Evolution, 52(6), 1998, pp. 1705–1712

APPEARANCE AND SWEEP OF A GENE DUPLICATION: ADAPTIVE RESPONSE AND POTENTIAL FOR NEW FUNCTIONS IN THE MOSQUITO *CULEX PIPIENS*

THOMAS LENORMAND,[1,2] THOMAS GUILLEMAUD,[1] DENIS BOURGUET,[1,3] AND MICHEL RAYMOND[1]
[1]*Laboratoire Génétique et Environnement, Institut des Sciences de l'Evolution (UMR 5554), Université Montpellier II, 34095 Montpellier Cedex 5, France*
[2]*E-mail: lenorman@isem.univ-montp2.fr*

Abstract.—Evolution of a new gene function is a fundamental process of adaptation. Gene duplication followed by divergence due to relaxed selection on redundant copies has been viewed as the predominant mechanism involved in this process. At a macroevolutionary scale, evidence for this scenario came from the analysis of sequences of genes families. However, even if several genetic models have described the different potential microevolutionary scenario for a new function to evolve, little is really known about the initial evolutionary dynamics of such processes. We analyze such early dynamics in natural populations of the mosquito *Culex pipiens* polymorphic for a duplication at *Ace.1*, a locus involved in insecticide resistance. The date of occurrence and the selective advantages of the duplication were estimated using frequency data. We propose a scenario where the spread of a duplication is driven, from the very beginning, by selection due to insecticide treatment.

Key words.—Acetylcholinesterase, clines, *Culex pipiens*, fitness cost, gene duplication, insecticide resistance, resistance gene.

Received May 8, 1998. Accepted August 10, 1998.

Adaptation can be seen as a cascade of genetic changes that consist of either an improvement of an existing function or the creation of a new function. These two pathways involve different mechanisms. In the case of improvement of an existing function, two different processes have been distinguished (Fisher 1928; Haldane 1932) that depend on whether they involve epistasis or not, selection of epistatic modifiers or allele replacement, respectively. In the case of evolution of new gene function, gene duplication has been considered as the main mechanism (Haldane 1932; Ohno 1970; Kimura and Ohta 1974; Ohta 1989), although exon shuffling or alternative splicing are potentially relevant (Smith et al. 1989; Crameri et al. 1998). A duplication can invade a population by drift, and a new function can be selected for after the divergence of two identical copies has arisen during a period of relaxed selection on one of the redundant copies (Ohta 1987, 1988; Walsh 1995). Invasion and maintenance of a duplication can also be driven by natural selection: it can be directly selected to mask deleterious mutations (Clark 1994), for the sake of producing more of the same molecule (such as GST [Wang et al. 1991], esterase [Mouchès et al. 1986], or metallothionein [Lange et al. 1990]) or two different molecules (gene sharing or overdominance, Spofford 1969; Hugues 1994). However, only the latter situation corresponds directly to creation of a new function.

Evolution by gene duplication is not straightforward, however, because multiple copies of the same gene can disrupt gene dosage. In such a case, dosage compensation or various regulatory mechanisms (such as tissue or ontogenic specific expression) may evolve as well (e.g., Batterham et al. 1984) or may preexist (e.g., prior selection for induction/repression in response to variable environments—bacterial operons).

Even if the two processes of improvement have been clearly identified in cases of resistance to insecticides, antibiotics or viruses (Clarke and McKenzie 1987; Lenski 1988a,b; Cohan et al. 1994; Guillemaud et al. 1998), processes leading to the creation of a new function have only been inferred from a posteriori observations on sequence data (Ohta 1991, 1994; Long and Langley 1993; Hugues 1994; Syvanen et al. 1996), that is, once the process of evolution is completed. This paper illustrates the occurrence and the sweep of a gene duplication in the mosquito *Culex pipiens* as an adaptive response to insecticide selection.

Two main mechanisms conferring resistance to organophosphate (OP) insecticides have been characterized in the mosquito *C. pipiens*. One corresponds to overproduction of nonspecific esterases that detoxify OPs. Such overproduction often results from gene amplification (Mouchès et al. 1986; Guillemaud et al. 1997). A second mechanism corresponds to modifications of acetylcholinesterase (AChE), the OP's target (Fournier and Mutero 1994). This enzyme has a crucial function in cholinergic synapses of the central nervous system, where it hydrolyzes the acetylcholine neurotransmitter. Due to intensive insecticide use against *C. pipiens*, less sensitive AChE to OPs have been found in numerous strains and natural populations from Europe, Africa, and Central and South America (Raymond et al. 1986; Bourguet et al. 1996a, 1997a). Among diptera, the mosquito *C. pipiens* presents an unusual situation: it possesses two distinct AChE enzymes, AChE1 and AChE2. Only AChE1, which is coded by the *Ace.1* locus, is involved in insecticide toxicology (Bourguet et al. 1996d). In addition, a duplication of the *Ace.1* locus, $Ace.1^{RS}$, which involves a copy of the susceptible ($Ace.1^{S}$) and resistance allele ($Ace.1^{R}$) has recently been found in strains and natural populations from the Caribbean (Bourguet et al. 1996c) and from France (Lenormand, unpubl. data).

In this paper, we address the following questions. What is the frequency pattern of the duplication in natural populations; is it possible to date its first occurrence in Southern France? Can the selective advantage of the duplication be identified and quantified? Finally, how might this duplicated gene diverge to a new functional gene and what mechanisms can lead to its emergence?

[3] Present address: Station de recherche de lutte Biologique, INRA La Minière, 78285 Guyancourt Cedex, France.

TABLE 1. Nomenclature.

Genotype	Coding rules	
	Genotype	Phenotype
$Ace.1^{R}Ace.1^{R}$	(RR)	[RR]
$Ace.1^{R}Ace.1^{S}$	(RS)	[RS]
$Ace.1^{RS}Ace.1^{S}$	(RSS)	[RS]
$Ace.1^{RS}Ace.1^{R}$	(RSR)	[RS]
$Ace.1^{RS}Ace.1^{RS}$	(RSRS)	[RS]
$Ace.1^{S}Ace.1^{S}$	(SS)	[SS]

MATERIALS AND METHODS

Identification of Resistance Genes

In Southern France, two different alleles have been identified at the *Ace.1* locus: $Ace.1^{R}$, which codes for an insensitive AChE1 (or AChE1R) and $Ace.1^{S}$, which codes for a sensitive AChE1 (or AChE1S) (Raymond et al. 1986; Bourguet et al. 1996d). Duplication of the *Ace.1* locus involves closely linked (i.e., less than 1% of recombination, Lenormand and Raymond, unpubl. ms.) functional copies of both the susceptible ($Ace.1^{S}$) and the resistance allele ($Ace.1^{R}$). Natural populations are therefore polymorphic for three haplotypes, $Ace.1^{R}$, $Ace.1^{S}$, and $Ace.1^{RS}$. For simplicity, the recombination between the two copies of $Ace.1^{RS}$ haplotype will be neglected in the computations. The presence of AChE1R was detected using a biochemical test on individual mosquitoes (Bourguet et al. 1996b). This test allows us to discriminate individuals producing only the susceptible ([SS] phenotype), only the resistant ([RR] phenotype) or both types of AChE1 ([RS] phenotype, see Table 1).

Field Sampling

Transect.—Populations were collected along a 50-km south-north transect across the treated (the first 20 km) and untreated areas already studied by Chevillon et al. (1995), Guillemaud et al. (1998), and Lenormand et al. (1998). Along this transect, pupae from 10 breeding sites and overwintering females from 10 caves were collected on 5 July 1995 and during November 1995, respectively. Pupae were reared until emergence and all adults were stored at −20°C. Each mosquito was subsequently analyzed at the *Ace.1* locus.

Maurin Population.—Sixty-nine mated females (referred to as the Maurin population) were collected in April 1996 around a breeding site in Maurin, south of Montpellier, France. These females were isolated and blood-fed in the laboratory, and their progeny reared. Each female and a fraction of their offspring were stored at −20°C before determining their genotype at the *Ace.1* locus.

Populations from the Treated Area.—Different datasets (from 1984 to 1997) of populations sampled within the treated area were used: August 1984 (EID, unpubl. data); June 1986 (Magnin 1986); June 1991 (Chevillon et al. 1995); April 1993, July 1995 (Lenormand et al. 1998); winter 1995–1996, May 1996, and winter 1996–1997. Winter samples were obtained between November and February.

Segregation in the Maurin Population

The biochemical test described by Bourguet et al. (1996b) does not allow discrimination of individuals that carry at least one $Ace.1^{RS}$ haplotype (RSS), (RSR), and (RSRS) from individuals that are standard (RS) heterozygotes. Thus, the frequency of the $Ace.1^{RS}$ haplotype cannot be directly determined. However, the occurrence of this haplotype is expected to create a phenotypic excess of "apparent heterozygotes" in natural populations. Thus, one may compute the duplication's frequency in a population by assuming that the deviation from Hardy-Weinberg proportions is only due to the presence of the duplication. However, this estimation is biased if the population is not at Hardy-Weinberg equilibrium. This bias may be avoided by estimating the frequency of each genotype independently using the phenotypes of the females and their offspring. The probability of observing the phenotype of a female (Y) and its offspring phenotypes ($\mathbf{X}$) assuming both a single father and the same genotypic frequencies in males and females in the population, is:

$$\begin{aligned}\Pr(\mathbf{X} = \mathbf{x} \text{ and } Y = y) &= \sum_{f=1}^{6}\sum_{m=1}^{6} \Pr(\mathbf{X} = \mathbf{x} \mid M = m, F = f) \\ &\quad \times \Pr(Y = y \mid F = f)\Pr(M = m)\Pr(F = f), \quad (1)\end{aligned}$$

where F and M stands for the genotype of the female and the male, respectively. The summations are over the six possible genotypes (see Table 1). $\Pr(F = f)$ is the frequency of the genotype f. $\Pr(Y = y|F = f)$ equals one if the observed phenotype of the female y corresponds to the genotype f and equals zero otherwise.

$$\Pr(\mathbf{X} = \mathbf{x} \mid M = m, F = f) = \frac{\left(\sum_{i=1}^{3} x_i\right)!}{\prod_{i=1}^{3} (x_i)!}\prod_{i=1}^{3} p_i(f, m)^{x_i}, \quad (2)$$

where x_i are the elements of $\mathbf{x}$, that is, the number of [SS], [RS] and [RR] individuals in the progeny, $p_i(f, m)$ is the Mendelian proportion of [SS], [RS], and [RR] individuals expected in the progeny of a female of genotype f mated to a male of genotype m. For instance, if the genotype of both the male and the female is (RSS), then $p_1 = 1/4$, $p_2 = 3/4$, and $p_3 = 0$.

The maximum-likelihood estimates of genotypic frequencies in the Maurin population were computed by maximization of the sum of the log-likelihood over the 69 observed families,

$$\sum_{i=1}^{69} \ln[\Pr(\mathbf{X} = \mathbf{x}_i \text{ and } Y = y_i)], \quad (3)$$

using a Metropolis algorithm adapted from N. H. Barton (Szymura and Barton 1986). Two hypotheses were tested using likelihood-ratio tests: (1) whether the frequency of the $Ace.1^{RS}$ haplotype is different from zero, and (2) whether the genotypic frequencies are in Hardy-Weinberg proportions.

Frequency Pattern in Natural Populations

To know whether clines of *Ace.1*^{RS} and *Ace.1*^{R} haplotypes occur along the transect, data were fit to descriptive clinal models. Haplotypic frequency clines were fit according to a scaled negative exponential:

$$\begin{cases} f_R(x) = h_1 e^{-a_R \times x^2} \\ f_{RS}(x) = h_2(1 - h_1)e^{-a_{RS} \times x^2} \\ f_S(x) = 1 - f_R(x) - f_{RS}(x), \end{cases} \quad (4)$$

where a_R, a_{RS}, h_1, and h_2 are parameters describing the clines and $f_R(x)$, $f_{RS}(x)$, and $f_S(x)$ describe the frequencies of *Ace.1*^{R}, *Ace.1*^{RS}, and *Ace.1*^{S} at distance x to the coast. Phenotypic distributions in population i were computed as

$$\begin{cases} f_{[RR]i} = f_R(x_i)^2 + F_i f_R(x_i)(1 - f_R(x_i)) \\ f_{[SS]i} = f_S(x_i)^2 + F_i f_S(x_i)(1 - f_S(x_i)) \\ f_{[RR]i} = 1 - f_{[RR]i} - f_{[SS]i}, \end{cases} \quad (5)$$

where Fi measures the departure from Hardy-Weinberg proportions in population i (F_i varies between -1 and 1). This complete model, A, was simplified in two different ways. First, Hardy-Weinberg proportions ($F_i = 0$) were assumed in model B_1. This model was further simplified by setting $a_R = 0$, $a_{RS} = 0$, or $a_R = a_{RS}$ in models B_2, B_3, and B_4, respectively. The complete model was also simplified by assuming the absence of the duplication ($h_2 = 0$) in model C_1 and further simplified by assuming Hardy-Weinberg proportions ($F_i = 0$) in model C_2. Comparisons between models B_1–B_4 allow us to detect frequency gradients for *Ace.1*^{R} and *Ace.1*^{RS} haplotypes. Comparisons between the A and B models allow us to determine whether clinal variation of *Ace.1*^{RS} can explain departure from Hardy-Weinberg proportions in the different populations. Finally, comparisons between the B and C models indicate whether the data support the presence of the duplication at detectable frequency. The clinal variation of allelic frequencies and the genotypic association are the two main sources of deviance. Deviance due to genotypic association equals the difference of deviance between models C_2 and C_1; thus, an estimate of the fraction of deviance due to the genotypic association explained by the presence of the duplication under model B_i equals $(\text{dev}[B_i] - \text{dev}[C_1])/(\text{dev}[C_2] - \text{dev}[C_1])$.

The phenotype was considered to be a three-state random variable. The likelihood of a sample was computed from the phenotypic multinomial distribution. Maximum-likelihood estimates of parameters were computed jointly using the Metropolis algorithm. Deviances were scaled to the dispersion of the residual deviance of the complete model (Lebreton et al. 1992; Crawley 1993). The support limits of a particular parameter were defined as the range of values within χ^2 (1 df, $\alpha = 0.05$) units of scaled deviance from the maximum (Lebreton et al. 1992). Likelihood-ratio tests were computed between related models, and model selection was performed using a heuristic measure for comparing distinct models, the Akaike Information Criterion (AIC = scaled deviance + 2 × degrees of freedom) (Akaike 1974; Lebreton et al. 1992; Anderson et al. 1994).

TABLE 2. Segregation models. For the complete model, five genotypic frequencies were fit independently. For the first simplified model (H-W), Hardy-Weinberg proportions were assumed and two allelic frequencies were fit independently. For the second simplified model ($f_{RS} = 0$), the frequency of the duplication was assumed to be zero and two genotypic frequencies were fit independently. *P*-values of likelihood-ratio tests against the complete model are indicated for each simplified model testing H-W proportions or the presence of a duplication, respectively.

		Simplified models	
	Complete model	H-W	$f_{RS} = 0$
Genotypes			
(SS)	0.085	0.076	0.111
(RR)	0.145	0.177	0.257
(RS)	0.286	0.233	0.632
(RSS)	0.092	0.167	0
(RSR)	0.263	0.255	0
(RSRS)	0.130	0.092	0
Alleles			
Ace.1^{R}	0.274	0.276	0.573
Ace.1^{S}	0.419	0.421	0.427
Ace.1^{RS}	0.307	0.303	0
Deviances	424.2	427.5	1029.5
df	5	2	2
P-values		0.349	$\ll 10^{-8}$

Appearance of the Duplication

To estimate the date of appearance of the duplication in Southern France, eight datasets from 1984 to 1997 were analyzed. Only populations from the treated area were included in the analysis. In each population, the excess of heterozygotes was calculated by assuming that only *Ace.1*^{R} and *Ace.1*^{S} alleles were segregating and using Weir and Cockerham's (1984) estimate of F_{IS}. Moreover, the excess of heterozygotes was tested globally for each dataset by using the multisample score test provided by the GENEPOP software (ver. 3.1b, Rousset and Raymond 1995) and a sequential Bonferroni correction (Holm 1979).

RESULTS

Segregation Analysis

Family Segregation.—Among the 69 females of the Maurin sample, 53 were [RS], 10 were [RR] and six were [SS]. A total of 1174 offspring were analyzed, that is, 17 offspring per female on average. Twenty-six of the 53 [RS] females had only [RS] offspring. A further analyses of 110 offspring of one of these females also revealed that all of her progeny consisted of [RS] individuals. These observations are very unlikely under the hypothesis of a Mendelian inheritance of only *Ace.1*^{R} and *Ace.1*^{S} alleles. This is unlikely even assuming that [RR] and/or [SS] individuals died during the larval stage because mortality from the eggs to adults was 32%, a rate typical of other *C. pipiens* strains (Bourguet et al. 1996c).

Duplication Frequency.—Frequency of the duplication in the Maurin population was estimated to be 0.31 and was significantly different from zero ($P \ll 10^{-8}$, Table 2). Genotypic frequencies were not significantly different from Hardy-Weinberg proportions ($P = 0.35$, Table 2). Estimating the duplication frequency directly from the heterozygote phe-

TABLE 3. Model selection for the summer and winter clines. Model refers to the models described in the text. They correspond to different simplifications (see the parameters that are used) of the complete model, A. Dev, residual scaled deviance for each model; AIC, the Akaike Information Criterion; and %TD, the part of the total deviance explained by each model. The lowest AIC is indicated in bold characters.

Model	Parameters					July 1995			November 1995		
						Dev	AIC	%TD	Dev	AIC	%TD
A	h_1	h_2	a_R	a_{RS}	F_i	6	34	0.96	6	30	0.95
B_1	h_1	h_2	a_R	a_{RS}	0	15.14	23.14	0.89	11.95	**15.95**	0.9
B_2	h_1	h_2	0	a_{RS}	F_i	58.66	64.66	0.58	43.35	45.35	0.62
B_3	h_1	h_2	a_R	0	0	18.91	24.91	0.86	22.14	24.14	0.81
B_4	h_1	h_2	$a_R=a_{RS}$		0	15.86	**21.86**	0.89	15.86	17.86	0.86
C_1	h_1	0	a_R	0	F_i	9.12	33.12	0.93	6.23	26.23	0.95
C_2	h_1	0	a_R	0	0	50.6	54.6	0.64	46.27	46.27	0.6

notype excess, at least in the Maurin population, is therefore a good approximation.

Frequency Patterns

Models B_4 and B_1 display the lowest AIC for the summer and winter samples and account for almost 90% of the total deviance (Table 3). This indicates a clinal pattern for both *Ace.1*^{R} and *Ace.1*^{RS} alleles for both sampling periods, and that Hardy-Weinberg proportions are a good approximation. The presence of *Ace.1*^{RS} clines explain about 85% of the deviance due to genotypic association for both summer and winter samples.

In both seasons, frequency of *Ace.1*^{RS} on the coast had a lower estimated value than the frequency of *Ace.1*^{R}. From the coast to the inland, frequencies of *Ace.1*^{R} and *Ace.1*^{RS} haplotypes decreased differently during the winter ($a_R < a_{RS}$) but similarly during the summer ($a_R = a_{RS}$). Finally, both *Ace.1*^{R} and *Ace.1*^{RS} frequencies were estimated to be lower in winter than in summer (Table 4, Fig. 1).

Appearance of the Duplication

Apparent heterozygote excess, which is consistent with the presence of the duplication, was not detected in the populations sampled between 1984 and 1993. In contrast, heterozygote excess was strongly supported for populations collected between 1995 and 1997 (Table 5). Although no global heterozygote excess was detected in the 1993 dataset, one population presented a low F_{IS} estimate, which could be interpreted as the possible presence of the duplication. These results are congruent with a rapid increase of the duplication from a very low frequency in 1993 to a high frequency (~ 30%) in 1995. Such a rapid replacement of *Ace.1*^{R} by *Ace.1*^{RS} presumably reflects the different fitness values associated with each haplotype.

DISCUSSION

Frequency Patterns

Progeny analysis of female mosquitoes sampled in Maurin allowed us to estimate the frequency of the duplication without assuming Hardy-Weinberg equilibrium. This estimation indicates that the duplication reached around 30% near the coast (where Maurin is located), and it further indicates that genotypic frequencies were close to Hardy-Weinberg proportions. This is also consistent with an *Ace.1*^{RS} frequency of 26.5% on the coast, estimated from the summer cline data. Additionally, the pattern of excess of phenotypic "heterozygotes" [RS] through the transect was well explained by a cline for the duplication in both summer and winter samples. The *Ace.1*^{R} and *Ace.1*^{RS} clines are likely to be maintained by a selection-migration balance, which generates heterozygote deficits (Wahlund effect) especially for populations located at the boundary between treated and nontreated areas (Lenormand et al. 1998). These deficits may slightly bias the clinal-parameter estimates through an underestimation of the *Ace.1*^{RS} frequency in some populations, which may explain why the estimation found in Maurin (30%) is slightly higher than the maximum frequency given by the summer cline (26.5%).

Although winter and summer clines were qualitatively similar, the former showed lower frequencies of *Ace.1*^{R} and *Ace.1*^{RS} haplotypes. This may be the mere consequence of seasonal variation in insecticide exposure, because pesticide treatments occur only during the spring and the summer. In the absence of insecticide treatments, clines of resistance haplotypes are only affected by fitness costs and gene flow, two factors that should decrease the maximum frequencies of the resistance haplotypes. A similar seasonal oscillation of the clines was detected in the spring and winter samples collected in 1996 (see Table 4).

TABLE 4. Descriptive fit estimates. Estimated parameters for each allelic cline, where the gene frequency is a function of distance to the coast $f(x) = f(0)\ e^{-ax^2}$, with $f(0)$ being the maximum frequency. SL indicates the support limits.

Clines	Parameters	July 1995		November 1995	
		Estimates	SL	Estimates	SL
Ace.1^{R}	$f(0)$	0.506	0.454–0.578	0.433	0.41–0.455
	a_R ($\times 10^4$)	8.55	6.92–10.6	3.50	2.9–4.14
Ace.1^{RS}	$f(0)$	0.265	0.198–0.333	0.148	0.108–0.187
	a_{RS} ($\times 10^4$)	8.55	5.15–13.85	12.4	8.1–20.0

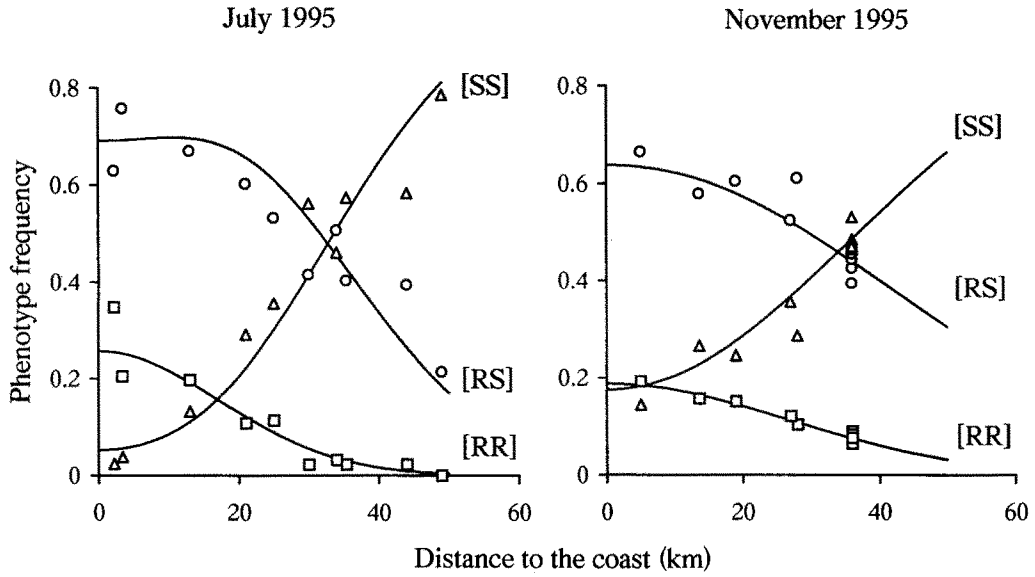

FIG. 1. Observed frequencies of *Ace.1* phenotype across the treated (between zero and 20 km from the coast) and the non-treated area in July and November 1995. Triangles, [SS] frequencies; circles [RS] frequencies; and squares, [RR] frequencies. The expected frequency of each phenotype is indicated for July (using B_4 model estimates) and November (using B_1 model estimates; see Table 4).

Duplication Sweep

The presence of $Ace.1^{RS}$ haplotype was first detected in summer 1995, with a frequency around 26% near the coast. Although not detected in 1993 (perhaps because of a low sample size), the duplication was probably already present at a low frequency, which can be estimated by maximum likelihood at 0.055 (support bounds 0.000–0.155). Using the upper bound of the 1993 frequency and the estimated frequency in 1995 and by assuming conservatively a maximum of 32 generations between 1993 and 1995 samples, it is possible to compute the minimum fitness difference between the $Ace.1^{R}$ and $Ace.1^{RS}$ haplotypes. By using numerical iterations in a stepping stone model described in Lenormand et al. (1998) and assuming additive effect of alleles within locus, the fitness of $Ace.1^{RS}$ relative to $Ace.1^{R}$ can be conservatively estimated as 1.028. A maximum difference can be computed by assuming that $Ace.1^{RS}$ conferred the same advantage in presence of insecticide as $Ace.1^{R}$, but without any fitness cost (i.e., as fit as $Ace.1^{S}$ in the absence of insecticide). Using selection coefficients estimated by Lenormand et al. (1998), this assumption gives a fitness of $Ace.1^{RS}$ relative to $Ace.1^{R}$ of 1.06. We thus conclude that the advantage in fitness of $Ace.1^{RS}$ relative to $Ace.1^{R}$ is between about 3% and 6%.

Duplication Characteristics

The evolution and spread of the $Ace.1^{RS}$ haplotype depends strongly on the level of insecticide resistance and fitness cost that it confers. Because $Ace.1^{RS}$ codes for both AChE1R and AChE1S, it can be seen as an intermediate haplotype between $Ace.1^{R}$ and $Ace.1^{S}$. Does that means that it possesses intermediate characteristics? This has been directly demonstrated for insecticide resistance in the laboratory where (RSRS) individuals from Martinique were found less resistant to insecticide than (RR) individuals (Lenormand and Raymond, unpubl. ms.). This also appears indirectly to be true for the fitness cost, because $Ace.1^{RS}$ has partially replaced $Ace.1^{R}$ in natural populations. However, the fitness cost of $Ace.1^{RS}$ is probably not zero: the presence of a frequency cline of $Ace.1^{RS}$ across treated and nontreated areas and its seasonal oscillations both support the view that $Ace.1_{RS}$ is not as fit as $Ace.1^{S}$ in the absence of insecticide.

Mortality due to OP insecticides can be caused either by an excess or a deficit of acetylcholine (ACh) in synapses through disruption of the balance between activities of AChE, which degrades ACh, and choline acetyltransferase, which synthesizes ACh (Bourguet et al. 1997b). This underlines the fact that the pivotal step for insecticide toxicology is not the

TABLE 5. Appearance of the duplication. Each dataset consists of different populations collected in the treated area around Montpellier on the same date. For each dataset, the number of populations (Nb pop) and the total sample size are indicated. The apparent excess of heterozygotes was globally tested for each dataset (see text). The corresponding *P*-value and a global F_{IS} estimation are indicated.

Dataset	August 1984	June 1986	June 1991	April 1993	July 1995	Winter 1995–1996	May 1996	Winter 1996–1997
Global F_{IS}	−0.005	0.0125	−0.072	−0.094	−0.426	−0.259	−0.420	−0.215
P-value	0.593	0.629	0.2834	0.186	< 0.0001	< 0.0001	< 0.0001	< 0.0001
Nb pop	1	16	3	4	4	5	2	8
Sample size	87	505	94	116	494	765	159	411

AChE activity, but the amount of ACh present in synapses. In that context, the proximate explanation for fitness cost associated with *Ace.1*^{R} (Chevillon et al. 1997; Lenormand et al. 1998) could be an excess of ACh in synapses. AChE1R is indeed approximately four times less efficient than AChE1S in degrading ACh (Bourguet et al. 1997a). The increased AChE1 dosage provided by the duplicated haplotype probably restores the steady-state level of ACh, thereby reducing the severity of the fitness cost. Additional improvements that independently regulate the expression of each duplicated copy could be subsequently selected. In particular, modifiers that switch the proportion of AChE1R from high levels during the larval stage (when insecticide treatments occur) to low levels during the adult stage, would increase insecticide resistance while decreasing fitness costs. More generally, the importance of gene dosage within a group of interrelated genes has been pointed out in many cases (e.g., dosage compensation for X-linked genes for example; Parkhurst and Meneely 1994), and the *Ace.1*^{RS} duplication would be an unusual case where the duplication in itself restores gene dosage instead of disrupting it.

Comparison with Esterases

Ace.1^{RS} is the second modification that has appeared at the *Ace.1* locus since 1969 in response to OP insecticide treatments along the French Mediterranean coast. This situation can be compared to the allele replacement that occurred at the esterase locus in the mosquito *C. pipiens* in the same area (Guillemaud et al. 1998). At this locus, an original resistance allele (A1) appeared in 1972 and started to be replaced by a second resistance allele (A4-B4) in 1986, a process that is not yet complete. This replacement was interpreted as the consequence of a lower fitness cost associated with A4-B4 compared to A1, because both alleles conferred very similar OPs resistance in the laboratory (Guillemaud et al. 1998). In our case, the duplication confers a reduced fitness cost, but also a lower resistance than the original resistance allele (i.e., *Ace.1*^{R}). Such a trade-off raises the possibility that the two resistance alleles might coexist indefinitely.

Maintenance of a Polymorphism?

The spread of *Ace.1*^{RS} seems to occur much more rapidly than does allele replacement at the esterase locus. This observation may be explained by changes in mosquito control. In 1991 OP insecticides used to control *C. pipiens* populations were largely replaced by *Bacillus sphaericus* toxins (Anonymous 1991–1996). AChE1R does not confer resistance to this toxin (N. Pasteur, pers. comm.). Reduction of OP treatments may have favored the spread of the *Ace.1*^{RS} haplotype. To clearly illustrate this point, consider the very simplified situation where c is the fitness cost associated with *Ace.1*^{R}, s the decrease in survival associated with *Ace.1*^{S} in the presence of insecticide, and t the probability that a treatment occurs. Fitness associated with *Ace.1*^{R}, *Ace.1*^{S}, and *Ace.1*^{RS} can we written $w_r = 1 - c$, $w_s = 1 - ts$, and $w_{rs} = 1 - \alpha ts - \beta c$, respectively. Parameters α and β indicate the survival in the presence of insecticide and the fitness cost associated with *Ace.1*^{RS}, respectively. If $\alpha + \beta < 1$, there is a range of values for t ($\beta c/s[1 - \alpha] < t < c[1 - \beta]/\alpha s$) where *Ace.1*^{RS} haplotype replaces both *Ace.1*^{R} and *Ace.1*^{S} (because $w_r < w_{rs} > w_s$). A decrease in OP treatment in the treated area (i.e., a decrease of t) makes the right inequality more likely, a condition for an initial increase of *Ace.1*^{RS} frequency, which is obviously fulfilled because *Ace.1*^{RS} has increased in frequency. However, the validity of this condition across the whole treated area is not known, thus it is too early to state whether *Ace.1*^{RS} will replace *Ace.1*^{R} in the whole treated area or if both haplotypes will coexist. The frequency estimated in 1996 is consistent with both hypotheses, but further data on *Ace.1*^{RS} frequencies in the forthcoming years may help to settle this issue.

Evolution and Maintenance of a New Function

Our example is completely opposite to the prevailing model for the evolution of a new function in which selection is initially relaxed at the duplication event. Instead, natural selection is present from the very beginning of the process. The *Ace.1*^{RS} haplotype is likely to have been created in a single step through unequal crossing-over between homologous chromosomes in an (RS) individual. Its extension would be due to lower fitness cost than *Ace.1*^{R} in conjunction with low insecticide selection conditions.

Describing the evolution of a new function from a molecular perspective requires that we define what is a function. A convenient "molecular" definition could be a class of identical RNA encoded by a haploid genome whose presence increases the organism's fitness. This definition is consistent with the different genetic models describing the evolution of a new function (Spofford 1969; Kimura and Ohta 1974; Ohta 1987, 1988, 1989; Hugues 1994; Walsh 1995). Thus, in the case of evolution of a new function through gene duplication, three criteria discriminate an n-function state from an $(n + 1)$-function state: (1) a gene duplication; (2) a divergence of the sequence of the two copies; and (3) a selective advantage associated with this divergence. The case of *Ace.1*^{RS} presented here satisfies these criteria and illustrates such an evolution. Molecular data has not been obtained to study this duplication at the sequence level (the *Ace.1* gene in *Culex* has not yet been cloned, despite intensive work by several teams). However, precise information about the beginning of the process first require population genetic data that are hardly available when gene duplication is detected by sequence analyses far after its occurrence.

As emphasized by Haldane, there is an immediate advantage to duplicating an overdominant locus. This advantage may even explain the rarity of such loci (Haldane 1954). However, this view may not be totally correct because the characteristics of a duplication are not exactly those of a standard heterozygote. In particular, a duplication doubles the number of gene copies and thus may modify the total amount of proteins produced. This may disrupt enzymatic activity dosage and outweigh the heterotic advantage (Ohno 1970). In our case, the duplication may benefit from both the heterotic advantage of producing AChE1R and AChE1S in low insecticide treatment conditions and the intrinsic advantage of the restoration of AChE and choline acetyltransferase activities to more favorable dosages.

Among the different processes by which a duplication can

evolve a new function, all are not equally detectable a priori. For instance, a process initially driven by selection rather than by drift is likely to occur on a much shorter time scale and thus may be studied in natural populations through completion. Similarly, the present situation—where two distinct alleles are involved in the occurrence of a new duplication—can be easily detected because it biases the expected Mendelian transmission of the original alleles. Unfortunately, the extent of this situation in the evolution of a new function cannot be evaluated given the absence of data on the onset of the process. Nevertheless, this is an opportunity to analyze the dynamic of the process. The fixation of such a new function may occur only at the scale of a local adaptation. Its further extension may require either environmental changes or further genetic improvements.

Acknowledgments

We are grateful to C. Chevillon, P. David, M. Kirkpatrick, R. Lenski, and F. Rousset for helpful comments and discussion on the manuscript, and to C. Bernard, M. Marquine, and G. Pistre for technical assistance. We would like to thank the people who have contributed to the sampling, P.-O. Cheptou, C. Chevillon, D. Heyse, M. Magnin, E. Marboutin, M. Michaud, A. Pasteur, N. Pasteur, and the Entente Interdépartementale pour la Démoustication du littoral Méditerranéen (EID). This work was financed in part by GDR 1105 du programme Environnement, Vie et Sociétés du CNRS, a grant from the Région Languedoc-Roussillon (no. 963223) and by an ACC SV3 (no. 9503037). TL was supported by an ASC from INRA. This is contribution 98.109 of the Institut des Sciences de l'Evolution de Montpellier (UMR CNRS 5554).

Literature Cited

Akaike, H. 1974. A new look at the statistical model identification. IEEE Trans. Auto. Control 19:716–723.

Anderson, D. R., K. P. Burnham, and G. C. White. 1994. AIC model selection in overdispersed capture-recapture data. Ecology 75:1780–1793.

Anonymous. 1990–1996. Rapport d'activité technique et scientifique. Entente interdépartementale pour la démoustication du littoral méditerranéen, Montpellier.

Batterham, P., G. K. Chambers, and W. T. Starmer. 1984. Origin and expression of an alcohol dehydrogenase gene duplication in the genus *Drosophila*. Evolution 38:644–657.

Bourguet, D., R. Capela, and M. Raymond. 1996a. An insensitive acetylcholinesterase in *Culex pipiens* L. mosquitoes from Portugal. J. Econ. Entomol. 89:1060–1066.

Bourguet, D., N. Pasteur, J. Bisset, and M. Raymond. 1996b. Determination of *Ace.1* genotypes in single mosquitoes: toward an ecumenical biochemical test. Pest. Biochem. Physiol. 55:122–128.

Bourguet, D., M. Raymond, J. Bisset, N. Pasteur, and M. Arpagaus. 1996c. Duplication of the *Ace.1* locus in Culex pipiens from the Caribbean's. Biochem. Genet. 34:351–362.

Bourguet, D., M. Raymond, D. Fournier, D. Malcolm, P. Toutant, and M. Arpagaus. 1996d. Existence of two acetylcholinesterases in the mosquito *Culex pipiens* complex (Diptera: Culicidae). J. Neurochem. 67:2115–2123.

Bourguet, D., T. Lenormand, T. Guillemaud, V. Marcel, D. Fournier, and M. Raymond. 1997a. Variation of dominance of newly arisen adaptive genes. Genetics 147:1225–1234.

Bourguet, D., M. Raymond, S. Berrada, and D. Fournier. 1997b. Interaction between acetylcholinesterase and choline acetyltransferase: an hypothesis to explain unusual toxicological responses. Pestic. Sci. 51:276–282.

Chevillon, C., N. Pasteur, M. Marquine, D. Heyse, and M. Raymond. 1995. Population structure and dynamics of selected genes in the mosquito *Culex pipiens*. Evolution 49:997–1007.

Chevillon, C., D. Bourguet, F. Rousset, N. Pasteur, and M. Raymond. 1997. Pleiotropy of adaptive changes in populations: comparisons among insecticide resistance genes in *Culex pipiens*. Genet. Res. Camb. 70:195–204.

Clark, A. G. 1994. Invasion and maintenance of a gene duplication. Proc. Natl. Acad. Sci. USA 91:2950–2954.

Clarke, G. M., and J. A. McKenzie. 1987. Developmental stability of insecticide resistant phenotypes in blowfly; a result of canalizing selection. Nature 325:345–346.

Cohan, F. M., E. C. King, and P. Zawadzki. 1994. Amelioration of the deleterious pleiotropic effects of an adaptive mutation in *Bacillus subtilis*. Evolution 48:81–95.

Crameri, A., S. Raillard, E. Bermudez, and W. P. C. Stemmer. 1998. DNA shuffling of a family of genes from diverse species accelerates directed evolution. Nature 391:288–291.

Crawley, M. J. 1993. GLIM for ecologists. Blackwell Scientific, Oxford, U.K.

Fisher, R. A. 1928. The possible modification of the responses of wild type to reccurrent mutations. Am. Nat. 62:115–126.

Fournier, D., and A. Mutero. 1994. Modification of acetylcholinesterase as a mechanism of resistance to insecticide. Comp. Biochem. Physiol. 108:19–31.

Guillemaud, T., N. Makate, M. Raymond, B. Hirst, and A. Callaghan. 1997. Esterase gene amplification in *Culex pipiens*. Insect Mol. Biol. 6:319–327.

Guillemaud, T., T. Lenormand, D. Bourguet, C. Chevillon, N. Pasteur, and M. Raymond. 1998. Evolution of resistance in *Culex pipiens*: allele replacement and changing environment. Evolution 52:443–453.

Haldane, J. B. S. 1932. The causes of evolution. Harper, New York.

———. 1954. The biochemistry of genetics. George Allen and Unwin, London.

Holm, S. 1979. A simple sequentially rejective multiple test procedure. Scand. J. Statist. 6:65–70.

Hugues, A. L. 1994. The evolution of functionally novel proteins after gene duplication. Proc. R. Soc. Lond. B Biol. Sci. 256: 119–124.

Kimura, M., and T. Ohta. 1974. On some principle governing molecular evolution. Proc. Natl. Acad. Sci. USA 71:2848–2852.

Lange, B. W., C. H. Langley, and W. Stephan. 1990. Molecular evolution of *Drosophila* metallothionein genes. Genetics 126: 921–932.

Lebreton, J.-D., K. P. Burnham, J. Clobert, and D. R. Anderson. 1992. Modeling survival and testing biological hypotheses using marked animals: a unified approach with case studies. Ecol. Monogr. 62:67–118.

Lenormand, T., T. Guillemaud, D. Bourguet, and M. Raymond. 1998. Evaluating gene flow using selected markers: a case study. Genetics 169:1383–1392.

Lenski, R. E. 1988a. Experimental studies of pleiotropy and epistasis in *Escherichia coli*. I. Variation in competitive fitness among mutant resistant to virus T4. Evolution 42:425–432.

———. 1988b. Experimental studies of pleiotropy and epistasis in *Escherichia coli*. II. Compensation for maladaptive effects associated with resistance to virus T4. Evolution 42:433–440.

Long, M., and C. H. Langley. 1993. Natural selection and the origin of *jingwei*, a chimeric processed functional gene in *Drosophila*. Science 260:91–95.

Magnin, M. 1986. Résistance aux insecticides organophosphorés: détection, caractérisation, génétique et dynamique dans les populations naturelles. Université de Paris VI, Paris.

Mouchès, C., N. Pasteur, J. B. Bergé, O. Hyrien, M. Raymond, B. Robert de Saint Vincent, M. De Silvestri, and G. Georghiou. 1986. Amplification of an esterase gene is responsible for insecticide resistance in a California *Culex* mosquito. Science 233:778–780.

Ohno, S. 1970. Evolution by gene duplication. Springer-Verlag, Berlin.

OHTA, T. 1987. Simulating evolution by gene duplication. Genetics 115:207–213.

———. 1988. Evolution by gene duplication and compensatory advantageous mutations. Genetics 120:841–847.

———. 1989. Role of gene duplication in evolution. Genome 31: 304–310.

———. 1991. Multigene families and the evolution of complexity. J. Mol. Evol. 33:34–41.

———. 1994. Further examples of evolution by gene duplication revealed through DNA sequence comparisons. Genetics 138: 1331–1337.

PARKHURST, S. M., AND P. M. MENEELY. 1994. Sex determination and dosage compensation: lessons from flies and worms. Science 264:924–932.

RAYMOND, M., D. FOURNIER, J.-M. BRIDE, A. CUANY, J. BERGÉ, M. MAGNIN, AND N. PASTEUR. 1986. Identification of resistance mechanisms in *Culex pipiens* (Diptera: Culicidae) from southern France: insensitive acetylcholinesterase and detoxifying oxidases. J. Econ. Entomol. 79:1452–1458.

ROUSSET, F., AND M. RAYMOND. 1995. Testing heterozygote excess and deficiency. Genetics 140:1413–1419.

SMITH, C. W. J., J. G. PATTON, AND B. NADAL-GINARD. 1989. Alternative splicing in the control of gene expression. Annu. Rev. Genet. 23:527–577.

SPOFFORD, J. B. 1969. Heterosis and the evolution of duplications. Am. Nat. 103:407–432.

SYVANEN, M., Z. ZHOU, J. WHARTON, C. GOLDSBURY, AND A. CLARK. 1996. Heterogeneity of the glutathione transferase genes encoding enzymes responsible for insecticide degradation in the housefly. J. Mol. Evol. 43:236–240.

SZYMURA, J. M., AND N. H. BARTON. 1986. Genetic analysis of a hybrid zone between the fire-bellied toads, *Bombina bombina* and *Bombina variegata*, near Cracow in southern Poland. Evolution 40:1141–1159.

WALSH, J. B. 1995. How often do duplicated genes evolve new functions? Genetics 139:421–428.

WANG, J. Y., S. MCCOMMAS, AND M. SYVANEN. 1991. Molecular cloning of a glutathione S-transferase overproduced in an insecticide-resistant strain of housefly (*Musca domestica*). Mol. Gen. Genet. 227:260–266.

WEIR, B. S., AND C. C. COCKERHAM. 1984. Estimating *F*-statistics for the analysis of population structure. Evolution 38:1358–1370.

Corresponding Editor: W. T. Starmer