The evolutionary fate of heterogeneous gene duplications: A precarious overdominant equilibrium between environment, sublethality and complementation

Pascal Milesi, Benoît S Assogba, Célestine Atyame Nten, Nicolas Pocquet, Arnaud Berthomieu, Sandra Unal, Patrick Makoundou, Mylene Weill, Pierrick Labbé

Corresponding author:
Pierrick Labbé
Institut des Sciences de l'Evolution de Montpellier (CNRS-UM-IRD-EPHE)
Université de Montpellier
Place Eugène Bataillon
CC065
34095 Montpellier
France
pierrick.labbe@umontpellier.fr

The evolutionary fate of heterogeneous gene duplications: a precarious overdominant equilibrium between environment, sublethality and complementation.

Short title: Evolution of heterogeneous gene duplications

Pascal Milesi[1,2], Benoît S. Assogba[1,3], Célestine M. Atyame[1,4], Nicolas Pocquet[5,6], Arnaud Berthomieu[1], Sandra Unal[1], Patrick Makoundou[1], Mylène Weill[1], Pierrick Labbé[1]*

1- ISEM Institut des Sciences de l'Evolution de Montpellier (CNRS-Université de Montpellier-IRD-EPHE), Montpellier, France.

2- Present address: Department of Ecology & Genetics, Evolutionary Biology Centre, Science for Life Laboratory, Uppsala University, Uppsala, Sweden

3- Present address: Department of Disease Control and Elimination, Medical Research Council, Unit The Gambia, Banjul, Gambia

4- Present address: University de La Réunion, UMR PIMIT (Processus Infectieux en Milieu Insulaire Tropical, CNRS-Université de La Réunion-INSERM-IRD) Sainte-Clotilde, Ile de La Réunion, France

5- UMR MIVEGEC Maladies infectieuses et vecteurs: écologie, génétique, évolution et contrôle (CNRS-Université de Montpellier-IRD), Montpellier, France.

6- Present address: URE-EM Unité de Recherche et d'expertise en Entomologie Médicale (Institut Pasteur de Nouvelle-Calédonie), Nouméa, New Caledonia.

* Corresponding author

Abstract

Gene duplications occur at a high rate. Although most appear detrimental, some homogeneous duplications (identical gene copies) can be selected for beneficial increase in produced proteins. Heterogeneous duplications, which combine divergent alleles of a single locus, are seldom studied due to the paucity of empirical data. We investigated their role in an ongoing adaptive process at the *ace-1* locus in *Culex pipiens* mosquitoes. We assessed the worldwide diversity of the *ace-1* alleles (single-copy, susceptible S and insecticide-resistant R, and duplicated D that pair one S and one R copy), analyzed their phylogeography, and measured their fitness to understand their early dynamics using population genetics models. It provides a coherent and comprehensive evolutionary scenario.

We show that D alleles are present in most resistant populations, and display a higher diversity than R alleles (27 *vs*. 4). Most appear to result from independent unequal crossing-overs between local single-copy alleles, suggesting a recurrent process. Most duplicated alleles have a limited geographic distribution, probably resulting from their homozygous sublethality (HS phenotype). In addition, heterozygotes carrying different HS D alleles showed complementation, indicating different recessive lethal mutations. Due to mosaic insecticide control practices, balancing selection (overdominance) plays a key role in the early dynamics heterogeneous duplicated alleles; it also favors a high local polymorphism of HS D alleles in natural populations (overdominance reinforced by complementation).

Overall our study shows that the evolutionary fate of heterogeneous duplications (and their long-term role) depends on finely balanced selective pressures due to the environment and to their genomic structure.

Keywords: acetylcholinesterase, *Culex pipiens*, recurrent adaptation, natural populations, overdominance, genome evolution, balancing selection.

Introduction

New generation sequencing revealed that copy-number variations (CNVs) are ubiquitous in genomes (reviewed in Schrider & Hahn, 2010). The rate of gene duplication per gene and per generation ranges from the nucleotide substitution rate up to substantially higher orders of magnitude (Lynch et al. 2008; Lipinski et al. 2011; Katju & Bergthorsson 2013; Schrider et al. 2013). However, most gene duplications are deleterious (up to 99%, Schrider et al. 2013), both for structural (gene function) or biochemical (gene-dosage) disruptions (reviewed in Kondrashov & Kondrashov 2006), and thus exposed to purifying selection (Emerson et al. 2008; Itsara et al. 2009; Reams et al. 2010; Langley et al. 2012; Katju & Bergthorsson 2013; Schrider et al. 2013).

The few duplications that remain are likely to be under positive selection (through various scenarios, Labbé et al. 2007a; Hahn 2009; Innan & Kondrashov 2010), although the underlying mechanisms are usually poorly documented. Nevertheless, many examples showed that identical repetitions of one gene (homogeneous duplications, also called amplifications) can be selected because they increase the quantity of produced protein (e.g. resistance to xenobiotics through increased detoxification genes, Maroni et al. 1987; Guillemaud et al. 1998; adaptation to novel diet through increased amylase genes, in human, Perry et al. 2007, and dog Axelsson et al. 2013).

The case of heterogeneous duplications, *i.e.* duplications that combine already-functionally-divergent alleles of a single locus, is far less documented. Nevertheless, Haldane suggested in 1954 that such duplications should be favored in cases of overdominance (*i.e.* when the heterozygote is the fittest genotype), because they cancel the heterozygote segregation burden

and thus allow the fixation of both alleles (Haldane 1954; Spofford 1969; Lewontin 1974). Due to the paucity of empirical examples (Hahn 2009; Remnant et al. 2013), their importance in the adaptive process has however been generally overlooked. Here we took advantage of the numerous heterogeneous duplications of the *ace-1* gene described in natural populations of mosquitoes to investigate their role in adaptation and their early dynamics.

In mosquitoes, the *ace-1* gene codes for an essential enzyme involved in the nervous system, the synaptic acetylcholinesterase (AChE1), which is the target of organophosphate (OPs) and carbamate (CXs) insecticides (Weill et al. 2002). It is thus submitted to intense purifying selection, so that resistance is reduced to few options: the same amino-acid substitution (G119S) has been selected independently in many mosquito species (Weill et al. 2003, 2004a, 2004b; Djogbénou et al. 2007; Alout & Weill 2008; Liebman et al. 2015). It results from a single-nucleotide substitution in the *ace-1* gene (R allele, Tab. 1) and hinders the insecticide-AChE1 binding (Weill et al. 2003). However, in the West-Nile virus vector *Culex pipiens* mosquito and the malaria vector *Anopheles gambiae* mosquito, the G119S substitution was shown to reduce the affinity of the resistant enzyme for the neurotransmitter by more than 60% compared to its susceptible (S) version (Bourguet et al. 1997; Alout et al. 2008). This reduction is most probably the cause of the strong selective cost associated to resistance in both species (Lenormand et al. 1999; Bourguet et al. 2004; Duron et al. 2006; Djogbénou et al. 2010; Assogba et al. 2015).

Heterogeneous duplications, pairing one susceptible and one resistance copies of the *ace-1* gene (hereafter D alleles, Tab. 1), have been found in several mosquito species (Bourguet et al. 1996; Lenormand et al. 1998; Labbé et al. 2007a; Djogbénou et al. 2008; Alout et al. 2010; Osta

et al. 2012), with 13 distinct D alleles in the *Cx. pipiens* species complex (this species complex comprises two widespread species, *Cx. pipiens* and *Cx. quinquefasciatus*, respectively in temperate and tropical regions, and other more localized species as, for example, *Cx. molestus* and *Cx. pallens,* Farajollahi et al. 2011). These D alleles provide alternatives to the resistance/cost evolutionary trade-off associated with the G119S mutation: by restoring the protein activity while maintaining a significant resistance level, D carriers display a phenotype similar to the RS heterozygotes (Labbé et al. 2014; Assogba et al. 2015, 2016). Two sets of D alleles have been identified by population surveys over time. D_1 which spread and invaded rapidly the Caribbean island of Martinique (Yébakima et al. 2004), and D_2 and D_3, which first invaded mosquito populations from Southern France but then reached a stable and low frequency equilibrium (~20%, Labbé et al. 2007b). Recently, we showed that D_1D_1 homozygotes carry little selective cost, and that D_1 is selected over both the susceptible S and the resistant R alleles and reached fixation in experimental evolution (Labbé et al. 2014; Milesi et al. 2017). By contrast, D_2D_2 and D_3D_3 are sublethal (homozygote sublethality, hereafter HS phenotype), although D_2D_3 heterozygotes do not endure this HS phenotype (D_2 and D_3 complement, Labbé et al. 2007b).

The present study offers a worldwide survey of the diversity of *ace-1* susceptible S and resistance R and D alleles in the *Cx. pipiens* complex, and investigated their geographic distribution and phylogenic structure. Focusing on the duplicated D alleles, we then analyzed their fitness (experimental evolution and life history traits) and potential complementation when associated in heterozygotes. Finally, we modeled their evolution to understand their early dynamics. This large-scale and thorough investigation provides a coherent and comprehensive

evolutionary scenario explaining how heterogeneous duplications are generated in natural populations, why some spread over large geographic areas while others remain more restrained, and how several D alleles can stably co-segregate in the same population. Overall, our study shows that the evolutionary fate of heterogeneous duplications is dependent on the fine-tuning of balancing selection between environment, sublethality and complementation.

Materials and Methods

ace-1 diversity in natural populations

69 samples from the *Cx. pipiens* species complex, collected worldwide (Tab. S1), were analyzed to establish the diversity of *ace-1* alleles, e.g. single-copy susceptible S and resistance R alleles, as well as duplicated D alleles.

For each sample, DNA was extracted from single mosquitoes, as described by Roger & Bendich (1988). A ~580 bp fragment of the *ace-1* gene, including intron 2 and most of exon 3 (with the resistance G119S mutation), was amplified using two generalist primers, Intron2dir1 and CpEx3rev (G-PCR), according to Labbé et al. (2007a) (Tabs. S2a and S3). The G119S mutation creates an *AluI* restriction site (Weill et al. 2004b), so that three phenotypes can be discriminated with this PCR-RFLP test (*AluI* test): [SS], corresponding to the SS genotype, [RR], corresponding to the RR genotype, and [RS], corresponding to four genotypes: RS, DD, DS and DR (Tab. 1).

Single-copy alleles. The diversity of S and R alleles was established respectively from SS or RR individuals: the PCR product (G-PCR) was purified (Qiagen® Purification Kit) and directly sequenced (BigDye Terminator Kit, Applied Biosystems, Foster City, CA; ABI Prism 310

sequencer). When the sequence revealed individuals carrying different S or different R copies, the PCR product was cloned (TOPO® TA cloning Kit, Invitrogen, Paisley, UK); six clones per individual were analyzed, to ensure sequencing both copies and to avoid *Taq* misincorporation error.

Duplicated alleles. Heterogeneous duplications can only be characterized using extensive crossing experiments, as they confer a phenotype similar to RS individuals using current available biochemical or molecular tests. Twenty-one samples were collected at the larval stage (> 500 individuals) and reared in the laboratory until adulthood (Tab. S1). Their progenies were exposed to 2 ppm of propoxur (CX insecticide, Baygon®), a dose that kills all susceptible individuals. In populations where resistance was detected, we applied Labbé et al.'s protocol (2007a) to detect females carrying *ace-1*^{D} alleles and, when possible, establish strains carrying only one D allele (D strains).

When females carrying D alleles were detected, we pooled and backcrossed their progeny at least eight times with SLAB (up to 14 times for most pools). All pools thus shared the same genetic SLAB background (> 99%), and differed from each other almost only by their *ace-1* locus (although recombination around the *ace-1* gene is not total, most of the background effects would be eliminated, Assogba et al. 2015). In parallel, the susceptible and resistance copies (respectively, D(S) and D(R)) of the D alleles carried by the females were characterized through PCR-amplification (G-PCR), cloning, clone screening using the AluI test and PCR product sequencing (six S and six R clones). A duplicated allele was considered as new if i) a specific D(S) and D(R) sequence association had never been described ii) either D(S) and/or D(R) sequences

were different from any already-known sequence (at least one mutation) and found in several individuals.

When new duplicated alleles were found, specific molecular tests were designed to differentiate them from each other (Tab. S2b). D alleles were then isolated in different D strains, polymorphic at the *ace-1* locus only with the reference susceptible SLAB allele, S_{SLAB}. Each generation was exposed to 2 ppm of propoxur to increase D frequency. To check the purity of each strain, DNA was regularly extracted from pools of first instar larvae (~200 L1 per pool). The partial *ace-1* fragment (Intron 2–Exon 3) was then amplified (G-PCR) and directly sequenced. Sequences were then compared to the expected ones (*i.e.* a mix of S_{SLAB} and a given D allele). Three strains, fixed for different D alleles and sharing the same SLAB genetic background were also used as references: DUCOS-DFix (D_1D_1), MAURIN-DFix (D_2D_2) and BIFACE-DFix (D_3D_3) (Labbé et al. 2007b, 2014).

Extended sequences. In the case of resistance alleles (R or D alleles) identical over the ~580 bp fragment (*i.e.* no mutation), but sampled in distant geographic areas, a larger fragment of *ace-1* (Intron1-Exon3) was sequenced to get access to more nucleotide diversity and confirm if these alleles actually came from a unique mutational origin that then spread by migration (Fig. S4). A PCR using primers Intron1dir2 and CpEx3rev was performed (~20 ng of genomic DNA, 10 pmol of each primer, 2.5 units of *Taq* polymerase in 1X reaction mix (GoTaq® Long PCR Master Mix), in a final volume of 50 µl; 30 cycles: 93°C for 30 s, 55°C for 30 s, and 72°C for 5 min). It generated fragments of ~5.5 kb for *Cx. pipiens* (with a ~4.5 kb Intron1) and of 3 kb for *Cx. quinquefasciatus* (with a ~2 kb Intron1). Because of the fragment sizes, internal primers were used to sequence each fragment (Fig. S4). For the *Cx. pipiens* species, the intron 1 contained

repeated sections that prevented sequencing it entirely. However, we were able to recover the first (5') and the last (3') ~2 kb of the amplified fragment, thus a total of ~4 kb (Fig. S4).

Diversity analyses. All *ace-1* sequences (D(S), D(R), S and R) were aligned using the Mega software (Tamura et al. 2013). Their phylogeny was assessed using the Jukes-Cantor model with a rate of variation among sites modeled with a gamma distribution (shape parameter = 5). The G119S (*i.e.* resistance) mutation has been removed to consider only variations that are not implicated in resistance. The geographical structure in each taxon (*Cx. pipiens* and *Cx. quinquefasciatus*) was analyzed using a Mantel test (Genepop v.4.2, Rousset 2008). For each pair of sequences, the number of substitutions and the linear geographic distance (km) between the sampling sites were computed.

***ace-1* allele's segregation between taxa along a hybrid zone**

To study the *ace-1* allele's segregation between the *Cx. pipiens* and *Cx. quinquefasciatus*, Dr Dina M. Fonseca (Rutgers University) kindly provided us with 43 individuals from eight populations (Tab. S1) from a stable hybrid zone between these two taxa in northern USA (Mattingly et al. 1951). For each individual, she also provided us with a hybridization index, measuring, from eight microsatellites loci, its probability to belong to the *Cx. pipiens* species (multilocus Bayesian clustering analysis, Fonseca et al. 2004; Strickman and Fonseca 2012). We then sequenced the *ace-1* locus of these individuals to assess the number of *ace-1* alleles (0, 1 or 2) belonging to *Cx. pipiens*, and thus determine whether these alleles were able to pass from one taxon to the other.

D alleles' fitnesses

Experimental evolution. To measure the relative fitness of DD and DS individuals, females of each D strain (DD or DS) were crossed with SLAB males (SS). The progeny was exposed to 2 ppm of propoxur to kill all SS, so that only DS individuals remained. The initial D frequency was thus $f(D)_1 = 0.5$. The larvae were reared until adulthood under standard conditions (25°C, > 60% humidity, 12:12 hours light:dark) and adults were released in a new cage to mate freely and reproduce. The new offspring was exposed to 2 ppm propoxur and the adult survivors released in a new cage. The process was repeated five times and at each generation 96 second-instar larvae (L2) were genotyped prior selection (Tab. S2c) to estimate the frequency of SS individuals (*f(SS)*) and deduce *f(D)* frequency in their parents assuming panmixia.

D frequencies at the 6th generation ($f(DD)_6$) were compared between the various strains (STRAIN) using the generalized linear model (GLM) $f(DD)_6$ = STRAIN + ε, with ε the error parameter (following a binomial distribution). We used likelihood ratio tests (LRT) to test the significance of the STRAIN effect and of the differences between strains (Crawley 2007). Proportion tests were computed for each strain to assess the significance of the differences between $f(DD)_6$ and $f(DD)_1$. Sequential Bonferroni correction for multiple testing was used to identify potential false positive.

Relative fitnesses were estimated using a deterministic genetic model (reproduction-selection, five cycles, no drift), which was adjusted to the data and optimized using a maximum likelihood approach. For the reproduction step, the frequency of each genotype in the larvae of generation *i* was computed from the allelic frequencies in the gametes of the previous generation, assuming panmixia. Selection was then computed between larval and adult stages

using the following genotype fitnesses: w_{SS} = 0 (all SS individuals were killed by propoxur exposure), w_{DS} = 1 and w_{DD} = 1 + *s,* with *s* the selection coefficient varying between -1 and 1. The genotypic frequencies after selection were used to calculate the allelic frequencies in the gametes produced by the surviving adults. 100 000 simulations provided the likelihood profile of the *s* coefficient; its support limits, *i.e.* rough equivalents of 95% confidence intervals, were determined as its maximum and minimum values for the maximum likelihood minus 1.96. The code (R script) is available from: https://doi.org/10.5061/dryad.b31g6.

Life-history traits performances. *a-D strains:* Two life history traits, the pre-imaginal mortality and development time, were investigated to compare in each D strain the performances of the DD and DS genotypes (no strain was DD homozygous). For each D strain, female oviposition was synchronized. L_2 larvae were exposed to 2 ppm of propoxur during 24h to kill SS larvae. The surviving larvae were sorted in three pools: i) 48 were genotyped to estimate the initial DD frequencies ($f(DD)_{L2}$) (Tab. S2d), ii) 96 were used to analyze pre-imaginal mortality and iii) the remaining larvae were used to study the development time.

b-Pre-imaginal mortality: 96 L_2 were isolated in standard hemolysis tubes in 2ml of mineral water, with food provided once at the beginning of the experiment (2mg.l^{-1} of TetraMin® powdered fish food/larva). Tubes were stored on racks (two 48 tubes racks per strain), randomly distributed on a single shelf and regularly shuffled. Rearing took place in standard conditions. Differences between D strains in overall pre-imaginal mortality (number of emerging adults over initial number of L_2) were tested using the following GLM: MORT = STRAIN + ε, with MORT the pre-imaginal mortality for each strain (STRAIN) and ε the error parameter (following a

binomial distribution). Each emerging adult was then genotyped using the appropriate molecular test (Tab. S2d) to estimate the DD frequencies in adults ($f(DD)_{ad}$). Proportion tests were computed to assess the significance of $f(DD)_{L2}$ and $f(DD)_{ad}$ differences (sequential Bonferroni correction was applied).

c-Development time: The remaining L_2 larvae of each strain were reared until adulthood (standard conditions). Emerging adults were collected each day and the 48 first and 48 last were genotyped. Proportion tests were computed to assess the significance of the differences between DD frequencies among the first ($f(DD)_{first}$) and the last ($f(DD)_{last}$) emerging adults (sequential Bonferroni correction was applied).

d-Complementation tests: When several D alleles were found in geographically close populations, we assessed the performances of the heterozygotes carrying two different D alleles (D_xD_y). We crossed individuals from two different D strains. After exposure to 2ppm of propoxur (killing all SS), the resulting progeny was a mix of D_xD_y, D_xS and D_yS genotypes. 48 surviving L_2 were genotyped using the appropriate molecular tests (Tab. S2d) to estimate the D_xD_y frequency ($f(D_xD_y)_{L2}$). Larvae were reared in standard conditions and adults were collected. 96 adults were genotyped to estimate the global D_xD_y frequency in adults ($f(D_xD_y)_{ad}$), 48 among the first and 48 among the last emerged. Proportion tests were computed to assess the significance of the differences between $f(D_xD_y)_{L2}$ and $f(D_xD_y)_{ad}$ (pre-imaginal mortality) and between $f(D_xD_y)_{first}$ and $f(D_xD_y)_{last}$ (development time; sequential Bonferroni correction was applied).

Invasion dynamics of homozygous-sublethal (HS) D alleles

To investigate D alleles invasion dynamics, we computed a two-steps (reproduction, selection) deterministic (*i.e.* no drift) population genetic model in an infinite population with discrete generations.

1) Reproduction: the frequency $f(g)_i$ of each genotype *g* in generation *i* was computed from the allele frequencies $f(a)_{i-1}$ in the gametes of the previous generation, assuming panmixia: $f(g_{homo})_i = f(a)_{i-1}^2$ and $f(g_{hetero})_i = 2f(a_1)_{i-1}f(a_2)_{i-1}$, respectively for homozygotes and heterozygotes.

2) Selection: the genotypic frequencies after selection, $f'(g)_i$, were computed as $f'(g)_i = (f(g)_i\, w_g)/\Sigma(f(g)_i\, w_g)$, where w_g is the fitness of the genotype *g*. The allelic frequencies in the gametes produced by the adults of generation *i* were then calculated as $f(a)_i = f'(g_{homo})_i + \Sigma(0.5f'(g_{hetero})_i)$.

The model starts from a treated population containing R and S (single-copy alleles) at overdominance equilibrium ($w_{RR} < w_{RS} > w_{SS}$): with $w_{SS} = 0$ (susceptible individuals are killed by insecticide) and $w_{RS} = 1$, the R allele initial frequency is thus $f_0(R) = 1/(2\text{-}w_{RR})$. At generation one, a homozygous-sublethal (HS) D allele was introduced at a frequency $f(D) = 0.001$ with $w_{DD} = 0$ and $w_{DS} = w_{DR} = w_{RS}$ (the frequencies of the alleles already present in the population were proportionally decreased); different values of w_{RR} were used to assess the effect of the R/D relative fitnesses. A new HS D allele was then introduced every fifty generations, similarly to the first one; different values of w_{DxDy} were used to assess the effect of the complementation degree. The code (R script) is available from https://doi.org/10.5061/dryad.b31g6.

All statistical analyses and models were realized using the R software (v.2.15.1 http://www.R-project.org/).

Results

ace-1 allelic diversity is structured

To understand the evolutionary dynamics of the *ace-1* duplicated alleles, we first analyzed the worldwide diversity at this locus in 69 samples from the *Cx. pipiens* species complex (Fig. 1 and Tab. S1). A ~580bp fragment of the *ace-1* locus (intron2 and exon3) was sequenced for R or S single-copy alleles, and resistant D(R) and susceptible D(S) copies in duplicated D alleles (Tab. 1). Among the *ca.* 500 investigated mosquitoes, we found 78 distinct S and D(S) sequences and 7 distinct R or D(R) sequences, differing by 1 to 29 mutations (most of them in intron2; the few in exon3 were synonymous or far from the active site; data not shown).

The *ace-1* diversity (excluding the G119S position involved in resistance) appears strongly structured into two clades (Fig. 2), corresponding to the *Cx. pipiens* and *Cx. quinquefasciatus* taxa (Jukes-Cantor model, 10,000 bootstraps: 85%). In particular, no resistance allele (R or D) is shared between the two taxa (Fig. 2). The analysis of the stable northern USA hybrid zone revealed a significant correlation between the two taxa hybridization index (based on eight microsatellites) and the number of *Cx. pipiens ace-1* alleles (Pearson's product moment correlation, $r = 0.58$, $t = 4.5$, $df = 40$, $p < 0.001$, Fig. S5). This indicates that the *ace-1* diversity remains structured even at a small geographic scale, due to limited introgression between the taxa. Finally, sequences within each taxon appeared geographically structured (Figs. 1 and 2): sequences from a same geographic zone tend to gather in the tree, and the genetic distances appear correlated with the geographic distances (Mantel test, 100,000 permutations, $p = 0.01$ and $p = 0.007$, respectively for *Cx. pipiens* and *Cx. quinquefasciatus*, Fig. 2).

D alleles are more diverse than R alleles and they are found in most resistant natural populations.

Only four single-copy R alleles were identified. Two were found in only one sample: R_3 (Pa-R in Labbé et al. 2007a) in *Cx. quinquefasciatus* from the Philippines sample (#49), and a new allele, R_4, in a single *Cx. pipiens* sample from Israel (Tel-Aviv #16, Fig. 1, Tabs. S1 and S6). On the contrary, R_1 and R_2 (respectively G-R and Du-R in Labbé et al. 2007a) spread over large geographic areas: R_1 was found in all western Mediterranean countries, and R_2 from Americas to Africa, the Indian Ocean and Australia (Tabs. 2 and S1, Fig. 1). For both R_1 and R_2, longer *ace-1* sequences were obtained: they were identical over 3919bp in *Cx. pipiens* R_1 (samples # 4, 7, 8, 9, 17 and 26) and over 2787bp in *Cx. quinquefasciatus* R_2 (samples # 31, 32, 38, 40, 50, 66 and 69).

D alleles displayed a very high diversity: among the 21 new live mosquito samples collected for the present study (Tab. S1), D alleles were found in all samples, except two (Tel-Aviv, Israel #16, and St-Denis, La Reunion #39) and 14 new D alleles were described (Tabs. 2 and S6), so that the total number of D alleles known in the *Cx. pipiens* complex is now 27.

Most D alleles shared a similar resistance copy D(R): 15 had a copy identical to R_1, six to R_2, and two to R_4; three displayed D(R) copies different from each other or from any known single-copy R allele (D_5, D_{22} and D_{25}) (Tab. 2). By contrast, only four pairs of D alleles shared their susceptible copy D(S), so that most of D diversity is due to D(S) variability (Tab. 2). Finally, five D alleles displayed strictly identical D(S) and D(R) copies, save for the G119S mutation (over 2787bp in the case of D_1; NB: the sequences were different between the five D alleles). In the others, the D(S) and D(R) copies differed by several mutations (Tab. 2).

Many local duplications events

Unlike single-copy R alleles, the vast majority of the duplicated alleles (23/27) were only found in few and nearby localities (Figs. 1 and 2). However, four presented regional (D_{13}, eastern Mediterranean Sea and D_{25} USA) or worldwide distributions (D_1, Martinique and Mayotte, no mutation over 5574bp, and D_5, Cuba, Togo and Mayotte; Figs. 1 and 2, Tab. 2).

Moreover, D alleles were generally composed of D(S) and D(R) copies similar to single-copy alleles found in the same or a geographically close sample (Figs. 1 and 2). Thus 17 D(S) copies were identical or similar (< 3 mutations) to an S allele found nearby (Figs. 1 and 2, Tab. 2). Similarly, while D(R) copies were mostly identical to the widespread R_1 and R_2 alleles, D_4(R) was identical to R_3 found in the same Philippines sample, and D_{13}(R) from Lebanon was identical to R_4 from Israel.

Surprisingly, in samples where resistant alleles were found, only one single-copy R allele was present, while most samples were polymorphic for D alleles: e.g. five segregated in a sample from Tunisia and four in a sample from Mayotte (Fig. 1 and Tab. 2).

Most duplicated alleles present a homozygous-sublethal (HS) phenotype, but they complement

We then characterized the fitness conferred by the duplicated D alleles. We tried to isolate these alleles from 11 available live samples, and identified 12 different D alleles (Tab. 3).

However, we only managed to isolate 8 of them in independent strains (D_1 from Mayotte, D_5 from Togo, D_{13} and D_{27} from Greece, D_{15}, D_{16} and D_{19} from Tunisia, and D_{24} from Martinique);

the others were lost during the strain establishment process (Tab. 3). These eight alleles were introgressed in the S_{LAB} background (susceptible reference strain) to compare their fitness to that of three alleles previously characterized (Tab. 3): D_1 (Martinique, Labbé et al. 2014), D_2 and D_3 (Southern France, Labbé et al. 2007b).

Experimental evolution. The evolution of the D frequency in competition with the susceptible reference S_{Slab} allele was surveyed for six discrete insecticide-treated generations in experimental evolution cages (Fig. 3). This allowed an integrative and quantitative assessment of the fitness taking the whole life cycle into account. Starting from an initial frequency $f(D)_1 = 0.5$, two statistically different groups of D alleles emerged (Generalized Linear Model, Likelihood Ratio Test, $\chi^2 = 39.6$, $\Delta df = 8$, $p < 0.001$, Fig. 3): i) D_1 and D_{24} significantly increased in frequency ($f(D)_1$ vs $f(D)_6$, proportion tests, both $p < 0.001$) to reach a frequency of ~0.75 at the sixth generation; ii) the frequency of the other D alleles remained close to 0.5, with non-significant (D_3, D_{13}, D_{15}, D_{16}, D_{19} and D_{27}, $f(D)_1$ vs $f(D)_6$, proportion tests, all $p > 0.05$) or limited $f(D)$ increases ($f(D_2)_6 = 0.58$ and $f(D_5)_6 = 0.59$; $f(D)_1$ vs $f(D)_6$, proportion tests both $p < 0.001$; Fig. 3).

We used a deterministic genetic model to estimate the fitness of the D_xD_x genotype ($w_{DD} = 1 + s$) relatively to the D_xS_{SLAB} genotype ($w_{DS} = 1$; NB: all SS are killed by insecticide exposure so $w_{SS} = 0$); drift did play a role in the allele dynamics (Fig. 3), but had a much smaller impact than selection, and was thus not considered in this model. Although the D_1 and D_{24} homozygotes displayed limited costs (w_{DD} = 0.73±0.1 and 0.76±0.13, respectively Tab. 4), the other alleles were sublethal at the homozygous state (HS phenotype): fitnesses were not different from 0 for D_3,

D_{13}, D_{16}, D_{19} and D_{27}, slightly higher for D_2 and D_5 (0.27±0.1 and 0.32±0.2, resp.; Tab. 4), but still much lower than is usually found for RR individuals (~0.6, Milesi et al. 2017).

Life-history traits. To investigate whether the origin of sublethality was the same for the different D alleles, we analyzed two life history traits, pre-imaginal mortality and development time, for three HS alleles (D_{15}, D_{16} and D_{19}) and for D_1 and D_{24}. The overall pre-imaginal mortality was significantly different between the strains and two groups emerged: i) D_{15}, D_{16} and D_{19} showed a strong and similar mortality (Tab 4, GLM, LRT, $\chi^2 = 1$, $\Delta df = 1$, $p > 0.05$), close to that previously observed for D_2 and D_3 (Labbé et al. 2007b), ii) D_1 and D_{24} displayed a similar (GLM, LRT, $\chi^2 = 0.17$, $\Delta df = 1$, $p > 0.05$), but significantly lower mortality (GLM, LRT, $\chi^2 = 65.4$, $\Delta df = 4$, $p < 0.001$, Tab. 4), as previously observed for the Slab reference (Labbé et al. 2014). In D_1 and D_{24} harboring strains, the frequency of the DD genotype (*f(DD)*) was similar between L_2 and adults, while *f(DD)* tended to decrease for D_{15}, D_{16} and D_{19}, although this decrease was significant only for D_{15} (Tab. 4). The same two groups appeared when assessing the development time: the adult emergence of DD individuals was delayed compared to DS individuals for D_{15}, D_{16} and D_{19}, while no significant difference was observed between the first and last emerged adults for D_1 and D_{24} (Tab. 4).

Complementation tests. To assess complementation, we compared the pre-imaginal mortality and development time of D_xD_y heterozygotes (*i.e.* carrying two different D alleles) to the D_xS heterozygotes, for the alleles D_{13}, D_{15}, D_{16} and D_{19}, using D_2 and D_3 as references. All these alleles present the HS phenotype and segregate in Mediterranean populations, some in a same locality

(Fig. 1). No significant differences in pre-imaginal mortality or development time between the D_xD_y and DS (D_xS or D_yS) individuals were found in any of the tested combinations (proportion tests, all $p > 0.05$, Fig. S7), except for $D_{15}D_{16}$ and $D_{13}D_{19}$, which appeared to develop slower (proportion tests, $p = 0.04$ and $p = 0.03$, respectively, but did not pass sequential Bonferroni correction).

Multiple overdominance maintain HS D alleles polymorphism

We modeled the dynamics of HS D alleles introduced in a population containing only single-copy alleles R and S at overdominance equilibrium. We first investigated the influence of the homozygous RR fitness (w_{RR}) on the invasion dynamics of a single HS D allele for 50 generations ($f_1(D) = 0.001$, $w_{SS} = w_{DD} = 0$, $w_{DS} = w_{DR} = w_{RS} = 1$). Invasion is faster when w_{RR} decreases (Fig. 4A): equilibrium is reached at *ca.* generation 15, 25 and 50 for $w_{RR} = 0.2$, 0.5 and 0.8, respectively. Similarly, the mean population fitness μW is higher at equilibrium when w_{RR} decreases.

We then investigated 1) the dynamics of several HS D alleles appearing successively in a single population and 2) the influence of the degree of complementation (*i.e.* the value of w_{DxDy} relatively to w_{DxDx} and w_{DyDy}) on these dynamics. In that effect, a new HS D allele was introduced every fifty generations. w_{RR} was fixed to 0.5 and we tested two degrees of complementation, with $w_{DxDy} = 1$ or $w_{DxDy} = 0.5$. It first appears that each new D allele invades more slowly (the slope of the curve decreases for each new allele, Figs. 4B1 and 4B2). Moreover, while the cumulated frequency of the D alleles $f_{tot}(D)$ increases with each new allele, the increase is each time smaller and the frequency of each allele $f_i(D)$ decreases (Fig. 4B1). Nevertheless, it clearly appears that several HS D alleles can successively invade and coexist in a single population, as

multiple overdominance allows stable polymorphism at equilibrium. Higher degrees of complementation result in faster invasions and higher frequencies at equilibrium (Fig. 4B1 vs B2).

Discussion

In contrast with homogeneous duplications, where the increased number of identical copies of a locus provides a clear quantitative advantage, contemporary examples of heterogeneous duplications are rare and their importance in the adaptive process remains generally poorly understood. In the *Cx. pipiens* complex, such duplications, pairing one S and one R copies of the *ace-1* gene, were repeatedly identified in natural populations over the last 40 years, in programs following the adaptation to insecticide mosquito control. They thus provide a rare opportunity to study their early evolutionary dynamics.

Heterogeneous *ace-1* gene duplications are recurrent

Our worldwide study of the *ace-1* locus diversity in mosquitoes of the *Cx. pipiens* complex (69 samples, Fig. 1) revealed that the genetic diversity was strongly structured: two clear clades emerged corresponding to the *Cx. pipiens* and *Cx. quinquefasciatus* taxa (despites regions of introgression, as in the hybrid zone in the USA), and a significant geographic structure was detected within each taxon (Fig. 2).

Focusing on the resistance alleles, the patterns were contrasted between R and D alleles. Only four single-copy R alleles were identified. Two of them, R_1 in *Cx. pipiens* and R_2 in *Cx. quinquefasciatus*, have spread from a unique origin to a large geographic area (as confirmed by

extended sequences >2.5kb), while the two others were each found in only one sample (Fig. 2). Mosquitoes are known to travel easily and widely thanks to human activities (Asahina 1970; Tatem et al. 2006; Benedict et al. 2007), and resistance alleles have been shown to spread rapidly between distant insecticide-treated populations thanks to their selective advantage (e.g. *Ester*^{2}, a resistance allele of the *Ester* locus spread worldwide in a few years, both in *Cx. pipiens* and *Cx. quinquefasciatus*, Raymond et al. 1991; Guillemaud et al. 1996; Labbe et al. 2005).

The diversity and distribution of the duplicated alleles (D) came at a striking contrast (Fig. 1, Tab. 2). We indeed detected 14 new D alleles (Tab. 2), raising their number to a total of 27 in the *Cx. pipiens* species complex (Labbé et al. 2007a; Alout et al. 2010; Osta et al. 2012, this study). This number is probably largely underestimated due to the complex protocol allowing their identification (Labbé et al. 2007a). These duplications were present in most resistant populations (19 out of the 21 live samples) and often two to five alleles were found in the same sample. However, most of them were restricted to a limited geographic area (Fig. 1, Tab.2).

Furthermore, while most D alleles (22/27) carried different D(S) and D(R) sequences, these sequences were identical to single-copy S and R alleles found in the same populations (Fig. 2). Thus they were probably formed independently, after the R allele spread, through unequal crossing-overs in heterozygous individuals carrying local S and R alleles. This conclusion of independent duplication events is further supported by the fact that homozygous-sublethal (HS) D alleles complement (see below).

The D(S) and D(R) copies were however strictly identical in five duplicated alleles (e.g. over ~3 kb between D_1(S) and D_1(R)). They were likely generated by a replication slippage affecting a R allele, followed by the S119G reverse mutation in one of the copies (Labbé et al. 2007a). Note

that as several D alleles share an identical *ace-1* copy (usually D(R)), we cannot rule out that some of them could have originated from secondary recombination in heterozygotes (Labbé et al. 2007a). However, the geographic distribution of D alleles, the fact that they complement, and the inhibition of the recombination due to the duplication itself suggest that secondary recombination events are probably rare.

The striking differences in diversity and spread between the R and D alleles is puzzling: as *ace-1* single-copy resistance alleles (R_1 and R_2) have spread over the World, why do most D alleles remain confined geographically, with several of them co-existing in natural populations?

Most D alleles are sublethal when homozygotes

The fitness of seven new D alleles (D_5, D_{13}, D_{15}, D_{16}, D_{19}, D_{24}, D_{27}) were assessed through experimental evolution and/or life-history trait experiments, using D_1, D_2 and D_3 as references (Labbé et al. 2007b, 2014). Only one allele (D_{24}) displayed the same phenotype as D_1, while the rest displayed HS phenotypes, similar to D_2 and D_3 (Fig. 3 and Tab. 4).

The large number of D alleles analyzed in this study allowed a deeper understanding of the origin of these HS phenotypes. First, they are independent from the original duplication process: D_1 and D_5 were probably generated through replication slippage, D_{24} and D_3 through unequal crossing-over, but while D_1 and D_{24} carried little cost, D_5 and D_3 were sublethal when homozygous. HS phenotypes are also independent from phylogeny (e.g. D_1 is fine but D_5 is HS in *Cx. quinquefasciatus*). Finally, the complementation experiment showed that every tested pairs compensated each other's deficiencies in D_xD_y heterozygotes, *i.e.* D_xD_y are as fit as DS (Figs. S7A and S7B). This demonstrates that the HS phenotypes are independent from the G119S

mutation; nor are they the consequences of disruptions in some biochemical equilibrium due a gene-dosage effect.

More importantly, the complementation experiment demonstrates that the mutations responsible for these sublethal-homozygote phenotypes were different for the different alleles. Development time and survival analyses further showed diverse patterns for the different alleles, suggesting variations in the degree of recessivity of these costs (Tab. 4). These recessive sublethal mutations could be generated by the chromosomic re-arrangements disrupting functional genes at the break points of the amplicons. Another hypothesis, non-exclusive but more likely to explain the high diversity of mutations identified in this study, is that the recessive sublethal mutations were captured by hitchhiking in the amplicons. The genomic resources to test these hypotheses are not yet available for *Culex*. However, two cases of heterogeneous duplications have been analyzed at the genomic level: in *D. melanogaster*, the *Rdl* gene (resistance to the insecticide dieldrin) is part of a duplicated region of 113kb containing five genes with the 5' and 3' break-points within two other genes (Remnant et al. 2013), and in *An. gambiae* the *ace-1* gene is part of a duplicated region of 200kb containing 12 genes (Assogba et al. 2016). Such large genomic rearrangements, encompassing several genes, offer ample opportunities to hitchhike deleterious mutations affecting different genes, thus allowing complementation. Moreover, such mutations would be difficult to purge, as the recombination is reduced in the vicinity of the duplication.

Studying the fitness consequences of the *ace-1* heterogeneous duplications thus revealed that most are deleterious at the homozygous state due to independent mutations. This probably explains why most of D alleles have a limited geographical range (Fig. 1). By contrast, D_1, which

did not display the HS phenotype, spread widely: the same allele (identical sequences over 5574 bp), conferring the same phenotype (Tab. 4), was found more than 12,000 km apart, in Mayotte (Indian Ocean) and in Martinique (Caribbean). However, if most of D alleles are deleterious when homozygous, how can we explain their striking diversity, both locally (up to five D in one sample) and globally (27 known D alleles)?

Costly duplications are maintained in the populations by balancing selection reinforced by complementation

We used a simple determinist population genetics model (one infinite population, no drift, no migration) to investigate the invasion dynamics of HS D alleles (note that, while qualitatively relevant for comparisons, the model simplicity prevents accurate quantitative predictions). It shows that, if insecticide selection is intense enough to kill most SS while preserving most heterozygotes (DS, DR or RS), even a D allele displaying HS phenotype will be selected for: the DS and DR genotypes indeed provide intermediate resistance levels, but their cost is much lower than RR homozygotes (Labbé et al. 2007b, 2014). However, as lethal homozygotes become more frequently generated, D frequency soon reaches a stable equilibrium (which depends on the relative fitnesses of the RR genotype, Fig. 4A).

Moreover, as it first appears at the advantageous heterozygous state, each new HS D allele will be selected for (Fig. 4B). When the HS alleles segregating in the same population complement, it results in a complex situation of balancing selection (multiple overdominance), where all the heterozygotes (RS, DS, DR and D_xD_y) are fitter than the homozygotes (SS are killed by the insecticides, RR endure a high fitness cost, and all D_xD_x are sublethal).

In these conditions, the more D alleles present, the higher their cumulated frequency, but the lower the frequency of each (Fig. 4B). This makes them more prompt to be lost through drift, all the more so that the time to reach equilibrium increases with the number of D alleles (as shown by the decreasing slopes, Fig. 4B), which indicates that the strength of the selection favoring new duplicated alleles decreases with their number. This model thus shows that several HS D alleles can segregate in a same population, but that their number will remain limited (although a higher degree of complementation would allow more alleles to co-segregate by limiting the effects of drift, as they more rapidly reach a higher frequency, Figs. 4B1 *vs* 4B2). The presence of several D alleles (up to five) found in most samples (Fig. 1) thus suggests that the mosquito control practices generate complex patterns of balancing selection: they result in a mosaic of treated and untreated areas and vary along the seasons, which generates antagonistic selective pressures (Lenormand et al. 1998; Labbé et al. 2007b, 2014; Assogba et al. 2015; Milesi et al. 2017).

This situation is however probably transitory: D alleles can indeed migrate and as soon as a D allele similar to D_1 would arrive, it would eliminate the others (note that it would also eliminate the R and S alleles and reach fixation, Milesi et al. 2017). This is indeed what we observed in the series of backcrosses performed while trying to isolate the various D alleles present in Mayotte samples (D_1, D_5, D_{20}, D_{21}, and D_{26}): at the end, only D_1 was retrieved, the other alleles being lost during the process. Because of its low cost, D_1 allele was probably selected over the HS D alleles.

In conclusion, our study of natural populations revealed that heterogeneous gene duplications were frequently generated, through different molecular mechanisms. However, while they are frequent, the evolutionary outcome is often messy. The majority of these D alleles are indeed

defective when homozygous (HS), probably due to ingrained hitchhiking lethal recessive mutations. This prevents these alleles from reaching fixation in natural populations. However, together with the heterogeneous insecticide control practices, it generates a complex situation of balancing selection, which favors both their early spread (permanent association of overdominant alleles) and the maintenance of their polymorphism (multiple overdominance situations favored by complementation). Only the rare D alleles that escape that fate (e.g. D_1) can then spread over large distances and reach fixation in populations. Once established, those could provide new genetic material for long-term evolution: as they associate already functionally-divergent copies of a same gene, these duplications are indeed likely to further evolve through neo- or subfunctionalization, and should thus be studied with a higher scrutiny.

Acknowledgments

We are very grateful to N Pasteur for helpful comments on the manuscript, and to DM Fonseca, BJ Johnson, M Raymond, J Catalan, A Yébakima and J Vontas for providing samples. All sequence data were obtained on the ISEM Environmental Genomic Platform. This work was funded by French ANR program (project "SilentAdapt", ANR-13-ADAP-0016). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. This study is the contribution ISEM 2017-258 SUD of the Institut des Sciences de l'Evolution de Montpellier (UMR 5554, CNRS-UM-IRD-EPHE).

References:

Alout, H., Djogbénou, L., Berticat, C., Chandre, F. & Weill, M. (2008). Comparison of *Anopheles gambiae* and *Culex pipiens* acetycholinesterase 1 biochemical properties. *Comp. Biochem. Physiol. B-Biochemistry Mol. Biol.*, 150, 271–277.

Alout, H., Labbé, P., Pasteur, N. & Weill, M. (2010). High incidence of *ace-1* duplicated haplotypes in resistant *Culex pipiens* mosquitoes from Algeria. *Insect Biochem. Mol. Biol.*, 41, 29–35.

Alout, H. & Weill, M. (2008). Amino-acid substitutions in acetylcholinesterase 1 involved in insecticide resistance in mosquitoes. *Chem. Biol. Interact.*, 175, 138–141.

Asahina, S. (1970). Transoceanic flight of mosquitoes on the northwest Pacific. *Japanese J. Med. Sci. Biol.*, 23, 255–258.

Assogba, B.S., Djogbénou, L.S., Milesi, P., Berthomieu, A., Perez, J., Ayala, D., Chandre, F., Makoutodé, M., Labbé, P. & Weill, M. (2015). An *ace-1* gene duplication resorbs the fitness cost associated with resistance in *Anopheles gambiae*, the main malaria mosquito. *Sci. Rep.*, 5, 14529.

Assogba, B.S., Milesi, P., Djogbénou, L.S., Berthomieu, A., Makoundou, P., Baba-Moussa, L.S., Fiston-Lavier, A.-S., Belkhir, K., Labbé, P. & Weill, M. (2016). The *ace-1* locus is amplified in all resistant *Anopheles gambiae* mosquitoes: fitness consequences of homogeneous and heterogeneous duplications. *PLOS Biol.*, 14, e2000618.

Axelsson, E., Ratnakumar, A., Arendt, M.-L., Maqbool, K., Webster, M.T., Perloski, M., Liberg, O., Arnemo, J.M., Hedhammar, A. & Lindblad-Toh, K. (2013). The genomic signature of dog domestication reveals adaptation to a starch-rich diet. *Nature*, 495, 360–4.

Benedict, M.Q., Levine, R.S., Hawley, W.A. & Lounibos, L.P. (2007). Spread of the tiger: global risk of invasion by the mosquito *Aedes albopictus*. *Vector Borne Zoonotic Dis.*, 7, 76–85.

Bourguet, D., Guillemaud, T., Chevillon, C. & Raymond, M. (2004). Fitness costs of insecticide resistance in natural breeding sites of the mosquito *Culex pipiens*. *Evolution*, 58, 128–35.

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

Bourguet, D., Roig, A., Toutant, J.P. & Arpagaus, M. (1997). Analysis of molecular forms and pharmacological properties of acetylcholinesterase in several mosquito species. *Neurochem. Int.*, 31, 65–72.

Crawley, M.J. (2007). *The R book*. John Wiley & Sons Ltd., Chichester, UK.

Djogbénou, L., Dabiré, K.R., Diabaté, A., Kengne, P., Akogbéto, M.C., Hougard, J.M. & Chandre, F. (2008). Identification and geographic distribution of the ace-1R mutation in the malaria vector Anopheles gambiae in south-western Burkina Faso, West Africa. *Am. J. Trop. Med. Hyg.*, 78, 298–302.

Djogbénou, L., Noel, V. & Agnew, P. (2010). Costs of insensitive acetylcholinesterase insecticide

resistance for the malaria vector *Anopheles gambiae* homozygous for the G119S mutation. *Malar. J.*, 9, 12.

Djogbénou, L., Weill, M. & Hougard, J. (2007). Characterization of insensitive acetylcholinesterase (*ace-1^R*) in *Anopheles gambiae* (Diptera : Culicidae): resistance levels and dominance. *J. Med. Entomol.*, 44, 1–7.

Duron, O., Labbé, P., Berticat, C., Rousset, F., Guillot, S., Raymond, M. & Weill, M. (2006). High *Wolbachia* density correlates with cost of infection for insecticide resistant *Culex pipiens* mosquitoes. *Evolution*, 60, 303–14.

Emerson, J.J., Cardoso-Moreira, M., Borevitz, J.O. & Long, M. (2008). Natural selection shapes genome-wide patterns of copy-number polymorphism in *Drosophila melanogaster*. *Science*, 320, 1629–1631.

Farajollahi, A., Fonseca, D.M., Kramer, L.D., Kilpatrick, A.M., Marm Kilpatrick, A. & Kilpatrick, A.M. (2011). "Bird biting" mosquitoes and human disease : A review of the role of *Culex pipiens* complex mosquitoes in epidemiology. *Infect. Genet. Evol.*, 11, 1577–1585.

Fonseca, D.M., Keyghobadi, N., Malcolm, C.A., Mehmet, C., Schaffner, F., Mogi, M., Fleischer, R.C. & Wilkerson, R.C. (2004). Emerging vectors in the *Culex pipiens* complex. *Science*, 303, 1535–1538.

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

Guillemaud, T., Rooker, S.J., Pasteur, N. & Raymond, M. (1996). Testing the unique amplification event and the worldwide migration hypothesis of insecticide resistance genes with sequence data. *Heredity (Edinb).*, 77, 535–543.

Hahn, M.W. (2009). Distinguishing among evolutionary models for the maintenance of gene duplicates. *J. Hered.*, 100, 605–617.

Haldane, J.B.S. (1954). *The biochemistry of genetics*. George Allen and Unwin, Ltd, London.

Innan, H. & Kondrashov, F. (2010). The evolution of gene duplications: classifying and distinguishing between models. *Nat. Rev. Genet.*, 11, 97–108.

Itsara, A., Cooper, G.M., Baker, C., Girirajan, S., Li, J., Absher, D., Krauss, R.M., Myers, R.M., Ridker, P.M., Chasman, D.I., Mefford, H., Ying, P., Nickerson, D.A. & Eichler, E.E. (2009). Population analysis of large copy number variants and hotspots of human genetic disease. *Am. J. Hum. Genet.*, 84, 148–161.

Katju, V. & Bergthorsson, U. (2013). Copy-number changes in evolution: rates, fitness effects and adaptive significance. *Front. Genet.*, 4, 1–12.

Kondrashov, F.A. & Kondrashov, A.S. (2006). Role of selection in fixation of gene duplications. *J. Theor. Biol.*, 239, 141–51.

Labbé, P., Berthomieu, A., Berticat, C., Alout, H., Raymond, M., Lenormand, T. & Weill, M. (2007a). Independent duplications of the acetylcholinesterase gene conferring insecticide resistance in the mosquito *Culex pipiens*. *Mol. Biol. Evol.*, 24, 1056–1067.

Labbé, P., Berticat, C., Berthomieu, A., Unal, S., Bernard, C., Weill, M. & Lenormand, T. (2007b). Forty years of erratic insecticide resistance evolution in the mosquito *Culex pipiens*. *PLoS Genet.*, 3, e205.

Labbe, P., Lenormand, T., Raymond, M., Labbé, P., Lenormand, T. & Raymond, M. (2005). On the worldwide spread of an insecticide resistance gene: a role for local selection. *J. Evol. Biol.*, 18, 1471–84.

Labbé, P., Milesi, P., Yébakima, A., Pasteur, N., Weill, M. & Lenormand, T. (2014). Gene-dosage effects on fitness in recent adaptive duplications: *ace-1* in the mosquito *Culex pipiens*. *Evolution*, 68, 2092–2101.

Langley, C.H., Stevens, K., Cardeno, C., Lee, Y.C.G., Schrider, D.R., Pool, J.E., Langley, S.A., Suarez, C., Corbett-Detig, R.B., Kolaczkowski, B., Fang, S., Nista, P.M., Holloway, A.K., Kern, A.D., Dewey, C.N., Song, Y.S., Hahn, M.W. & Begun, D.J. (2012). Genomic variation in natural populations of *Drosophila melanogaster*. *Genetics*, 192, 533–598.

Lenormand, T., Bourguet, D., Guillemaud, T. & Raymond, M. (1999). Tracking the evolution of insecticide resistance in the mosquito *Culex pipiens*. *Nature*, 400, 861–864.

Lenormand, T., Guillemaud, T., Bourguet, D. & Raymond, M. (1998). Appearance and sweep of a gene duplication: adaptive response and potential for new functions in the mosquito *Culex pipiens*. *Evolution*, 52, 1705–1712.

Lewontin, R.C. (1974). *The genetic basis of evolutionary change*. Columbia University Press, New York, USA.

Liebman, K. a, Pinto, J., Valle, J., Palomino, M., Vizcaino, L., Brogdon, W. & Lenhart, A. (2015). Novel mutations on the *ace-1* gene of the malaria vector *Anopheles albimanus* provide evidence for balancing selection in an area of high insecticide resistance in Peru. *Malar. J.*, 14, 1–10.

Lipinski, K.J., Farslow, J.C., Fitzpatrick, K.A., Lynch, M., Katju, V. & Bergthorsson, U. (2011). High spontaneous rate of gene duplication in *Caenorhabditis elegans*. *Curr. Biol.*, 21, 306–310.

Lynch, M., Sung, W., Morris, K., Coffey, N., Landry, C.R., Dopman, E.B., Dickinson, W.J., Okamoto, K., Kulkarni, S., Hartl, D.L. & Thomas, W.K. (2008). A genome-wide view of the spectrum of spontaneous mutations in yeast. *Proc. Natl. Acad. Sci. U. S. A.*, 105, 9272–9277.

Maroni, G., Wise, J., Young, J.E. & Otto, E. (1987). Metallothionein gene duplications and metal tolerance in natural populations of *Drosophila melanogaster*. *Genetics*, 117, 739–744.

Mattingly, P.F., Rozeboom, L.E., Knight, K.L., Laven, H., Drummond, F.H., Christophers, S.R. & Schute, P.G. (1951). The *Culex pipiens* complex. *Trans. R. Entomol. Soc. London*, 102, 331–342.

Milesi, P., Weill, M., Lenormand, T. & Labbé, P. (2017). Heterogeneous gene duplications can be adaptive because they permanently associate overdominant alleles. *Evol. Lett.*, 1, 169–180.

Osta, M.A.M.M.A., Rizk, Z.Z.J., Labbé, P., Weill, M. & Knio, K. (2012). Insecticide resistance to organophosphates in *Culex pipiens* complex from Lebanon. *Parasit. Vectors*, 5, 132.

Perry, G.G.H., Dominy, N.N.J., Claw, K.K.G., Lee, A.A.S., Fiegler, H., Redon, R., Werner, J., Villanea, F. a, Mountain, J.L., Misra, R., Carter, N.P., Lee, C. & Stone, A.C. (2007). Diet and the evolution of human amylase gene copy number variation. *Nat Genet*, 39, 1256–1260.

Raymond, M., Callaghan, A., Fort, P. & Pasteur, N. (1991). Worldwide migration of amplified insecticide resistance genes in mosquitoes. *Nature*, 350, 151–153.

Reams, A.B., Kofoid, E., Savageau, M. & Roth, J.R. (2010). Duplication frequency in a population of *Salmonella enterica* rapidly approaches steady state with or without recombination. *Genetics*, 184, 1077–1094.

Remnant, E.J., Good, R.T., Schmidt, J.M., Lumb, C., Robin, C., Daborn, P.J. & Batterham, P. (2013). Gene duplication in the major insecticide target site, *Rdl*, in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. U. S. A.*, 110, 14705–10.

Roger, S.O. & Bendich, A.J. (1988). Extraction of DNA from plant tissues. In: (eds. Gelvin, S.B. & Schilperoort, R.A.). Kluwer Academic Publishers, Boston, pp. 1–10.

Rousset, F. (2008). genepop'007: a complete re-implementation of the genepop software for Windows and Linux. *Mol. Ecol. Resour.*, 8, 103–106.

Schrider, D.R. & Hahn, M.W. (2010). Gene copy-number polymorphism in nature. *Proc. Biol. Sci.*, 277, 3213–3221.

Schrider, D.R., Houle, D., Lynch, M. & Hahn, M.W. (2013). Rates and genomic consequences of spontaneous mutational events in *Drosophila melanogaster*. *Genetics*, 194, 937–954.

Spofford, J.B. (1969). Heterosis and the evolution of duplications. *Am. Nat.*, 103, 407–432.

Strickman, D. & Fonseca, D.M. (2012). Autogeny in *Culex pipiens* complex mosquitoes from the San Francisco Bay area. *Am. J. Trop. Med. Hyg.*, 87, 719–726.

Tamura, K., Stecher, G., Peterson, D., Filipski, A. & Kumar, S. (2013). MEGA6: Molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.*, 30, 2725–2729.

Tatem, A.J., Rogers, D.J. & Hay, S.I. (2006). Estimating the malaria risk of African mosquito movement by air travel. *Malar. J.*, 5, 57.

Weill, M., Berthomieu, A., Berticat, C., Lutfalla, G., Nègre, V., Pasteur, N., Philips, A., Leonetti, J.P., Fort, P. & Raymond, M. (2004a). Insecticide resistance: a silent base prediction. *Curr. Biol.*, 14, R552-3.

Weill, M., Fort, P., Berthomieu, A., Dubois, M.-P.P., Pasteur, N. & Raymond, M. (2002). A novel acetylcholinesterase gene in mosquitoes codes for the insecticide target and is non-homologous to the ace gene in Drosophila. *Proc. R. Soc. B-Biological Sci.*, 269, 2007–2016.

Weill, M., Lutfalla, G., Mogensen, K., Chandre, F., Berthomieu, A., Berticat, C., Pasteur, N., Philips, A., Fort, P. & Raymond, M. (2003). Insecticide resistance in mosquito vectors. *Nature*, 423, 423–426.

Weill, M., Malcolm, C., Chandre, F., Mogensen, K., Berthomieu, A., Marquine, M. & Raymond, M. (2004b). The unique mutation in *ace-1* giving high insecticide resistance is easily detectable in mosquito vectors. *Insect Mol. Biol.*, 13, 1–7.

Yébakima, A., Marquine, M., Rosine, J., Yp-Tcha, M.M. & Pasteur, N. (2004). Evolution of resistance under insecticide selection pressure in *Culex pipiens quinquefasciatus* (Diptera, Culicidae) from Martinique. *J. Med. Entomol.*, 41, 718–725.

Data Accessibility Statement

GenBank accession numbers for *ace-1* resistance allele sequences are listed in Table 2. All the data (life history traits, *ace-1* susceptible sequences) and R scripts are available from the Dryad Digital Repository: https://doi.org/10.5061/dryad.b31g6.

Authors' contributions

Conceptualization, P.L., P.Mi. and M.W.; Methodology, P.L., P.Mi. and M.W.; Formal analysis, P.L. and P.Mi.; Investigation, P.L., P.Mi., S.U., A.B., C.M.A., P.Ma. and B.S.A.; Ressources, C.M.A., B.S.A. and N.P.; Writing - Orignal Draft, P.L. and P. Mi.; Writing - Review &Editing, M.W, B.S.A., C.M.A. and N.P; Supervision: P.L. and M.W.; Funding Acquisition: P.L. and M.W.

Competing interests

The authors declare that they have no competing interest.

Table 1: Nomenclature of the different alleles and genotypes at the *ace-1* locus.

Alleles and copies		Code	Allele structure
Single-copy susceptible allele	*ace-1*^{Sx}	S_x	- S -
Single-copy resistance allele	*ace-1*^{Rx}	R_x	- R -
Duplicated allele	*ace-1*^{Dx}	D_x	- R – S -
Susceptible copy in a duplication		D_x(S)	- R – **S** -
Resistance copy in a duplication		D_x(R)	- **R** – S -

Genotypes		Code	Phenotype
Homozygotes	*ace-1*^{Sx} / *ace-1*^{Sx}	S_xS_x	[SS]
	ace-1^{Rx} / *ace-1*^{Rx}	R_xR_x	[RR]
	ace-1^{Dx} / *ace-1*^{Dx}	D_xD_x	[RS]
Heterozygotes	*ace-1*^{Rx} / *ace-1*^{Sx}	R_xS_x	[RS]
	ace-1^{Dx} / *ace-1*^{Sx}	D_xS_x	[RS]
	ace-1^{Dx} / *ace-1*^{Rx}	D_xR_x	[RS]
	ace-1^{Dx} / *ace-1*^{Dy}	D_xD_y	[RS]

For each considered allele, copy or genotype, the code used in the paper is presented. The structure of each allele is schematized (note that the relative position of the S and R copies in a duplicated D allele is unknown). The molecular assay phenotype of each genotype is indicated; it is identical to the resistance phenotype. Note that the phenotype of an individual carrying a duplicated allele is always heterozygote [RS].

Table 2: *ace-1* resistance alleles.

Culex pipiens					
	ace-1^{R}		Sample	Ref.	Accession #
R_1	-	-	4, 5, 7-9, 17-29	Weill et al. 2003	KT002464 / 65
R_4	-	-	16	This study	KT002468
ace-1^{D}	D_i(S)	D_i(R)	Sample	Ref.	Accession #
ace-1^{D2}	D_2(S)	D_2(R) = R_1	4, 5, 6	Labbé et al. 2007a	JX007768 / 69
ace-1^{D3}	D_3(S)	D_3(R) = R_1	4	Labbé et al. 2007a	JX007770 / 71
ace-1^{D6}	D_6(S)	D_6(R) = R_1	5	Labbé et al. 2007a	JX007776 / 77
ace-1^{D7}	D_7(S)=D_{13}(S)	D_7(R) = R_1	29	Alout et al. 2010	JX007778 / 79
***ace-1*^{D8}**	**D_8(S)**	**D_8(R) = R_1**	**28**	**Alout et al. 2010**	**JX007780 / 81**
ace-1^{D9}	D_9(S)	D_9(R) = R_1	27, 28	Alout et al. 2010	JX007782 / 83
ace-1^{D10}	D_{10}(S)	D_{10}(R) = R_1	27	Alout et al. 2010	JX007784 / 85
ace-1^{D11}	D_{11}(S)	D_{11}(R) = R_1	26	Alout et al. 2010	JX007786 / 87
ace-1^{D12}	D_{12}(S)	D_{12}(R) = R_1	24, 26	Alout et al. 2010	JX007788 / 89
ace-1^{D13}	D_{13}(S)=D_7(S)	D_{13}(R) = R_4	12, 14	Osta et al. 2012	JX007790 / 91
ace-1^{D15}	D_{15}(S)	D_{15}(R) = R_1	20, 22-25	This study	KT002448 / 65
ace-1^{D16}	D_{16}(S)=D_{22}(S)	D_{16}(R) = R_1	21-25	This study	KT002449 / 65
ace-1^{D17}	D_{17}(S)	D_{17}(R) = R_1	22, 24	This study	KT002450 / 65
ace-1^{D18}	D_{18}(S)	D_{18}(R) = R_1	18, 19, 22	This study	KT002451 / 65
ace-1^{D19}	D_{19}(S)	D_{19}(R) = R_1	18-22, 24, 25	This study	KT002452 / 65
***ace-1*^{D22}**	**D_{22}(S) = D_{16}(S)**	**D_{22}(R)**	**24**	**This study**	**KT002455 / 56**
ace-1^{D23}	D_{23}(S)	D_{23}(R) = R_1	5	This study	KT002457 / 65
ace-1^{D27}	D_{27}(S)	D_{27}(R) = R_4	12	This study	KT002462 / 63
Culex quinquefasciatus					
	ace-1^{R}		Sample	Ref.	Accession #
R_2	-	-	31, 32, 35-41, 50, 66, 69	Weill et al. 2003	KT002466
R_3	-	-	49	Labbé et al. 2007	KT002467
ace-1^{D}	D_i(S)	D_i(R)	Sample	Ref.	Accession #
***ace-1*^{D1}**	**D_1(S)**	**D_1(R) = R_2**	**35-38, 66**	**Labbé et al. 2007**	**JX007766 / 67**
ace-1^{D4}	D_4(S) = D_5(S)	D_4(R) = R_3	49	Labbé et al. 2007	JX007772 / 73
***ace-1*^{D5}**	**D_5(S) = D_4(S)**	**D_5(R)**	**33, 35, 65**	**Labbé et al. 2007**	**JX007774 / 75**
ace-1^{D14}	D_{14}(S)	D_{14}(R) = R_2	30	This study	KT002447 / 66
ace-1^{D20}	D_{20}(S)	D_{20}(R) = R_2	35, 38	This study	KT002453 / 66
ace-1^{D21}	D_{21}(S)	D_{21}(R) = R_2	38	This study	KT002454 / 66
ace-1^{D24}	D_{24}S)	D_{24}(R) = R_2	66	This study	KT002458 / 66
***ace-1*^{D25}**	**D_{25}(S)=D_{26}(S)**	**D_{25}(R)**	**54, 64**	**This study**	**KT002459 / 60**
ace-1^{D26}	D_{26}(S)=D_{25}(S)	D_{26}(R) = R_2	35	This study	KT002461 / 66

All resistance alleles described so far at the *ace-1* locus are listed, both for single-copy (*ace-1*^{Rx}) and duplicated alleles (*ace-1*^{Dx}, with D_x(S) and D_x(R), respectively the susceptible and resistance copies). D_x(R) = R_x means that the D(R) copy of the *ace-1*^{Dx} allele is identical to an *ace-1*^{Rx} allele. For each resistance allele, are indicated: the sample(s) where it was found (numbers as in Tab. S1), the associated reference (Ref.) and the GenBank accession numbers. D alleles where the D(S) and D(R) copies are strictly identical (except for the G119S mutation) are bolded.

Table 3: *ace-1^D* detection, new strain establishment and reference strains.

Taxa	Geographic Area	Sampled populations	Map #	Duplicated alleles		
				Detected	Lost	Isolated (new strain)
Cx. quinquefasciatus	Martinique	Pool	66	D_1, D_{24}	-	D_{24}
	Mayotte	Acoua	35	D_1, D_5, D_{20}, D_{26}	D_5, D_{20}, D_{26}	-
		M'Tsamoudou	38	D_1, D_{20}, D_{21}	D_{20}, D_{21}	-
		Tsoundsou	36	D_1		D_1*
	Togo	Baguida	33	D_5		D_5
Cx. pipiens	Tunisia	Al Battan	23	D_{15}, D_{16}	-	D_{16}
		Utique	20	D_{15}, D_{19}	-	D_{15}
		Hamra	24	D_{15}, D_{16}, D_{19}	-	D_{19}
		Djedaida	22	D_{15}, D_{16}, D_{18}, D_{19}	D_{18}	-
	Greece	Heraklion	12	D_{13}, D_{27}	-	D_{13} D_{27}

Reference strains

Taxa	Geographic Area	Sampled populations	Map #	Isolated	Strains	References
Cx. quinquefasciatus	Martinique	Ducos	66	D_1*	DUCOS-DFix	Labbé et al. 2014
Cx. pipiens	France	Maurin	5	D_2	MAURIN-DFix	Labbé et al. 2007b
		Ganges	4	D_3	BIFACE-DFix	Labbé et al. 2007b

* Two stains carry the same D_1 allele (DUCOS-DFix isolated from Martinique, Labbé et al. 2007a, and Mayotte, this study).

Several populations were sampled and analyzed to find duplicated D alleles and establish pure strains. For each sample, we give the taxa, the geographic area, the name of the sampled locality and its number (Tab. S1), the various identified D alleles and those lost during the isolation process.

Table 4: Phenotypic and fitness consequences of different *ace-1* duplicated alleles.

Alleles	W_{DD} (SL)	Development time		Preimaginal mortality		
		First / Last	χ^2 (*df*)	*Overall (%)*	f_{DD} ±95% CI (*N*) *L_2 / adults*	χ^2 (*df*)
D_1 [a]	0.73 (0.63-0.83)	-	-	-	-	-
D_1 [b]	-	0.44 / 0.47	0.0 (1) [n.s.]	13.5	0.5 ±0.14 (48) / 0.52 ±0.11 (83)	3.10^{-4} (1) [n.s.]
D_2	0.27 (0.17-0.35)	-	-	-	-	-
D_3	0.14 (0.00-0.29)	-	-	-	-	-
D_5	0.32 (0.22-0.42)	-	-	-	-	-
D_{13}	0.00 (0.00-0.09)	-	-	-	-	-
D_{15}	-	0.00 / 1.00	92 (1) ***	53.1	0.5 ±0.14 (48) / 0.13 ±0.1 (45)	12.7 (1) ***
D_{16}	0.13 (0.00-0.28)	0.17 / 0.56	12 (1) ***	45.8	0.45 ±0,14 (47) / 0.38 ±0.13 (52)	0.18 (1) [n.s.]
D_{19}	0.02 (0.00-0.19)	0.17 / 0.53	11 (1) ***	47.9	0.33 ±0,13 (48) / 0.2 ±0.11 (50)	1.6 (1) [n.s.]
D_{24}	0.76 (0.63-0.89)	0.85 / 0.80	0.1 (1) [n.s.]	15.6	0.71 ±0.13 (45) / 0.84 ±0.08 (81)	2.2 (1) [n.s.]
D_{27}	0.04 (0.00-0.13)	-	-	-	-	-

[a] Martinique; [b] Mayotte

For each D_x allele, W_{DD} is the fitness of the homozygote D_xD_x ($W_{DD} = 1 + s$), relative to the fitness of the heterozygote D_xS_{SLAB} ($W_{DS} = 1$), with the associated support limits (SL), estimated from the population cage dynamics (for the estimation method see text).

For the development time, *First* and *Last* are the frequencies of D_xD_x individuals respectively among the 48 first and 48 last emerged individuals; they were compared using a proportion test, the χ^2 statistics of which is given with the associated degrees of freedom *df* and significance ([n.s.], $p > 0.05$; *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$).

For the preimaginal mortality, the overall percentage is indicated. The frequencies of the DD genotype (f_{DD}) in L_2 larvae and surviving adults is given ± its 95% confidence interval (95% CI), with the number of genotyped individuals (*N*) in brackets. For each D allele, the two frequencies were compared using a proportion test (χ^2 statistics, *df* and significance are indicated).

Figure legends:

Figure 1: Worldwide diversity of *ace-1* sequences.

The *ace-1* diversity, single-copy alleles (R_x and S_x) and duplicated copies (D_x(S) and D_x(R)), is presented using Jukes-Cantor distance (ClustalW). The G119S mutation (inducing resistance) has been removed to consider only neutral variations. *Cx. pipiens* and *Cx. quinquefasciatus* alleles are respectively in the blue and red clades. Bootstraps are indicated as percentages (10,000 bootstraps), only if > 50%. For each allele/copy, the sample(s) where it was found is indicated; colors refer to different geographic zones: Western Europe (green), Middle-East (yellow), Maghreb (red), sub-Saharan Africa (brown), Asia/Oceania (blue) and the Americas (violet) (see also Tab. S1).

[a] see Tab. 2 for a detailed list of the populations harboring the D alleles.

Figure 2: Worldwide distribution of *ace-1* alleles.

Each sample is indicated by its number (see Tab. S1 for details) and a symbol corresponding to the taxon (see legend). Resistance alleles are indicated (D_x and R_x, respectively for *ace-1*^{D} and *ace-1*^{R}) in red (dashed line) for *Cx. quinquefasciatus* and in blue (dotted line) for *Cx. pipien*s. Resistance alleles found in multiple samples are bolded.

Figure 3: D alleles experimental evolution.

Left panel: Dashed lines represent the dynamics of the various duplicated D alleles over the six discrete generations under insecticide selection (D_1, dark-blue empty circles; D_2, green triangles; D_3, pink plus; D_5, brown full circles; D_{13}, light-blue full squares; D_{16}, red diamonds; D_{19}, orange

downward triangles; D_{24}, dark-green crosses and D_{27}, violet stars). The dotted line represents the initial frequency ($f(DD)_1 = 0.5$). **Right panel:** The final frequencies ($f(DD)_6$) with their associated 95% confidence intervals are represented for each duplicated allele. The dotted line represents the initial frequency ($f(DD)_1 = 0.5$). The $f(DD)_6$ statistical groups among the tested strains are also indicated (^{ns}, $p > 0.05$; ***, $p < 0.001$).

Figure 4: Homozygous-sublethal (HS) D alleles dynamics.

The HS D alleles invasion dynamics was computed using a population genetic model. It starts from a treated population containing only R and S at overdominance equilibrium; at generation one and every fifty generations a new HS D allele is introduced. HS D allele frequencies, cumulated ($f_{tot}(D)$, solid lines) and for each allele ($f_i(D)$, colored dotted lines), as well as the population average fitness (μW, dashed lines) are represented over the generations. **A:** single HS D allele invasion dynamics depending on the RR genotype relative fitness w_{RR}. **B:** invasion dynamics of several successive HS D alleles fully ($w_{DxDy} = 1$, B1) or only partially ($w_{DxDy} = 0.5$, B2) complementing.

Supporting information

Table S1: Sampled mosquito populations.

Table S2: Molecular tests

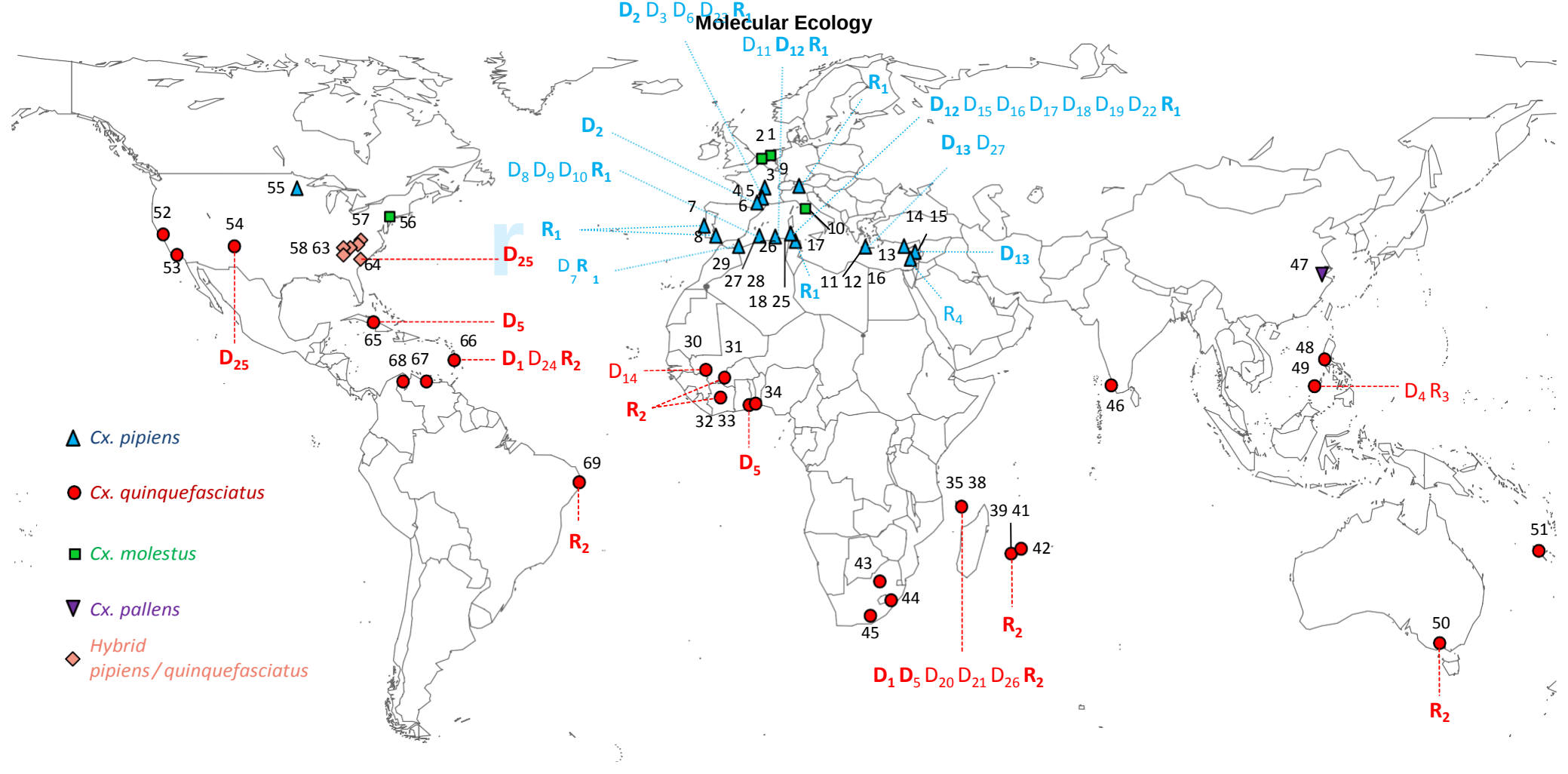
Table S3: Primer sequences

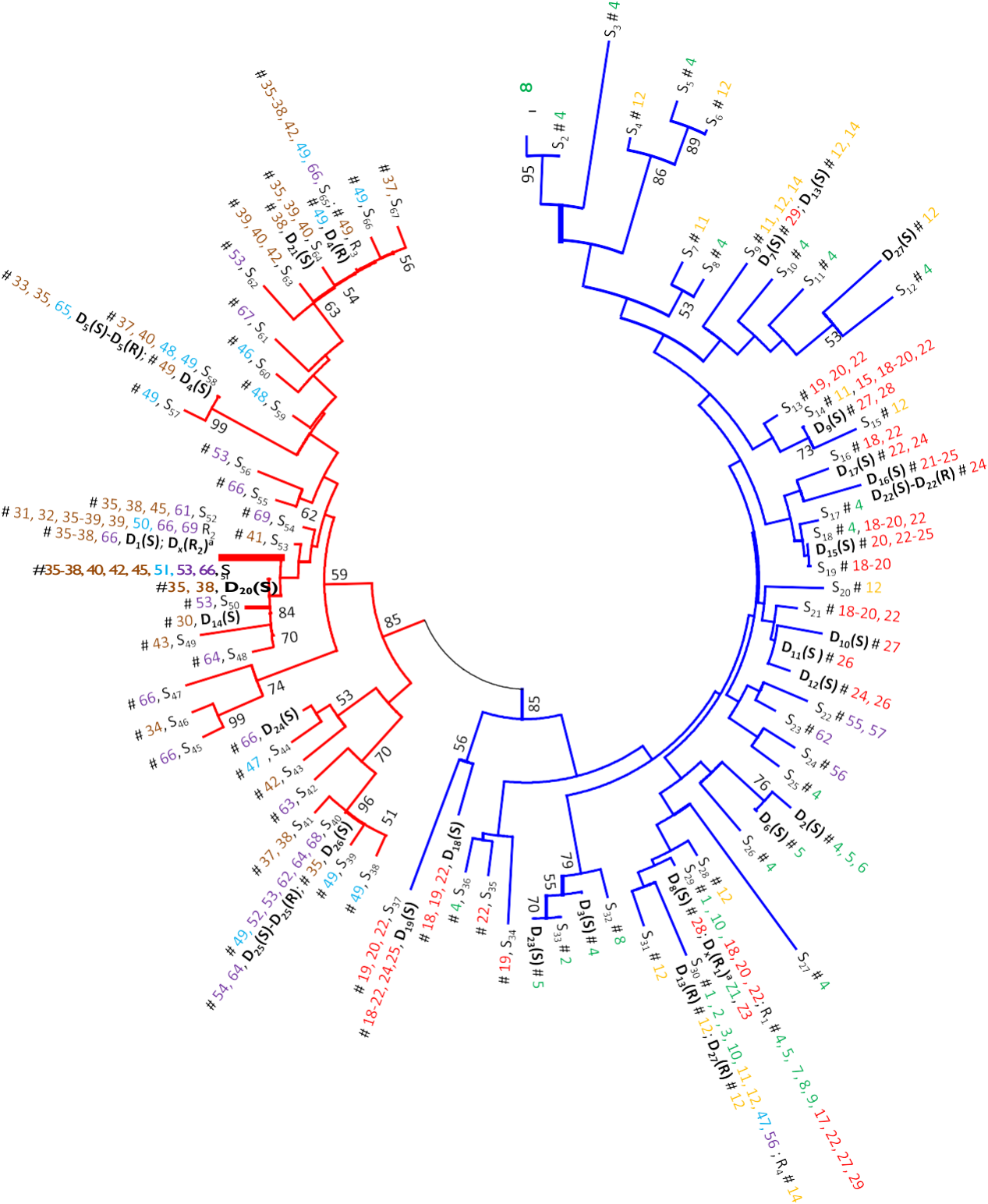
Figure S4: *Cx. quinquefasciatus* and *Cx. pipiens ace-1* large fragment sequencing.

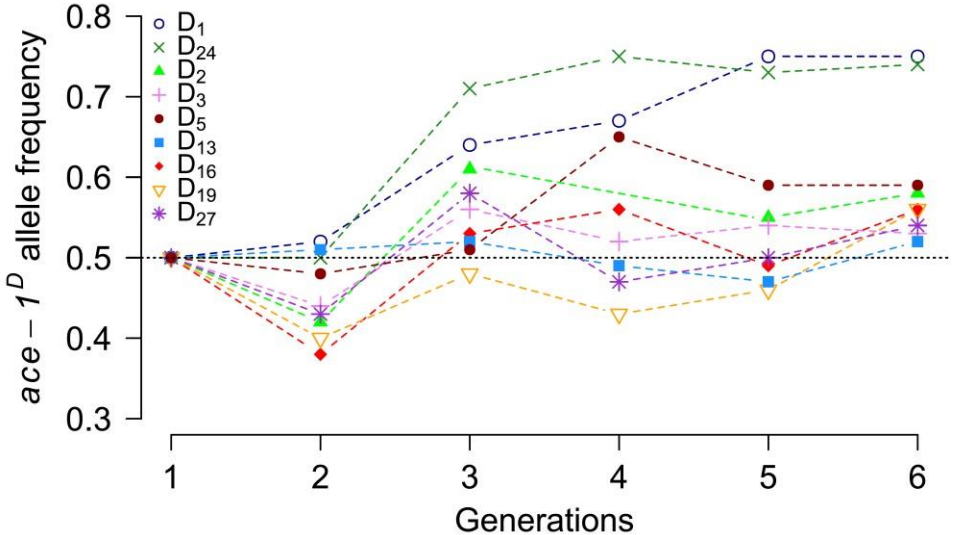
Figure S5: Correlation between nuclear hybridization index and taxa-specific *ace-1* alleles.

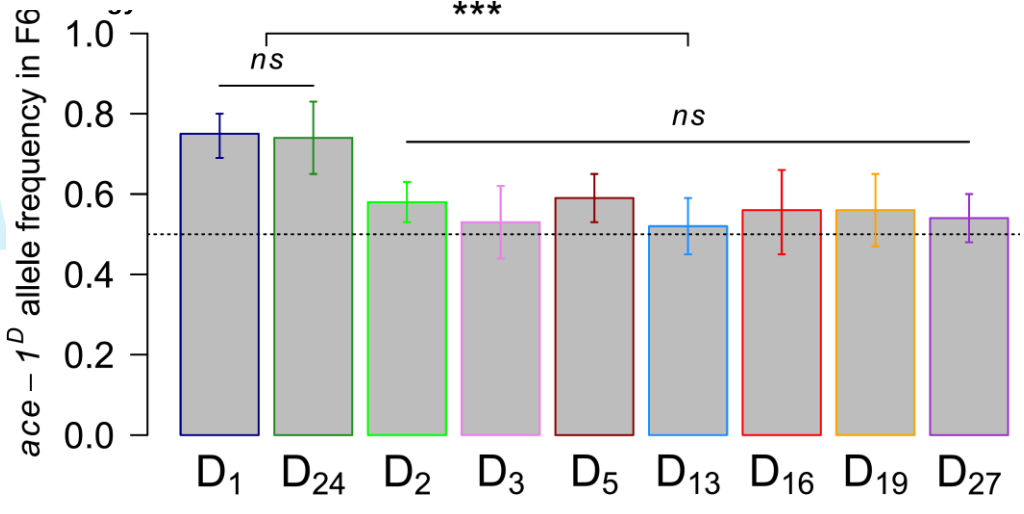
Table S6: Diversity of *ace-1* single-copy R alleles and of each copy of the duplicated D alleles.

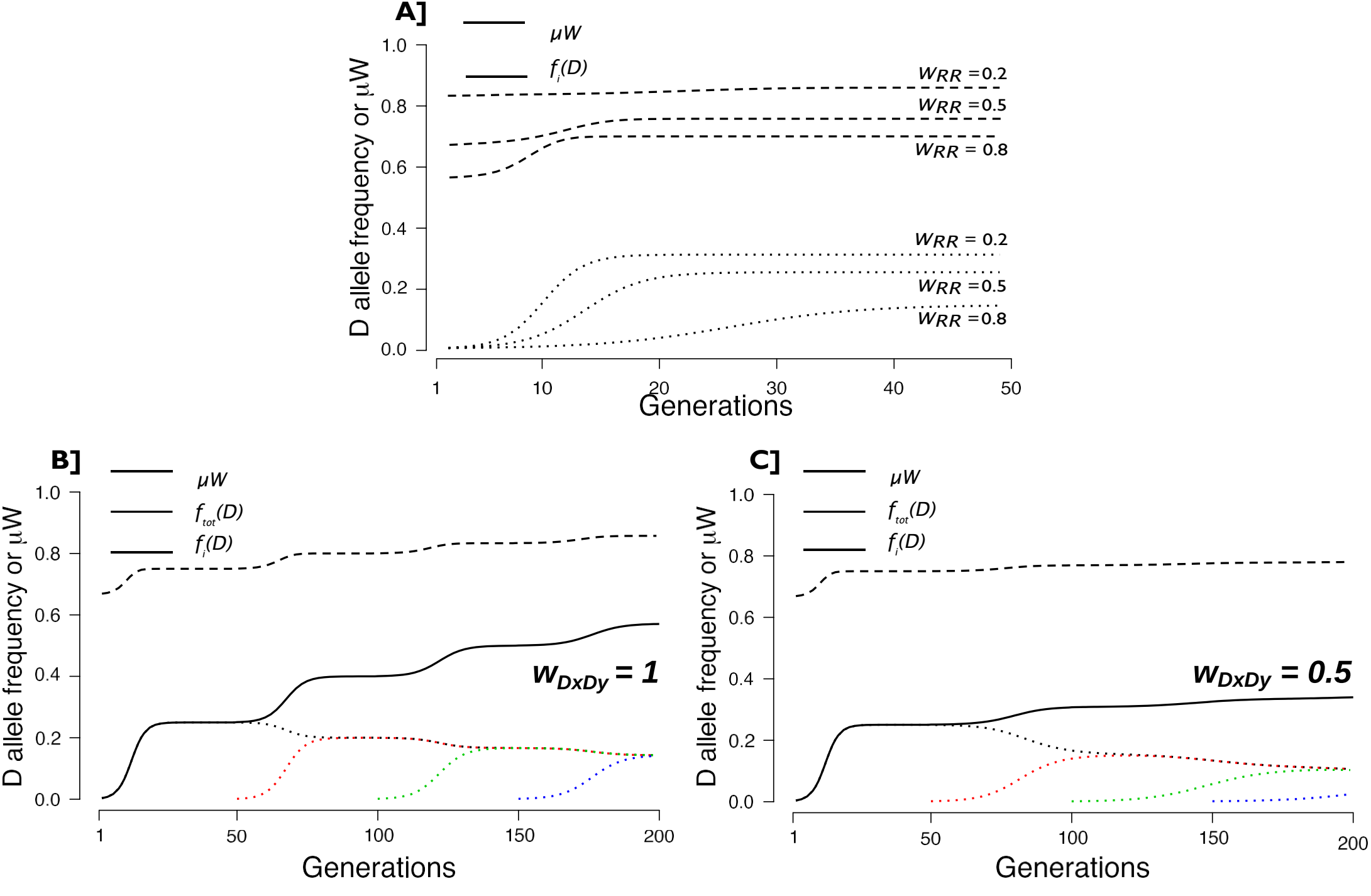
Figure S7: Complementation experiment.


Cx. pipiens
Cx. quinquefasciatus
Cx. molestus
Cx. pallens
Hybrid
pipiens / quinquefasciatus
D2 D3 D6 D23 R1
D11 D12 R1
R1
D12 D15 D16 D17 D18 D19 D22 R1
D2
D13 D27
D8 D9 D10 R1
D13
R1
D25
D7 R1
R1
R4
D5
D25
D1 D24 R2
D14
R2
D5
D4 R3
R2
D1 D5 D20 D21 D26 R2
R2
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 25 26 27 28 29 30 31 32 33 34 35 38 39 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 63 64 65 66 68 67 69


S3 # 4
S2 # 4
S4 # 12
S5 # 4
S6 # 12
S7 # 11
S8 # 4
S9 # 11, 12, 14
D7(S) # 29; D13(S) # 12, 14
S10 # 4
S11 # 4
D27(S) # 12
S12 # 4
S13 # 19, 20, 22
S14 # 11, 15, 18-20, 22
D9(S) # 27, 28
S15 # 12
S16 # 18, 22
D17(S) # 22, 24
D16(S) # 21-25
D22(S)-D22(R) # 24
S17 # 4
S18 # 4, 18-20, 22
D15(S) # 20, 22-25
S19 # 18-20
S20 # 12
S21 # 18-20, 22
D10(S) # 27
D11(S) # 26
D12(S) # 24, 26
S22 # 55, 57
S23 # 62
S24 # 56
S25 # 4
D2(S) # 4, 5, 6
D6(S) # 5
S26 # 4
S27 # 4
S28 # 12
S29 # 1, 10, 18, 20, 22; R1 # 4, 5, 7, 8, 9, 17, 22, 27, 29
D8(S) # 28; Dx(R1)a 21, 23
S30 # 1, 2, 3, 10, 11, 12, 47, 56; R4 # 14
D13(R) # 12; D27(R) # 12
S31 # 12
S32 # 8
D3(S) # 4
S33 # 2
D23(S) # 5
S34
19, S35
22, S36
4, S36
18, 19, 22, D18(S)
19, 20, 22, S37
18-22, 24, 25, D19(S)
49, S38
49, S39
35, D26(S)
49, 52, 53, 62, 64, 68, S40
54, 64, D25(S)-D25(R); # 37, 38, S41
63, S42
42, S43
47 , S44
66, D24(S)
66, S45
34, S46
66, S47
64, S48
43, S49
30, D14(S)
53, S50
#35, 38, D20(S)
#35-38, 40, 42, 45, 51, 53, 66, S51
31, 32, 35-39, 39, 50, 66, 69 R2
35-38, 66, D1(S); Dx(R2)a
35, 38, 45, 61, S52
41, S53
69, S54
66, S55
53, S56
49, S57
33, 35, 65, D5(S)-D5(R); # 49, D4(S)
48, S59
46, S60
67, S61
53, S62
39, 40, 42, S63
35, 39, 40, S64
38, D21(S)
49, R3
49, D4(R)
49, S66
35-38, 42, 49, 66, S65
37, S67
95
8
86
89
53
53
99
73
62
59
85
84
70
74
99
53
70
96
51
56
79
55
70
76
63
54
56
85


0.8
0.7
0.6
0.5
0.4
0.3
ace – 1^D allele frequency
D1
D24
D2
D3
D5
D13
D16
D19
D27
1
2
3
4
5
6
Generations


***
ns
ns
1.0
0.8
0.6
0.4
0.2
0.0
ace – 1^D allele frequency in F6
D1
D24
D2
D3
D5
D13
D16
D19
D27


A]
μW
$f_i(D)$
w_{RR} =0.2
w_{RR} =0.5
w_{RR} = 0.8
w_{RR} = 0.2
w_{RR} =0.5
w_{RR} =0.8
D allele frequency or μW
Generations
B]
μW
$f_{tot}(D)$
$f_i(D)$
W_{DxDy} = 1
D allele frequency or μW
Generations
C]
μW
$f_{tot}(D)$
$f_i(D)$
W_{DxDy} = 0.5
D allele frequency or μW
Generations