA/GAG is 11/6 for *Amy4N* and 3/15 in *D. malanogaster*.

The noncoding flanking regions of *Amy4N* and *Amyi5* are strikingly similar (Fig. 4C shows the 3′ region). The 5′ regions are almost identical (99% similarity) within the last 860 bp before ATG. The putative TATA box lies at −50 and the putative CAAT box at −121. The putative transcription start is at −19. A putative midgut regulatory element, GATAAGAT, is at −909 in the *Amy4N* clone, but it was not found in *Amyi5*. The 3′ regions are also highly conserved, up to the end of the *Amy4N* clone (*Sal* I site), except the presence of a 255-bp indel and the expansion of the motif TCTG (or TCCG) in *Amyi5,* raising the quadruplet from 3 to 11 repeats.

Despite their similarity, *Amy4N* and *Amyi5* are clearly not alleles. The first evidence is that they code for AMY4 and AMY3, respectively, and that these two proteins are produced by different genes in the strain Taï 13-1610. Another evidence has been supplied by single-fly Southern hybridizations (not shown) on *Sal* I digests, which revealed that the bands of *Amy4N* (4 kb) and *Amyi5* (5 kb) were always present (20 flies tested). A third evidence is that the similarity in the flanking regions does not span the entire clones.

At the 5′ end of the *Amy4N* clone, there is a complex region made of four repeats of a motif of about 150 bp. This region is also present downstream to *Amyi5,* with a better conservation of the repeat units. The dot-plot (Fig. 6) *Amy4N* versus *Amyi5* illustrates the complexity of the region. A search in sequence databases has shown that this sequence had been found in the flanking region of a *D. ananassae mariner* element (Robertson and Lampe 1995). This fragment is likely responsible for the multiple labelings observed in *Amy4N* in situ hybridization.

The Amyc1 *Gene.* The *Amyc1* gene lies on the chromosomal arm 3L, the same as the *Amy4N* cluster. There is no clue that this *Amy* copy is duplicated. Its coding

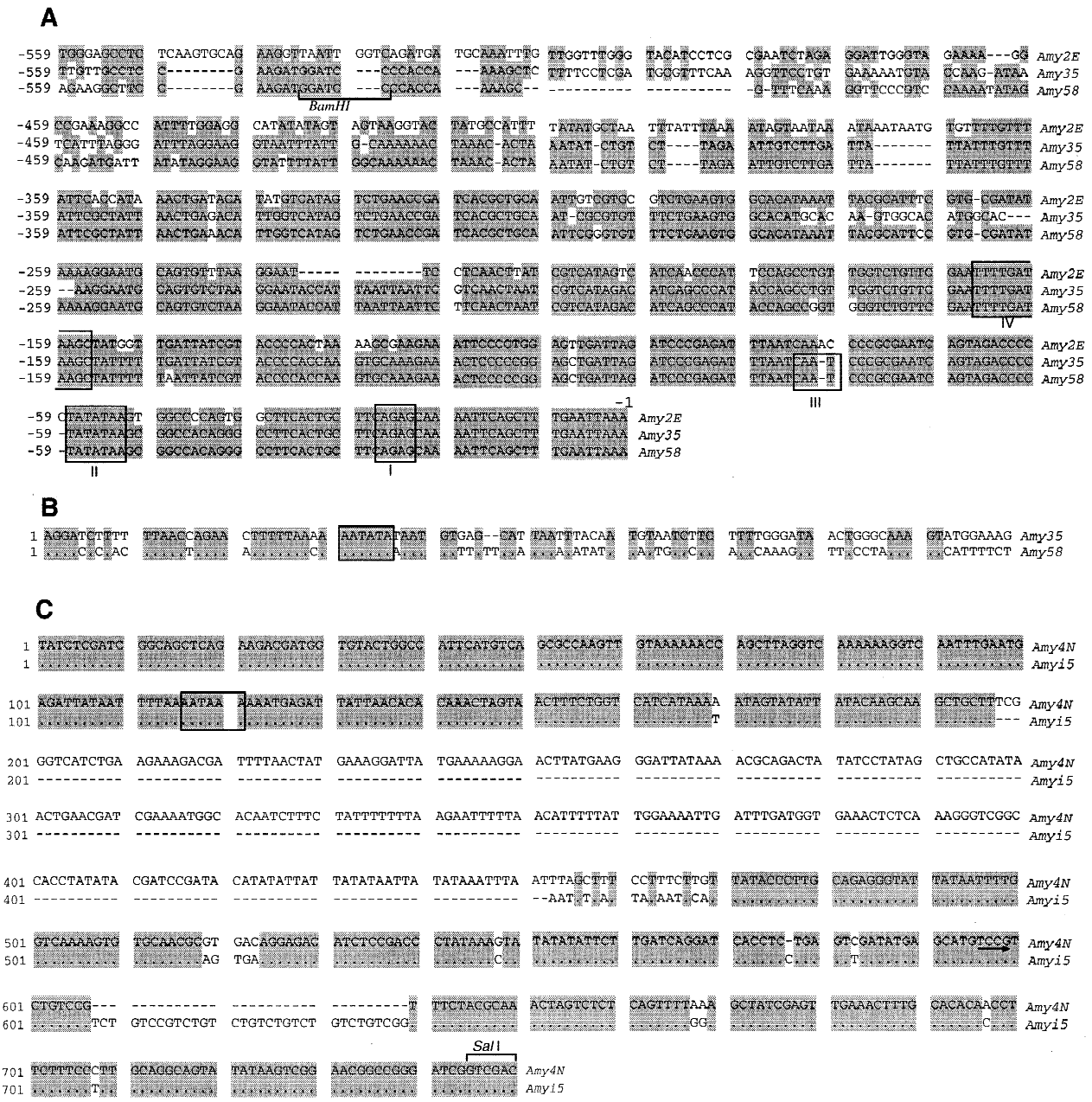

Fig. 4. Alignments (with CLUSTAL W) of the flanking regions of duplicated *Amy* genes. **A:** Upstream regions of the three copies of the *Amy35* group, numbered from the ATG; box I: putative transcription start; box II: putative TATA box; box III: putative CAAT box; box IV: putative midgut regulatory element. **B:** Downstream region of the three copies of the *Amy35* group. The putative polyadenylation site is boxed. **C:** Downstream region of the *Amy4N* group. The putative polyadenylation site is boxed. The repeated tetranucleotide (see text) is underlined by an arrow. The restriction site *Sal* I, at the end of the 4N clone, is indicated.

sequence is 1485 bp long and is interrupted by a short intron (56 bp) at the ancestral position. The putative protein encoded by *Amyc1* has not been detected but is expected to migrate in the reverse way, given the score of charged residues. No transcript has been detected yet, and the transient expression assay was unsuccessful. Unlike most amylase genes studied until now in *Drosophila* (Da Lage et al. 1998; Inomata et al. 1997; Moriyama and Gojobori 1992; Popadic and Anderson 1995), *Amyc1* has a low codon bias. The GC_3 is 55.5%, which is much lower than in *D. melanogaster*, for instance. As generally observed in *Drosophila* (Nakamura et al. 1997), *Amyc1* uses little or not at all the codons TTA (Leu) and GGG (Gly). On the other hand, there is no preference for the phenylalanine codon TTC versus TTT (10/16), which is unusual in *Drosophila* amylases. Other codon families are also unbiased or biased toward the unusual preference for T or A: Asn (AAT/AAC = 20/23); Thr (ACT/ACC = 10/5); Asp (GAT/GAC = 14/11); Glu (GAA/GAG = 11/9). These data suggest that *Amyc1* is potentially active since it exhibits a correct and full-size coding sequence; but the analysis of codon preference

Table 3. Haplotype variants found by sequencing in various *D. ananassae* strains

Strain (one fly)	Number of clones sequenced	Region -283/-258			Region -137/-124		
		Type normal	Type Taï	Type Bangalore	Type 1	Type 2	Type 3
5′-Amy35							
Takapoto (Tuamotu)	4	+			+		
371-1 (Mexico)	3	+				+	
R1061 (Réunion)	2	+				+	
Beruwala (Sri Lanka)	1	+			+		
Bangalore (India)	3			+	+		
Cuba	3	+			+		
Mexico	4	+	+		+		
Mauritius	2	+	+		+		
Bouaké (Ivory Coast)	7	+	+		+	+	+
Djeffa (Benin)	2	+	+		+		
Brazzaville (Congo)	3	+	+		+		
K3422 (Thailand)	3	+	+		+		
Korat (Thailand)	2	+			+		
Guadeloupe	5	+			+		+
Sao Paulo (Brazil)	1	+			+		
Lambir (Borneo)	2	+			+	+	
Colombo (Sri Lanka)	3	+			+	+	
Porto Rico	4	+					+
Yaoundé (Cameroon)	3	+	+		+		
Taï 13-1610	3		+		+		
5′-Amy58							
Takapoto (Tuamotu)	ND						
371-1 (Mexico)	DS	+			+		
R2061 (Réunion)	DS	+			+		
Beruwala (Sri Lanka)	2	+			+		
Bangalore (India)	NA						
Cuba	ND						
Mexico	DS	+			+		
Mauritius	DS	+			+		
Bouaké (Ivory Coast)	6	+			+		+
Djeffa (Benin)	NA						
Brazzaville (Congo)	DS	+			+		
K3422 (Thailand)	DS	+			+		
Korat (Thailand)	ND						
Guadeloupe	DS	+			+		
Sao Paulo (Brazil)	DS	+			+		
Lambir (Borneo)	DS	+			+		
Colombo (Sri Lanka)	DS	+			+		
Porto Rico	DS	+			+		
Yaoundé (Cameroon)	NA						
Taï 13-1610	DS	+			+		

One single fly per strain was assayed. For cloned PCR products, several clones were sequenced when possible. ND: not done; NA: no amplification; DS: direct sequencing of PCR products.

indicates that the level of expression should be low. We have found no typical regulatory sequence in the upstream region available for *Amyc1* (563 bp).

The Amyc6 *Gene. Amyc6* is identified as the *Amyrel* gene of *D. ananassae* (see its branching with *Amyrel* of *D. malanogaster* on Fig. 2). It has been extensively described previously (Da Lage et al. 1998). We may recall that it is located on the 3L chromosomal arm at position 76C and is thought to be single-copy. *Amyrel* has a specific characteristic, which is the presence of a unique and short intron in position 655. The putative encoded protein (493 aa) has not been detected by the usual electrophoresis method (even in transient expression assay), but is expected to migrate fast. A transcript has been detected by RT-PCR in third-instar larval midguts (not shown).

In the 5′ region of *Amyc6*, a putative TATA box has been found in −45 and the putative regulatory element GCGATAAGATT in −66. However, we found no clear CAAT box.

Intergroup Comparisons

A general alignment of the proteins (except *Amy2E*, not known completely) is shown in Figure 7.

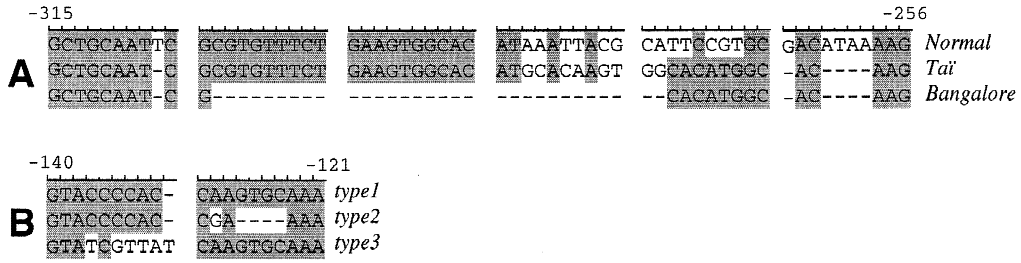

Fig. 5. Polymorphism in the 5′ region of Amy35. **A:** Region −315 to −256. The "Bangalore" type seems derived from the "Taï" type by a single 32-bp deletion. **B:** Region −140 to −121: type 1 and type 2 differ by a single indel; type 3 has multiple substitutions.

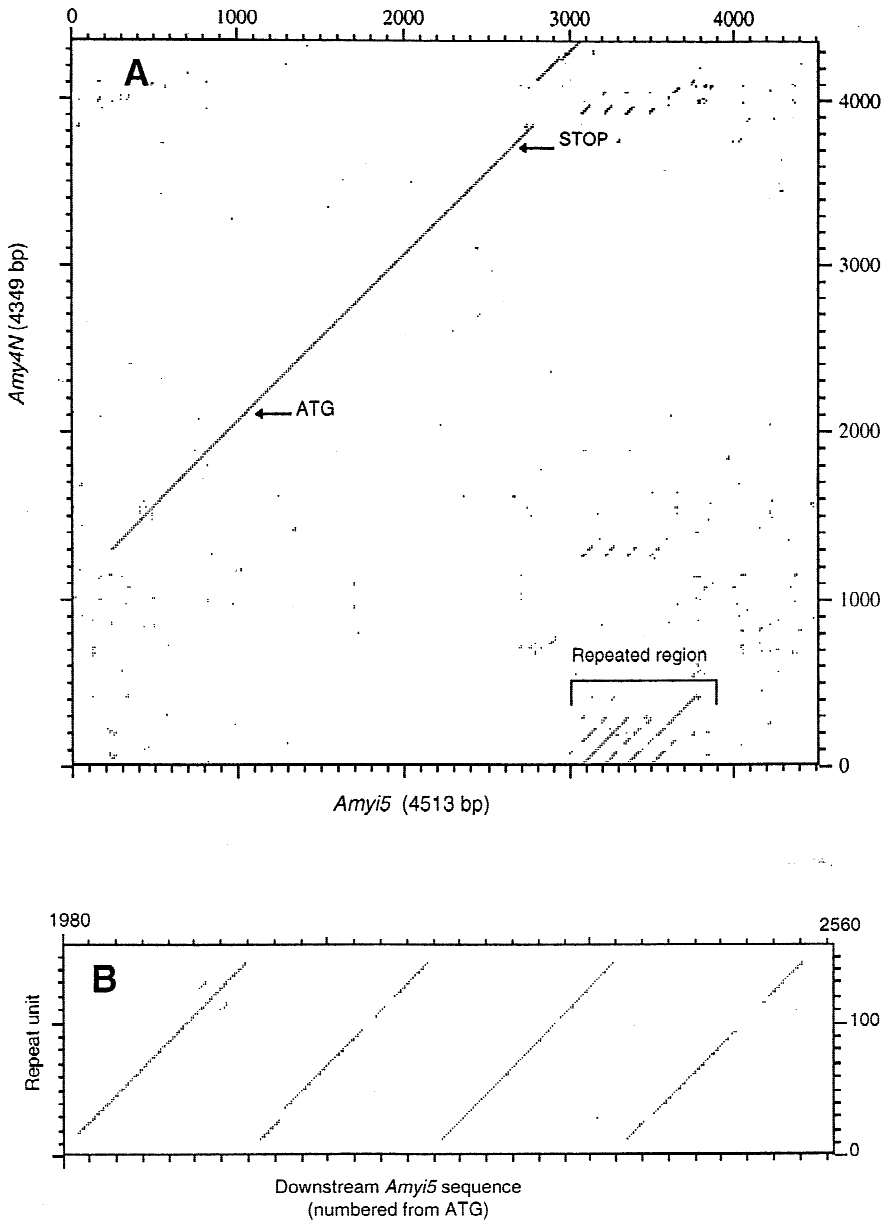

Fig. 6. **A:** Dot plot (drawn with SEQAPP) of the *Amy4N* region against *Amyi5*. The stringency is 17 matches in a window of 25 nucleotides. The repeated region is indicated. **B:** Dot plot (same stringency) of the repeated region within the *Amyi5* sequence.

Comparison Between the Two Gene Clusters Amy35 *and* Amy4N. *Amy35*, *Amy58*, *Amyi5*, and *Amy4N* encode AMY1, AMY2, AMY3, and AMY4 respectively, and are therefore considered as the "classical genes." Since *Amy35* and *Amy58* (and *Amy2E*, as far as we know) are very similar to each other and *Amy4N* and *Amyi5* are almost identical, we will use one gene from each cluster for comparisons, namely, *Amy35* and *Amy4N*.

Amy35 and *Amy4N* are located on different chromosomes. *Amy35* is intronless while *Amy4N* has an intron at

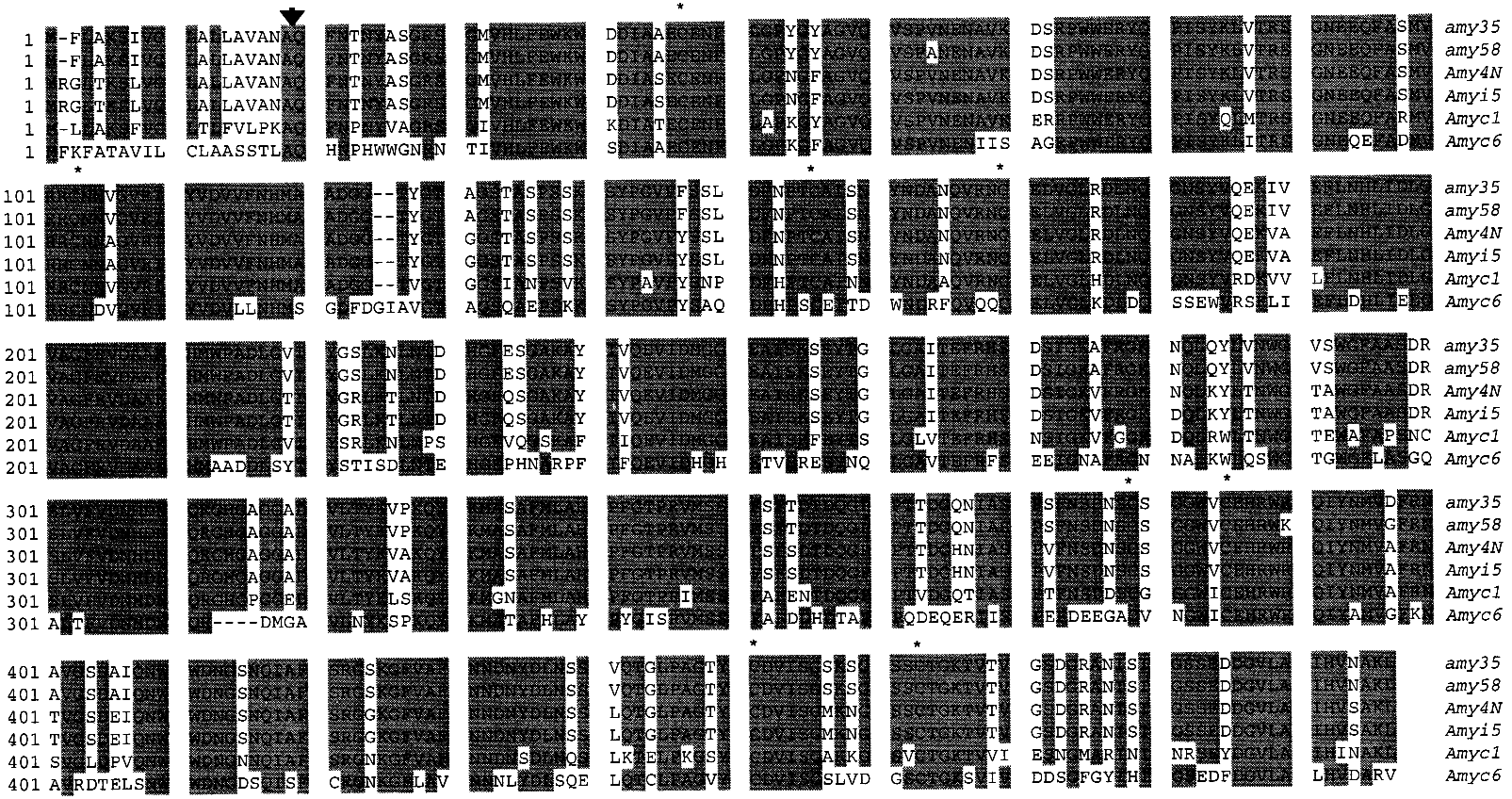

Fig. 7. Alignment of the six complete amylase protein sequences of *D. ananassae* (made with CLUSTAL W and edited with SEQAPP). Shading is for 5/6 consensus. The black arrowhead marks the cleavage site of the signal peptide. Asterisks mark the conserved cysteins involved in disulfide bonds in the pig and putatively in other animals.

the ancestral position. *Amy4N* has an additional amino acid (Arg) in second position, in the signal peptide, which is made of the first 18 residues (cleavage site between A and Q). The nucleotide divergence between these genes is 18% and the amino acid divergence is 6.1% (see also Table 1). The GC_3 contents are 74.9% and 64.9%, respectively (significantly different: $\chi^2 = 11.8$; $p < 0.001$), which is indicative of a higher codon bias in *Amy35*. These data show that a significant structural evolution occurred after duplication from the ancestral gene, with a probable intron loss in the *Amy35* group (see discussion). However, both proteins remain classical alpha-amylases and their biochemical differences remain unknown.

The major differences are observed in the flanking regions. Indeed, it is difficult to identify homologies, and it is impossible to align the 5′ or 3′ regions from *Amy35* and *Amy4N*. The TATA boxes are not identical, nor the putative CAAT boxes. A GATAAG motif has been found in both genes. This motif might be involved in midgut regulation, as shown by Magoulas et al. (1993). But it is much more upstream in *Amy4N* (−909) than in *Amy35* (−162) and it is absent in *Amyi5*, the "twin" companion of *Amy4N*. On the other hand, the putative transcription start site TCAGAG remains unchanged between the two genes, and in both, it is 20 bp downstream to the TATA box.

Comparison Between the Classical Genes and the Highly Divergent Copies. *Amy35* and *Amy4N* may be considered as not very different from each other when the divergent genes *Amyc1* and *Amyc6* are considered, as suggested by the tree in Fig. 2. The amino acid divergence between *Amyc1* and *Amy4N* or *Amy35* is 21.6%. Like *Amy4N*, *Amyc1* has an intron (with no sequence similarity), but like *Amy35*, it lacks the extra residue in second position. The amino acid replacements are spread along the sequence, but the typical motifs of animal alpha-amylases (Janecek 1994) are conserved in *Amyc1*, including the crucial eight cysteins (not counting the cystein in the signal peptide).

Amyc6 (i.e., *Amyrel*) was compared previously to the classical *Amy35* and *Amy4N* (Da Lage et al. 1998). The divergence between *Amyrel* and the classical genes is about 41% in amino acids, with an additional pair of cysteins that could build a fifth disulfide bridge. The divergence with *Amyc1* reaches 43%. As already mentioned, *Amyrel* has a specific intron.

A global comparison of the nucleotide sequences shows that the base usage is not significantly different between the different members of the family at the first two codon positions ($\chi^2 = 4.85$ and 4.08, respectively, df = 9), which reflects the conservation of a majority of amino acids. In contrast, at the third position, the composition is very different between copies ($\chi^2 = 78.2$, $p < 0.001$; df = 9). This is due to the differential codon usage. The GC_3 content is highest in the *Amy35* group and *Amyrel* (around 75%), it is somewhat lower in the *Amy4N* group (65%) and almost equilibrated in *Amyc1* (55%). On the other hand, we know that *Amy35* and *Amy4N* groups have comparable amylase productions. Thus, except for *Amyc1*, which has a fairly low value, the expression level is not clearly correlated to the GC_3 content. However, for *Amy*, this parameter is also variable among species. Inomata et al. (1997) mentioned that the values range from 69.8% (*D. virilis*) to 96.2% (*D. fuyamai*). In *D. ananassae*, all copies are in the lower range.

We can summarize that *Amy35* and *Amy4N* are active, classical amylase genes, as demonstrated by transient transformations. Their sequences remain close to those of other *Drosophila* species. In contrast, *Amyc1* and *Amyc6* are strikingly divergent. Their protein products, if any, and their function are unknown.

Discussion

D. ananassae is known to produce numerous amylase variants, which have been assigned to four loci by genetics experiments (Da Lage et al. 1989, 1992). In this model, the strain Taï 13-1610, which was used in the present study, is supposed to have four active genes organized in two independent clusters [*Amy1, Amy2*] and [*Amy3, Amy4*]. Molecular cloning of seven different *Amy* copies from this strain has confirmed that numerous duplicates were present in its genome. Two of them, *Amy35* and *Amy58*, were evidenced as a compact tandem that encodes AMY1 and AMY2 enzymes. *Amy2E* belongs to the same locus, at some unknown distance. The sequence of this copy is not known completely, but in the available region, negative charges are dominant. *Amy2E* might therefore code for a faster known amylase variant, namely, AMY-1, which is observed in some larvae from the Taï population and sometimes in individuals from Taï 13-1610. From the other clones we can state that *Amy4N* and *Amyi5* belong to another locus on a different chromosome, but the phage inserts (Fig. 1) show that the two copies are less close to each other than in the former cluster.

These four (or five) genes are classical amylase genes, in that they produce enzymes that are revealed by the usual techniques of detection of starch degradation. The divergence between the *Amy35* and *Amy4N* groups is moderate within the coding sequence, except the intron loss in *Amy35*. The biochemical differences of the proteins are not known, and should be studied in vitro, as it has been done for *D. virilis* and *D. repleta* (Prigent et al. 1998). Parameters such as optimal pH and temperature may be different and could be indicative of evolutionary adaptation to broader environmental conditions.

The main differences are in the regulatory sequences,

which control the spatial and temporal expression. It has been observed that AMY1 and AMY2 are often larval, and AMY3 and AMY4 more specifically adult products. Some tissue-specificity has also been shown in larval midgut (Da Lage et al. 1996a). We suggested that the differential expression of these gene clusters might be beneficial and thus could explain that the duplicated structure was retained by selection. At the molecular level, the sequence motifs involved in these regulations are not known, nor are the putative regulatory proteins. For instance, we may wonder if the repeated element upstream of *Amy58* may have been involved in the regulation of this gene cluster. In human, a retroviral insertion 5′ to an amylase gene has triggered the switch of this gene toward a salivary-specific expression (Ting et al. 1992; Samuelson et al. 1996). A large field of investigation is thus open.

Within each gene cluster, it is likely that the regulation is identical because the 5′ regions are very well conserved. We have no evidence to date that the polymorphism described in the regulatory regions of *Amy35* modifies the gene expression. Moreover, previous observations have shown that AMY1 and AMY2 are coordinately expressed, as well as AMY3 and AMY4 (Da Lage et al. 1996a). Therefore, we may consider the presence of several quite identical genes as a way for increasing amylase production. However, we cannot definitely exclude some biochemical differences. In this respect, duplications could be regarded as a mean for fixing an advantageous heterozygosity (Ohno 1970).

The molecular events leading to duplications have encompassed sufficient flanking sequence to allow the new copies to be active in the two gene clusters: mostly upstream sequences in the *Amy35* cluster, but also sufficient 3′ sequence has been retained to conserve a polyadenylation site; in the *Amy4N* cluster, hundreds of nucleotides 5′ and 3′ of the genes were duplicated. Such conservation in flanking regions was not observed in *D. melanogaster,* in which there are only 57% similarity upstream of the TATA box (Boer and Hickey 1986).

We have pointed out that within each gene cluster, the sequence divergence is very low, as generally observed in closely linked copies (see, e.g., Wang et al. 1999). The high conservation between duplicates may be the result of concerted evolution, as well as of a recent duplication. To check whether concerted evolution occurred, it was necessary to get the corresponding sequence in other populations/species. At the intraspecific level, we tried to find some information from the upstream regions of *Amy35* and *Amy58,* because coding regions were not variable enough. In a single population only (Bouaké), a stretch of changed nucleotides was found to be shared by the two regions, which may be indicative of gene conversion. Older events of concerted evolution might be detected by studying related species, but we failed to amplify these regions in other species. On the other hand, we were not able to check for concerted evolution using the coding sequences because we could not amplify each copy separately in other species.

A recent study (Inomata and Yamazaki 2000) shows a parallel situation in *D. kikkawai,* with two divergent clusters of two closely related copies located on distinct chromosomal arms. Concerted evolution was evidenced in coding sequences, but flanking regions have diverged. The situation is rather more complicated in *D. ananassae* and seems unrelated, in spite of similar aspects. To date, it is a challenge to understand the evolutionary history of the *Amy* family in *D. ananassae,* and it is difficult to determine which gene was ancestral. *D. ananassae* is known to bear a number of inversions and translocations (Singh 1985; Tobari 1993), and this may be a mechanism for spreading the duplicate copies at remote chromosomal loci. It was suggested (Sturtevant and Novitski 1941) that the 3L arm of *D. ananassae* (that bears the *Amy4N* group, *Amyc1* and *Amyrel*) is homologous to the 2R of *D. melanogaster* (that bears *Amy* and *Amyrel*), which could have been a clue in finding the original locus. However, because of the numerous rearrangements in both species, we should be cautious in inferring from a correspondence between the chromosomal arms of *D. ananassae* and *D. melanogaster*. Also, other data support this hypothesis: we had previously suggested that AMY3 was ancestral because in several species of the *D. ananassae* complex, and in *D. varians* (less related), a single amylase is expressed and migrates like AMY3 (encoded by *Amyi5*) (Da Lage et al. 1989); perhaps more convincing is the presence of an intron in the *Amy4N* group, which is certainly ancestral (Da Lage et al. 1996b). However, alternatively, we observe that the *Amy35* group shares clear sequence similarity with *D. melanogaster* (see Fig. 2), including the intron loss. But the intron loss was probably independent in the *D. melanogaster* and *D. ananassae* lineages. An argument for a specific intron loss in the *D. melanogaster* lineage is that in *D. takahashii,* which is very close to the *D. malanogaster* subgroup, we did not find intronless classical genes (unpublished). On the other hand, we cannot explain the sequence similarity of *D. melanogaster* with *Amy35,* except by assuming that *Amy35* is the ancestor, which has lost its intron after a duplication that gave rise to *Amy4N*, and that this latter gene has undergone an accelerated divergence. Later in the evolution of the subgroup, each of them would have been duplicated, in at least some species. The presence of repeated sequences in the vicinity of each gene cluster might have helped in duplication events (see, e.g., Cross and Renkawitz 1990). Actually we suspect that the number of classical *Amy* genes may be variable even between *D. ananassae* populations, as suggested by our results on the Bouaké population. This has already been found in rodents and primates

(Groot et al. 1989; Nielsen 1977). Indeed, tandem arrangements are favorable to subsequent unequal crossing over (Ohno 1970). However, the electrophoretic monomorphism of some populations is most likely due to regulatory events rather than to gene loss (Da Lage et al. 1996a). The evolution of the *Amy4N* gene cluster may have been influenced by its peculiar cytological localization, at a very basal position near the centromere, in a region suspected of low recombination and genetic variation (Stephan and Langley 1989). In this respect, further investigations on the amylase family within and between *D. ananassae* populations would be very interesting.

The evolutionary status of the divergent genes *Amyc1* and *Amyrel* and the hypothetical advantage they might confer to the bearer are puzzling. The case of *Amyrel* (*Amyc6*) has been discussed in detail (Da Lage et al. 1998). However, we have now good indication that this gene predated the *Drosophila* radiation (unpublished), so that its presence in *D. ananassae* is not surprising and its function would be common to all *Drosophila* species. On the other hand, this function is unknown, and it was not possible to observe the protein in transient transformations. *Amyc1* has been detected in several species of the *D. ananassae* subgroup, including *D. varians*, which is considered the most distant species inside the *D. ananassae* subgroup (unpublished). Thus, it is reasonable to consider *Amyc1* as a "permanent" member of the *Amy* family in the subgroup. On the other hand, no *Amyc1* homologous was found in the complete *D. melanogaster* genome. According to the moderate ancientness of these two genes, we can suspect accelerated divergence to have occurred, which is often observed after duplications through relaxation of constraints or direct positive selection (Ohta 1991; Wu et al. 1986). Shibata and Yamazaki (1995) have shown such a process to have worked in the evolution of *Amy* within the *D. melanogaster* subgroup, in the *D. erecta* lineage. This species is specialized on Pandanus tree and it is clear that adaptation to new feeding resources through digestive enzyme modification should be of interest for the species.

Concerning the function of *Amyc1*, we can only hypothesize from the sequence data. The transient transformations with this gene were negative, in both polarities of migration (but no positive control was coinjected in the same eggs). The low codon and compositional bias (55% GC at third codon position) may be indicative of a low level of expression (Moriyama and Hartl 1993), and indeed the sequence does not suggest that it is a pseudogene: correct reading frame, low usage of some codons typical of *Drosophila*.

In *D. ananassae*, the *Amy* gene duplications had visible effects in terms of electrophoretic polymorphism. Four active genes may produce a number of allelic combinations the expression of which is modulated by regulatory sequences. We have shown that the stage- and tissue-specificity of *Amy* expression could be due to the high divergence of flanking regions, while coding sequences remain similar. The multicopy structure of the classical genes has favored the occurrence of potential physiological differences and a possibility of a high yield of the enzyme, with a high capacity of variation, which could be useful, for instance in escaping from amylase inhibitors from food. More divergent copies may have gained different biochemical properties, perhaps adapted to other dietary carbohydrates. Remarkably, it is worth noting that no pseudogene was found among the seven genes of *D. ananassae.*

Acknowledgments. We are pleased to thank N. Chaminade, P. Santamaria, and C. Maisonhaute for their help in experimental work and C.H. Langley for fruitful discussion.

References

Baker JE, Halliday WR, Lum PTM (1990) Genetics of alpha amylase in *Sitophilus oryzae* (L.) (Coleoptera: Curculionidae). J Stored Prod Res 26(1):7–10

Boer PH, Hickey DA (1986) The alpha-amylase gene in *Drosophila melanogaster* nucleotide sequence, gene structure and expression motifs. Nucleic Acids Res 14:8399–8411

Brown CJ, Aquadro CF, Anderson WW (1990) DNA sequence evolution of the amylase multigene family in *Drosophila pseudoobscura*. Genetics 126:131–138

Cariou M-L, Da Lage J-L (1993) Isozyme polymorphisms. In Tobari, YN (ed) *Drosophila ananassae*, genetical and biological aspects. Tokyo:Japan Scientific Societies Press, pp 160–171

Comeron JM (1995) A method for estimating the number of synonymous and nonsynonymous substitutions per site. J Mol Evol 41: 1152–1159

Cross M, Renkawitz R (1990) Repetitive sequence involvement in the duplication and divergence of mouse lysozyme genes. EMBO J 9(4):1283–1288

Da Lage J-L, Cariou M-L, David JR (1989) Geographical polymorphism of amylase in *Drosophila ananassae* and its relatives. Heredity 63:67–72

Da Lage J-L, Lemeunier F, Cariou M-L, David JR (1992) Multiple amylase genes in *Drosophila ananassae* and related species. Gen Res Camb 59:85–92

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

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 Natl Acad Sci USA 95(12):6848–6853

Gemmill RM, Levy JN, Doane WW (1985) Molecular cloning of alpha-amylase genes from *Drosophila melanogaster*. I. Clone isolation by use of a mouse probe. Genetics 110:299–312

Groot PC, Bleeker MJ, Pronk JC, Arwert F, Mager WH, Planta RJ, Eriksson AW, Frants RW (1989) The human alpha amylase multigene family consists of haplotypes with variable number of genes. Genomics 5:29–42

Hickey DA, Benkel BF, Boer PH, Genest Y, Abukashawa S, Ben-David G (1987) Enzyme-coding genes as molecular clocks: the molecular evolution of animal alpha-amylases. J Mol Evol 26:252–256

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(4):601–615

Inomata N, Tachida H, Yamazaki T (1997) Molecular evolution of the *Amy* multigenes in the subgenus *Sophora* of *Drosophila* Mol Biol Evol 14(9):942–950

Janecek S (1994) Sequence similarities and evolutionary relationships of microbial, plant and animal alpha-amylases. Eur J Biochem 224: 519–524

Janecek S (1997) Alpha-amylase family: molecular biology and evolution. Prog Biophys Mol Biol 67(1):67–97

Kumar S, Tamura K, Nei M (1993) MEGA, molecular evolutionary genetics analysis. The Pennsylvania State University

Laulier M (1988) Génétique et systématique évolutives du complexe d'espèces *Sphaeroma hookeri* Leach, *Sphaeroma levii* Argano et *Sphaeroma rugicauda* Leach (Crustacés Isopodes Flabellifères). 1. Génétique formelle de onze locus enzymatiques. Genet Sel Evol 20:63–74

Magoulas C, Loverre-Chyurlia A, Abukashawa S, Bally-Cuif L, Hickey DA (1993) Functional conservation of a glucose-repressible amylase gene promoter from *Drosophila virilis* in *Drosophila melanogaster*. J Mol Evol 36:234–242

Meisler MH, Ting C-N (1993) The remarkable evolutionary history of the human amylase genes. Crit Rev Oral Biol Med 4:503–509

Meisler MH, Antonucci TK, Treisman LO, Gumucio DL, Samuelson LC (1986) Interstrain variation in amylase gene copy number and mRNA abundance in three mouse tissues. Genetics 113:713–722

Moriyama EN, Gojobori T (1992) Rates of synonymous substitution and base composition of nuclear genes in *Drosophila.* Genetics 130:855–864

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

Moriyama EN, Powell JR (1997) Codon usage bias and tRNA abundance in *Drosophila.* J Mol Evol 45:514–523

Nakamura Y, Gojobori T, Ikemura T (1997) Codon usage tabulated from the international DNA sequence databases. Nucleic Acids Res 25:244–245

Nielsen JT (1977) Variation in the number of genes coding for salivary amylase in the bank vole *Clethrionomys glareola.* Genetics 85: 155–169

Ohno S (1970) Evolution by gene duplication. Berlin: Springer

Ohta T (1991) Multigene families and the evolution of complexity. J Mol Evol 33:34–41

Payant V, Abukashawa S, Sasseville M, Benkel BF, Hickey DA, David J (1988) Evolutionary conservation of the chromosomal configuration and regulation of amylase genes among eight species of the *Drosophila melanogaster* species subgroup. Mol Biol Evol 5(5): 560–567

Popadic A, Anderson WW (1995) Evidence for gene conversion in the amylase multigene family of *Drosophila pseudoobscura*. Mol Biol Evol 12(4):564–572

Popadic A, Norman RA, Doane WW, Anderson WW (1996) The evolutionary history of the amylase multigene family in *Drosophila pseudoobscura.* Mol Biol Evol 13(6):883–888

Prigent S, Matoub M, Rouland C, Cariou M-L (1998) Metabolic evolution in alpha-amylases from *Drosophila virilis* and *D. repleta,* two species with different ecological niches. Comp Biochem Physiol 119B(2):407–412

Robertson HM, Lampe DJ (1995) Recent horizontal transfer of a mariner transposable element among and between Diptera and Neuroptera. Mol Biol Evol 12:850–862

Samuelson LC, Phillips RS, Swanberg LJ (1996) Amylase gene structures in Primates: retroposon insertions and promoter evolution. Mol Biol Evol 13(6):767–779

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

Shibata H, Yamazaki T (1995) Molecular evolution of the duplicated *Amy* locus in the *Drosophila melanogaster* species subgroup: concerted evolution only in the coding region and an excess of nonsynonymous substitutions in speciation. Genetics 141:223–236

Singh BN (1985) *Drosophila ananassae*—a genetically unique species. Nucleus 28(3):169–176

Stephan W, Langley CH (1989) Molecular genetic variation in the centromeric region of the X chromosome in three *Drosophila ananassae* populations. I. Contrasts between the *vermilion* and *forked* loci. Genetics 121:89–99

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

Tadlaoui-Ouafi A (1993) Evolution structurale et moléculaire de la famille multigénique Amylase chez quelques Drosophilidae. Thesis, Université Pierre et Marie Curie, Paris, 142 pp

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

Ting C-N, Rosenberg MP, Snow CM, Samuelson LC, Meisler MH (1992) Endogenous retroviral sequences are required for tissue-specific expression of a human salivary amylase gene. Genes Dev 6:1457–1465

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

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

Van Wormhoudt A, Sellos D (1996) Cloning and sequencing analysis of three amylase cDNAs in the shrimp *Penaeus vannamei* (Crustacea decapoda): evolutionary aspects. J Mol Evol 42:543–551

Wang S, Magoulas C, Hickey D (1999) Concerted evolution within a trypsin gene cluster in *Drosophila*. Mol Biol Evol 16(9):1117–1124

Wu CI, Li WH, Shen JJ, Scarpulla RC, Limbach KJ, Wu R (1986) Evolution of cytochrome c genes and pseudogenes. J Mol Evol 23:61–75