A Computer Simulation Model of Gene Replacement in Vector Populations

Mauricio Guevara and Edgar E. Vallejo

***Abstract*— In this paper, we conduct a series of computer simulations on the genetic modification of disease vectors. In our model, a population of mosquitoes is modeled as DNA sequences. A gene drive mechanism based on simulated transposable elements is inserted into DNA sequences to confer immunity to individuals. Different scenarios on the spread and prevalence of modified genes are simulated using the evolutionary game theory. Preliminary results suggest that the proposed model could be useful for identifying a set of population-level conditions that a disease control strategy based on genetic modification of vectors should satisfy to be effective.**

I. INTRODUCTION

The development of effective strategies for controlling vector borne diseases such as malaria and dengue is a major epidemiology concern worldwide. In effect, the malaria parasite *Plasmodium falciparum* which is transmitted by mosquitoes kills over two million people per year –mostly children, in Africa alone [1][2].

Earlier malaria disease control strategies relied on the extensive use of insecticide to reduce vector populations. Although initially effective, the abuse of insecticide contributed rapidly to the emergence of resistant strains. Further, owing to the subsequent propagation of insecticide resistance these disease control strategies are no longer considered effective. For instance, the general use of the DDT pesticide has been already banned in the US a few decades ago [1].

Recent advances in molecular biology and bioengineering have contributed widely to our increasing understanding of the machinery of life at the molecular level [3][4]. As a result, our abilities to manipulate fundamental biological processes have improved greatly. For instance, the genetic modification of organisms to confer disease refractoriness is now within the reach of these technological advances [5].

A disease control strategy based on the genetic modification of mosquitoes to confer disease immunity is considered promising [6]. For instance, the Gates Foundation has recently provided generous funding for the development and implementation of a disease control strategy based on the synthesis of disease resistant strains of the mosquito *Anopheles gambiae*, the principal vector of the malaria parasite.

Although the creation of individual disease resistant strains is now possible, in principle, there exist additional necessary conditions that must be satisfied by this disease control strategy to be effective. For example, the refractory gene should confer some fitness advantage with respect to their wild counterpart, or else it will likely be removed from the gene pool within a few generations. Moreover, the fitness differential should be sufficiently important so as to increase the possibilities of gene replacement in wild populations [7].

Besides the uncertainty on the fixation of the refractory gene, there is much concern on the environmental consequences of the release of genetically engineered mosquitoes [1]. As a result, the implementation of this strategy often comprises a preliminary stage of controlled experiments conducted within cages. These experiments are not only expensive but most importantly, the controlled environment provided by these cages could lead to spurious conclusions related to fitness and thus replacement potential of the modified strains.

Moreover, multiple scenarios on the spread and prevalence of resistant genes in populations would be difficult to study directly given the life span of *Anopheles gambiae* (about a month). Both mathematical and computer simulation models on population genetics have proven to be useful tools for overcoming the inherent limitations of experimental studies on genetic variation [8][9][10].

Particularly, computer simulation represents a complementary approach for the study of population dynamics [11]. Despite the obvious limitations of working with abstract models, the simulation approach is often more tractable than experimental population studies. Further, computer simulations are excellent platforms for testing alternative scenarios on relevant population variables under consideration [12][13]. Moreover, simulation results may be useful as confirmatory evidence of the predictions derived from the population genetics theoretical models [14].

In this work, we propose a computer simulation model to study the genetic modification of disease vectors at the population level. In our model, a population of vectors is modeled as DNA sequences. A gene drive mechanism based on simulated transposable elements is introduced in DNA sequences to confer immunity to individuals. Different scenarios for the spread and prevalence of modified genes are simulated using the evolutionary game theory. Simulation results suggest that this model could be useful for identifying a set of conditions at the population level that a disease control strategy based on the genetic modification of vectors should satisfy in order to be effective.

This work was supported by a Consejo Nacional de Ciencia y Tecnología (CONACYT) Award number SEP-204-C01-47434.

Mauricio Guevara and Edgar E. Vallejo are with the Computer Science Department, ITESM, Campus Estado de México, Atizapán de Zaragoza, México {A00456476,vallejo}@itesm.mx

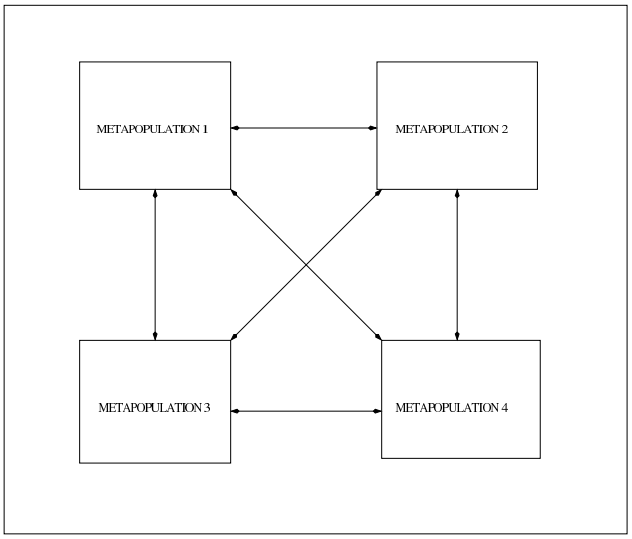

Fig. 1. Mosquito population structure

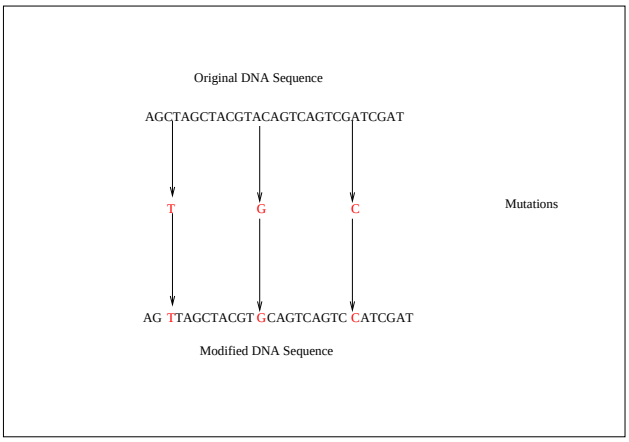

Fig. 2. Mutation of DNA sequence

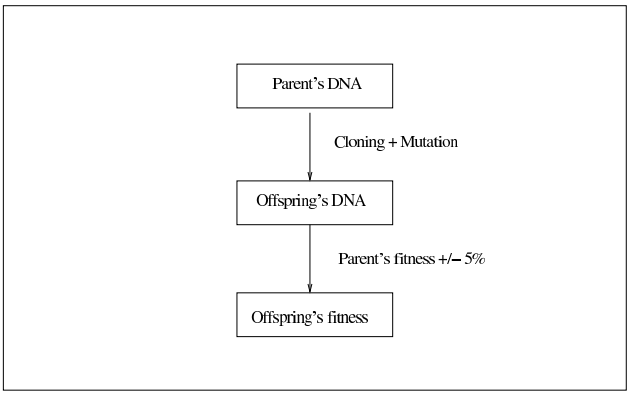

Fig. 3. Reproduction of individuals

II. MODEL DESCRIPTION

A. Population variables simulated in the computer model

The proposed computer model simulates a subset of the most important variables that explain genetic variation in populations. Particularly, we considered the following parameters in the simulations: fitness, age, migration, reproduction, mutation, insertion and duplication of a transposable element, and the ability of competition for resources.

The population of mosquitoes is organized as a collection geographically dispersed metapopulations as shown in figure 1. Each metapopulation is generated at random and then, an arbitrary fitness value is ascribed uniformly to each individual in the metapopulation.

The migration occurred once the mosquitoes have reached two weeks old of simulated age. For this purpose, a probability distribution was created to determine whether a mosquito remains in the same metapopulation or migrate to a neighbor metapopulation. The probability distribution is intended to model the geographical distance among metapopulations.

The mutation process consists of the sparse substitution of nucleotides in the DNA sequence as shown in figure 2. For the simulations reported here, we set the mutation probability to 0.01. We used a simple montecarlo algorithm to determine whether or not a mutation event occurs at a particular locus.

Reproduction is simulated as a cloning-mutation process. We use the DNA sequence of a randomly selected parent and mutate it to obtain the DNA sequence of the offspring as shown in figure 3. The fitness of the offspring is calculated from the fitness of the parent adding an small perturbation to that. Once the reproduction is completed, the parent is removed from the population.

To insert a transposable element, we choose a random position in the ADN sequence, cut the sequence, insert the transposon at the selected position and then paste the remaining sequence to the right of the transposon (see Figure 4). After the insertion is completed, the fitness of the mosquito is slightly reduced.

The main feature of a transposon is that it can reproduce itself and change its position in the genome [6]. This behavior is simulated in the computer model, and each time the mosquito obtains a new copy of the transposon, its fitness is slightly reduced.

The proposed model also incorporates a turnover process. In particular, we predefined a maximum number of individuals in the population. If the population size surpasses this limit, the oldest individuals are removed from the population until its size is less or equal to the maximum. The turnover was considered due to the bounded availability of computer resources for the simulations. Similarly, populations do not grow unboundedly in the wild because they are tied to the availability of natural resources [15].

The competition for resources is simulated using the evolutionary game theory [16]. At each simulation step, the fitness of all individuals is penalized simulating the need of a particular resource. Then, each individual competes for a resource with another randomly selected individual.

According to the evolutionary game theory, we considered two types of behavior: the *hawk* and the *dove*. The interaction between these two types of behavior is as follows:

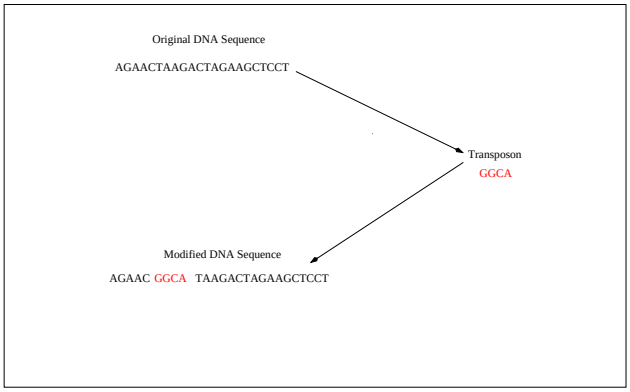

Fig. 4. Insertion of the transposable element

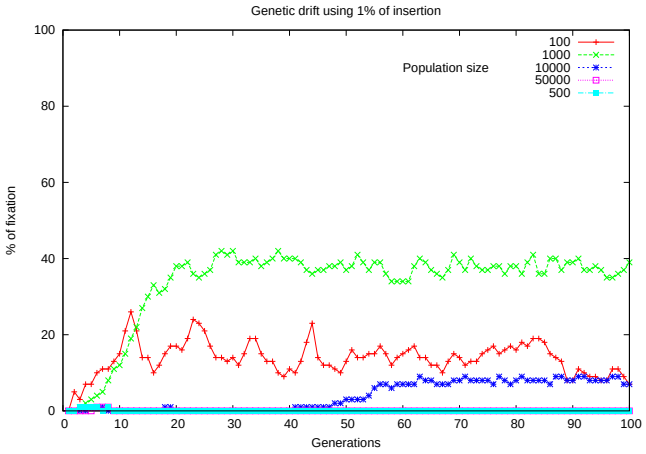

Fig. 5. Effects of genetic drift with 1% of transgenic mosquitoes

- **Hawk-hawk:** When two individuals with hawk behavior compete for a resource, the fitness of both individuals is decreased. The hawks are aggressive and fight for the resource, so both get hurt as a result of the competition.
- **Hawk-dove:** The dove behavior is a cautious one. When a dove and a hawk behavior compete for a resource, the dove escapes to avoid the consequences. In this type of interaction, the hawk gets the resource to its own and his fitness is increased.
- **Dove-dove:** When two doves compete for a resource they share it, so every dove obtains half of the resource and both fitnesses are increased.

B. Simulated scenarios

The computer model maintains a data structure to describe each individual of the population. Each individual consists of a 100-nucleotide long DNA sequence, an ID of its parent, its own ID, its genealogy, geographical origin, age, transposons acquired during its life time.

In all of the simulations described here we used four metapopulations. Each metapopulation was inicialized with one individual that serves as the common ancestor. At each simulation step, all individuals are subject to the aforementioned transformation processes. A total of 100 generations were simulated in all of the experiments.

Initially, there exist two different populations in the model: (1) a population of wild individuals and (2) a population of transgenic individuals. Each population is evolved separately and when the wild individuals reach a predefined population size we combine both populations to observe their interaction in terms of gene flow.

Finally, at the end of each simulation step, we enumerate the number of individuals that have acquired the transposon and the number of individuals that does not have it in order to quantify the fixation rate of the transposon.

III. EXPERIMENTS AND SIMULATION RESULTS

In this paper we conducted four types of different experiments. The first experiment is aimed at estimating the population size required to sufficiently reduce the deleterious effect of genetic drift so as to increase the possibility of the fixation of the refractory gene. The second experiment is intended to determine the number of transgenic mosquitoes that should be inserted in the native population in order to achieve prevalence of the transposon over a number of generations. The third experiment simulates the effect of the fitness differential of transgenic mosquitoes on the fixation of the modified gene. Finally, we use the evolutionary game theory to simulate the interactions among native and transgenic mosquitoes during resource competition. We decided to conduct all of the experiments serially as it would be extremely complex to analyze all the relevant variables simultaneously.

A. Estimation of native population size

The objective of this experiment is to determine the population size needed to minimize the deleterious effect of genetic drift for the consistent fixation of the transposon.

We considered two different populations, the native one and the transgenic one. In this experiment, the occurrence of the transposon does not change the fitness of the mosquito.

We observed the fixation of the transposon at different population sizes. We used native population of 100, 500, 1,000, 10,000 and 50,000 individuals. In addition, we used 1%, 5% and 10% as the proportion of transgenic mosquitoes inserted in the native population. Simulation results for different combination of these parameters are shown in figures 5, 6 and 7.

In our simulations, a small proportion of transgenic mosquitoes did not produce the fixation of the transposon in large populations. However, the fixation of the transposon was sparsely observed in small populations as a result of genetic drift (see figure 5). The inconsistency observed on the preservation of the modified gene suggested that a larger proportion of transgenic mosquitoes should be inserted in the native populations.

Figure 6 shows the effect of genetic drift when 5% of transgenic mosquitoes are inserted in the native population. The fixation of the transposon was consistently observed

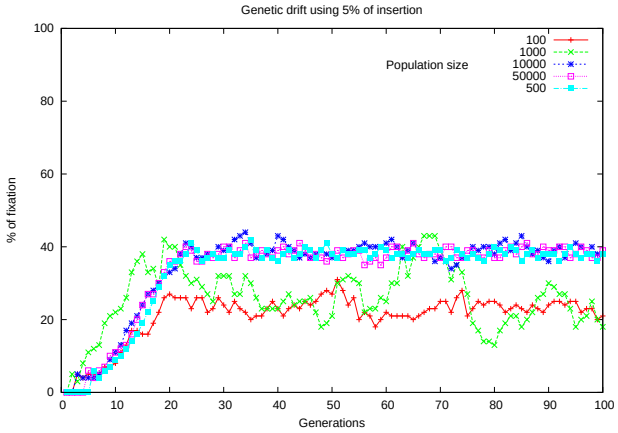

Fig. 6. Effects of genetic drift with 5% of transgenic mosquitoes

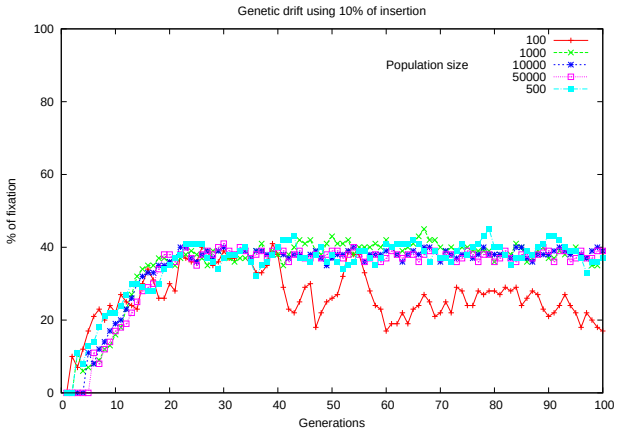

Fig. 7. Effects of genetic drift with 10% of transgenic mosquitoes

at various population sizes. However, the prevalence of the modified gene seems to stabilize as the population size increases. In addition, we observed a high variability on the frequency of he transposon when small native populations are considered in repeated experiments.

Figure 7 shows an even more consistent fixation of the transposon when 10% of transgenic mosquitoes are inserted in the native population. However, the modified gene prevalence was inconsistent when small native populations were considered.

Theoretical results of genetic drift indicate that the deleterious effect of genetic drift is inversely proportional to population size [17]. The simulation experiments presented here confirmed the predictions derived from the theoretical results.

B. Estimation of transgenic population size

This experiment was intended to determine the proportion of transgenic mosquitoes required for the consistent fixation of the modified gene in the native population.

The scenario was similar to the previous experiment. Again, we assume that the occurrence of the transposon have

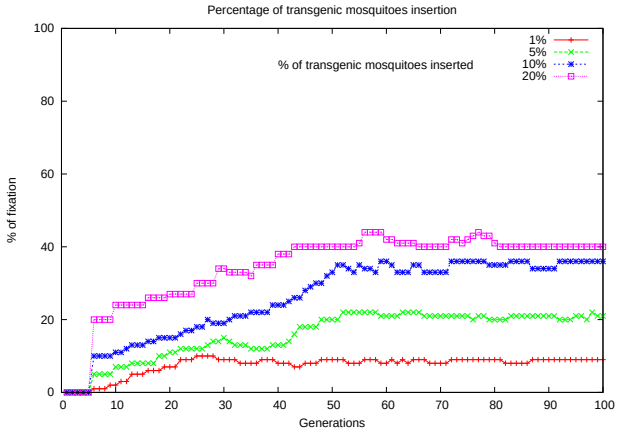

Fig. 8. Proportion of transgenic mosquitoes inserted in the native population

a silent effect on fitness of the mosquitoes. We used a population size of 100,000 individuals as suggested from the previous experiment. The proportion of transgenic mosquitoes inserted in the native population was 1%, 5%, 10% and 20%, respectively. Figure 8 shows the simulation results obtained in this experiment.

As predicted, the frequency of the modified gene increased monotonically with the proportion of transgenic mosquitoes inserted in the native population. However, no obvious difference was observed when more than 10% of this proportion was considered. Moreover, the frequency of transposons obtained with a 5% was relatively consistent, so this would be a better alternative than producing larger quantities of modified mosquitoes without apparent benefits.

C. Effects of fitness differential on the transgenic population

The objective of this experiment was to observe effects of the fitness differential conferred by the occurrence of the transposon on the prevalence of modified mosquitoes. We considered an decrement/increment of -10%, -5%, 0%, 5% and 10% on fitness each time the transposon was acquired/duplicated.

According to theoretical results on population genetics, it is reasonable to predict that deleterious genes would likely be removed from the gene pool as a result of natural selection [18]. For this experiment we used a population of 100,000 individuals and the percentage of transgenic mosquitoes was 5%.

Figure 9 shows that the frequency of the transposon was proportional to the fitness differential. In nature, the possession of transposable elements often decrease the fitness of an individual. However, fitness is very difficult to assess as it is a function of not only the individual's genotype but also a function of the environment [15][19].

As shown in figure 9 a transposon that decreases the fitness of the carrier in 5% could reach a fixation of 20% of the total population, which is a reasonable percentage of fixation. We can conclude that even if the transposon diminishes the

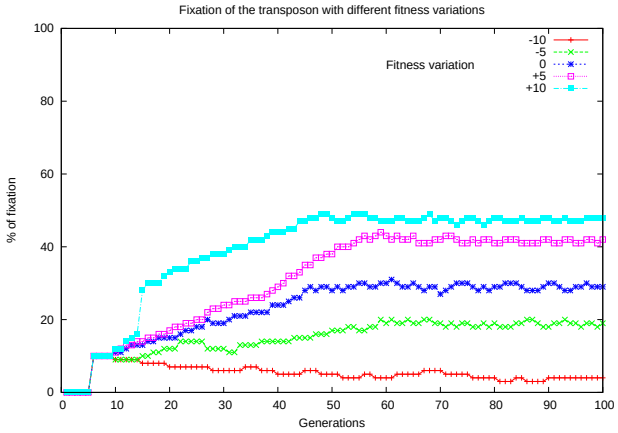

Fig. 9. Effects of various fitness differentials on transposon frequency

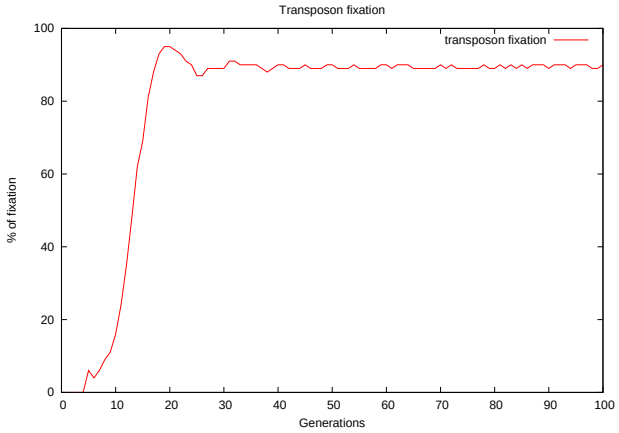

Fig. 10. Frequency of the transposon, scenario 1

	hawk	dove
hawk	-5%	20%
dove	0%	5%

TABLE I
SCENARIO 1

	hawk	dove
hawk	-20%	5%
dove	0%	25%

TABLE II
SCENARIO 2

fitness of the mosquitoes, it could be fixed in the native population.

D. Evolutionary game theory for modeling competition for resources

In this experiment, we observed the effect of the modified gene on the interaction of transgenic and native mosquitoes when competing for resources. To simulate the behavior among individuals of the population we used the evolutionary game theory [16]. The behavior of each mosquito with respect to another will be different depending on whether or not it possesses the transposon. We simulated two types of behavior, the hawk and the dove. In our simulations, the native mosquitoes behave like doves and the transgenic mosquitoes behave like hawks.

We considered the following scenarios to observe the degree of prevalence of the modified gene.

1) Scenario 1: Table 1 shows the interaction matrix that we used for this scenario.

As observed in this matrix we know that when two hawks compete for the resource, both hawks will be penalized with 5% of their fitness. If a hawk and a dove are selected for the contest, the hawk will increase his fitness in 20% while the doves fitness remains the same. Finally, when two doves are competing for the resource, both of their fitness will increase in 5%.

The following equations are used to determine if a population has reached an evolutionary stable state:

$$W(J,I) < W(I,I)$$

$$W(J,I) = W(I,I)$$

By solving the equations for this scenario we obtain:

$$W(hawk, dove) = 20$$

$$W(dove, dove) = 5$$

Using the framework provided by the evolutionary game theory we can observe that the native population would not reach an evolutionary stable state, meaning that it could be invaded by the transgenic population.

As can be seen in figure 10 the fixation of the transposon increases rapidly in early generations until it stabilizes at about generation 25 where it remains constant. This result confirms the predictions derived from the evolutionary game theory. In effect, as the native population was not in an evolutionary stable state, it was invaded by the transgenic population.

This behavior is a consequence of the favorable environmental conditions defined for the hawk as it receives little damage when encounters another hawk and a great reward when faces a dove. In contrast, the interaction among doves is relatively favorable, but very deleterious when interacting with a hawk.

2) Scenario 2: The conditions of this experiment were similar to those of the previous experiment with the exception of the interaction matrix as shown in Table II.

Again, by solving the equations for this scenario we obtain:

$$W(hawk, dove) = 5$$

$$W(dove, dove) = 25$$

The equations are solved so as the native population would reach an evolutionary stable state. Therefore, it could not be invaded by the transgenic population, in principle.

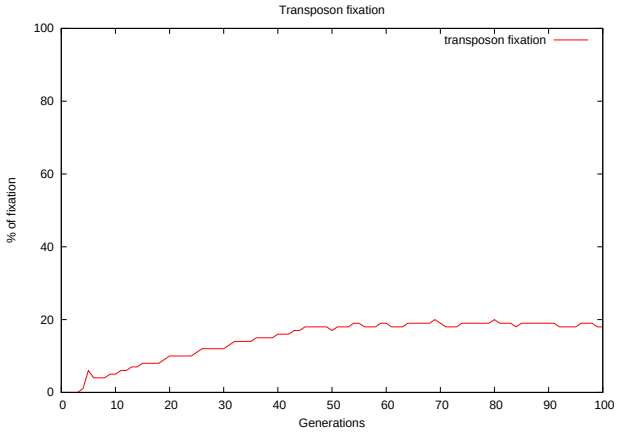

Fig. 11. Frequency of the transposon scenario 2

As shown in Figure 11, the frequency of the transposon was relatively small over generations; although it was not removed from the native population. This result is consistent with the predictions derived from the evolutionary game theory when a population achieves an evolutionary stable state.

Particularly, the transgenic mosquitoes almost experienced extinction due to the unfavorable fitness consequences of behaving like hawks. On the one hand, the fitness penalty of competing with another hawk was considerably damaging. On the other hand, the fitness reward when competing with a dove was relatively small. In contrast, a dove got an important reward when competing with another dove and since the native population was larger than the transgenic population, the prevalence of transgenic population under this scenario seemed unlikely.

IV. DISCUSSION

There is much hope in developing a disease control strategy based on the genetic modification of disease vectors within the next decade. Unfortunately, insect borne pathogens have had devastating effects on African populations over the last decades [20].

Although promising in principle, there is still much uncertainty that such a disease control strategy would be effective. The reason is that many uncontrolled biotic and abiotic variables would come into play during its deployment in the wild. Moreover, there is no reason to believe that the genetic modification of organisms would not have catastrophic environmental consequences (such as the emergence of highly efficient multiple disease vectors, for example). Careful considerations on ethical, social and legal issues will be crucial for the realization of such a scientific feat [22].

Computer simulations are useful tools for studying the potentials of such a disease control strategy and for establishing a set of conditions under it would be effective. There is, by all means, an extensive gap between the reality and the abstractions presented here. Still, we believe computer simulations are capable of representing fundamental aspects of many biological phenomena and will be increasingly useful for supporting the study of fundamental questions on the biology of organisms [21].

V. ACKNOWLEDGMENTS

We thank Professor Charles E. Taylor of the Ecology and Evolutionary Biology Department, University of California, Los Angeles for all of his invaluable comments and advise during the conception and development of this work.

REFERENCES

[1] Levy, S., “Mosquito modifications: New approaches to controlling malaria”, *BioScience*, vol. 57 pp. 816-821, 2006.
[2] Zhong, D., “Dynamics of gene introgression in the malaria vector *Anopheles gambiae*”, *Genetics: Published Articles Ahead of Print*, pp. 1-27, 2006.
[3] Brown, T., *Genomes*, Wiley-Liss, 1994.
[4] Wang, R., “Microsatellite markers and genotyping procedures for *Anopheles gambiae* parasitology”, *Today*, 2004.
[5] Deceliere, G., “The Dynamics of transposable elements in structured populations”, *Genetics Society of America*, vol. 169 pp. 467-474, 2005.
[6] Kidwell, M., “Transposable elements as sources of variation in animals and plants”, *Proceedings of the National Academy of Sciences*, vol. 94 pp. 7704-7711, 1997.
[7] Deceliere, G., “TESD: A transposable element dynamics simulation environment ”, *Bioinformatics*, vol. 22 pp. 2702-2703, 2006.
[8] Le Rouzic, A., “Models of the population genetics of transposable elements”, *Genetic Research*, pp. 171-181, 2005.
[9] Clough, J., “Computer simulation of transposable elements evolution: Random template and strict master model”, *Molecular Evolution.*, vol. 42 pp. 52-58, 1996.
[10] Manoukis, N., “Detecting recurrent extinction in a metapopulation of *Anopheles gambiae*: Preliminary results using simulation”, *WSEAS Transactions on Systems.*, 2004.
[11] Langton, C., *Artificial Life*, The MIT Press, 1997.
[12] Taylor, C. E., Jefferson, D. R., Turner, S. R., and Goldman, S. R., “RAM: Artificial life for the exploration of complex biological systems”. In Langton, C. G., (ed.) *Artificial Life*, pp. 275-316, 1989, Addison-Wesley. pp. 1-26, 2004.
[13] Carnahan, J., Li, S. Constantini, C., Touré, Y. and Taylor, C. E., “Computer simulation of dispersal by Anopheles gambiae in West Africa”. In Langton, C. G., and Shimohara, K., (eds.) *Artificial Life V*, pp. 387-394, 1997, The MIT Press.
[14] Wilson, W., *Simulating Ecological and Evolutionary Systems*, Cambridge University Press, 2000.
[15] Hartl, D., *Principles of Population Genetics*, Sinauer, 1997.
[16] Smith, J., *Evolution and the Theory of Games*, Cambridge University Press, 1982.
[17] Guillespie, J. H., *Populaton Genetics. A concise Guide. Second Edition*, The Johns Hopkins University Press, 2004.
[18] Metter, L., *Population Genetics and Evolution*, Prentice Hall, 1998.
[19] Menge, D., “Fitness consecuences of *Anopheles gambiae* population hybridization,” *Malar Journal*, 1977.
[20] Chen, C., “A synthetic maternal-effect selfish genetic element drives population replacement in *Drosophilia*” *Science*, vol. 316 pp. 597-600, 2007.
[21] Krane, D., *Fundamental Concepts of Bioinformatics*, San Francisco, 2003.
[22] Taylor, C. E. *Personal communication.*