Unusual Variability of the *Drosophila melanogaster* *ref(2)P* Protein Which Controls the Multiplication of Sigma Rhabdovirus

Philippe Dru, Françoise Bras, Sybille Dezélee, Pierre Gay, Anne-Marie Petitjean, Anne Pierre-Deneubourg, Danielle Teninges and Didier Contamine

Laboratoire de Génétique des Virus, CNRS, 91198 Gif sur Yvette cedex, France

Manuscript received September 17, 1992

Accepted for publication December 21, 1992

ABSTRACT

The *ref(2)P* gene of *Drosophila melanogaster* was identified by the discovery of two alleles, *P^o* and *P^p*, respectively, permissive and restrictive for sigma rhabdovirus multiplication. A surprising variability of this gene was first noticed by the observation of size differences between the transcripts of permissive and restrictive alleles. In this paper, another restrictive allele, *Pⁿ*, clearly distinct from *P^p*, is described: it exhibits a weaker antiviral effect than *P^p* and differs from *P^p* by its molecular structure. Five types of alleles were distinguished on the basis of their molecular structure, as revealed by S1 nuclease analysis of 17 *D. melanogaster* strains; three alleles were permissive and two restrictive. Comparison of the sequences of four haplotypes revealed numerous point mutations, two deletions (21 and 24 bp) and a complex event involving a 3-bp deletion, all affected the coding region. The unusual variability of the *ref(2)P* locus was confirmed by the high ratio of amino acid replacements to synonymous mutations (7:1), as compared to that of other genes, such as the *Adh* (2:42). Nevertheless, nucleotide sequence comparison with the *Drosophila erecta ref(2)P* gene shows that selective pressures are exerted to maintain the existence of a functional protein. The effects of this high variability on the *ref(2)P* protein are discussed in relation to its specific antiviral properties and to its function in *D. melanogaster*, where it is required for male fertility.

NATURAL populations of *Drosophila melanogaster* are endemically infected by the sigma rhabdovirus. This virus shares the capacity to induce a symptom of sensitivity to carbon dioxide with several vesiculoviruses: infected flies are killed after an exposure of 15 min to pure CO₂, a very unlikely situation in natural conditions. Sigma virus is not cytopathogenic to *D. melanogaster* cells and establishes a persistent infection at the cellular level (OHANESSIAN and ECHALIER 1967; RICHARD-MOLARD 1975). In nature, its only known mode of propagation is via gametes. Sigma virus is able to infect oogonial cells and, during oogenesis, the number of viral genomes increases with the cytoplasmic volume (BRUN and DIATTA 1973). As a result, most of the cells of the embryo are infected, especially primordial germ cells. When the embryo is a female, all its progeny are infected. Cytoplasmic transmission is an efficient mode of propagation: a survey of *D. melanogaster* natural populations has shown that 10 to 20% of flies are infected (FLEURIET 1988).

This mode of transmission implies that the viral syntheses is kept within limits, so as not to disturb any critical steps in development, such as gametogenesis or embryogenesis. Although not cytopathogenic, sigma virus is slightly deleterious: a higher sensitivity of infected flies to the winter conditions and a lower viability of eggs laid by infected females were observed

(FLEURIET 1981a,b). Thus, the least deleterious genotypes of sigma virus should be selected in nature. Viral multiplication is partly under host control; five *D. melanogaster* genes, *ref(1)H*, *ref(2)M*, *ref(2)P*, *ref(3)D* and *ref(3)O* were identified from the restrictive effect of at least one of their alleles, which specifically interferes with the multiplication of certain sigma virus strains (GAY 1978). Two of these *refractory* genes are located on chromosome 2: *ref(2)P* in 37EF (NAKAMURA, GAY and CONTAMINE 1986) near the dominant marker *Bristle* (*Bl* 54.8) and *ref(2)M* closely linked to the marker *Jammed* (*J* 41). Alleles restrictive for sigma virus multiplication, *P^p* and *M^m*, found in the Paris and *sp2* strains, respectively, are easily distinguishable: for instance, raising the temperature increases the restrictive effect of *M^m*, but not of *P^p*.

The biological effects of a permissive *P^o* and a restrictive *P^p* allele have been described in detail. In particular, among its numerous effects on viral multiplication, the *P^p* allele reduces the efficiency of the hereditary transmission and the probability of initiating infection. Whereas permissive and restrictive alleles differ at the molecular level by the size of their transcripts (CONTAMINE, PETITJEAN and ASHBURNER 1989), flies carrying either allele are otherwise phenotypically indistinguishable. Loss-of-function alleles have been obtained by mutagenesis. The only phenotype of *ref(2)P^{null}* homozygous flies is male sterility;

homozygous *ref(2)^{P^{null}}* females are viable and fertile (CONTAMINE, PETITJEAN and ASHBURNER 1989). This male sterility is associated with mitochondrial degeneration, which only occurs in the spermatids (DEZÉLÉE *et al.* 1989). The expression of *ref(2)P* has been studied by the antiviral activity of the restrictive allele. First, this expression was shown to be autonomous: when imaginal discs from larvae were serially transplanted into adult hosts, permissive infected discs in restrictive hosts remained infected and vice versa (BERNARD 1968). Second, expression was noticeable during all the developmental stages of the fly, from the ovarian cysts through the larva to the imago (GAY and OZOLINS 1968) and in various tissues such as imaginal discs, thoracic ganglia, ovaries, and cells of the circulatory system (GAY 1978; NAKAMURA 1978). This wide range of expression is very surprising in view of the only known host function of *ref(2)P*: its requirement for male fertility.

The permissive *P^o* allele has been cloned (CONTAMINE, PETITJEAN and ASHBURNER 1989) and sequenced (DEZÉLÉE *et al.* 1989). It encodes a hydrophilic protein of 599 amino acids in which several remarkable motifs have been recognized (amphiphilic helices, zinc-finger, phosphorylation sites, PEST regions), but no homology has been found with any other known protein sequence.

To gain further insight into the function of the *ref(2)P* gene in infected and uninfected flies, we have extended our studies to other restrictive *D. melanogaster* strains. First, a restrictive effect was found to be linked to chromosome 2 of the Nagore strain. In this report, we show that a restrictive allele at the *ref(2)P* locus is responsible for this and that this allele is clearly different from *P^o*, since it exhibits a weaker antiviral effect. Secondly, the transcripts of permissive and restrictive alleles were shown to differ in length, *P^o* transcripts being about 50 nucleotides smaller than those produced by the *P^o* allele (CONTAMINE, PETITJEAN and ASHBURNER 1989). This suggests that some allelic differences may be detected using molecular techniques, such as S1 nuclease analysis. We report here that five classes of alleles, differing at the molecular level, could be defined among the seventeen *D. melanogaster* strains studied. Nucleotide sequence comparison revealed a surprisingly high number of amino acid replacements, as well as insertions or deletions, in the predicted proteins of four haplotypes (including the *P^o* allele). To better estimate *ref(2)P* gene variability, we wished to extend our analysis to another species. This species had to be distant enough from *D. melanogaster* to substantially increase the total number of mutations, but sufficiently related to avoid saturation. Therefore, we chose to clone and sequence the *ref(2)P* gene of *D. erecta*. Sequence com-

parison provided evidence for some selective constraint on the protein.

MATERIALS AND METHODS

***D. melanogaster* strains:** The OM and Paris flies, where the permissive *ref(2)P^o* and the restrictive *ref(2)P^o* alleles respectively were found, were the reference strains. The permissive *ebony* strain was used to construct the standard host which was an heterozygote *OM/ebony* (*OM/E*) (CONTAMINE 1981).

The structural variability of the *ref(2)P* locus was analyzed in different *D. melanogaster* strains. Four of them, *pink ebony*, *sp²*, *Cyd26* and Nagore, were studied in our laboratory because they are of interest for the study of the sigma virus. They carry either a restrictive allele for one of the five known *ref* genes or a suppressor gene of the *ref(2)P^{null}* homozygous male sterility. The strain *Cyd26* (BRITTNACHER and GANETZKY 1983) has a balancer carrying the *Cy* marker heterozygous with *Df(2L)pr-d26*, which is a *ref(2)P⁻* deficiency. Nagore was constructed with its chromosomes 2 originating from the Nagasaki strain (OHANESSIAN-GUILLEMAIN 1963) and chromosomes 1 and 3 from the Oregon strain. Five strains, *ovo⁺*, *ovoD1* (BUSSON *et al.* 1983), Canton-S, Harwich and II2 (BINGHAM, KIDWELL and RUBIN 1982), originated from other laboratories. Biziat, Menetreol and Tübingen were wild flies collected, respectively, in 1968, 1972 and 1988 in the locality indicated by their name, in France for the first and the second, in Germany for the third. These natural populations were polymorphic at the *ref(2)P* locus, as determined by appropriate crosses (FLEURIET 1976). From these populations, homozygous lines either permissive (1) or restrictive (2) were created.

The chromosomes bearing the deficiency *Df(2L)E55* and the *l(2)37Fc^{E42}* and *l(2)37Fb^{E146}* lethals, which arose from the same mutagenesis experiment (WRIGHT, HODGETTS and SHERALD 1976), were derived from a permissive chromosome (NAKAMURA, GAY and CONTAMINE 1986). All these second chromosomes were kept balanced over the *In(2LR)CyO* chromosome. In short, all the permissive second chromosomes from the Oregon, OM and *CyO* strains were named O.

Virus strains: The sigma virus strains used in this study were 23DAa, A3, A3O2 and A3V. Sigma 23DAa was a subclone of the wild-type viral stock 23DA which has been previously described (CONTAMINE 1981). This subclone was obtained from the stabilized progeny of an inoculated female. Sigma A3, A3O2 and A3V were subclones of a wild-type virus strain from the United States; A3 was the original stock (GAY 1978), A3O2 was multiplied twice in Oregon flies infected by inoculation and A3V came from the CO₂ sensitive offspring of a cross between males stabilized with the A3 strain and Oregon females. Sigma A3, A3O2 and A3V multiplication is restricted in the Paris strain (*P⁻* type). The 23DAa clone multiplication in this strain was only partly impaired (*p⁺* type). All the viral clones and stocks were stored at -80°.

Sigma assay methods: In contrast to uninfected flies, those carrying the sigma virus remained paralyzed after a 15 min exposure to pure CO₂ at 12° and subsequently died. The sigma virus was introduced into flies by two different ways: (i) the virus was transmitted by a stabilized female to its progeny. The effect of the host genotype was analyzed by studying the kinetics of appearance of CO₂ sensitivity in the offspring and (ii) a viral suspension (supernatant of homogenized infected flies) was inoculated by abdominal injection. The probabilities of initiating infection were com-

pared with those of a reference host. The ratio of the two probabilities is the ratio of the viral titers assayed in these two hosts; the titers were estimated by the end point dilution method: the average number of infectious units per injected volume (n) was related to the fraction of uninfected flies (P) by the first term of the Poisson equation: $n = -\ln P$ (BRUN and PLUS 1980). This ratio was expressed in logarithmic units. Assay methods and calculations have been previously described and discussed (CONTAMINE 1981; COULON and CONTAMINE 1982). All statistical tests were performed with a confidence coefficient of 0.95.

Introducing the *ref(2)P* loci of *P* and *N* chromosomes within an *OM* context: *P2* and *P3*, *N1* and *N2* chromosomes were constructed by successive rounds of recombination between an *OM* second chromosome and either *P* or *N* chromosomes. Females heterozygous for an *O* chromosome carrying the dominant markers *Jammed* (*J*) and *Bristle* (*Bl*) (LINDSLEY and GRELL 1968) and either *P* or *N* chromosome were crossed with *In(2LR)CyO/Pm* males. The *Bl* marker locus (2-54.8) is closely linked to the *ref(2)P* locus (2-54) (NAKAMURA, GAY and CONTAMINE 1986). Among the progeny, males carrying a *J Bl*⁺ recombinant chromosome were isolated and individually crossed successively with females stabilized for a sigma P⁻ (in order to test for the presence of a *P*^p allele on the recombinant chromosome) and with *OM* females. The *J Bl*⁺ daughters of this last cross were mated with *OM* males; this procedure was repeated three times to allow exchanges between the distal parts of the chromosomes. The *J Bl*⁺ females were then mated with *In(2LR)CyO/J*⁺ *Bl* males (the *J*⁺ *Bl* chromosome was from a recombination between *O* and *J Bl* chromosomes) and from this cross the *J Bl*⁺/*J*⁺ *Bl* daughters were selected and mated with *In(2LR)CyO/Pm* males. Individual recombinant restrictive chromosomes were selected among non-*J* non-*Bl* chromosomes.

DNA cloning and sequencing: All DNA manipulations were done as previously described (CONTAMINE, PETITJEAN and ASHBURNER 1989; DEZÉLÉE *et al.* 1989) or according to established procedures (MANIATIS, FRITSCH and SAMBROOK 1982).

A recombinant phage library was constructed in λ GEM11 (*Bam*HI half arms; Promega) using partially digested *Mbo*I chromosomal DNA fragments from *D. melanogaster* (a strain having chromosomes 1 and 2 of the Paris strain and chromosome 3 of the reference *ebony* strain). This library was successively screened by hybridization with three different ³²P-nick translated probes from *P*^p allele DNA: (i) the *Hind*III-*Acc*I restriction fragment containing the promoter region (from position -584 to +191); (ii) the *Eco*RI restriction fragment including the second and third exons (from +944 to +3294); (iii) a DNA fragment outside the *ref(2)P* structural gene downstream from the 3' end (from +3295 to +4175). Nucleotide positions correspond to the published sequence of *P*^p (DEZÉLÉE *et al.* 1989) (EMBL ACX16993). The first probe revealed numerous plaques giving weak signals, probably due to its binding to AT-rich regions, and intense positive signals which coincided with those detected with the two other probes. No background was observed with these two last probes. Two phages selected from the library were isolated and their inserts were partially mapped. The *Hind*III-*Xho*I restriction fragment which covers the whole structural gene was subcloned into the pEMBL19 vector (*Hind*III and *Sal*I restricted).

Cloning of *ref(2)P* cDNAs has been previously described (DEZÉLÉE *et al.* 1989). These cDNAs originate from a cDNA library of *D. melanogaster* imaginal discs provided by N. BROWN. They were subcloned in the pEMBL vector.

A genomic DNA library of *D. erecta* in the λ EMBL3

vector (BALLY-CUIF *et al.* 1990) was provided by V. PAYANT. This library was screened by hybridization with the *Eco*RI restriction fragment (+944 to +3294) from *P*^p allele DNA. The inserts of positive phages were mapped and a *Sac*I-*Sac*I fragment containing the whole *D. erecta ref(2)P* gene was subcloned in the pUC18 vector in both orientations. Serial unidirectional deletions were created in the insert, using the double-stranded nested deletion kit of Pharmacia.

DNA sequencing was performed with the Amersham M13 sequencing kit or with the Pharmacia T7 polymerase kit, using a set of synthetic oligonucleotides derived from the sequence of the *P*^p allele. The template was either the single-stranded phagemid DNA (PEG precipitated and phenol extracted) or the double-stranded plasmid DNA (obtained by alkaline lysis and isopropanol precipitation). The regions of sequence divergence between cDNAs or *P*^p and *P*^p were run in parallel on the same sequencing gels and on gels containing 7 M urea plus 40% formamide. For the *D. erecta* gene, both strands were sequenced.

RESULTS

Evidence for the existence of another restrictive allele, *P*ⁿ, different from *P*^p at the *ref(2)P* locus of *D. melanogaster*: The Nagore strain exhibited a restrictive phenotype with regard to sigma virus, a phenotype which was associated with chromosome 2 (OHANESSIAN-GUILLEMAIN 1963). To identify the gene(s) implicated, a simple allelism test was not sufficient since the second chromosomes were not isogenic; they could carry different alleles of *ref(2)P*, *ref(2)M* as well as other genes which could be involved in the restrictive phenotype. Accordingly, we have located the gene responsible for the restrictive phenotype in several stages, including the use of a *ref(2)P*⁻ deficiency and the construction of recombinant chromosomes. We have shown that the Nagore strain carries a restrictive allele of *ref(2)P* which expresses a phenotype distinct from that of the Paris allele. In short, the second chromosomes of the Nagore, Paris and Oregon strains will be named *N*, *P* and *O*, respectively. The *O* chromosome carries permissive alleles at the *ref(2)P* and *ref(2)M* loci. The chromosomes 1 and 3 will not be mentioned, since they are always identical in the analyzed and control strains. (i) Flies heterozygous for the second chromosome, *N/P*, behaved like *P/P* or *N/N* homozygotes, when infected via maternal inheritance with a sigma virus susceptible to the *ref(2)P* restrictive allele, *P*^p. When tested with CO₂ just after imaginal emergence, they were resistant (Figure 1). They became progressively CO₂ sensitive over time, but the delay required to obtain 50% sensitive flies depended on the genotype: 10, 14 and 25 days for *N/N*, *N/P* and *P/P* flies respectively. On the contrary, *P/O* or *N/O* heterozygotes were sensitive from imaginal emergence and for the rest of their lives. These results indicated a gradient of restrictive effect depending on the second chromosome: *P/P P/N N/N P/O = N/O*. (ii) In these different genetic contexts, three sigma virus strains were inoculated to

FIGURE 1.—Kinetics of the appearance of CO₂ sensitivity depending on genotype. The appearance of CO₂ sensitivity was studied in the offspring of individual crosses between females heterozygous for the second chromosome *O/P* or *O/N*, stabilized with the sigma virus P^- A3, and heterozygous males *O/P* or *O/N* (the first and third chromosomes of both parents were from the Oregon strain). Independent samples of 40–50 flies were taken after emergence at the time indicated and exposed to CO₂. The percentage of flies which did not recover after CO₂ exposure was determined for each class bearing second chromosomes as follows: *O/N* ($\diamond$); *O/P* (∇); *P/N* ($\blacktriangle$); *N/N* ($\blacksquare$); *P/P* ($\star$).

determine their relative probabilities of initiating infection (Table 1). The ordering of the viruses, on the basis of their susceptibility to the P^b allele, was the same in *N/N*, *P/P* or *N/P* flies. Again, the restrictive effect was less severe with the *N* than with the *P* chromosomes. (iii) Using recessive markers (*nub*, *black*, *purple* and *engrailed*), the restrictive locus was located to the left of the *purple* marker, as is *ref(2)P*. (iv) The *N* and *P* chromosomes were made heterozygous either with the chromosome carrying the *ref(2)P* deficiency *Df(2L)E55* or with the control chromosomes (*l(2)37Fc^{E42}* and *l(2)37Fb^{E146}*) obtained during the mutagenesis experiment from which *Df(2L)E55* arose (WRIGHT, HODGETTS and SHERALD 1976). The same weak decrease in the relative probability of initiating infection of a p^+ virus was observed in both heterozygotes constructed with the control chromosomes (Table 2). This decrease was noticeable in the heterozygotes bearing the deleted chromosome, but less so with *N/Df(2L)E55* (about 10-fold if compared to the control) than with *P/Df(2L)E55* (about 80-fold).

All these results supported the idea that the *N*, like the *P* chromosome, was restrictive at the *ref(2)P* locus. The distinct phenotypes observed with the *N* and the *P* chromosomes resulted either from two different alleles at this locus or from the same allele, the function of which should be modulated by another gene present in different allelic forms on the *N* and *P* chromosomes. A candidate was obviously *ref(2)M*. To eliminate the implication of *ref(2)M* or other genes, the *ref(2)P* region from *N* and *P* chromosomes was introduced by recombination into the second chro-

TABLE 1

Probabilities of infection of flies of different genotypes by various sigma virus strains

Host genotype ^a	Viral strains ^b		
	P^- A302	P^- A3V	P^+ 23DAa
<i>P/O</i>	-1.25 ± 0.25		-0.5 ± 0.3
<i>N/O</i>	-0.7 ± 0.2	-0.2 ± 0.2	-0.3 ± 0.2
<i>P/P</i>	-3.56 ± 0.01	-4.4 ± 0.3	-2.2 ± 0.2
<i>N/N</i>	-3.28 ± 0.15	-4.0 ± 0.1	-1.5 ± 0.2
<i>P/N</i>	-3.28 ± 0.05		-1.3 ± 0.1

Results are expressed as the logarithm of the ratio between the viral infectious units in the studied and the reference host (Oregon strain). Variance takes intra- and inter-experimental variations into account.

^a Only chromosomes 2 are indicated; chromosomes 1 and 3 are from the Oregon strain.

^b P^- and P^+ indicate the type of sigma virus inoculated: σP^- are the most susceptible viruses to the effect of the P^b allele of the Paris strain.

TABLE 2

Location of the Nagore restrictive effect in the *ref(2)P* locus

Reference host ^a	Host genotype ^a	Infection probability ^b
<i>OM/Df(2L)E55</i>	<i>P/Df(2L)E55</i>	-2.2 ± 0.3
	<i>N/Df(2L)E55</i>	-1.4 ± 0.2
<i>OM/l^{E42} or l^{E146}</i>	<i>P/l^{E42} or l^{E146}</i>	-0.3 ± 0.2
	<i>N/l^{E42} or l^{E146}</i>	-0.4 ± 0.2
<i>OM/OM</i>	<i>P2/P2^d</i>	-2.7 ± 0.3
	<i>P3/P3^d</i>	-2.2 ± 0.1
	<i>N1/N1^d</i>	-1.4 ± 0.1
	<i>N2/N2^d</i>	-1.8 ± 0.1

Results are expressed as the logarithm of the ratio between the viral effective units obtained in the studied and reference hosts.

^a Only chromosomes 2 are indicated; chromosomes 1 and 3 were from the standard permissive host in the upper panel and from the *OM* strain in the lower.

^b The sigma virus was P^+ , 23DAa strain.

^c *l^{E42}* and *l^{E146}* are the control chromosomes *l(2)37Fc^{E42}* and *l(2)37Fb^{E146}* respectively, which arose from the same mutagenesis experiment that *Df(2L)E55*.

^d *P2*, *P3* and *N1*, *N2* are recombinant between *OM* and *P* or *N* chromosomes containing the *ref(2)P* region from *P* or *N*, respectively.

mosome of the *OM* strain, which carries a *ref(2)M* permissive allele. Two independent recombinant chromosomes of each type, *N1*, *N2* and *P2*, *P3*, were obtained. The relative probability of initiating infection of a p^+ virus was studied in homozygous flies (Table 2). All four of the recombinant strains were clearly restrictive and again the virus multiplication was more impaired by the P^b than by the P^n allele. These data argued for the existence of a restrictive P^n allele on the Nagore second chromosome, different from the previously described P^b allele.

The structural variability of the *ref(2)P* locus within the *D. melanogaster* species: Besides the biological polymorphism described above, the existence

of a structural polymorphism was indicated by the difference in length of the P^o and P^p transcripts (CONTAMINE, PETITJEAN and ASHBURNER 1989). In order to define more precisely the structural variability at the *ref(2)P* locus, a molecular analysis was undertaken of seventeen *D. melanogaster* strains, permissive or restrictive, from natural populations or from laboratories. The *ref(2)P* mRNAs of these strains were compared with those of the reference *OM* strain using the S1 mapping technique. The fragments resulting from the S1 digestion were separated by electrophoresis and identified using different probes. Their position relative to the 5' and 3' ends of exons was determined in parallel experiments using *Bgl*I restricted genomic DNA, as previously described (DEZÉLÉE *et al.* 1989). The results obtained on alkaline gels with three *D. melanogaster* strains are shown in Figure 2. Several differences were detected: two deletions (D1 and D2) and three divergences (d1, d2, d3). All these differences were located in the coding part of RNAs. The A probe revealed two fragments corresponding to the long and short exons 1 (E1L and E1S) in the *OM* and *pink ebony* strains. In the latter strain, two shorter products were also detectable in low amounts, revealing the existence of the weak divergence d2. In the Paris strain, only two short fragments (pL and pS) were generated. They were due to the systematic cut of both strands in the hybrids at the level of the strong divergence d1, as revealed by both alkaline and neutral gels. With the B probe which revealed the second exon (E2) in the *OM* strain, two fragments ($k2'$ and $k2''$) were detected in the *pink ebony* strain, indicating the presence of the deletion D2; the intensity of the short fragment signal was weak, since the B probe only partially overlapped it. A third fragment present in low amounts was due to the weak divergence d3. In the Paris strain, the deletion D2 and another deletion, D1, led to the appearance of three fragments ($p2'$, $p2''$ and $p2'''$); a fourth fragment in low amounts was due to a divergence which was located at the same place as d3 and which had the same susceptibility to S1 nuclease. With the C probe, the fragments $k2''$ and $p2'''$ were clearly revealed; no sequence variation was observed in the third exon (E3).

The data obtained with all the studied strains are summarized in Table 3. At least another divergence (d4) was detected in one of these strains. According to the observed differences, five allele types, P^o , P^m , P^k , P^n and P^p , can be distinguished. All these alleles were functional in the flies: the homozygous males were fertile even though the third chromosome was replaced by an *OM* chromosome, a genetic background in which P^{null} homozygous males are sterile (P. GAY and D. CONTAMINE, in preparation). Flies of each strain were infected by the A3 sigma virus via a maternal inheritance and their CO₂ sensitivity deter-

FIGURE 2.—Determination of the differences between the P^o , P^p and P^k alleles by S1 nuclease mapping. Poly(A)⁺ RNAs from *OM*, Paris and *pink ebony* strains (3 µg/track) were hybridized to the plasmid DNA containing the *ref(2)P^o* HindIII-XhoI fragment. The hybrids were treated with S1 nuclease and analysed by electrophoresis in nondenaturing or in alkaline agarose gels. After blotting onto nylon membranes, the protected fragments were detected by hybridization with one of the three labeled probes (A, B or C) and autoradiographed. After exposure, the membrane was successively hybridized with the second and third probes. Only the autoradiograms of the alkaline gels are presented (lower panel). E1L, E1S, E2 and E3 correspond to the exons detected in the *OM* strain. The fragments generated with the *pink ebony* and Paris strains were named k and p, respectively. The experiment is schematically represented (upper panel). The restriction sites used to generate the probes A, B and C are indicated: EcoRI (Eco) and BamHI (Bam), as well as the *Bgl*I sites (Bg) used in parallel experiments (see in the text). In the middle section, the structure of the *ref(2)P* gene is presented: the three exons are represented as shaded boxes and the introns are connecting lines. The positions of the differences are indicated (the d4 divergence, only detected in the *sp²* strain, was found to be located in the same place as the weak d3 divergence, but different since both strands were systematically cleaved by S1 nuclease). The fragments obtained with the transcripts from the *OM*, Paris and *pink ebony* strains are schematically represented below. The dashed lines in the upper and lower panels represent minor digestion products which were detectable on alkaline and neutral gels.

mined. This allowed us to determine whether these alleles were also functional with regard to the sigma virus. They were all functional, since the phenotypes of homozygotes and heterozygotes with P^p did not

TABLE 3

Differences between the *ref(2)P* transcripts of the *OM* strain and various *D. melanogaster* strains as revealed by S1 nuclease protection experiments

Strains	Differences observed						Allele type
	d1	d2 ^a	D1	d3 ^b	d4	D2	
Permissive							
<i>OM</i>	—	—	—	—	—	—	<i>P^o</i>
<i>ebony</i>	—	—	—	—	—	—	
Canton-S	—	—	—	—	—	—	
<i>ovo D1</i>	—	—	—	—	—	—	
$\pi 2$	—	—	—	—	—	—	
Harwich	—	—	—	—	—	—	
Menetreol 1	—	—	—	—	—	—	
Biziat 1	—	—	—	—	—	—	
Tübingen 1	—	—	—	—	—	—	
<i>sp2</i>	—	—	—		+	—	<i>P^m</i>
<i>pink ebony</i>	—	+	—	+	—	+	<i>P^k</i>
<i>ovo^r</i>	—	+	—	+	—	+	
Restrictive							
Nagore	+		—	+	—	+	<i>Pⁿ</i>
<i>Cy d26</i>	+		—	+	—	+	
Menetreol 2	+		—	+	—	+	
Biziat 2	+		—	+	—	+	
Tübingen 2	+		—	+	—	+	
Paris	+		+	+	—	+	<i>P^p</i>

^a The d2 divergence was undetectable in the presence of d1.

^b The d3 divergence was undetectable in the presence of d4.

correspond to those previously described for strains carrying loss-of-function alleles: respectively, CO₂ sensitive and CO₂ resistant (CONTAMINE, PETITJEAN and ASHBURNER 1989). In fact, the two genotypes were either CO₂ sensitive (permissive alleles: *P^o*, *P^m* and *P^k*) or CO₂ resistant (restrictive alleles: *Pⁿ* and *P^p*). Considering the alleles present in wild flies, *P^o* and *Pⁿ* would be the most frequent types in nature, at least in Europe.

Sequence comparison of four *ref(2)P* haplotypes:

The above results indicate that the *ref(2)P* gene tolerates numerous variations without affecting *Drosophila*. One of these divergences (d1) could be correlated to the restrictive phenotype, but this phenotype could also be the result of point mutations undetectable by S1 nuclease mapping experiments. To further analyze the differences between restrictive and permissive alleles, we sequenced several haplotypes. First, a genomic library in λ GEM11 was prepared from *D. melanogaster* adults of the restrictive Paris strain. Clones which contained the *P^p* allele were selected. The *HindIII-XhoI* restriction fragment which covered the whole structural gene was subcloned into pEMBL19 and sequenced. Secondly, *ref(2)P* cDNA clones were isolated from a cDNA library prepared from *D. melanogaster* imaginal discs (BROWN and KAFATOS 1988).

The sequences of two cDNAs which differed in length were established.

These new sequence determinations permitted us to correct the previously published sequence which is a *P^o* type allele, named *P^{o1}* in this paper (DEZÉLÉE *et al.* 1989); a G must be inserted at position 1774 and a C must be deleted at position 1877. The frameshift changes 35 amino acids in the deduced protein sequence and modifies the net charge of the large basic region between amino acids 250 and 300 which is, in fact, mainly acidic. But the other features of the deduced protein, in particular the region including putative zinc fingers, are not altered.

The sequence variations between *P^{o1}*, *P^p*, cDNA1 and cDNA2 are listed in Table 4. Four large variations between the permissive *P^{o1}* allele and the restrictive *P^p* allele were observed. (i) In the first exon, two codons, CAG-AAT (glutamine-asparagine), were replaced by GGA (glycine). This corresponded to the d1 divergence detected by S1 nuclease analysis in the Paris strain (Table 3). A single event cannot explain such a complex change which may involve deletions and substitutions. (ii) In exon 2, a 21-bp deletion (between positions 1556 and 1578) corresponded to the D1 deletion. (iii) In exon 2, another deletion of 24 bp was observed between positions 2196 and 2221; it corresponded to the D2 deletion. (iv) In the first intron of the restrictive allele, a 7-bp insertion was the perfect repeat of a motif present 23 bp upstream. Imperfect inverted repeats of this motif were present in this intron, suggesting that they could be involved in its secondary structure. All the large sequence variations which affected the exons did not disrupt the open reading frame, but shortened the restrictive transcripts and the protein. Numerous point mutations were located all along the sequence: one of these affected one of the three perfect repeats present in the promoter region of the *P^{o1}* allele (GTAAAAATA was changed to ATAAAAATA). Four mutations affected non coding sequences (three in the first intron and one in the trailer). Six concerned the coding region, five of them resulted in an amino acid change. The point mutation T $\rightarrow$ A in 2076 could be the d3 divergence (Table 3), since this mismatch was located in an AT-rich context (AAT(A/T)TTT) and would be susceptible to the S1 nuclease.

The two *ref(2)P* cDNA clones differed in length by about 120 bp. Both corresponded to copies of long transcripts: cDNA 2 started at position +137 and was an incomplete copy, whereas cDNA 1 started at position +25 which corresponded exactly to the initiation point of one of the long transcripts (DEZÉLÉE *et al.* 1989). Unexpectedly, these two cDNAs differed in sequence. cDNA 1 was similar to a transcript of the *P^{o1}* permissive allele, except for four point mutations; three of them were located in the coding region and

TABLE 4
Sequence comparison of four *ref(2)P* haplotypes of *D. melanogaster*

	Nucleotide position	Permissive allele <i>P^{o1}</i>	Restrictive allele <i>P^o</i>	cDNA clone 1 <i>P^{o2}</i>	cDNA clone 2 <i>Pⁿ</i>	
	-101	g	a	—	—	
Exon 1	145	t	t	t	c	
	453-458	CAGAAT	GGA	CAGAAT	GGA	d1
		Q N	G	Q N	G	
	465	ATA	GTA	ATA	GTA	
		I	V	I	V	
Intron 1	499	CAG	CTG	CAG	CTG	
		Q	L	Q	L	
	606-607	ca	c<tggttc>a insertion			
	652	a	t			
Exon 2	897	t	c			
	1045	g	a			
	1557-1577	21bp LGSRSGR	>21bp< deletion in the deletion	21bp LGRRSGR	21bp LGRRSGR	D1
	1564	AGC		CGC	CGC	
		S		R	R	
	1693	ACT	TCT	ACT	TCT	
		T	S	T	S	
	1993	AGT	CGT	AGT	CGT	
		S	R	S	R	
	2076	ATT	ATA	ATT	ATA	d3
		I	I	I	I	
	2197-2220	24bp ANQSNVPS	>24bp< deletion	24bp ANQSNVPS	>24bp< deletion	D2
	2628	GAC	GAG	GAG	GAG	
Intron 2		D	E	E	E	
	2781	ACC	ACC	AGC	ACC	
		T	T	S	T	
	2986	c	c	a	c	
Exon 3	3004	g	a	g	g	

Nucleotide positions correspond to the published sequence of *P^{o1}* (DEZÉLÉE *et al.* 1989). *P^o* was sequenced from -260 to +3110; *P^{o2}* from +25 to +3104; *Pⁿ* from +137 to 3104.

resulted in the change of a serine to an arginine, an aspartic acid to a glutamic acid and a threonine to a serine. cDNA1 may correspond to a permissive allele, so we named it *P^{o1}*. cDNA 2 was identical to a restrictive *P^o* transcript with one exception: the D1 deletion was not observed and the CGC codon present in this region in *P^{o2}* was conserved. As cDNA2 exhibited the d1, d3 and D2 sequence variations, it belonged to the *Pⁿ* allele class. These results indicate that the flies used to prepare the cDNA library were polymorphic at the *ref(2)P* locus. This is not surprising since permissive and restrictive alleles are nearly equally represented in nature, indicating that no selective pressure is exerted to exclude one of them.

Nucleotide sequence comparison of these four haplotypes indicates: (i) a mutation frequency in the coding region (4.4×10^{-3}) as high as in the noncoding region excluding introns (4.8×10^{-3}); (ii) among the 8 point mutations observed in the coding region, only

one (at position 2076) resulted in a synonymous codon. This contrasts with what has been observed in other genes. For example, the analysis of the nucleotide polymorphism at the *Adh* locus showed that only 2 out of 42 substitutions in the coding region resulted in amino acid changes (McDONALD and KREITMAN 1991).

Comparison between the *ref(2)P* genes of *D. melanogaster* and *D. erecta*: These results raise the question of the absence of selective constraint on the *ref(2)P* protein. This question could be only addressed by additional sequence, thus the analysis was extended to another species. According to BODMER and ASHBURNER (1984) who have studied the *Adh* locus of various *Drosophila* species, the *yakuba* complex, which includes the *D. oreana* and *D. erecta* species, appeared appropriately distant from *D. melanogaster*. Thus, the *ref(2)P* gene of *D. erecta* was cloned and sequenced. The *P^{o1}* and the *D. erecta* sequences were aligned

(Figure 3) and gaps were inserted in order to maximize the similarity of the two sequences. The overall organization of the *D. melanogaster* gene seemed to be maintained in *D. erecta*, as deduced from the conservation of the consensus sites for splicing, polyadenylation and translational initiation and termination. Of the 23 deletions or insertions introduced into the two sequences, only 4 (which affected one or several triplets) were located in the coding region, but none of them corresponded to the d1, D1 and D2 differences observed between the *D. melanogaster* haplotypes. In *P^{o1}*, the first AUG was in frame with a 1797 nucleotide long open reading frame (ORF). In *D. erecta*, in addition to this AUG which was in frame with a 1782 nucleotide long ORF, another AUG was present 211 bp upstream and could initiate a 32-amino acid long polypeptide; however, this AUG was flanked by a sequence which does not fit with the start signal consensus sequence for *Drosophila* (CAVENER 1987).

Using the *D. erecta* *ref(2)P* sequence as an outgroup, a relationship tree between the four sequenced haplotypes of *D. melanogaster* was drawn (Figure 4). The construction of this tree was based upon the principle of parsimony. Three of the point mutations (1564, 2628, 2781) and the D1 deletion allowed to define the upper branches of the tree, since each of them was present in only one haplotype. For the other nucleotide variations, comparison with the *D. erecta* sequence allowed the direction of the changes to be deduced and thus, to locate on which branch of the tree these mutations have occurred in the presumed ancestral *D. melanogaster* sequence, *P^a*. The mutations at 1693, 1993 and 2076 were unambiguously located on the permissive branch, while those at 465 and 499 and the d1 and D2 differences on the right branch from which restrictive alleles arose. The S1 nuclease experiments (Table 3) revealed sequence variations which behaved like d3 and D2 and the absence of the d1 divergence in the *P^k* permissive allele type; this indicated that D2 cannot be responsible for the restrictive phenotype. The two point mutations at 465 and 499 were assigned to the restrictive branch, but their precise location cannot be determined, since this *P^k* allele has not been sequenced. Nevertheless the d1 and/or one or both of these point mutations could be implicated in the restrictive phenotype, if the ancestral sequence corresponded to a permissive allele. On the contrary, if it represented a restrictive allele, the mutation at 1693 and/or at 1993 on the left branch of the tree could be responsible for the permissive phenotype (the mutation at 2076 corresponding to d3 is not taken into account since it is conservative).

In the absence of selective constraint, one would expect that all nucleotide mutations are equally likely, whether a substitution causes an amino acid replacement or is silent. For the *D. erecta* and the *P^a* *D.*

melanogaster sequences, the frequency of effective silent sites (*i.e.*, the frequency of silent mutations among the nine possible substitutions at each codon all along the sequence) was estimated according to NEI and GOJOBORI (1986). The same values were obtained: 0.23. From these values, the replacement substitution frequency expected would be 3.3 times greater than the frequency of silent substitutions: $(1 - 0.23)/0.23$. Figure 5 shows the amino acid sequence alignment of the *ref(2)P* deduced proteins of *D. erecta* and *P^a* (*P^a* was chosen instead of *P^{o1}* to overcome the *D. melanogaster* intraspecies polymorphism). Silent mutations affected 84 of the 590 common codons; among the 81 amino acid changes observed, 80 were produced by 82 substitutions. Thus the number of replacements was not three times greater than the number of silent mutations, demonstrating that the *ref(2)P* protein was constrained by natural selection. However the amino acid replacement frequency was unusually high, 0.135, as compared to the values obtained for other genes with similar genetic distances: 0.036 for the amylase of *D. melanogaster* and *D. erecta* (BALLY-CUIF *et al.* 1990) and 0.039 for the *Adh* of *D. melanogaster* and *D. orena* (BODMER and ASHBURNER 1984). Moreover, the variability of the *ref(2)P* protein was underestimated, since the insertions/deletions of amino acids were not taken into account. It is noticeable that, in spite of this high variability, the *D. erecta* gene is functional, as indicated by preliminary transformation experiments: it is able to restore the fertility of *D. melanogaster* *ref(2)P^{null}* homozygous males. The numerous mutations in the *D. erecta* protein did not affect the main features described for the *ref(2)P* protein (amphiphilic helices, zinc finger) as shown in Figure 5.

DISCUSSION

Each *D. melanogaster* *refractory* gene has been identified because two allelic forms coexist in nature, one permissive and the other restrictive for the multiplication of the rhabdovirus sigma (GAY 1978). Until now, the *ref(2)P* gene was defined by the existence of the permissive *P^o* and the restrictive *P^p* alleles. A first indication that other alleles exist in nature is given by the discovery of the *Pⁿ* allele of the Nagore strain which is clearly different from the *P^p* allele, although restrictive. The *Pⁿ* allele recognized the same viral strains as *P^p*, but its antiviral action was clearly weaker. As inferred from the structural molecular study of the *ref(2)P* transcripts within numerous *D. melanogaster* strains, the *Pⁿ* allele type would be the most frequent restrictive type in nature, at least in Europe.

At the molecular level, numerous large differences were observed between *P^o* and the alleles present in various *D. melanogaster* strains either from nature or from laboratories. Nevertheless, the overall gene

FIGURE 4.—Relationship tree between the four sequenced haplotypes of *D. melanogaster*. This hand-constructed tree was based on the assumption that: (i) a mutation present in only one haplotype has occurred recently; (ii) a particular sequence present in two *D. melanogaster* haplotypes and not in the *D. erecta* sequence has been introduced into the unknown ancestral sequence, P^a , by mutation; (iii) the mutations present on the same branch are not ordered. The d3 divergence between the mRNAs of the restrictive and *OM* strains detected by S1 nuclease mapping can be attributed to the point mutation $A \rightarrow T$ (2076).

structure was conserved, since none of these sequence variations was due to alternative splicing. Two deletions were detectable in the exon 2 of P^b . The first was present only in this strain, the second in all the restrictive strains, but also in two permissive ones. Thus, shorter transcripts were not characteristic of the restrictive alleles. A sequence divergence in the

exon 1 was only found in each restrictive strain and may thus be involved in the restrictive phenotype. Variability was observed among the permissive strains, since three allele types were distinguished: P^o , P^m and P^k .

Cloning and sequencing of four haplotypes were performed. The two deletions present in the P^b allele did not disrupt the open reading frame. The first one involved a region which contained a potential phosphorylation site (or amidation site if the protein is processed) in the permissive protein. Since the permissive protein is present as different phosphorylated forms (F. WYERS, P. DRU, B. SIMONET and D. CONTAMINE, in preparation) and if this particular site is phosphorylated *in vivo*, the P^b protein could be less phosphorylated than the permissive P^o protein and its function partly modulated in this way. The second deletion eliminated an eight amino acid motif, ANQSNVPS, from a repetitive region of the protein. The strong divergence, which was only present in the exon 1 of all the restrictive alleles, involved a 3-bp deletion and nucleotide substitutions. It was located near the N-terminal of the protein; it altered the secondary structure predicted according to GARNIER, OSGUTHORPE and ROBSON (1978) at the level of a turn and excluded the existence of the amphiphilic A2

FIGURE 5.—Sequence alignment of the deduced proteins of *D. erecta* and *D. melanogaster* presumed ancestor, P^a . The sequence in single letter code starts with the first methionine of the P^a deduced protein and extends to amino acid 599 (middle line); the stretch between brackets corresponds to the 35 amino acids which differ from the previously published sequence (DEZÉLÉE *et al.* 1989). For the *D. erecta* sequence (lower line), the dashed lines indicate that the amino acids are the same as in P^a . The four regions, where gaps are introduced in order to align the nucleotides, are indicated by arrows. Stars indicate the amino acids which have diverged in the four sequenced *D. melanogaster* haplotypes; the d1, D1 and D2 sequence variations are also shown. Characters in bold correspond to the features of the P^o deduced protein which have been previously described (DEZÉLÉE *et al.* 1989): (i) A1 and A2 are the two amphiphilic helices predicted according to the method of EISENBERG *et al.* (1984); the changes of glutamine to histidine (residue 11), of glycine to valine (residue 12) and of glutamine to leucine (residue 28) in *D. erecta* protein enhanced the amphiphilic character of both helices; (ii) F corresponds to the potential finger region from residues 127–167; the cysteine and histidine residues which have been tentatively aligned with the TFIIIA consensus sequence were all conserved; (iii) R1, R2, R3 correspond to the repeated sequences SANQs(x)(x)P; all the serine residues in these motifs were conserved; (iv) P2 and P3 correspond to the PEST regions which are conserved in *D. erecta*; they display scores >5 calculated according to RODGERS, WELLS and RECHSTEINER (1986).

helix which was predicted in the *P^o* protein (DEZÉLÉE *et al.* 1989) according to the method of EISENBERG *et al.* (1984). This potential structure appeared uncertain, since the permissive and restrictive proteins were equally functional in *D. melanogaster*. In addition to these important sequence variations, numerous point mutations were detected. Transformation of germ line cells with *in vitro* mutagenized *ref(2)P* genes are in progress to identify the sequence variation(s) which led to the restrictive phenotype.

The most striking features of the *ref(2)P* gene were the high level of amino acid replacements (7R), a low silent polymorphism (1S) and many in-frame insertions or deletions (3 in/del), without any phenotype. A single silent mutation in the four *D. melanogaster* alleles sequenced is a low value as compared to what was described for other genes. For example, among the 11 alleles of the *Adh* gene sequenced, 13 synonymous mutations were observed for 192 effective silent sites (KREITMAN 1983). The frequency of synonymous mutations, $0.006 [13/(192 \times 11)]$, is 10 times higher than the frequency obtained for the *ref(2)P* gene, $0.0006 [1/(405 \times 4)]$. Actually, a low value of silent polymorphism is what might be expected for a gene located near the proximal heterochromatin of 2L, region of reduced recombination which undergoes periodic sweeps (BEGUN and AQUADRO 1992; BERRY, AJIOKA and KREITMAN 1991). The variability (7R and 3 in/del) within the *D. melanogaster* species appears relatively high, but the number of alleles sequenced was low. A between-species comparison could allow us to obtain a better estimate of this variability, which might indicate a locus either under relatively loose constraint or undergoing rapid protein evolution.

Cloning and sequencing the *ref(2)P* gene of *D. erecta* led to the following results: (i) the overall gene organization was conserved; (ii) the insertions or deletions observed in the coding region within *D. melanogaster* and *D. erecta* did not disrupt the open reading frame; (iii) none of the sequence variations introduced a translational stop signal; (iv) the mutation rate was significantly higher in the effective silent sites (0.24) than in the non silent positions (0.06) when estimated according to NEI and GOJOBORI (1986). All these data clearly demonstrate that the selection exerts a constraint to maintain the *ref(2)P* protein functional. As observed in preliminary transformation experiments, the *ref(2)P* gene of *D. erecta* was able to restore the fertility of the *ref(2)P^{null}* homozygote males of *D. melanogaster*. Whether the implication of the gene in male fertility corresponds to the only function of the protein is not sure, since, in *D. melanogaster*, (i) the *ref(2)P* gene was transcribed at a high level: the *ref(2)P* mRNAs corresponded to about 10^{-3} of the total poly(A)⁺ RNAs, even in the female ovaries (D. CONTAMINE, unpublished results); (ii) it exhibited an ubiqui-

itous expression which appeared to largely exceed its apparent requirement for male fertility; (iii) a dominant suppressor of the male sterility existed in one third of the *D. melanogaster* strains studied (P. GAY and D. CONTAMINE, in preparation), suggesting that other suppressors of other putative *ref(2)P* phenotypes might exist and mask still unknown function(s).

The variability of the *ref(2)P* gene within *D. melanogaster* species (7R, 3 in/del and 1S) is significantly higher than between species (82R, 4 in/del and 84S); the probability that an equal or higher difference might be observed is $P = 0.0109$, when estimated according to FISHER (1958). This indicates that selection must be acting to preserve the polymorphism within the *D. melanogaster* species. This implies that advantageous mutations have been recently pulled up in frequency in the population. Selection by the virus itself is an attractive hypothesis, since the event(s) responsible for the restrictive phenotype are most likely advantageous for the infected flies, allowing them to overcome the deleterious effect of the virus. The selection of these adaptive mutations could have led to the concomitant fixation of some mutations on the *ref(2)P* gene by hitchhiking. The other mutations, especially in the permissive alleles, may have been responsible for a restrictive effect against earlier sigma virus strains and, then, may be remnants of the same selection process than that of the mutation responsible for the actual restrictive phenotype. This assumption is based on the observation that, in an infected fly, the frequency of sigma virus spontaneous mutants, adapted to grow in the presence of a restrictive allele, could reach 10^{-3} . Thus, with rapid virus evolution, a restrictive allele may become equally rapidly a permissive allele. The validity of this hypothesis, selection by the virus, could be tested by the study of the variability at the *ref(2)P* locus within another *Drosophila* species which has never been infected by sigma virus in nature. One might expect a lower variability than in the *D. melanogaster* species; otherwise, this would imply that the polymorphism is required for the gene function and even for multiple functions in the fly.

We thank ANNIE FLEURIET for providing the natural permissive and restrictive lines of *D. melanogaster*. We thank also NICOLAS BROWN and VÉRONIQUE PAYANT for providing, respectively, the *D. melanogaster* cDNA library and the *D. erecta* DNA library. We are grateful to FLORENCE LAFAY for critical reading of the manuscript and to MARTIN KREITMAN for especially helpful discussions. P.D. was a recipient of a fellowship of the Ministère de la Recherche et de la Technologie. This work was supported by the Centre National de la Recherche Scientifique through the UPR A2431. and by a grant 861009 from the Institut National de la Santé et de la Recherche Médicale.

LITERATURE CITED

- BALLY-CUIF, L., V. PAYANT, S. ABUKASHAWA, B. F. BENKEL and D. A. HICKEY, 1990 Molecular cloning and partial sequence

- characterization of the duplicated amylase genes from *Drosophila erecta*. *Genet. Sel. Evol.* **22**: 57–64.
- BEGUN, D. J., and C. F. AQUADRO, 1992 Levels of naturally occurring DNA polymorphism correlate with recombination rates in *D. melanogaster*. *Nature* **356**: 519–520.
- BERNARD, J., 1968 Contribution à l'étude du virus héréditaire de la *Drosophila*, sigma. *Exp. Cell. Res.* **50**: 117–126.
- BERRY, A. J., A. JIOKA, J. W. and M. KREITMAN, 1991 Lack of polymorphism on the *Drosophila* fourth chromosome resulting from selection. *Genetics* **129**: 1111–1117.
- BINGHAM, P. M., M. G. KIDWELL and G. M. RUBIN, 1982 The molecular basis of the P-M dysgenesis: the role of the P element; a P-strain-specific transposon family. *Cell* **29**: 995–1004.
- BODMER, M., and M. ASHBURNER, 1984 Conservation and change in the DNA sequences coding for alcohol dehydrogenase in sibling species of *Drosophila*. *Nature* **309**: 425–430.
- BRITTNACHER, J. G., and B. GANETZKY, 1983 On the components of segregation distortion in *Drosophila melanogaster*. II. Deletion mapping and dosage analysis of the SD locus. *Genetics* **103**: 659–673.
- BROWN, N. H., and F. C. KAFATOS, 1988 Functional cDNA Libraries from *Drosophila* Embryos. *J. Mol. Biol.* **203**: 425–437.
- BRUN, G., and F. DIATTA, 1973 Guérison germinale de femelles de *Drosophila* stabilisées pour le virus sigma, à la suite d'un traitement par la température ou par l'éthylméthane-sulfonate. *Ann. Microbiol.* **124A**: 421–438.
- BRUN, G., and N. PLUS, 1980 The viruses of *Drosophila*, pp. 625–702 in *The Genetics and Biology of Drosophila*, Vol. 2d, edited by M. ASHBURNER and T. R. F. WRIGHT. Academic Press, New York.
- BUSSON, D., M. GANS, K. KOMITOPOULOU and M. MASSON, 1983 Genetic analysis of three dominant female-sterile mutations located on the X chromosome of *Drosophila melanogaster*. *Genetics* **105**: 309–325.
- CAVENER, D. R., 1987 Comparison of the consensus sequence flanking translational start sites in *Drosophila* and vertebrates. *Nucleic Acids Res.* **15**: 1353–1361.
- CONTAMINE, D., 1981 Role of the *Drosophila* genome in sigma virus multiplication. I. Role of the *ref(2)P*: selection of host-adapted mutants at the non-permissive allele *ref(2)P^o*. *Virology* **114**: 474–488.
- CONTAMINE, D., A. M. PETITJEAN and M. ASHBURNER, 1989 Genetic resistance to viral infection : the molecular cloning of a *Drosophila* gene that restricts infection by the rhabdovirus sigma. *Genetics* **123**: 525–533.
- COULON, P., and D. CONTAMINE, 1982 Role of the *Drosophila* genome in sigma virus multiplication. II. Host spectrum variants among the *haP* mutants. *Virology* **123**: 381–392.
- DEZÉLÉE, S., F. BRAS, D. CONTAMINE, M. LOPEZ-FERBER, D. SEGRETAIN and D. TENINGES, 1989 Molecular analysis of *ref(2)P*, a *Drosophila* gene implicated in sigma rhabdovirus multiplication and necessary for male fertility. *EMBO J.* **8**: 3437–3446.
- EISENBERG, D., E. SCHWARZ, M. KOMAROMY and R. WALL, 1984 Analysis of membrane and surface protein sequences with the hydrophobic moment plot. *J. Mol. Biol.* **179**: 125–142.
- FISHER, R. A., 1958 TEST OF GOODNESS OF FIT, INDEPENDENCE AND HOMOGENEITY. THE TREATMENT OF 2 × 2 TABLES, PP. 96–97 IN *Statistical Methods for Research Worker*, Ed. 13, revised. Oliver and Boyd Ltd., Edinburgh, U.K.
- FLEURIET, A., 1976 Presence of the hereditary rhabdovirus sigma and polymorphism for a gene for resistance to this virus in natural populations of *Drosophila melanogaster*. *Evolution* **30**: 735–739.
- FLEURIET, A., 1981a Effect of overwintering on the frequency of flies infected by the rhabdovirus sigma in experimental populations of *Drosophila melanogaster*. *Arch. Virol.* **69**: 253–260.
- FLEURIET, A., 1981b Comparison of various physiological traits in flies (*Drosophila melanogaster*) of wild origin, infected or uninfected by the hereditary rhabdovirus sigma. *Arch. Virol.* **69**: 261–272.
- FLEURIET, A., 1988 Maintenance of a hereditary virus, the sigma virus, in populations of its host *Drosophila melanogaster*. *Evol. Biol.* **23**: 1–30.
- GARNIER, J., D. J. OSGUTHORPE and B. ROBSON, 1978 Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins. *J. Mol. Biol.* **120**: 97–120.
- GAY, P., 1978 Les gènes de la *Drosophila* qui interviennent dans la multiplication du virus sigma. *Mol. Gen. Genet.* **159**: 269–283.
- GAY, P., and C. OZOLINS, 1968 Transmission héréditaire du virus sigma par des drosophiles porteuses du gène "réfractaire." I. Etude d'une souche virale P⁺ g⁺. *Ann. Inst. Pasteur* **114**: 29–48.
- KREITMAN, M., 1983 Nucleotide polymorphism at the *Alcohol dehydrogenase* locus of *Drosophila melanogaster*. *Nature* **304**: 412–417.
- LINDSLEY, D. L., and E. H. GRELL, 1968 *Genetic Variations of Drosophila melanogaster*. Carnegie Inst. Wash. Publ. 627.
- MANIATIS, T., E. F. FRITSCH and J. SAMBROOK, 1982 *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- MCDONALD, J. H., and M. KREITMAN, 1991 Adaptive protein evolution at the *Adh* locus in *Drosophila*. *Nature* **351**: 652–654.
- NAKAMURA, N., 1978 Influence du dosage de l'allèle non permissif du gène *ref(2)P* de la drosophile sur les souches sensibles du virus sigma. *Mol. Gen. Genet.* **159**: 285–292.
- NAKAMURA, N., P. GAY and D. CONTAMINE, 1986 Etude du locus *ref(2)P* de *Drosophila melanogaster*. I. Localisation cytogénétique de *ref(2)P*. *Biol. Cell.* **56**: 227–238.
- NEI, M., and T. GOJOBORI, 1986 Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Mol. Biol. Evol.* **3**: 418–426.
- OHANESSIAN, A., and G. ECHALIER, 1967 Multiplication of *Drosophila* hereditary virus (sigma virus) in *Drosophila* embryonic cells cultivated *in vitro*. *Nature* **213**: 1049–1050.
- OHANESSIAN-GUILLEMAIN, A., 1963 Etude de facteurs génétiques contrôlant les relations du virus sigma et de la drosophile son hôte. *Ann. Genet.* **5**: 1–64.
- RICHARD-MOLARD, C., 1975 Isolement de lignées cellulaires de *Drosophila melanogaster* de différents génotypes et étude de la multiplication de deux variants du rhabdovirus sigma dans ces lignées. *Arch. Virol.* **47**: 139–146.
- RODGERS, S., R. WELLS and M. RECHSTEINER, 1986 Amino acid sequences common to rapidly degraded proteins: the PEST hypothesis. *Science* **234**: 364–368.
- WRIGHT, T. R. F., R. B. HODGETTS and A. F. SHERALD, 1976 The genetics of dopa decarboxylase in *Drosophila melanogaster*. I. Isolation and characterization of deficiencies of the structural locus and the alpha methyl dopa hypersensitive locus. *Genetics* **84**: 267–285.

Communicating editor: M. T. FULLER