Nucleotide Polymorphism at the RpII215 Gene in Drosophila subobscura: Weak Selection on Synonymous Mutations

Ana Llopart and Montserrat Aguadé

Departament de Genètica, Facultat de Biologia, Universitat de Barcelona, 08071 Barcelona, Spain

Manuscript received November 23, 1998
Accepted for publication March 29, 2000

ABSTRACT

Nucleotide variation in an 8.1-kb fragment encompassing the RpII215 gene, which encodes the largest subunit of the RNA polymerase II complex, is analyzed in a sample of 11 chromosomes from a natural population of Drosophila subobscura. No amino acid polymorphism was detected among the 157 segregating sites. The observed numbers of preferred and unpreferred derived synonymous mutations can be explained by neutral mutational processes. In contrast, preferred mutations segregate at significantly higher frequency than unpreferred mutations, suggesting the action of natural selection. The polymorphism to divergence ratio is different for preferred and unpreferred changes, in agreement with their beneficial and deleterious effects on fitness, respectively. Preferred and unpreferred codons are nonrandomly distributed in the RpII215 gene, leading to a heterogeneous distribution of polymorphic to fixed synonymous differences across this coding region. This intragenic variation of the polymorphism/divergence ratio cannot be explained by different patterns of gene expression, mutation, or recombination rates, and therefore it indicates that selection coefficients for synonymous mutations can vary extensively across a coding region. The application of nucleotide composition stationarity tests in coding and flanking noncoding regions, assumed to behave neutrally, allows the detection of the action of natural selection when stationarity holds in the noncoding region.

SYNONYMOUS mutations do not imply changes in the amino acid sequences of proteins and hence they have been considered to behave in the neighborhood of neutrality (Kimura 1983), according to a mutation-selection-drift model (Li 1987; Bulmer 1991). The G + C content at synonymous sites in Drosophila melanogaster varies extensively among genes (Shields et al. 1988) and this variation cannot be explained by only mutational biases (Shields et al. 1988; Kliman and Hey 1994), suggesting the action of natural selection. Accordingly, in Drosophila, the departure from the neutral expectations in the synonymous codon usage (codon bias) is negatively correlated with the rate of synonymous substitution (Sharp and Li 1989). The overall degree of codon bias is associated with the preferential usage of a set of synonymous codons (preferred codons). These preferred codons in Drosophila are more frequent in genes with a high level of expression, estimated by expressed sequence tag abundance (Duret and Mouchiroud 1999), and limited evidence suggests that they are related to iso-accepting tRNA availability (Moriyama and Powell 1997). All these previous results are in agreement with those obtained in enterobacteria and yeast where at the translational level there is extensive evidence for natural selection action shaping synonymous codon usage (reviewed in Li 1997).

Two different approaches have been proposed to contrast whether natural selection actually discriminates among synonymous mutations and to uncover their effects on fitness: (i) the examination of the frequency distribution of two distinct categories of mutations with respect to their different effects on fitness, in our case preferred and unpreferred changes (Hartl et al. 1994; Akashi and Schaeffer 1997) and (ii) the comparison of the number of polymorphisms segregating in a natural population with the number of fixed differences between two closely related species (McDonald and Kreitman 1991) for both preferred and unpreferred changes (Ballard and Kreitman 1994; Akashi 1995). The ratio of polymorphic to fixed changes can be tested even when the assumption of mutation-drift equilibrium does not hold. Population genetics studies revealed that in D. simulans unpreferred changes show a significantly higher ratio of polymorphism to divergence than preferred changes (Akashi 1995), consistent with their deleterious effect on fitness compared to preferred mutations. In the closely related species D. melanogaster, preferred and unpreferred changes exhibit a similar ratio of polymorphism to divergence, suggesting a reduced effect of selection (Akashi 1995, 1996), congruent with a smaller effective population size, $N_e$ (Aquadro et al. 1988; Moriyama and Powell 1996). In agreement with the previous observations, preferred mutations segregate at a
higher frequency than unpreferred mutations in D. simulans (Akashi and Schaeffer 1997; Kiman 1999) and in the only region studied in D. pseudoobscura (Adh-Adhr; Akashi and Schaeffer 1997).

The comparative study of the RpII215 gene in several species of Drosophila (Llopart and Aguadé 1999) revealed not only a very low rate of amino acid replacements, congruent with the critical role of this gene in the transcription process, but also a high rate of synonymous substitutions, consistent with the length (5667 bp) of the coding region (Comeron and Aguadé 1996; Comeron et al. 1999). In that interspecific study the comparison of synonymous substitution rates between the insular species D. guanche and the wider-distributed D. subsobcra showed a faster accumulation of synonymous changes in the D. guanche lineage. This lineage also showed a significant excess of unpreferred changes. Both results suggest a nearly neutral behavior of synonymous mutations at the RpII215 gene of Drosophila and they can be attributed to a reduction of the effectiveness of selection in the D. guanche lineage due to a smaller $N_e$ (Ohta and Kimura 1971; Ohta 1972).

Herein our aim is to explore the effects of natural weak selection on inter- and intraspecific synonymous variation at the RpII215 gene in D. subsobcra. The high rate of synonymous substitutions for the RpII215 gene together with its large number of codons (1889 amino acids in D. subsobcra) provides a good opportunity to study synonymous mutations, with the advantage of excluding possible sources of heterogeneity such as different levels of gene expression and rates of mutation and recombination. Moreover, the long coding region of the RpII215 gene will allow for testing the hypothesis of a homogeneous distribution for the effectiveness of selection on synonymous mutations across the gene. In addition, the studied region is affected in D. subsobcra by a polymorphic inversion (A2) leading to two different chromosomal arrangements, [A0] and [A2], and hence we discuss the effect of population subdivision on synonymous preferred and unpreferred mutations.

MATERIALS AND METHODS

Drosophila lines: Thirty-seven isofemale lines of D. subsobcra were established upon collection in Collcerola (Barcelona, Spain) in November 1995. For each line, one male was selected for study and crossed with virgin females from a laboratory stock homozygous for the [A0] chromosomal arrangement. Polytene chromosomes of one female larva of the offspring were examined to establish the arrangement of the A chromosome carried by the selected male, which was frozen for further analysis.

DNA preparation and sequencing: For the sequencing study, 11 males were randomly chosen within each chromosomal arrangement, but keeping the same proportions of the [A0] and [A2] arrangements as in the natural population. Genomic DNA from the single adult male was extracted using a standard procedure (Ashburner 1989) with minor modifications. A region of 8.1 kb was divided in five overlapping fragments of 2.0, 1.9, 1.6, 2.1, and 2.3 kb that were PCR amplified using primers designed on the published D. subsobcra sequence (Llopart and Aguadé 1999). The nucleotide sequence of each template was obtained for both strands with internal primers and using the cycle sequencing method and an ABI377 automatic sequencer (Perkin Elmer, Norwalk, CT). Sequences were aligned with the Clustal W (v.1.4) program (Thompson et al. 1994).

Preferred and unpreferred synonymous codons: The composition of synonymous codons at monomorphic sites of the RpII215 gene was determined with the GCG package (v.7.3; Devereux et al. 1984). Preferred and unpreferred codons were classified according to Akashi and Schaeffer (1997) using the information available in D. pseudoobscura. The hypothesis of random distribution of preferred and unpreferred codons across the coding region of the RpII215 gene was contrasted by calculating the probability associated with the observed number of runs. Because the numbers of preferred and unpreferred codons are large, a normal approximation was used (runs test for dichotomized data; Sokal and Rohlf 1995). Alternatively, a random permutation test was also performed to calculate the probability of having a number of runs equal to or smaller than that observed in the data (J. M. Comeron, personal communication). The possible departure from the neutral frequency spectrum for preferred and unpreferred changes was tested by Tajima's test (Tajima 1989) on the basis of the $D$ statistic. The significance of the observed $D$ value was achieved by computer simulations (J. Rozas, personal communication) using Hudson’s coalescence algorithm with recombination (Hudson 1990).

Genetic flow between chromosomal arrangements: The level of genetic flow between chromosomal arrangements, in our case [A0] and [A2], can be inferred from the ratio of shared to total polymorphisms (S/T). Because the detectable amount of genetic flow between different chromosomal arrangements depends on the sample size, we performed Monte Carlo simulations (10,000 replicates) to test whether this amount of genetic transfer, as estimated by the S/T fraction, was the same in the RpII215 region as in the rp49 region (Rozas and Aguadé 1994) after correcting for different sample sizes (J. M. Comeron, personal communication). For each replicate, 6 and 5 rp49 sequences were randomly chosen from the [O0] and [O2] classes, respectively, and the S/T ratio was calculated. We considered the [O2] class equivalent to the [A2] class because they both were the least polymorphic classes.

Molecular population analyses were done using the version 2.98 of the DnaSP program (Rozas and Rozas 1997). The number of synonymous substitutions per site ($K_s$) was estimated using version 5.3 of the K-estimator program (Comeron 1999). Sequences newly reported in this study are deposited in EMBL, GenBank, and DDBJ database libraries under accession nos. AF272643–AF272653.

RESULTS

A region of 8.1 kb (7824 bp, excluding sites with alignment gaps) was sequenced in a random sample of 11 A chromosomes isolated from a natural population of D. subsobcra. The sequenced region encompasses the entire RpII215 gene (7025 bp) and also its 5' and 3' untranslated ends. Twenty-six insertion/deletion events that range from 1 to 27 bp were detected in the sample, all of them located in noncoding regions. Two variable microsatellites were found: a tetranucleotide (TCCG)
repeated from 8 to 12 times in the first intron and a trinucleotide (GGA) repeated from 4 to 8 times in the 3' untranslated end. Both microsatellites showed point mutations in the repeated unit. The observed 157 nucleotide polymorphisms in the entire RpII215 region are presented in Figure 1. No amino acid replacement was detected among the 57 segregating sites observed in this species is the ancestral state. A G-test of independence revealed an equivalent ratio of preferred to un-coding region (5667 bp long in D. subobscura). Sequence variation in the RpII215 region is summarized in Table 1. Analysis of synonymous codons: Our first goal was to determine whether natural selection acting on preferred and unpreferred synonymous mutations at the RpII215 gene of D. subobscura was strong enough to affect their presence as polymorphic variants in the population. The number of preferred and unpreferred codons (p and u, respectively) was determined for monomorphic sites and for the ancestral variant of polymorphic sites (Table 2). The ancestral nucleotide at each polymorphic site was determined by comparison with the homologous sequence of the closely related species D. madeirensis considering that the variant presented in Figure 1. No amino acid replacement was detected among the 57 segregating sites observed in the coding region (5667 bp long in D. subobscura). Sequence variation in the RpII215 region is summarized in Table 1.
matrix among the four nucleotides inferred from the polymorphisms in Rpl215 noncoding regions ($G = 1.71$, $P = 0.19$).

The ancestral polymorphic and monomorphic codons reflect the particular codon usage bias at the Rpl215 gene. If we compare the numbers of preferred and unpreferred codons (ancestral polymorphic and monomorphic) of the Rpl215 gene, a reflection of its codon usage bias, with the numbers of preferred and unpreferred derived mutations, we essentially test whether the codon usage bias and the mutational tendencies observed in this gene are compatible. The ratio of preferred to unpreferred newly arisen mutations can be explained by mutational processes, but it differs significantly from that observed in the ancestral polymorphic and monomorphic codons ($P = 0.0014$ and $P < 1 \times 10^{-5}$, respectively). This result suggests different probabilities of fixation for preferred and unpreferred mutations, determined by unequal segregating frequencies in the population. Consistently, preferred mutations at the Rpl215 gene segregate at significantly higher frequency in the population than unpreferred mutations (average frequency of 0.338 and 0.195, respectively; Mann-Whitney $U$ test, $z = -2.416$, $P = 0.016$).

Although the presence of two chromosomal arrangements ([A$_S$] and [A$_J$]) in the population might have an impact on the analysis of the frequency distribution of polymorphisms in the global population (see discussion), it should have affected both preferred and unpreferred mutations in the same way and therefore it cannot explain our results. We have analyzed the frequency distribution of preferred and unpreferred mutations within each chromosomal class. For the [A$_S$] class, which is considered the ancestral chromosomal arrangement, the 9 preferred mutations segregate at significantly higher frequency (Mann-Whitney $z = -3.43$, $P < 0.0001$) than the 31 unpreferred mutations. In contrast, no significant result is obtained for the derived chromosomal arrangement, [A$_J$] (Mann-Whitney $z = -0.32$, $P = 0.71$), in which only 3 out of the 23 observed mutations are preferred variants (see discussion). When unpreferred mutations derived from unpreferred codons are excluded from the analyses, the difference in the segregating frequencies between preferred and unpreferred mutations remains unchanged ($P = 0.01$, $P = 0.001$, and $P = 0.84$ for the random, [A$_S$] and [A$_J$] samples, respectively).

The possible unequal effects on fitness of preferred and unpreferred changes were contrasted using the modification of the McDonald-Kreitman test (McDon-

### Table 1

Sequence variation at the Rpl215 region of D. subobscura

<table>
<thead>
<tr>
<th>Positions (bp)</th>
<th>[A$_S$]</th>
<th>[A$_J$]</th>
<th>Pop</th>
</tr>
</thead>
<tbody>
<tr>
<td>Noncoding region</td>
<td>0.014</td>
<td>0.011</td>
<td>0.013</td>
</tr>
<tr>
<td>Coding region (synonymous)</td>
<td>0.012</td>
<td>0.0073</td>
<td>0.012</td>
</tr>
<tr>
<td>Total</td>
<td>0.0059</td>
<td>0.0043</td>
<td>0.0057</td>
</tr>
</tbody>
</table>

Nucleotide diversity ($\pi$) and number of segregating sites ($S$) were estimated separately for the [A$_S$] and [A$_J$] chromosomal classes (with sample sizes 6 and 5, respectively) and for the random sample [population (pop), $n = 11$].

### Table 2

Preferred and unpreferred codons at polymorphic and monomorphic sites of the Rpl215 coding region of D. subobscura

<table>
<thead>
<tr>
<th>Polymorphic</th>
<th>Monomorphic</th>
<th>Polymorphic</th>
<th>Ancestral</th>
<th>Derived observed</th>
<th>Derived expected</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preferred</td>
<td>873</td>
<td>27</td>
<td>11</td>
<td>17.1</td>
<td></td>
</tr>
<tr>
<td>Unpreferred</td>
<td>870</td>
<td>30</td>
<td>46</td>
<td>39.9</td>
<td></td>
</tr>
</tbody>
</table>

$G^a = 0.16$, $P = 0.69$  
$G^b = 10.19$, $P = 0.0014$  
$G^c = 1.71$, $P = 0.19$

Williams’ correction has been applied to the G-tests of independence (Sokal and Rohlf 1995).

$^a$ G-test of independence for preferred and unpreferred monomorphic and ancestral codons.

$^b$ G-test of independence for preferred and unpreferred ancestral and derived codons.

$^c$ Goodness-of-fit test for preferred and unpreferred derived mutations. The expected numbers of preferred and unpreferred mutations were calculated according to the mutational matrix based on polymorphism data in the noncoding regions.
ald and Kreitman 1991; Ballard and Kreitman 1994; Akashi 1995). In this sense, polymorphic and fixed differences were classified as preferred (p → u) and unpreferred (u → p) changes. The number of fixed differences was determined by comparison to the homologous sequence of D. maderensis and the direction of the changes inferred using the D. guanche sequence as an outgroup. The McDonald and Kreitman test for preferred and unpreferred changes yielded a significant departure from neutral expectations (G = 4.11, P = 0.043) (Table 3). This departure is consistent with an excess of polymorphic unpreferred changes within species, or with a deficit of fixed unpreferred changes between species compared to the number of preferred changes, or both. The different effect on fitness of preferred and unpreferred changes is more conspicuous if we combine data from the inter- and intraspecific comparisons together with information about the frequency distribution of polymorphisms in the population, using the approach suggested by Templeton (1996) (Table 3). Preferred and unpreferred changes were equally distributed among fixed differences and mutations segregating more than once (nonsingleton) in the population. In contrast, most (18 out of 21) of the singleton (unique) mutations fell in the category of unpreferred changes (Figure 2).

Distribution of synonymous codons along the RpII215 gene: Comparative analysis of the RpII215 gene between different species of Drosophila had revealed a heterogeneous distribution of synonymous divergence (Ks) across the coding region (Llopart and Aguadé 1999). Accordingly, a runs test for dichotomized data (see materials and methods; Sokal and Rohlf 1995) showed that the unpreferred codons (fixed or ancestral) were nonrandomly distributed along the RpII215 gene (P = 0.044). Equivalent results were obtained by a random permutation test. On the basis of the previous results and considering that preferred and unpreferred mutations segregate at different frequencies in the population, we tested the expected heterogeneous distribution of the ratio of synonymous polymorphism to divergence. McDonald’s (1996) runs test yielded no significant heterogeneity when the mere 19 synonymous fixed differences between D. subobscura and its sibling species D. maderensis were considered with the 57 polymorphisms. In contrast, when the more distantly related species D. guanche was used (71 fixed differences), the ratio of synonymous polymorphic to fixed differences was heterogeneously distributed across the coding region (Table 4).

DISCUSSION

In D. subobscura, our analyses in the RpII215 gene show that both preferred and unpreferred codons are equally polymorphic and that the newly arisen synonymous mutations are those expected by mutational process on ancestral codons. Nevertheless, the proportion of preferred and unpreferred mutations differs from that observed in the ancestral sequence. Also, a modification of the McDonald and Kreitman test reveals a significantly different ratio of polymorphism to divergence for preferred and unpreferred changes, disclosing a differential action of natural selection on preferred and unpreferred codons. Results from the Templeton test (Templeton 1996) suggest that unpreferred changes are “young” and have not achieved intermediate frequencies in natural populations, in agreement with their deleterious effects on fitness. In contrast, preferred changes produce translationally superior codons, consistent with their adaptive nature; and their average prevalence time in the populations would be longer. As revealed by the Mann-Whitney nonparametric test, preferred mutations in the RpII215 gene of D. subobscura segregate in the population at a different frequency than unpreferred mutations, confirming their different effects on fitness. Tajima’s D statistic (Tajima 1989) for preferred and unpreferred changes reveals that the frequency spectrum of each type of change exhibits opposite tendencies, presenting a negative and a positive value for unpreferred and preferred changes, respectively. Results from Tajima’s test with recombination show a significant excess of unpreferred changes present at low frequency in the sample, while no deviation from neutral expectations was detected for preferred changes (Table 3). The small number of preferred changes in the sample (11 in the 5667-bp-long coding region) could reduce the power of Tajima’s test (Simonsen et al. 1995; Akashi 1999).

We conclude, then, that in D. subobscura weak selection modulates the frequency of variants at synonymous polymorphic sites of the RpII215 gene. The intensity of selection is very small and therefore does not have an effect on the expected number of segregating sites. This allows us to detect young unpreferred mutations segregating in populations at low frequencies. It is, however, strong enough to ensure that some preferred mutations will achieve fixation faster than unpreferred mutations.

We have already shown that in D. subobscura the effect of natural selection on synonymous mutations at the RpII215 gene is different for preferred and unpreferred changes. The McDonald (1996) runs test was performed using only the coding region of the RpII215 gene and the significant results obtained denote that synonymous polymorphism and divergence are differentially distributed across the coding region (Figure 3). This significant result from the McDonald runs test suggests that selection coefficients of synonymous mutations vary across the RpII215 coding region as a consequence of the heterogeneous distribution of preferred and unpreferred codons. Clusters of preferred codons produce a high number of unpreferred mutations segregating in the population, thus generating a high ratio of polymorphism to divergence. This observation is con-
TABLE 3

Neutrality tests on preferred and unpreferred changes at the RpII215 coding region

<table>
<thead>
<tr>
<th></th>
<th>McDonald-Kreitman&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Templeton&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Tajima&lt;sup&gt;c&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Fixed Pol</td>
<td>&quot;Old&quot;</td>
<td>&quot;Young&quot;</td>
</tr>
<tr>
<td>u → p</td>
<td>8 12</td>
<td>16 3</td>
<td>0.24</td>
</tr>
<tr>
<td>p → u</td>
<td>5 26</td>
<td>14 18</td>
<td>-1.39</td>
</tr>
</tbody>
</table>

|         | G = 4.11 | G = 8.39 | P = 0.042 | P = 0.0038 |

Williams’ correction has been applied to the G-tests of independence (Sokal and Rohlf 1995).
<sup>a</sup>G-test of independence for fixed and polymorphic (pol) preferred and unpreferred changes (McDonald and Kreitman 1991).
<sup>b</sup>G-test of independence for preferred and unpreferred "old" and "young" changes (Templeton 1996).
<sup>c</sup>Tajima’s (1989) D statistic for preferred and unpreferred changes.
<sup>d</sup>Probability values from one-tailed tests achieved by computer simulation with 5000 iterations using Hudson’s coalescence algorithm with recombination (Hudson 1990). R = 3N, where r is the recombination rate per generation between the most distant sites (Hudson 1987), and Ne is the effective population size.

consistent with the observed heterogeneous distribution of synonymous divergence and the detected negative correlation between this synonymous divergence and codon bias across the gene (Llopart and Aguadé 1999). In the RpII215 gene, the distribution of preferred and unpreferred codons across the coding region can then reflect the subtle equilibrium between selection enhancing translational accuracy and selection enhancing translational elongation rate.

The direction of each synonymous mutation was inferred using a parsimony approach. Misclassifications of changes due to multiple hits at a site appear to be unlikely because of the low divergence among the three D. subobscura cluster species: synonymous divergence per site (K<sub>J</sub>) of 0.021 and 0.058 for the D. subobscura-D. madeirensis and for the D. subobscura-D. guanche comparisons, respectively (Llopart and Aguadé 1999). In the McDonald and Kreitman test, the lack of polymorphism data in the D. madáirensis species would result in a misclassification of some fixed differences. This situation would make our approach conservative because unpreferred mutations are more abundant than preferred mutations; i.e., the number of unpreferred fixed differences would have been overestimated.

The population under study is subdivided in two smaller subpopulations, [A<sub>S</sub>] and [A<sub>T</sub>], genetically differentiated as revealed by the permutation test (K<sub>s</sub> = 5.25, P = 0.005) proposed by Hudson et al. (1992). Nevertheless, there are evidences of genetic flow, by double crossover or gene conversion, between these two chromosomal classes because we detected 23 polymorphisms segregating for the same pair of nucleotides (shared polymorphisms) and no fixed nucleotide differences. The amount of genetic transfer between chromosomal arrangements, estimated by the S/T fraction, at the RpII215 region (S/T = 0.146) is larger (P = 0.0369; see materials and methods) than that observed at the extensively studied rp49 region (S/T = 0.132; Rozas and Aguadé 1994). This result is in agreement with their different locations with respect to the inverted region: the RpII215 region centered in the A<sub>S</sub> inversion and the rp49 located near a breakpoint of a complex polymorphic inversion. Note, however, that the fact that the RpII215 and rp49 genes are located in the sex and O chromosomes, respectively, could bias this comparison in the observed direction because sex chromosomes are more frequent in females, where recombination occurs. Putative gene conversion in males in Drosophila, on the other hand, would make the test conservative.

This complex situation of population subdivision would not explain, however, the difference in the frequency of preferred and unpreferred mutations in the population revealed by the Mann-Whitney test. Equiva-
Synonymous Mutations in *D. subobscura*

Figure 3. Sliding window plot of the ratio of synonymous polymorphism to divergence (Pol/Div) and of synonymous divergence (*K*<sub>s</sub>) across the *RpII215* coding region. (A) Pol/Div. The structure of the gene (four exons) is depicted below the graph.

In contrast, the significance of the observed Tajima's *D* values for preferred and unpreferred changes is clearly affected by population substructure, though no simple prediction can be made. The observed Tajima's *D* value for noncoding regions in the *RpII215* gene (5' untranslated ends and introns) is negative (*D* = -0.87) and therefore it would be consistent with a population subdivided in two subpopulations with different *N*<sub>e</sub> and a small migration rate between them (Tajima 1993).

Another possible source of bias in our analyses is departure from stationarity in base composition, which could be associated with the whole *D. subobscura* genome or only with the region surrounding the *RpII215* gene. Testing for stationarity of *G* + *C* content under neutrality is equivalent to testing if the number of mutations *G* → *A* or *G* → *T* and *C* → *A* or *C* → *T* is equal (Eyre-Walker 1994, 1999). In noncoding regions, we find 31 *G* → *A* and 32 *G* → *T* derived mutations at the polymorphic level, validating nucleotide composition stationarity. In coding regions, however, we find an excess of *G* → *A* compared to *G* → *T* derived mutations.

Table 4: Runs test on polymorphic and fixed synonymous differences of the *RpII215* coding region

<table>
<thead>
<tr>
<th>Sample</th>
<th>rt</th>
<th>Polymorphic</th>
<th>Fixed</th>
<th>No. of runs</th>
<th>P(R = 180)&lt;sup&gt;b&lt;/sup&gt;</th>
<th>P(R = 32)</th>
<th>P(R = 16)</th>
<th>P(R = 8)</th>
<th>P(R = 4)</th>
<th>P(R = 2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population</td>
<td>11</td>
<td>57</td>
<td>71</td>
<td>51</td>
<td>0.033</td>
<td>0.036</td>
<td>0.045</td>
<td>0.046</td>
<td>0.026</td>
<td>0.014</td>
</tr>
<tr>
<td>[A₂]</td>
<td>6</td>
<td>40</td>
<td>71</td>
<td>37</td>
<td>0.009</td>
<td>0.012</td>
<td>0.016</td>
<td>0.011</td>
<td>0.006</td>
<td>0.009</td>
</tr>
<tr>
<td>[A₃]</td>
<td>5</td>
<td>23</td>
<td>76</td>
<td>29</td>
<td>0.068</td>
<td>0.082</td>
<td>0.095</td>
<td>0.086</td>
<td>0.048</td>
<td>0.052</td>
</tr>
</tbody>
</table>

* Sample size.

<sup>b</sup> Probability values calculated with 1000 iterations (McDonald 1996). R = 3*N*<sub>r</sub>, where *r* is the recombination rate per generation between the most distant sites (Hudson 1987) and *N*<sub>e</sub> is the effective population size. The estimated R value for the *RpII215* coding region is 180 (Hudson 1987).
We are grateful to Josep M. Comeron for computer simulations, helpful suggestions, and comments on the manuscript, and J. Rozas for coalescence simulations and sharing the version 2.98 of the DnaSP program. Thanks to Manyou Long for useful comments on the manuscript. We thank Serveis Científico-Òtècnicos of the Universitat de Barcelona for automatic sequencing facilities. A.L.I. was a predoctoral fellow from Ministerio de Educación y Ciencia, Spain, during this study. This work was supported by grant PB94-0923 from Dirección General de Investigación Científica y Técnica, Ministerio de Educación y Ciencia, Spain, and by grant 1997 SGR-59 from Comissió Interdepartamental de Recerca i Innovació Tecnològica, Generalitat de Catalunya, to M.A.

**LITERATURE CITED**


Akashi, H., 1999 Inferring the fitness effects of DNA mutations from polymorphism and divergence data: statistical power to detect directional selection under stationarity and free recombination. Genetics 151: 221–238.


