Independent Duplications of the Acetylcholinesterase Gene Conferring Insecticide Resistance in the Mosquito Culex pipiens

Pierrick Labbé,* † Arnaud Berthomieu,* Claire Berticat,* Haoues Alout,* Michel Raymond,* Thomas Lenormand,† and Mylène Weill*†

*Institut des Sciences de l’Évolution (UMR CNRS 5554), University Montpellier II, Montpellier, France; and †Centre d’Ecologie Fonctionnelle et Evolutive (UMR CNRS 5175), CNRS campus, Montpellier, France

Gene duplication is thought to be the main potential source of material for the evolution of new gene functions. Several models have been proposed for the evolution of new functions through duplication, most based on ancient events (Myr). We provide molecular evidence for the occurrence of several (at least 3) independent duplications of the ace-1 locus in the mosquito Culex pipiens, selected in response to insecticidal pressure that probably occurred very recently (<40 years ago). This locus encodes the main target of several insecticides, the acetylcholinesterase. The duplications described consist of 2 alleles of ace-1, 1 susceptible and 1 resistant to insecticide, located on the same chromosome. These events were detected in different parts of the world and probably resulted from distinct mechanisms. We propose that duplications were selected because they reduce the fitness cost associated with the resistant ace-1 allele through the generation of persistent, advantageous heterozygosis. The rate of duplication of ace-1 in C. pipiens is probably underestimated, but seems to be rather high.

Introduction

Evolutionary potential is constrained by the number and type of genes present, but the nature of the constraints shaping the evolution of new functions remains a matter of debate. Gene duplication is thought to be a major feature of genome evolution and the main potential source of material for the origin of new evolutionary features, such as new gene functions (Haldane 1932; Ohno 1970). Two distinct phases can be distinguished in the evolution of a recent duplication: a polymorphic and a fixed period (Ohta 1988; Otto and Yong 2002). Most models of evolution following gene duplication concern the second phase—the fate of duplicated genes after fixation (for reviews see Zhang 2003; Lynch and Katju 2004), assuming that fixation is achieved by drift alone (Walsh 1995). However, several models have shown that selection can play a role in fixation (Ohta 1987; Clark 1994; Lynch et al. 2001), and increasing numbers of empirical studies have stressed the importance of selection in the early stages of duplication evolution (Hughes MK and Hughes AL 1993; Lynch and Conery 2000; Kondrashov et al. 2002). Focusing on the evolution of a new function by duplication, we can distinguish 4 scenarios with specific constraints (fig. 1; see also Otto and Yong 2002; Lynch and Katju 2004). In all these scenarios, selection favors the new function.

In the first scenario (Ohno 1970), no gene is available for the new function and no point mutation can solve this problem without altering existing (and presumably necessary) functions. In this case, only a redundant duplicated gene can accumulate the necessary mutations, with the original copy of the gene retaining its original function. An obvious constraint on this system is the number of deleterious mutations likely to disqualify the new duplicated gene before neofunctionalization (Walsh 1995; Lynch et al. 2001). Worse still, the duplicate is likely to be simply lost by drift before fixation, being initially neutral at best.

Key words: insecticide resistance gene, duplication, mosquito, Culex pipiens, selection, ace-1.

E-mail: weill@isem.univ-montp2.fr.

doi:10.1093/molbev/msm025
Advance Access publication February 5, 2007

© The Author 2007. Published by Oxford University Press on behalf of the Society for Molecular Biology and Evolution. All rights reserved.
For permissions, please e-mail: journals.permissions@oxfordjournals.org
The new function is represented in gray. Short-term and long-term events are distinguished and the evolutionary force implied is indicated (in bold). The new function is represented in gray. Short-term and long-term events are distinguished and the evolutionary force implied is indicated (in bold).

In this scenario, the duplication is less likely to be initially lost by drift (with a probability of $\sim 1 - 2s$, if $s$ is the heterotic advantage, vs. $\sim 1 - 1/2N_e$ in the 2 first scenarios) and the main evolutionary constraint is the frequency of occurrence of the duplication itself. This type of duplication requires an unequal recombination between homologous chromosomes (presumably at meiosis) and may therefore occur less frequently than duplication of a gene on a single chromosome (i.e., replication slippage, which can occur at each round of DNA replication (Chen et al. 2005)).

In all these scenarios, 2 time scales should be considered (fig. 1): 1) an initial short period in which the critical first mutation creating the new function appears, leading to the preservation of these duplicates by selection and 2) a longer time scale, in which new mutations may occur, redefining the new function and leading to further divergence of the 2 copies. The initial period is longer for the first 3 scenarios (time for fixation plus time for appearance of the mutation creating the new function) than for the fourth scenario, in which it is instantaneous.

These scenarios are complicated by the potential disruption of gene dosage by multiple copies of the same (or a similar) gene (Papp et al. 2003; Veitia 2005). Selection may also favor duplication due to the masking of deleterious mutations. However, this effect has been shown to be very weak and negligible in a first approximation (Clark 1994; Pal and Hurst 2000; Otto and Yong 2002). The relative importance of the 4 evolutionary scenarios described above is difficult to assess and may vary with population size (Ohta 1987; Clark 1994; Walsh 1995; Force et al. 1999; Lynch and Force 2000; Lynch et al. 2001) or particular events (e.g., polyploidization [Otto and Whitton 2000]). Understanding how duplications are fixed in natural populations is the first requirement—albeit a difficult empirical challenge (Zhang 2003). The early stages of duplication are hard to follow and most studies have focused on a posteriori analyses based on sequence data for fixed duplications (Long and Langly 1993; Hughes 1994; Ohta 1994; Syvanen et al. 1996; Lynch and Conery 2000; Gu et al. 2002; Moore and Purugganan 2003; Zhang 2003). The problem is that contemporary examples bearing witness to the evolution of a new function are extremely rare. The only case studied in detail concerns selection for insecticide resistance in the mosquito Culex pipiens (common house mosquito). The gene ace-I of C. pipiens populations exposed to organophosphate (OP) insecticides, at least one duplication—previously named ace-I<sup>R</sup> but referred to here as ace-I<sup>D</sup>—combining resistant and susceptible alleles of the ace-I locus has recently appeared (Bourguet, Raymond, et al. 1996; Lenormand et al. 1998). This locus encodes acetylcholinesterase (AChE1), the target of OP insecticides (Weill et al. 2002). The resistance allele, ace-I<sup>R</sup> is present worldwide and causes OP resistance in several mosquito species. It displays a single amino acid substitution, G119S, due to a mutation in the third exon of the ace-I gene, leading to the replacement of a glycine (GGC, susceptible alleles, ace-I<sup>S</sup>) by a serine (AGC [Weill, Lutfalla, et al. 2003; Weill et al. 2004]). This mutation is associated with reduced susceptibility to OP insecticide, modifications of the catalytic properties of AChE1, and a high fitness cost (for a review see Weill, Duron, et al. 2003). As no ace-I resistance alleles are detected in absence of OP insecticide, probably due to their high fitness cost, a duplication combining a resistant and a susceptible copies like ace-I<sup>D</sup> probably occurred very recently, that is, since OP insecticide treatments, less than 40 years ago anywhere in the world. The existence of ace-I<sup>D</sup> was inferred from enzymatic and genetic analyses, before cloning and sequencing of the ace-I gene.

Enzyme assays can be used to discriminate between individuals expressing only the susceptible (AChE1S, phenotype [SS]), only the resistant (AChE1R, phenotype [RR]), or both types (phenotype [RS]) of AChE1 (Test Propoxur Propoxur [TPP] test [Bourguet, Pasteur, et al. 1996]). The duplication of this gene was first suggested for Caribbean strains of C. pipiens (from Martinique and Cuba), which were mass selected in the laboratory and fixed with an [RS] phenotype (Bourguet, Raymond, et al. 1996). A similar duplication was next described in Southern France, where some [RS] individuals had only [RS] progeny, and an excess of the [RS] phenotype was observed in natural populations (Lenormand et al. 1998). The probable occurrence of this duplication in Southern France was traced back to 1993—15 years after ace-I<sup>R</sup> was first detected in the area. The duplication was shown to have gradually replaced ace-I<sup>R</sup> in treated areas (Lenormand et al. 1998).

The gene ace-I of C. pipiens has now been sequenced (GenBank accession numbers AJ489456 and AJ515147 [Weill, Lutfalla, et al. 2003; Weill et al. 2004]), providing molecular tools for more precise investigation of the generation and early evolution of ace-I duplications. The primary aim of this study was to produce molecular evidence for the existence of the ace-I<sup>D</sup> duplications. The secondary aim was to determine whether the Caribbean and South...
France duplications occurred independently or whether there was only one duplication that then spread worldwide. We also searched for ace-1 duplications in other populations. We then considered the mechanism by which the duplications occurred and spread and the likelihood of each of the possible scenarios for their evolution.

Materials and Methods

Nomenclature

The precision of ace-1 genotyping in mosquitoes and the nomenclature depend on the technique used (enzymatic assay or molecular analysis) for characterization (table 1). The enzymatic assay (TPP test [Bourguet, Pasteur, et al. 1996]) measures the susceptibility of AChE1 to an insecticide (propoxur) and detects the presence of AChE1S and AChE1R, thus generating 3 phenotypes, [SS], [RS], and [RR] (table 1). This test is limited as the [RS] phenotype comprises the undistinguishable (ace-1D/ace-1S), (ace-1D/ace-1R), (ace-1D/ace-1S), and (ace-1D/ace-1R) genotypes.

Molecular protocols generate 2 classes of fragments corresponding to susceptible or resistant copies. Fragments displaying the 119S mutation correspond either to ace-1D or to the resistant copy of ace-1D. These fragments, characteristic of resistant alleles, are collectively designated [R]. When additional sequence information is available, [R] fragments are attributed to an allele ace-1R (abbreviated to R) or to the resistant copy of an ace-1D haplotype (abbreviated to D(R)) (table 1). The second class of fragments generated by molecular protocols, displaying the 119G amino acid, corresponds to ace-1S or to susceptible copy of ace-1D. These fragments are collectively designated [S]. When additional sequence information is available, [S] fragments are attributed to an allele ace-1S (abbreviated to S) or to the susceptible copy of a ace-1D haplotype (abbreviated to D(S)) (table 1).

Mosquito Collection

Recent Strains

We analyzed 8 strains from the laboratory, searching for the presence of duplications (table 2). These strains were BIFACE, from a population sampled in Ganges (Southern France, July 2002); MAURIN, from a population sampled in Maurin (Southern France, May 2005) (for precise location, see fig. 2 in Labbé et al. 2005); DUCOS, from a population sampled in Martinique in 2003 (Duron et al. 2005); PALAWAN and MANILLE, from populations sampled in the Philippines in 2003 (Duron et al. 2005); KUNU, from a population sampled in Crete in 2002 (Duron et al. 2005); and COTONOU, from a population sampled in Cotonou City in 2005 (Benin). All these strains contain individuals resistant to OP insecticides due to the modification of AChE1.

Older Strains

Previous crossing experiments in several laboratory strains identified duplications. The mosquitoes concerned were preserved in liquid nitrogen for further analysis. These

Table 1

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Enzymatic Assay</th>
<th>Molecular Identification</th>
</tr>
</thead>
<tbody>
<tr>
<td>ace-1R</td>
<td>[R]</td>
<td>[R]</td>
</tr>
<tr>
<td>ace-1S</td>
<td>[S]</td>
<td>[S]</td>
</tr>
</tbody>
</table>

Note.—Diploid combinations of the 3 haplotypes—susceptible (ace-1D), resistant (ace-1R), and duplicated (ace-1D)—are given in the genotype column. For each genotype, the various phenotypes obtained with the different methods of identification are indicated. The enzymatic assay assesses the sensitivity or insensitivity to propoxur (an insecticide) of the AChE1 products (TPP test [Bourguet, Pasteur, et al. 1996]). Molecular identification is based on the presence ([R]) or absence ([S]) of a specific mutation in the first position of codon 119 (Weill et al. 2004), or the overall sequence of a large PCR fragment, allowing the attribution of [R] and [S] classes to specific alleles, ace-1D (R), ace-1S (S), or to specific copies of ace-1D haplotype, that is, susceptible (D(S)) or resistant (D(R)).

Table 2

<table>
<thead>
<tr>
<th>Mosquito Collection</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strain</td>
</tr>
<tr>
<td>---------</td>
</tr>
<tr>
<td>BIFACE</td>
</tr>
<tr>
<td>MAURIN</td>
</tr>
<tr>
<td>DUMONT</td>
</tr>
<tr>
<td>M-RES</td>
</tr>
<tr>
<td>MANILLE</td>
</tr>
<tr>
<td>KUNU</td>
</tr>
<tr>
<td>SLAB</td>
</tr>
<tr>
<td>COTONOU</td>
</tr>
</tbody>
</table>

Note.—Each strain used is indicated with the name of the field sample of origin (when available) and the year of collection. Taxonomic status (subspecies: Culex pipiens pipiens or Culex pipiens quinquefasciatus) is provided for each strain. The various ace-1 haplotypes present in the strain (considering the different susceptible alleles as a single class, ace-1D) are indicated.
was reared independently, and second instar larvae were

S) males (strain SLAB). The progeny of each female

Resistant females from each strain were crossed with (S/

possible to discard the confusing (R/S) genotype (fig. 2).

this problem by designing crosses and bioassays making

{R} copies were either R or D(R) sequences.

sequences were expected—1 susceptible, D(S), and 1 resis-

the reference susceptible strain SLAB (Georghiou et al.

containing strain was backcrossed for 5 generations with

Maurin in 1996 (Lenormand et al. 1998). This duplication-

available for detecting

ace-1

Characterization of the Duplications

Characterization of the Duplications

Protocol for the Detection of Females Carrying ace-1D

Protocol for the Detection of Females Carrying ace-1D

No specific test (enzymatic or molecular) is currently

strains were analyzed to determine the stability of duplica-

over time.

MARTINIQUE and M-RES were derived from popu-

lations collected in Martinique in 1994 and in Cuba in 1987,

respectively. These strains were assumed to be free of ace-1R

and ace-1D alleles (Bourguet, Raymond, et al. 1996).

DUMONT was derived from a population sampled in

Maurin in 1996 (Lenormand et al. 1998). This duplication-

containing strain was backcrossed for 5 generations with

the reference susceptible strain SLAB (Georghiou et al. 1966),

with selection at each generation with propoxur concen-

trations giving 90% mortality. Individuals displaying

only the duplicated haplotype ace-1D and the susceptible

allele ace-1S of SLAB were present in the frozen DUMONT

samples.

The taxonomic status of the mosquitoes of each strain

was determined, using the molecular test discriminating be-

tween the C. p. pipiens and C. p. quinquefasciatus

subspecies (Bourguet et al. 1998, table 2).

Identification of the Different Copies Present

The simplest way to obtain R and S sequences from a

strain is to sequence individuals displaying an [RR] or

[S] phenotype (i.e., in a genome without the ace-1D haplo-

type), identified with the TPP test (Bourguet, Pasteur,

et al. 1996).

We amplified part of exon 3 of the ace-1 gene from

females displaying duplication and from their progeny.

Polymerase chain reaction (PCR) products were then

cloned (to separate the different copies present), using the

TOPO® Cloning Kit (Invitrogen, Paisley, UK) according

to the manufacturer’s instructions. We expected a maximum

of 2 [S] and 2 [R] clone types. A first screen was applied to
discriminate [R] and [S] clones (PCR and AluI digestion,
as described by Weill et al. 2004).

If 2 [S] clone types are present at the same frequency,

the probability P of detecting both is

P = 1−1/2(n−1),

where n is the number of clones analyzed. Thus, to detect both
types of clones with a risk of less than 5% (i.e., P > 0.95), a minimum of n = 6 clones should be analyzed.

We followed the same reasoning for the resistant clones

[R]. Thus, at least 6 clones were sequenced for each class

([S] or [R]), ensuring with a 95% probability that all the
different copies present in an individual were detected. Fi-

nally, a minimum of 5 clones of each haplotype were an-

talyzed to avoid Taq misincorporation errors.

Susceptible Allele Variability

The S sequences were acquired from susceptible indi-

dividuals (S/S) from the field samples identified as [SS] with

the TPP test (table 2). PCR products were purified (Qagen

Purification Kit) and directly sequenced. For apparently

heterozygous individuals (with 2 different S sequences),

the PCR product was cloned and at least 6 clones were se-

quenced. Intron variability was assessed for the C. p. quin-

quefasciatus subspecies by analyzing the largest PCR

fragment obtained from several susceptible individuals from

the Ducos and Palawan field samples.

Sequences of the Exon 3 of the ace-1 Gene

DNA was extracted from single mosquitoes as de-

scribed by Roger and Bendich (1988). Part of exon 3 of

the ace-1 gene, including position 119, was amplified using

<table>
<thead>
<tr>
<th>Females</th>
<th>Progeny</th>
</tr>
</thead>
<tbody>
<tr>
<td>(R/S)</td>
<td>(S/D)</td>
</tr>
<tr>
<td>(D/D)</td>
<td>(R/S)</td>
</tr>
</tbody>
</table>

Fig. 2.—Duplicated haplotype detection protocol. Females of the

strain tested with an [RS] phenotype were crossed with SLAB males (ge-

notype (S/S)). Their progenies were then selected individually, using pro-

poxur insecticide at a concentration killing only susceptible individuals. If

no mortality was detected in the progeny, the female was identified as bear-

ing a duplicated ace-1 gene and was thus analyzed. If some or all of the

progeny died following exposure to insecticide, the corresponding female

was eliminated.

The simplest way to obtain R and S sequences from a

strain is to sequence individuals displaying an [RR] or

[S] phenotype (i.e., in a genome without the ace-1D haplo-

type), identified with the TTP test (Bourguet, Pasteur,

et al. 1996).

We amplified part of exon 3 of the ace-1 gene from

females displaying duplication and from their progeny.

Polymerase chain reaction (PCR) products were then

cloned (to separate the different copies present), using the

TOPO® Cloning Kit (Invitrogen, Paisley, UK) according

to the manufacturer’s instructions. We expected a maximum

of 2 [S] and 2 [R] clone types. A first screen was applied to
discriminate [R] and [S] clones (PCR and AluI digestion,
as described by Weill et al. 2004).

If 2 [S] clone types are present at the same frequency,

the probability P of detecting both is

P = 1−1/2(n−1),

where n is the number of clones analyzed. Thus, to detect both
types of clones with a risk of less than 5% (i.e., P > 0.95), a minimum of n = 6 clones should be analyzed.

We followed the same reasoning for the resistant clones

[R]. Thus, at least 6 clones were sequenced for each class

([S] or [R]), ensuring with a 95% probability that all the
different copies present in an individual were detected. Fi-

nally, a minimum of 5 clones of each haplotype were an-

talyzed to avoid Taq misincorporation errors.

Susceptible Allele Variability

The S sequences were acquired from susceptible indi-

dividuals (S/S) from the field samples identified as [SS] with

the TPP test (table 2). PCR products were purified (Qagen

Purification Kit) and directly sequenced. For apparently

heterozygous individuals (with 2 different S sequences),

the PCR product was cloned and at least 6 clones were se-

quenced. Intron variability was assessed for the C. p. quin-

quefasciatus subspecies by analyzing the largest PCR

fragment obtained from several susceptible individuals from

the Ducos and Palawan field samples.

Sequences of the Exon 3 of the ace-1 Gene

DNA was extracted from single mosquitoes as de-

scribed by Roger and Bendich (1988). Part of exon 3 of

the ace-1 gene, including position 119, was amplified using

$\text{P} = 1 - \frac{1}{2^{(n-1)}}$,

where $n$ is the number of clones analyzed. Thus, to detect both
types of clones with a risk of less than 5% (i.e., $P > 0.95$), a minimum of $n = 6$ clones should be analyzed.

We followed the same reasoning for the resistant clones

[R]. Thus, at least 6 clones were sequenced for each class

([S] or [R]), ensuring with a 95% probability that all the
different copies present in an individual were detected. Fi-

nally, a minimum of 5 clones of each haplotype were an-

talyzed to avoid Taq misincorporation errors.
2 specific primers: CpEx3dir 5′-CGA CTC GGA CCC ACT CGT-3′ and CpEx3rev 5′-GAC TTG CGA CAC GGT ACT GCA-3′, generating a 457 bp fragment. PCR was carried out with ~20 ng of genomic DNA, 10 pmol of each primer, 100 µM of each dNTP, 2.5 units of high fidelity Taq polymerase in 1 × reaction buffer (Tris–HCl [pH 9.0; 75 mM], (NH₄)₂SO₄ [20 mM], Tween 20 [0.1 %], and MgCl₂ [1.25 mM]), in a final volume of 50 µl. The PCR mixture was subjected to 30 cycles of amplification (93 °C for 30 s, 55 °C for 30 s, and 72 °C for 1 min).

A larger part of ace-1 comprising the extreme end of exon 2, the following intron (intron 2) and exon 3 were amplified as described above, using primers Intron2dir 5′-GAC TTG CGA CAC GGT ACT GCA-3′ and CpEx3rev, generating a fragment 588–597 bp in size, depending on intron size. Fragments were sequenced with an ABI Prism 310 sequencer (BigDye Terminator Kit, Applied Biosystems, Foster City, CA).

Sequence Analyses

Sequences were aligned with Multalin software (http://prodes.toulouse.inra.fr/multalin/multalin.html [Corpet 1988]). The similarity between the various sequences was assessed with ClustalW (Neighbor-Joining method, v1.83, http://www.ddbj.nig.ac.jp/search/clustalw-e.html [Thompson et al. 1994]).

Nucleotide variability was analyzed with DnaSP, v. 4.10.3 (Rozas et al. 2003). We calculated the number of polymorphic sites and nucleotide diversity, which was estimated using Nei’s π index (Nei 1987; Nei and Miller 1987).

Deduced amino acid sequences were obtained with deduce (Thompson et al. 1994) to determine whether the mutations identified were synonymous or nonsynonymous. When a nonsynonymous mutation was identified, the putative position of the mutated amino acid was sought on a 3-dimensional model of the Torpedo californica AChE (pdb 1EAS), using Swiss-PdbViewer v. 3.7 (Guex and Peitsch 1997), to determine whether this mutation was located at a site crucial for protein activity.

Results

Molecular Evidence for ace-1 Duplication

Females from several strains were analyzed using the “duplicated haplotype detection protocol” (fig. 2). This protocol allows identifying females of genotypes (D/R) or (D/D)—both display a [RS] phenotype and no mortality of their progeny in the presence of insecticide after mating with SLAB males (S/S). Such females were found in strains MAURIN, BIFACE, DUCOS, and PALAWAN, but not in strains MANILLE, KUNU, and COTONOU (see table 2 for strains description).

For each strain harboring duplication, a fragment of exon 3 of the ace-1 gene was cloned and sequenced for each female, to identify D(S) copies. The D(R) and R copies were identified by analyzing the sequences of the progeny of each female and of [RR] individuals respectively. Duplicated haplotypes were numbered in the order of their discovery and are summarized in table 3.

We analyzed 7 females carrying the duplication from the DUCOS strain collected in Martinique. All contained a single [S] allele of D(S) and a single [R] copy, identical in all females. The D(R) sequence was identical to the single R sequence present in this strain and to the previously described ace-1D1 allele of Culex p. quinquefasciatus (Weill, Lutfalla, et al. 2003). This duplication was named ace-1D1. Note that D1(S) and D1(R) differed only

Table 3

<table>
<thead>
<tr>
<th>Intron 2</th>
<th>Exon 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Copy</td>
<td>Strain</td>
</tr>
<tr>
<td>---------</td>
<td>--------</td>
</tr>
<tr>
<td>C, T</td>
<td>G, C</td>
</tr>
<tr>
<td>G, C</td>
<td>T, A</td>
</tr>
<tr>
<td>T, A</td>
<td>G, C</td>
</tr>
<tr>
<td>G, C</td>
<td>T, A</td>
</tr>
</tbody>
</table>

Note.—Mutations are indicated for the D(S) and D(R) copies of each duplicated haplotype ace-1D1. The strains in which the duplicated haplotype was found are also indicated. D(R) is the reference sequence. The dash (−) and the star (*) indicate identity or a deletion, respectively. The position from the first nucleotide of intron 2 or exon 3 is indicated at the top. The mutation at position 361 (boxed) is the only mutation conferring resistance (G119S, see Weill, Lutfalla, et al. 2003), discriminating between {R} copies (which possess an adenine at this position) and {S} forms (possessing a guanine at this position).
for the G119S mutation (table 3). For the South France strain MAURIN, we analyzed 9 females with duplication. We found 2 [S] and 1 [R] copy. The first [S] copy was present in 8 females. Seven mutations differentiated the 2 [S] copies for this partial exon 3 fragment (table 3). The same [R] copy was found in all females. The D(R) sequence was identical to the single R sequence found in this strain and to the previously described ace-1R allele of C. p. pipiens (Weill, Lutfalla, et al. 2003). Duplications containing the first and second [S] copy were named ace-1D2 and ace-1D3, respectively. Seven mutations differentiated D2(S) and D2(R) and 9 differentiated D3(S) and D3(R) (including G119S mutation; table 3). We analyzed 6 females with duplication from the South France strain BIFACE. Only one [S] and one [R] copy were found, identical in all females. The duplication in this strain was identical to the ace-1D3 duplication identified in MAURIN. For the PALAWAN strain from the Philippines, we tested 5 females and identified only one [S] and one [R] copy. The D(R) copy was identical to the single R in this strain, but different from the ace-1R allele described for C. p. quinquefasciatus (Weill, Lutfalla, et al. 2003). This duplication was named ace-1D4. Three mutations differentiated D4(S) and D4(R) (including the G119S mutation; table 3).

All the sequences of the [R] or [S] copies were more than 96% identical. D(S) sequences differed from each other by at least 3 mutations (table 3), with a higher level of divergence observed between than within subspecies. We therefore identified 4 different duplicated haplotypes. In all cases, the D(R) copy was identical to the single non-duplicated [R] copy found in [RR] individuals from the corresponding field sample.

We detected no recombinants in the progeny of any of the crosses between strains harboring duplications carried out in the laboratory (Lenormand T and Labbé P, Unpublished data). The 2 ace-1 copies therefore seem to be on the same chromosome for all the duplications detected in this study.

Variability of Susceptible Copies

In order to compare the different duplicated haplotypes and to elaborate a possible scenario for their occurrence, we measured the variability of a part of ace-1 exon 3 in susceptible individuals from each field sample where duplication was detected. For the Ducos field sample (Martinique), we analyzed 10 [SS] individuals, and characterized 7 S sequences, differing by 1–6 mutations, one being identical to D1(S) (supplementary table S1, Supplementary Material online). For the Palawan field sample, 10 [SS] individuals were analyzed, and only 4 different S sequences were identified, differing from each other by 1 mutation, 1 being identical to D4(S) (supplementary table S1, Supplementary Material online). For the South France samples, no susceptible individual was found in the Maurin2 field sample (intense insecticide treatment), so we sequenced [SS] individuals from the Ganges population (this locality is less than 35 km North of Maurin). Sixteen [SS] individuals were analyzed, leading to the description of 26 different S sequences, differing by 1–15 mutations, 1 being identical to D2(S) and 1 to D3(S) (supplementary table S2, Supplementary Material online and fig. 3). In all cases, a [S] copy identical to the D(S) copy was found in [SS] individuals from the corresponding field sample.


Duplications in Mosquitoes 1061

FIG. 3.—Diversity of the ace-1 copies in Culex pipiens pipiens samples. Variability of the different resistant D(R) and susceptible D(S) copies of each duplicated haplotype ace-1P and the single copy alleles R and S found in the Ganges sample (G) are presented (Jukes–Cantor distance, ClustalW 1.83 software [Thompson et al. 1994]), 1,000 bootstraps. The sequenced of ace-1 considered is a part of the exon 3, for each copy. The G119S mutation, selected for resistance to OP, has been removed to consider only the neutral variation. Bootstraps are presented in percentage.
of the corresponding duplicated haplotype were found respectively.

The intron sequences of the duplicated haplotypes diverged considerably between the *C. p. p. pipiens* and *C. p. quinquefasciatus* subspecies (table 3), with an 8 bp insertion detected in the sequences of individuals from the DU-COS and PALAWAN strains (*C. p. quinquefasciatus*) but not in strains from the Montpellier area (*C. p. p. pipiens*). After this extended sequence analysis, for the *ace-I* haplotype, D1(S) and D1(R) still differed only for the G119S mutation. Similarly, D2(R) and D3(R) stayed strictly identical (table 3). However, the D2(S) and D2(R) intron sequences from PALAWAN differed considerably, by 6 mutations and 1 insertion (table 3).

Stability of the Duplication Over Time

We analyzed individuals displaying duplication that had previously been collected from the same sample sites, to follow the evolution of duplicated haplotypes since their first detection.

In the Caribbean, duplications found in old strains from Martinique (MARTINIQUE, 1994) and Cuba (M-RES, 1987) were compared with the duplication found in recent samples from Martinique (DUCOS, 2003). In the MARTINIQUE strain, we found only one [S] and one [R] copy in 5 individuals. MARTINIQUE D(S) and D(R) sequences did not differ from DUCOS D1(S) and D1(R) sequences; these 2 strains therefore displayed the same haplotype, *ace-I* [H] (table 3 and supplementary table S3, Supplementary Material online).

Two sequences were identified in the 5 M-RES individuals analyzed: a susceptible and a resistant sequence, attributed to D(S) and D(R), respectively. The MARTINIQUE and M-RES strain duplications were very different, particularly in terms of the intron 2 sequences (table 3 and fig. 4). The M-RES duplication was thus named *ace-I* [H]. Note that D1(S) and D1(R) differed only for the G119S mutation. This analysis also showed that the D2(S) was identical to the D4(S) sequence found in PALAWAN (table 3 and fig. 4).

In the Montpellier area, the duplicated haplotype was analyzed in mosquitoes sampled in 1996 (DUMONT) and 2005 (MAURIN). As expected, 3 sequences, 2 [S] and 1 [R], were identified from DUMONT (strain backcrossed on SLAB, composed of only D/D and D(S)/D(S) individuals) in the 5 individuals analyzed. One of the [S] sequences was identical to the SLAB sequence, and the other corresponded to D(S). The [R] sequence was attributed to D(R). D(R) in this strain was identical to the D2(R) found in MAURIN. However, DUMONT D(S) was not identical to D2(S), differing by one insertion in intron 2 and a synonymous mutation in exon 3 (table 3 and fig. 3 and supplementary table S3, Supplementary Material online). The DUMONT haplotype was thus named *ace-I* [H].

Discussion

Duplications Mechanisms and Independence of Events

We compared the duplicated haplotypes by sequencing a part of the *ace-I* gene of their D(S) and D(R) copies,
and several nucleotide differences observed suggested different duplication events. The duplications described here are probably very recent: *ace-1R* allele is very costly for mosquitoes and cannot be detected in the absence of OP insecticides, and OP insecticides have been used against *C. pipiens* for only about 40 years in most parts of the world. Moreover, the probable occurrence of a duplication in Southern France was traced back to 1993—15 years after *ace-1R* was first detected in the area (Lenormand et al. 1998). This situation provides the first known contemporary example of ongoing evolution of a new function through duplication.

Different scenarios may account for the occurrence of these duplications (fig. 5). They correspond to scenarios 3 and 4 exposed in the introduction, as selection rather than drift seems necessary to explain the rapid emergence of the different duplications worldwide (mosquito population size is indeed certainly very large). The first scenario (3a) involves unequal crossing-over (or replication slippage [Chen et al. 2005]) in a resistant individual (i.e., of (R/R) genotype) followed by either a reversion (S119G) in one of the R copies or a supplemental recombination with an S allele. The second scenario (3b) involves unequal crossing-over (or replication slippage) in a susceptible individual (i.e., of (S/S) genotype) followed by mutation (G119S) in one of the S copies or a supplemental recombination with an R allele. A mutation step will generate duplication with very similar D(S) and D(R) copies, whereas a recombination step will generate distinct D(S) and D(R) copies, whose divergence will depend on the diversity present in natural populations (fig. 5). Scenarios 3a and 3b are similar to the third scenario proposed in the introduction, the duplication being selected first and the divergence between copies being acquired later. The scenario 4 involves unequal crossing-over in a heterozygous individual (i.e., of (R/S) genotype), resulting immediately in a new *ace-1D* haplotype (fig. 5). This scenario corresponds to that proposed by Haldane (1954) as the functionally divergent copies are already present before the duplication. In such a scenario D(S) and D(R) copies should display distinct nucleotides sequences, whose divergence will also depend on the diversity of alleles present in natural populations (fig. 5). The likelihood of these scenarios differs, as scenarios 3a and 3b require an intermediate step. Also, scenarios 4 and 3a appear more probable than 3b as the corresponding duplications (RS and RR, respectively) would confer resistance to insecticides. Duplication with 2 R copies (scenario 3a) could be advantageous by restoring a part of the normal AChE1 activity (see paragraph “Advantages and Costs of Duplications for Insecticide Resistance”). By contrast,
the duplicated susceptible haplotype (SS duplication, scenario 3b) would not confer any resistance to insecticide and might even disrupt protein dosage. Unfortunately, if duplications involving 2 [S] or 2 [R] copies exist, they cannot be detected with our protocol and require more powerful molecular tools.

These 3 scenarios involve an unequal crossing-over (or replication slippage) in individual carrying only single copy alleles. We will call these events duplication sensus stricto or type A event. However, recombination in (D/S), (D/R), or (D/D) genotypes may also generate new duplicated haplotypes without adding a new gene (fig. 5). Because of these secondary events of recombination, no phylogenetic analysis of the haplotypes can be conducted based on the point mutations as the copies of a given haplotype may have different phylogenetic history. The type B event involves a crossing-over in individual carrying a single copy allele and a duplicated haplotype ((D/S) or (D/R)), and type C event involves a crossing-over between 2 different duplicated haplotypes (D/D). Type B and type C secondary events require a preliminary event of duplication, but result from a simple recombination and thus could be more frequent than type A events. A possible signature of B and C events would be the occurrence of different duplicated haplotypes sharing a similar copy (either D(S) or D(R)).

A first question is how many independent duplication events (i.e., type A rather than B or C events) can be detected within our data set? The comparison of sequences of susceptible and duplicated alleles from different geographic origins showed that the level of variability was low (as shown previously in Weill, Lutfalla, et al. 2003). Furthermore, the stability of the duplication from Martinique described in populations sampled in 1994 (MARTINIQUE) and 2003 (DUCOS), for intron 2 and exon 3 (corresponding to ~13% of total gene length, using Anopheles gambiae [GenBank accession number BN000066 and AJ515148] as a reference [Weill et al. 2002]) confirms that mutations in this part of the ace-1 gene are rare. These observations increase the significance of any mutation found. First, the striking differences both of intron 2 and exon 3 between C. p. pipiens and C. p. quinquefasciatus sequences indicate that at least 2 types A events occurred, one in each subspecies (fig. 3 and 4).

In C. p. pipiens subspecies, 3 haplotypes (ace-1D2, ace-1D3, and ace-1D6) have been identified, sharing the same D(R) sequence but associated with different D(S) copies (fig. 3). These haplotypes may result from the same duplication sensus stricto event followed by recombination in D/S genotypes (type B (b) in fig. 5). Moreover, only 2 mutations were observed between duplicated haplotypes isolated from populations sampled in Maurin in 1996, ace-1D6 (DUMONT) and in 2005, ace-1D2 (MAURIN). This level of diversity is remarkably low. Thus, the duplicated haplotype ace-1D2 may derive (by mutation) from ace-1D6. Thus, for C. p. pipiens, the minimum duplication event (type A) number is one.

In C. p. quinquefasciatus subspecies, 3 haplotypes have been identified, ace-1D4 and ace-1D5 display strictly identical D(S) copies and one of them may result from a type B (a) secondary event (fig. 5). However, ace-1D1 from Martinique is highly divergent from ace-1D4 and ace-1D5, for both the resistant (D(R)) and the susceptible (D(S)) copies. This divergence suggests at least 2 independent sensus stricto duplications (fig. 4). Several arguments based on the distribution of alleles in the different populations corroborate this view: 1) in both Martinique and Palawan field samples, the local single R copy is identical to the local D(R) copy and these R (or D(R)) copies are different between the 2 populations. This differentiation on R and D(R) copies (they are private to each population) is a strong indication that these duplications (ace-1D1 and ace-1D2) occurred independently in different places; 2) in both Martinique and Palawan field samples, there is a private S copy corresponding to the local D(S) copy. The possibility that duplicated genes can revert back to single S copies is very unlikely because pesticide treatments are still used in all populations studied; 3) all S alleles tend to globally cluster per population (see fig. 4), which reinforces the view that the populations are structured at ace-1 locus so that the more likely explanation is again an independent origin of ace-1D1 and ace-1D4 in different places.

A process such as gene conversion between duplicated genes and alleles present in each population might blur the signal emerging from sequence data. However, this mechanism appears less likely than independent duplications for several reasons: 1) if gene conversion occurred, it should be active and the duplicated D1(R) and D1(S) alleles sampled in Martinique in 2003 would not have been strictly identical to the ones sampled in 1994; 2) if gene conversion occurred at such a fast rate, no within-population divergence among alleles should be found; 3) gene conversion would also affect nonduplicated genes but previous data (Weill, Lutfalla et al. 2003) showed that all R alleles found in distinct C. p. quinquefasciatus populations (Africa, USA, China, and South America) are strictly identical, and homogenization with single S alleles present in natural populations was never detected; 4) gene conversion would not have discriminated between the [R] and [S] copies: in Martinique, D1(R) and D1(S) are identical to the local single R allele (except of course the codon conferring the insecticide insensitivity), suggesting that only the single R allele and not any single S allele could have been used for the conversion. In Palawan, however, D2(R) and D2(S) are identical to the local single R allele and to one local single S allele, respectively, suggesting that only the single R could have been used to convert D2(S), and only one single S to convert D2(S).

Thus, for C. p. quinquefasciatus, we conclude that the minimum number of duplication event (type A) is 2.

Overall, at least 3 duplications sensus stricto are required to explain the 6 observed duplicated haplotypes. In this minimum scenario, recombination (type B) does not appear to be more frequent than duplication sensus stricto.

The second question is which scenario leads to the different duplication events? Scenario 3 could be the more likely explanation for the Martinique (ace-1D1) and the Cuba (ace-1D5) haplotypes. This conclusion stems from the striking observation that, in these duplications, the D(R) and D(S) copies are exactly identical if we except for the G119S site characterizing [S] versus [R] copies. This observation is striking because, in each case/population,
many S alleles differing from the local R allele are segregating (see fig. 4). With scenario 4 (unequal crossing-over in a (R/S) heterozygote), we do not expect D(R) and D(S) to be more similar than a R and S allele sampled randomly in the population and the exact D(R)–D(S) similarity seems therefore difficult to explain given the diversity of S alleles. On the contrary, with scenario 3, in particular 3a, where a RR duplication occurs first followed by a back (R to S) mutation at G119S site, we expect that R, D(R) and D(S) should be very similar compared with the divergence of S alleles in the population, which is in very close agreement with the data (this is particularly clear for Ducos alleles, see fig. 4). Alternatively, gene conversion between the 2 copies of these duplicated haplotypes could be involved (Teshima and Innan 2004), although this seems unlikely as these duplications are very recent.

As the D(R) and the D(S) copies of the other haplotypes differ by several mutations, they could result from either scenario 3a involving additional recombination steps or more simply from scenario 4 (fig. 5). Thus, duplications described in this study probably occurred by different mechanisms.

Advantages and Costs of Duplications for Insecticide Resistance

Gene dosage has been shown to be important in many cases and duplication may disrupt this balance (Kondrashov et al. 2002; Papp et al. 2003; Veitia 2005). In the case of ace-1, duplication could partly restore gene dosage, rather than disrupting it. OP insecticides are lethal to mosquitoes because they cause the accumulation of acetylcholine (ACh) in synapses, due to the inhibition of AChE1, which degrades ACh (Bourguet, Raymond, et al. 1997). The fitness cost associated with ace-1R probably results from the excess of ACh in synapses as the activity of the AChE1R is more than 60% lower than that of the AChE1S (Bourguet, Raymond, et al. 1996; Bourguet, Lenormand, et al. 1997). Thus, a duplicated haplotype ace-1D could be advantageous because it restores, at least in part, normal AChE1 activity. It should be noted that a duplication with 2 R copies (scenario 3a) could also restore a part of normal AChE1 activity, although at a lower level, and thus could be selected as an intermediate step to ace-1D haplotypes. AChE1 activity in ace-1D homozygotes has been shown to reach levels similar to or larger than those in susceptible homozygotes (Bourguet, Raymond, et al. 1996). The higher total AChE1 activity associated with ace-1D (15–30% higher than ace-1S) may also induce another type of fitness cost, resulting from ACh deficit. However, excess or deficit of ACh may have different fitness consequences so that the proximal reason for which duplicated haplotypes could be advantageous over ace-1R alleles in the field is still an open question. In addition, all duplicated haplotypes may not share the same level or pattern of AChE1 activity (as for instance MARTINIQUE and M-RES haplotypes [Bourguet, Raymond, et al. 1996]) so that a detailed biochemical analysis in the different haplotypes has to be performed, along with their fitness consequences, to settle this issue.

Field surveys in Caribbean islands (Yébakima et al. 2004) and in the Montpellier area (Lenormand et al. 1998) have provided insight into the relative advantages and costs of duplicated ace-1D haplotypes with respect to ace-1R and ace-1S. In Martinique (Yébakima et al. 2004), resistance due to insensitive AChE1 was extremely rare in 1990, whereas in 1999, half the population displayed a resistant phenotype (i.e., [RS] or [RR]). Moreover, the mean frequencies of each phenotype were 0.51, 0.49, and less than 0.01 for [SS], [RS], and [RR], respectively, with the frequency of [RS] reaching 0.76 in some populations. This corresponds to a very large departure from Hardy–Weinberg equilibrium and is certainly due to the high frequency of the duplicated haplotype (Lenormand et al. 1998). The ace-1R allele was extremely rare in Martinique (only one population displayed the [RR] phenotype, at a frequency of 0.03). As (D/D) individuals from MARTINIQUE are less resistant to insecticides than (R/R) individuals (Lenormand et al. 1998), the overall fitness advantage of the duplicated haplotype may result from a much lower fitness cost, but this hypothesis remains to be tested.

In the Montpellier area, the adaptive sweep (i.e., adaptive replacement) by a duplicated haplotype (characterized in this study as 2 haplotypes, ace-1D2 and ace-1D3) is consistent with the duplicated haplotypes being associated with a higher fitness than ace-1R (Lenormand et al. 1998). These duplicated haplotypes display a frequency cline across the treated and nontreated areas and seasonal oscillations in frequency, suggesting that they confer a lower fitness than ace-1S in the nontreated area.

Duplication Rate

Two studies estimated the number of duplications through sequence analyses in 3 model organisms (Lynch and Conery 2000; Gu et al. 2002). They reported mean rates of gene duplications of 0.002, 0.01, and 0.02 per gene per Myr for Saccharomyces cerevisiae, Caenorhabditis elegans, and Drosophila melanogaster, respectively, although slightly higher rates were reported for some genes, as shown by variability in gene family size (Gu et al. 2002). These rates are probably underestimated, as only fixed duplications were considered; whereas most duplications are likely to be lost by drift or selected against in the early stages (Otto and Yong 2002).

It is not possible to estimate precisely the frequency of duplication events in the mosquito C. pipiens as the genetic protocol used to detect duplication requires crosses and could not be applied extensively. Unfortunately, no conserved feature was identified in the duplicated haplotypes, precluding the design of a simple molecular detection test. However, it is very remarkable that at least 3 ace-1 duplications sensus stricto have appeared independently in such a short lapse of time. Furthermore, this number is certainly an underestimate because our geographic survey is limited. This is clearly not representative of the entire genome of C. pipiens, but demonstrates that some genes may have much higher duplication rates than estimated by comparing sequences between species or within a single genome.

Our study demonstrates the importance of duplication in the adaptive process and shows that selection may play an important role in the occurrence of such events, as processes
driven by selection, rather than by drift, are more likely to occur over such a short time scale. Insecticide resistance in mosquitoes provides us with a unique opportunity to analyze the dynamics of this process.

Supplementary Material

Supplementary tables S1–S3 are available at Molecular Biology and Evolution online (http://www.mbe.oxfordjournals.org/).

Acknowledgments

We would like to thank Janice Britton-Davidian, Mark Kirkpatrick, Thomas Galewski, Emmanuel Douzery, and Nicole Pasteur for helpful comments on the manuscript, C. Bernard, M. Marquine, and S. Ural for technical assistance, and V. Durand for the literature search. This work was funded in part by APR PNETOX 2001 (Ministère de l’Aménagement et du Territoire) and by ANR MOREVOL Sante-Environnement (Ministère délégué à la Recherche). Contribution 2007.017 of the Institut des Sciences de l’Évolution de Montpellier (UMR CNRS-UM2 5554).

Literature Cited


Haldane JBS. 1932. The causes of evolution. London: Longmans, Green, and Co.


