Bipartite Structure of the 5S Ribosomal Gene Family in a 
Drosophila melanogaster Strain, and Its Evolutionary Implications

Marie-Laure Samson and Maurice Wegnez

Center de Génétique Moléculaire, Centre National de la Recherche Scientifique, 91190 Gif-sur-Yvette, France and Laboratoire d'Embryologie Moléculaire et Expérimentale, Unité Associée 1134 du Centre National de la Recherche Scientifique, Université Paris XI, 91405 Orsay Cedex, France

Manuscript received August 10, 1987
Revised copy accepted January 2, 1988

ABSTRACT

Knowledge of multigenic family organization should provide insight into their mode of evolution. Accordingly, we characterized the 5S ribosomal gene family in the Drosophila melanogaster strain rY06. The 5S genes in this strain display a striking HindIII restriction difference compared to the "standard" D. melanogaster 5S genes. The sequence of three rY06 5S genes was determined. We show that the HindIII restriction site heterogeneity within the rY06 5S family most probably results from the same point mutation, suggesting that a single 5S variant was propagated into the 5S cluster of this strain. Furthermore, we demonstrate that the structural organization of the 5S genes in rY06 is a bipartite structure, i.e., that about 40% of the 5S genes constitute a HindIII +/HindIII - mixed cluster, while those remaining constitute an homogeneous HindIII - cluster. The events which might lead to such an heterogeneous pattern are discussed from an evolutionary point of view.

MULTIGENIC families constituted by repeated sequences are present in all eukaryotic species. Although unit structure and family organization are highly variable among species, a high level of homogeneity is maintained within each, the different units of the family evolving, apparently, in unison. This process is usually referred to as concerted evolution (ZIMMER et al. 1980).

Nonreciprocal exchanges, mainly unequal crossing-over and conversion events, are commonly thought of as phenomena affecting the evolution of repeated sequences (SMITH 1976; NAGYLA KY and PETES 1982). These events are able to modify the frequency of a variant in a multigene family, leading thus to its spread or elimination. For example, an experimental determination of the rate of unequal mitotic crossing-over in Saccharomyces cerevisiae shows that it would be high enough to homogenize the 140 tandemly arranged 18S-28S ribosomal units (SZOS-TAK and WU 1980). The authors calculate that there is a 0.5 probability that a ribosomal locus containing two equally abundant variants would be homogenized after 48,000 generations. However, in most cases, the rate of homogenization resulting from these events is too low to explain the observed levels of homogeneity (DOVER 1982). Thus, at least some of the phenomena affecting the structure of multigene families are faster, and/or are not strictly random, thus keeping only very similar copies as family members. Processes such as nonrandom recombination, transposition and RNA-mediated correction may be involved in concerted evolution (ARNHEIM 1983; STRACHAN, WEBB and DOVER 1985; MORZYCKA-WROBLEWSKA et al. 1985). For instance, the dispersed structure of the Neurospora crassa 5S ribosomal gene family is best explained by transposition (SELKER et al. 1981). The occurrence of directional gene conversion has been demonstrated at the Ascobolus immersus b2 locus, where it leads to a preferential conservation of determined alleles, depending on the molecular nature of the mutation (HAMZA et al. 1986; HAMZA, NICOLAS and ROSSIGNOL 1987).

Experimental approaches to the study of concerted evolution dynamics are difficult. Based on available structural criteria, multigenic families appear homogeneous and seem to have reached a state of equilibrium. Depending on the family, i.e., of nature of functional and structural constraints, different degrees of homogeneity are observed. Noncoding families, such as the 360 and 500 Drosophila satellites, display high levels of heterogeneity (STRACHAN, WEBB and DOVER 1985). This is not surprising since there are in this case few or no functional constraints. On the contrary, high levels of homogeneity are observed in the case of the coding families (ARNHEIM 1983).

The Drosophila melanogaster 5S ribosomal genes constitute a multigenic family clearly displaying concerted evolution. In this species, the haploid genome contains approximately 160 clustered 5S genes (PRO-

1 Present address: Department of Molecular Biology, Massachusetts General Hospital, Boston, Massachusetts 02114.
2 To whom reprint requests should be addressed.

The sequence data presented in this article have been submitted to the EMBl/GenBank Data Libraries under the accession number Y00612.
Drosophila strains: All strains were from the GIF collection. Several D. melanogaster isolates were used (Canton-S, p2, Charollies, Gruta, and Oregon-R), as well as the mutant strains min (PROCUNIER and TARTOF 1975), ry506 (COTÉ et al. 1986) and the balancer stock CyO; TM3/T(2;3)ap1xa males and the F1 males carrying CyO and TM3 balancers were mated with min virgin females. Eight phenotypically distinguishable classes of F2 flies were recovered. They respectively bore all possible combinations of the first, second and third chromosomes of the ry506 strain. All were tested by Southern analysis for the presence or the absence of ry506 specific 5s sequences.

In situ hybridizations were performed according to SPERER et al. (1983). The 5s DNA probe (pBR7A containing ten copies of D. melanogaster 5s genes, provided by V. PIRROTTA) was labeled with [3H]dGTP by nick translation.

Cloning and sequencing of HindIII restriction fragments containing 5s sequences: Fifty micrograms of genomic DNA from the ry506 strain were cut with HindIII. The resulting restriction fragments were fractionated on a 0.6% agarose gel. Two size classes of DNA fragments, respectively around 375 and 750 nucleotide long, were electroeluted into dialysis bags, and then cloned into the HindIII site of pUC8. Screening for clones containing 5s sequences was performed using a D. melanogaster 5s probe, purified from pBR7A.

Plasmid DNA was prepared according to the method described by Birnboim and Doly (1979). Drosophila sequences contained in the clones were purified on acrylamide gels (MAXAM and GILBERT 1980). A set of fragments with various 5' protruding ends were generated by cutting the inserts with appropriate restriction enzymes. Klenow labeling and additional restriction cuts provided fragments with a unique 3' labeled 3' end. These fragments were used in sequencing reactions as described by MAXAM and GILBERT (1980).

RESULTS

Heterogeneity of the D. melanogaster 5s cluster: To study the structural heterogeneity of the 5s locus, we digested the genomic DNA of different strains with restriction enzymes which do not cut standard 5s units, digestions with HindIII yielding the most informative patterns. In addition to a fragment of high Mr (>50 kb) corresponding to clustered standard 5s sequences, we found in all strains, with the exception of Charollies, other fragments hybridizing to 5s sequences (Figure 1). As shown in Figure 1, each strain has a specific HindIII 5s restriction pattern. The number of HindIII fragments usually is low (1-3), with the exception of ry506 (Figures 1 and 2).

The ry506 5s locus: As shown in Figure 2A, the sizes of the ry506 HindIII 5s fragments correspond...
Structural and functional analysis of Drosophila 5S ribosomal gene family

JUNAKOVIC (1982) in another 5S unit. The difference in the length of the spacer segment (nucleotides 1 and 316, respectively, for both clones. This modification is located at the 371 to 374 positions (Figure 3). The two cloned 5S sequences from p002 differ from the standard 5S sequence, showing three point mutations in the transcribed region. These mutations were also found by TSCHUDI, PIRROTTA and JUNAKOVIC (1982), and SHARP et al. (1984).

The HindIII restriction site which distinguishes the cloned 5S genes from the standard sequence results in a T to G point mutation (Figure 3, nucleotide 378) for both clones. This modification is located at the 5S cluster in D. melanogaster (PRENSKY, STEFFENSEN and HUGHES 1973). No additional signals were detected.

Structure and organization of “nonstandard” \(5S\) sequences: Clones of 5S genes bearing HindIII restriction sites were isolated for sequencing. Two clones, p002 and p003, respectively, containing 373- and 749-bp long inserts, were selected. The sequencing results are shown in Figure 3. The two cloned genomic inserts correspond to a whole 5S gene in p002, and to two tandemly arranged 5S units in p003. Their sequences are very similar to the D. melanogaster standard 5S sequence published by TSCHUDI and PIRROTTA (1980), but differ at some locations. The single 5S sequence contained in p002 differs from the standard 5S unit at two positions located in the spacer segment (nucleotides 1 and 316, respectively, Figure 3). The two 5S units in p003 show a total of five differences when compared with the standard 5S sequence. Three point mutations are found in the nontranscribed spacer of the gene (nucleotides 1, 403, and 692, Figure 3), in addition to a duplication of a preexisting trinucleotide (ATT, nucleotides 369-371 and 372-374, Figure 3). The point mutation at position 403 and the duplication of the trinucleotide have also been found by TSCHUDI, PIRROTTA and JUNAKOVIC (1982) in another 5S unit. The difference located at position 578 in one of the two transcribed regions of the p003 insert has previously been described as a variant position in D. melanogaster 5S genes (TSCHUDI and PIRROTTA 1980; TSCHUDI, PIRROTTA and JUNAKOVIC 1982; SHARP et al. 1984).

The HindIII restriction site which distinguishes the cloned 5S genes from the standard sequence results from a T to G point mutation (Figure 3, nucleotide 1) for both clones. This modification is located at the

### TABLE 1

<table>
<thead>
<tr>
<th>Chromosomes from (r_\gamma^{506}) strain</th>
<th>HindIII specific (5S) sequences</th>
</tr>
</thead>
<tbody>
<tr>
<td>Females, (CyO;TM3)</td>
<td>+</td>
</tr>
<tr>
<td>Females, (CyO)</td>
<td>+</td>
</tr>
<tr>
<td>Females, (TM3)</td>
<td>+</td>
</tr>
<tr>
<td>Females, wild type</td>
<td>+</td>
</tr>
<tr>
<td>Males, (CyO;TM3)</td>
<td>-</td>
</tr>
<tr>
<td>Males, (CyO)</td>
<td>-</td>
</tr>
<tr>
<td>Males, (TM3)</td>
<td>-</td>
</tr>
<tr>
<td>Males, wild type</td>
<td>-</td>
</tr>
</tbody>
</table>

**Legend:**
- +: present
- -: absent

**F2 flies were obtained in genetic crosses as described in MATERIALS AND METHODS.** The presence or absence of \(r_\gamma^{506}\) specific sequences was checked by Southern analysis. These are present only in \(r_\gamma^{506}\) DNA. They are absent in the two other strains which were involved in the crosses. The following symbols have been used: +: present, -: absent, U: unknown.
**FIGURE 3.**—Nucleotide sequences of the *r*$_{506}$ HindIII’ 5S units inserted in p002 and p003 (noncoding strand). The sequences corresponding to the transcribed regions are underlined. Arrows at positions 1 indicate the point mutations generating the HindIII site. The corresponding position is not changed in the middle of p003 (open arrow). The point differences with the standard X sequence (see text) are written above each sequence. Hyphens are missing in the standard sequence.

**FIGURE 4.**—Autoradiogram of DNA prepared from a population of *γ*$_{506}$ flies, digested with various amounts of HindIII and analyzed as in Figure 1. Lane 1, total HindIII digest (one unit of enzyme per µg of DNA). Lanes 2 to 4, partial HindIII digests (½, ¼ and ⅛ unit per µg of DNA, respectively).

same position in each case, i.e. 116 nucleotides upstream from the first transcribed nucleotide. On the basis of the multimeric pattern that we observe (Figures 1 and 2) and the primary structure of p002 and p003 inserts, we assume that the other multimeric fragments detected on Southern blots have the same overall structure, i.e. that they are composed of HindIII’ 5S genes surrounded by two HindIII’ 5S units.

To obtain insight into the organization of the HindIII’ 5S sequences, we performed partial HindIII digests of *r*$_{506}$ DNA. Results are shown in Figure 4. Partial restriction cuts lead to a range of 5S hybridizing fragments whose sizes correspond to multimers, up to approximately 20 × 375 bp units. This result indicates that the HindIII’ 5S blocks constitute a cluster of adjacent sequences. The 2.4-kb long fragment (Figure 2) detected in all HindIII digests of *r*$_{506}$ DNA, is not a multimer of a 5S unit and could be one of the outside borders of the...
**HindIII+/HindIII-** mixed cluster. Southern analyses of HindIII restriction digests show high portion of the HindIII- genes remain clustered can be interspersed with HindIII- genes, a significant ure 2A), indicating that while the HindIII+ 5s genes in addition to the smaller multimeric fragments (Fig- 2A, kb fragment appears as a shoulder of the hexamer.

HindIII sites, if all HindIII fragments are clustered, be determined to be approximately 160 per haploid genome HindIII fragment (nine 5s genes within the range of their 40% of the entire 5s cluster.

**Quantifying HindIII+ 5s units in ry®**: Densitometric analyses of the 5s pattern from several HindIII digests of ry® DNA were performed (Figure 5). Relative hybridization intensities of the six restriction fragments whose size is proportional to 375 bp (Figure 2) were determined by measuring the areas of the peaks (Table 2). As shown in Figure 2, relative hybridization intensities are identical, using both DNA from populations of flies or from single flies. Within the range of their M, (375–2250 bp), we can reasonably postulate that the relative hybridization intensities of the multimeric fragments are directly related to their relative proportions. The absolute copy number of HindIII+ 5s genes was thus calculated by considering that the less abundant multimers, i.e., the tetramers and the hexamers, are present only as one single copy (Table 2). This was determined using a ry® strain bearing a single copy of a 3.2-kb HindIII fragment (nine 5s genes of Drosophila teissieri) introduced by transformation (manuscript in prepara- tion). As shown in Table 2, the HindIII+/HindIII- cluster would consist of approximately 66–68 5s genes (Table 2). The corresponding number of HindIII sites, if all HindIII fragments are clustered, would be around 28–29 (Table 2).

The copy number of the 5s genes has been determined to be approximately 160 per haploid genome in D. melanogaster (PROCUNIER and DUNN, 1978). Although gene redundancy varies from strain to strain (PROCUNIER and DUNN 1978), we can deduce that the mixed HindIII+/HindIII- locus contains approximately 40% of the entire 5s cluster.

**DISCUSSION**

**Heterogeneity within the Drosophila 5s clusters:**

In D. melanogaster, 5s genes are known to be tandemly arranged (PROCUNIER and TARTOF 1976; HERSHEY et al. 1977; ARTAVANIS-TSAKONAS et al. 1977). Digestion of genomic DNA with enzymes which do not cut standard 5s units should thus give rise to a single fragment containing the clustered 5s genes (PROCUNIER and TARTOF 1976). However, 5s clusters can be split in some cases into one or more fragments by such restriction enzymes (JUNAKOVIC 1980; TSCHUDI, PIRROTTA and JUNAKOVIC 1982; SAMSON and WEGNEZ 1984; this report), shown to be due to the presence of rare restriction site variants or to the integration of the B104 element within the 5s cluster (TSCHUDI, PIRROTTA and JUNAKOVIC 1982). In this paper, we tested six D. melanogaster strains with HindIII, an enzyme which does not cut the standard 5s gene (TSCHUDI and PIRROTTA 1980). We found that HindIII splits the 5s locus in five of them, giving rise to a specific restriction pattern for each of the strains (Figure 1). In all cases except ry®, the number of HindIII restriction sites within the 5s cluster is low. In ry®, the HindIII restriction pattern includes six restriction fragments whose sizes correspond to 5s gene multimers (Figure 2). On the basis of frag-
ment sizes, this pattern is similar to the ladder resulting from partial digestion with a restriction enzyme whose recognition sequence is present in all 5S units. Sequencing data strongly suggest that all HindIII sites are located at the same position (Figure 3). By analysis of HindIII partial digestions (Figure 4) we found that the six types of multimeric blocks constitute a cluster of adjacent sequences.

**Origin and spread of the HindIII* sites:** The 5S locus in *ry*506 includes 28–29 HindIII* sites (Table 2). Only two point substitutions, one of them corresponding to that found in our sequencing, could generate a new HindIII site in a 5S unit (T to G or A to G substitutions, nucleotides 1 and 332, respectively, Figure 3). As shown in Figure 1, several Drosophila strains bear an occasional HindIII site. Thus, the probability that all of these sites are located at the same position as found in the *ry*506 5S genes is very high. This is supported by the work of TSCHUDI, PIRROTTA and JUNAKOVIC (1982), who located a HindIII variant site at this position by restriction analysis in the Oregon-R Yale strain. One may thus hypothesize that a point mutation leading to a HindIII site occurred in the 5S cluster in the common ancestor of all of the studied strains. Spreading of this variant did not occur, with the exception of the *ry*506 chromosome.

We can thus describe the *ry*506 5S gene cluster as a bipartite locus, with about 40% of the 5S genes constituting a HindIII+/HindIII− mixed cluster, while those remaining constitute a homogeneous HindIII− cluster. Similar clustering was already described for recurrent mutations localized in the non-transcribed spacers of ribosomal RNA genes (Dvorak, Jue and Lassner 1987).

Starting with a single 5S variant, how many steps were required to reach a situation in which 28–29 variants have invaded the 5S locus? This surely required several rounds of conversion or unequal crossing-over. An important point is the fact that HindIII* genes are clustered. If unequal crossing-over and/or conversion events are involved in the process, this means that genetic exchanges concern 5S genes not more than six units apart, since HindIII* fragments longer than six units were not observed in the 5S ladder (Figure 2). This result is surprising, but perhaps explains why the size of the 5S cluster does not vary to any large extent among Drosophila strains. Similar observations were reported in the case of the *S. cerevisiae* ribosomal locus, where the average displacement during unequal crossing-over involves only six to eight ribosomal units (Sztostak and Wu 1980).

An alternative model is provided by the finding of PONT, DEGROOTE and PICARD (1987) who reported the existence of nonchromosomal 5S sequences in Drosophila embryos. These 5S sequences are present in supercoiled DNA molecules whose size corresponds to multimers of 1–16 5S genes. The copy number of these molecules is between 200 and 1000 per embryo. The integration of such circles including several variants within the 5S cluster might increase the variant copy number and, through localized amplification, lead to an asymmetric 5S pattern as found in *ry*506.

**The future of the *ry*506 5S locus:** The main difficulty encountered when discussing phenomena related to concerted evolution is the lack of knowledge about absolute rates of evolution. What is the absolute rate of unequal crossing-over within the Drosophila 5S cluster, for example? The *ry*506 5S locus could be very informative in this respect. Its asymmetric organization is fortuitously revealed by the point mutation leading to a HindIII restriction site. The locus has been invaded by a 5S variant, and displays a characteristic HindIII 5S restriction pattern. What will happen in the future to that locus? As a consequence of concerted evolution, the number of the HindIII 5S variants might be either increased or decreased. Recalling that the 5S pattern is the same when DNA from populations or from single flies is analyzed, we can surmise that the rate at which the distribution of HindIII sites within the 5S cluster is changed is slow. However, the heterogeneous structure of the *ry*506 locus itself demonstrates that occasional exchanges do occur. A survey of the *ry*506 5S locus over several years should bring some interesting information about the absolute rate of evolution within Drosophila 5S genes.

We thank M. C. DAUGERON for performing some preliminary experiments and L. RABINOW for critical reading of the manuscript. During part of this work, M.L.S. was a fellow of the "Ligue Nationale Française Contre le Cancer.” This research was supported by a grant from the Centre National de la Recherche Scientifique (ATP: Organisation et Expression du Génome).

**LITERATURE CITED**


BIRNBOIM, H. C. and J. DOLY, 1979 A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. 7: 1513–1523.


HAMZA, H., A. NICOLAS and J.-L. ROSSIGNOL, 1987 Large heter-
Drosophila 5S Ribosomal Gene Family

691

ologies impose their gene conversion pattern onto closely linked point mutations. Genetics 116: 45–53.


