10

Effective population size in relation to genetic modification of *Anopheles gambiae sensu stricto*

Charles E. Taylor[#] *and Nicholas C. Manoukis*[#]

Abstract

Effective population size (N_e) is a single number which allows us to relate a large part of population-genetics theory built around ideal populations to real populations. In the case of *Anopheles gambiae*, N_e is important to interpreting the temporal and spatial distribution of genes. These patterns are in turn used to explore the structure of (typically non-ideal) natural populations. We discuss the complex structure of *An. gambiae s.s.* in and around Banambani, Mali, as it is currently understood, based on estimates of N_e. This reveals a population that is structured temporally between and within years, spatially between villages and non-dimensionally into chromosomal forms. We suggest that the subpopulations of this species might usefully be viewed as a metapopulation. Successful and efficient genetic modification of *An. gambiae* will require as complete an understanding of their population structure as possible, which we believe can be attained through the convergence of multiple population-genetic techniques and the application of new methods.

Keywords: effective population size; N_e; *Anopheles gambiae*; population structure; Mali

Introduction

The abstract ideal for most theory in population genetics and ecology is a panmictic population – a single, randomly mating group of individuals in which the population size stays constant, all offspring have an equal chance to reproduce and there is no geographic variation in gene or genotype frequencies. Like frictionless surfaces or perfect vacuums in theoretical physics, the ideal panmictic population plays an important role in theoretical population genetics, though no such ideal population actually exists. Rather, all real populations are in some way(s) structured. That is to say, they consist of subpopulations that are finite in size, may or may not mate randomly, may or may not share migrants among one another, and probably change size over time. Making inferences or predictions about these populations requires that their structure be understood and described in a manner commensurate with existing theory. The notion of effective population size, N_e, plays an especially central role for describing this population structure. N_e may be defined to be the size of an ideal population that exhibits the same rate of drift as the actual population it

[#] Department of Organismic Biology, Ecology and Evolution, University of California, Los Angeles, CA 90095-1606 E-mail: taylor@biology.ucla.edu; manoukis@ucla.edu

characterizes. As Futuyma (1998) puts it: "if we count 10,000 adults in a population but only 1,000 of them successfully breed, genetic drift proceeds at the same rate as if the population size were 1,000, and this is the effective size". Alternatively, as discussed later in this chapter, there are huge seasonal differences in the population size of *An. gambiae*. The harmonic mean of these sizes is N_e. N_e was introduced by Sewall Wright, with important subsequent contributions by G. Malécot, J. F. Crow, M. Kimura and others. See Wright (1969), Crow and Kimura (1970) and Kimura and Ohta (1971) for useful and accessible discussions.

In this chapter we are concerned with N_e as it can be used to describe and make inferences about the population structure of *Anopheles gambiae s.s.*, with an emphasis on its utility for guiding efforts to introduce a genetically modified vector to control malaria. After a few introductory remarks about taxonomic questions we will describe the population biology of *An. gambiae* in one location that has been studied for several decades. With these components to anchor the discussion we will examine how N_e and associated parameters of population structure are measured generally, and their values estimated specifically at our focal location. We then mention some new theoretical developments that are likely to assist our understanding. Finally, we conclude with some recommendations about what, in our opinion, are the outstanding research questions about population structure that must be investigated prior to an attempt to genetically modify natural *Anopheles* populations.

The population structure of *An. gambiae* is quite complex. The highest taxonomic level in the system is *An. gambiae sensu lato* (s.l.) that comprises at least 7 species, one of which is *An. gambiae sensu stricto* (s.s.). *An. gambiae* s.s. in turn has as many as 5 different "chromosomal forms." In some locations (e.g. Mali), distributions of chromosomal forms coincide with "molecular forms" which can be distinguished with polymerase chain reaction (PCR) assays (Favia et al. 1997; 2001). This is not, however, the case in all locations (Della Torre et al. 2002; 2001). Gene flow across the forms is limited, so this is an important part of the population structure. For a comprehensive survey of the chromosomal forms of *An. gambiae s.s.* see Touré et al. (1998b), and for a recent description of the population structure around Banambani, Mali, see Taylor et al. (2001).

Our principal focus in this paper is on *Anopheles gambiae* s.s. in the village of Banambani, Mali, and its surrounding area, where three chromosomal forms are present – termed *Bamako*, *Savanna* and *Mopti*. There are at least three types of structure that relate to N_e in this area. The first is the extent of population-size changes over the year and between years. The second is the structure imposed by chromosomal forms. Finally there is geographic structure because the species inhabits discrete patches (villages), among which gene flow is limited. A significant understanding of the complex structure in the study area is critical for the successful introduction of genetically modified *An. gambiae*. It is important to know, for example, how many individuals will need to be released in order to alter the population, how fast the introduced genetic element will spread from the site of release and if it will move from one form or species into another. As we will see below, even simple models for known features of population structure in this species lead to quite complex patterns of gene flow.

Structure of *Anopheles gambiae s.s.* at Banambani

Mark-release-recapture (MRR) studies conducted at peak seasonal abundance during a six-year period indicate that there is much variation in the size of local

populations of *An. gambiae s.s.* Annual abundance estimates varied at least 4-fold, as shown in Table 1, from Taylor et al. (2001). Compounding the variability of population size between years, the density of *An. gambiae s.s.* changes dramatically within a single year. For example, a survey in non-irrigated villages near the Niono irrigation projects (Dolo et al. 1999) found peak numbers of bites in villages during the wet season to be more than 1000 times the number recorded during the dry season.

Table 1. Population sizes from 1993 to 1998 estimated from mark-release-recapture experiments at the village of Banambani, Mali (from Taylor et al. 2001). All measurements were made at the peak of seasonal abundance (August-September).

Year	*Replicates*	*Released*	*Recaptured*	*Daily Survival*	*Estimated N*
1993	3	938	83	0.80	20,178
1994	4	1,913	57	0.80	64,002
1996	4	1,421	44	0.92	63,006
1997	2	1,002	24	0.97	53,400
1998	4	1,205	21	0.97	79,280

Seasonality seems to vary from place to place; near our study site in Banambani we have estimated dry-season population sizes to be 5 to 10% of the wet-season peak (Taylor et al. 2001, Touré, personal communication). Assuming a more or less exponential growth and decline between the seasonal low and highs, a typical annual cycle in the *An. gambiae s.s.* population at Banambani is depicted in Figure 1.

Incidence of the chromosomal forms of *An. gambiae s.s.* also vary during the year. All three forms, Bamako, Savanna and Mopti are about equally numerous during the wet-season peak, in late August to early September, but during the drier part of the year the Savanna and Bamako forms decrease in numbers, so Mopti individuals comprise nearly all of the mosquitoes that can be collected then. Based on the data of Touré et al. (1998b) the numbers of each form are approximately as shown in Figure 2.

The three forms found at each location are thought to exchange genes among themselves, though how much and the significance of this is still unclear. Nonetheless, some exchange certainly occurs. The extent of this is further discussed in Taylor et al. (2001). Based on the number of hybrids observed, the number of migrants from one form to another is estimated to be in the order of 0 - 11% depending on the forms exchanging genes and the manner in which gene flow was estimated (Touré et al. 1998b; Tripet et al. 2001). There is currently insufficient information to determine whether gene flow is symmetrical or not.

Finally, there appears to be some movement of individuals among villages. Touré et al. (1998a) and Dolo (2000) report on MRR studies where released females were captured in neighbouring villages, but the numbers were low and the confidence limits large. As would be expected, there appears to be less gene exchange among villages farther apart than among those closely together (Johnson 1969). Carnahan et al. (2002) collated all of the available microsatellite DNA studies and found a linear relation between distance and log gene flow (i.e. $Log\ N_e m$), as expected from theoretical considerations (see Table 2, below, and Kimura and Ohta 1971). In this case m is the {number of migrants entering a subpopulation}/{the size of that subpopulation}. For distances a few kilometres apart, the typical inter-village distance around Banamabani, m is about 0.008-0.039.

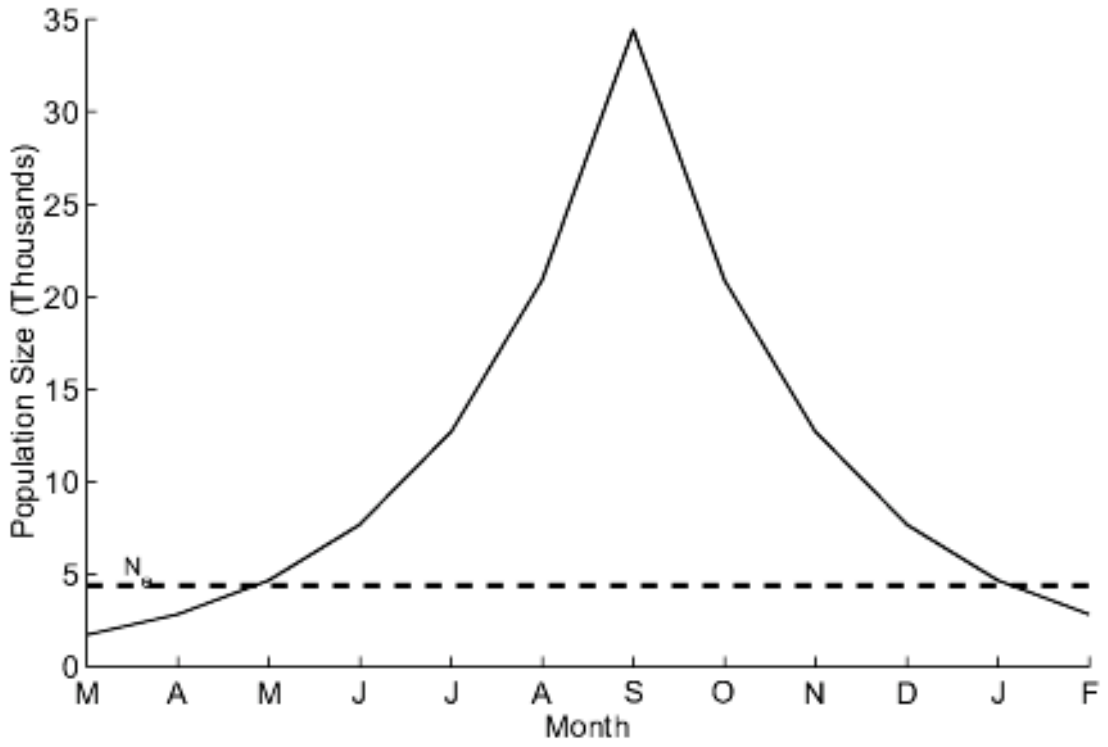

Figure 1. Estimated seasonal change in size of the adult population of *Anopheles gambiae s.s.* in Banambani, Mali. The broken line represents the value of N_e, calculated as per equation (1), below.

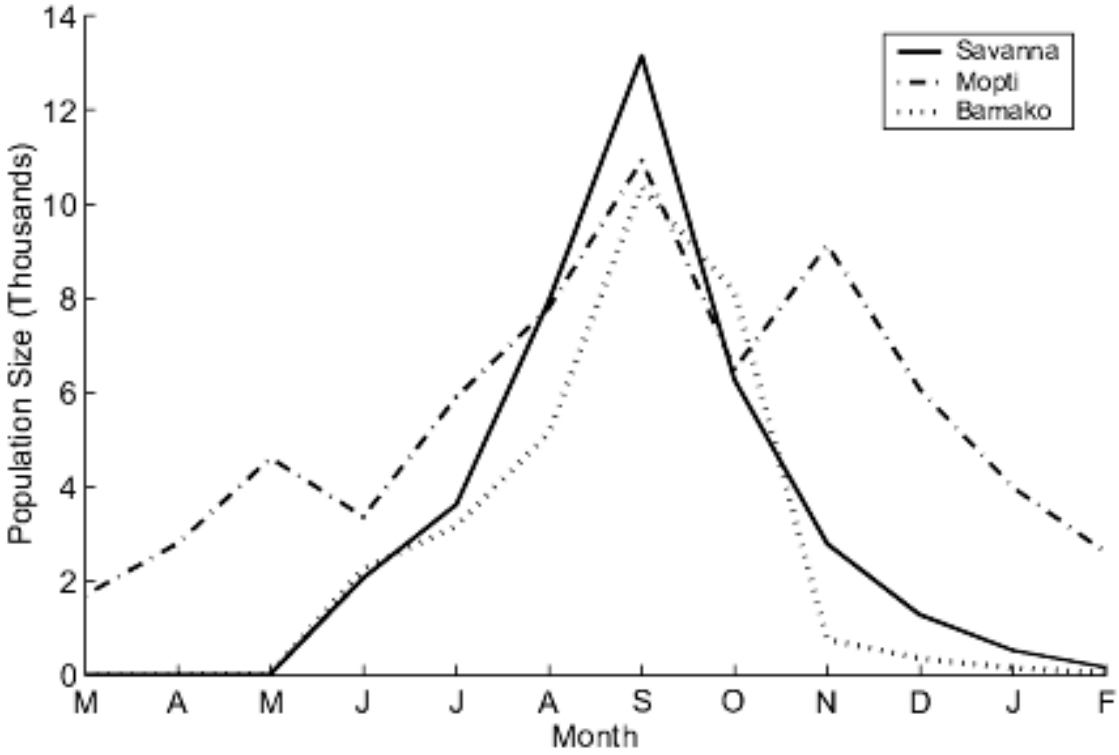

Figure 2. Estimated size of the adult populations of each of three chromosomal forms of *Anopheles gambiae s.s.* collected in and around Banambani, Mali.

Putting all this together, the full population structure around Banambani is summarized in Figure 3, according to the best information available.

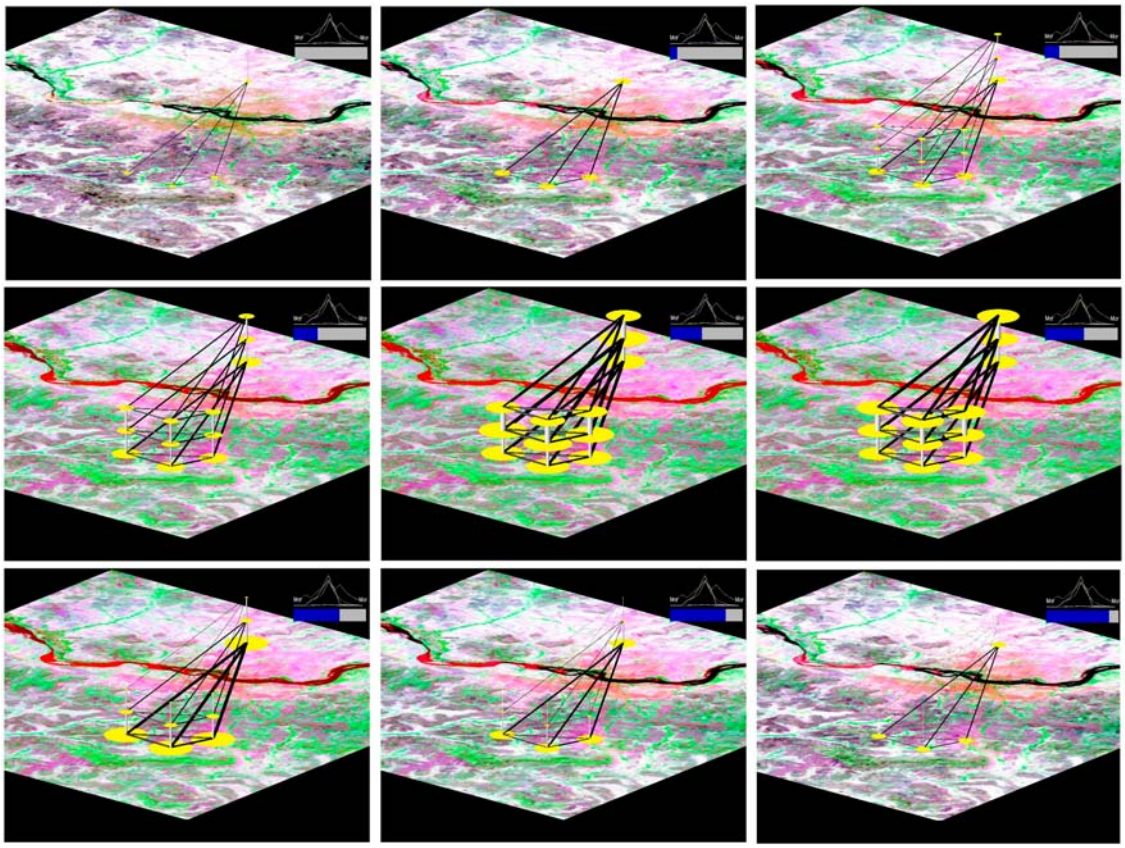

Figure 3. Summary of the population structure of the three chromosomal forms around Banambani, Mali. The series begins in the dry season (March) and proceeds through the following dry season (February). The population sizes of the three chromosomal forms are shown as discs, with the area proportional to the size of the population. The discs at each location represent (from bottom to top): Mopti, Savanna and Bamako forms. Locations in the map shown hosting populations are those sampled by Touré et al. (1998a). Villages between them are not shown. All connecting pipes represent gene flow; white is between chromosomal forms and black is between villages. The diameter of the pipe is proportional to the magnitude of migration. The time bar in the upper right corner of each frame represents one year from March to March. The full movie from which these frames were excerpted is available online at http://taylor0.biology.ucla.edu/~manoukis/Structure.

Measuring population structure

Several methods have been used to measure the principal features of population structure, population sizes and migration rates. These have been classed as “direct” or “indirect” by Slatkin (1985), and correspond to “ecological” and “genetic” by Taylor and Powell (1983). Here we follow Slatkin’s terminology.

Direct measures of population structure consist of those measures of population properties themselves. For example, several groups have conducted MRR studies of adult female *An. gambiae* (see Costantini et al. 1996; Touré et al. 1998a). Relative density can be estimated from how many bites a sedentary person receives during fixed periods (Dolo et al. 1999), and larval dippings have been used to estimate numbers of immatures (Service 1993). Movement and migration are inferred by

recaptures from distances or villages away from a release site and interbreeding between forms by the numbers of hybrids that can be detected. Service (1993) has an extended discussion about application of these methods for mosquitoes; more general discussions of these topics can be found in Blower, Cook and Bishop (1981), Johnson (1969) and Turchin (1998).

Indirect measures of population structure are the methods of primary interest here. A particular history of size change and gene flow should give rise to corresponding patterns of allele frequencies in space and/or time. If those patterns can be appropriately measured, then it is frequently possible to work backwards and infer the population structure that gave rise to the pattern. Some population-genetic patterns that have been measured to infer the structure and history of populations are shown in Table 2.

Exact specification of the models and precise definitions of the variables is not possible here; the references given should be consulted for these. Very briefly, n_e refers to effective number of alleles and θ is a measure of genetic diversity derived from allele frequencies by Watterson (1975), often used in coalescent theory. These both depend on the product of effective population size and neutral mutation rate, μ. There are several models of gene flow that have received theoretical attention; these differ about the importance of long-distance (m or m_∞), or short-distance (m_1) migration and whether there are discrete populations of size N_e or if they are distributed with a constant density. Distance, whether continuous (x) or discrete population steps (ρ), affects how correlated are populations to one another. F_{ST} is a correlation of allele frequencies across subpopulations, closely related to $r(\rho)$ and $\phi(x)$. (N_e is defined for these measures and could be incorporated into the equations, but we have chosen to leave them in their more traditional forms.) Because large populations are expected to drift more than small populations, change in allele frequency (p) from generation to generation, measured by $\sigma^2(p)$, will depend on effective population size. And finally, when gene trees are constructed, the time, t, to the most recent common ancestor, MRCA, will depend on effective population size and mutation rate.

Table 2. Some indirect methods used to study patterns of genetic variation. Key: HC = Hartl and Clark (1997); H = Hudson (1990); FL = Fu and Li (1999); CK = Crow and Kimura (1970); KO = Kimura and Ohta (1971); A = Anderson, Williamson and Thompson (2000); W = Waples (1989).

Measure	*Formulae*	*Reference*
1. Amount of polymorphism	$n_e = 4N_e\mu + 1$	HC
	$\theta \approx 4N_e\mu$	H; FL
2. Differentiation of populations		CK; KO
Island model	$F_{ST} = 1/(4N_e m + 1)$	KO
Isolation-by-distance model	$\phi(x) \propto \mathrm{Exp}[-x(2\mu)^{1/2}]/\sqrt{\sigma}$	KO
Stepping-stone model	$r(\rho) \propto \mathrm{Exp}[-\rho(4m_\infty/m_1)^{1/2}]/\sqrt{\rho}$	KO
3. Temporal changes in allele frequency	$\sigma^2(p) \approx p(1-p)[1-(1-1/2N_e)]$	A; W
4. Coalescence	$t(\mathrm{MRCA}) \approx N_e m$	H; FL

Patterns of detectible genetic variation shown in Table 2 are all sensitive to the structure and history of the population being studied. Taken together with estimates of N_e, these provide insight into how populations of *An. gambiae* are structured, though there may be more than one interpretation for any particular pattern.

Estimating N_e

In the equations in Table 2 the estimates of effective population size can be taken as similar to the actual population size, N, but because of the extensive structure already discussed, N is likely to be a poor approximation of the effective population size. Recall from above that N_e is defined to be the size of an ideal population that exhibits the same rate of drift as the actual population (Crow and Kimura 1970). Such drift might affect inbreeding, variance in gene frequencies, or rate of extinction of alleles. Consequently, one can distinguish *inbreeding*-effective population size, *variance*-effective population size, and *eigenvalue*-effective population size or *extinction*-effective population size (Crow 1956). In most cases these will be similar to one another and no distinction need be made, but in some instances they can differ substantially (see Kimura and Ohta 1971).

Figure 1 illustrates how an ideal population may drift at the same rate as the average drift of an actual population. When the population size is low, as at the beginning and end, then the allele frequencies will drift rapidly, but when the population size is large, then the drift will be slower. Between these extremes there must be an average rate that can be calculated by noting that the rate of drift at time t is proportional to $1/N_t$. The average rate of drift, described by $1/N_e$, is seen to be:

$$(1/N_e) = (1/k)\ (1/N_1 + 1/N_2 + \ldots + 1/N_k) \qquad (1)$$

Other adjustments can be made for other departures from the idealized, panmictic, population. For example when there are N_m males and N_f females then asymmetrical contributions to the next generation can be adjusted by:

$$N_e = (4\ N_m\ N_f)\ /(N_m + N_f) \qquad (2)$$

How important unequal sex ratios are for *Anopheles* is, at present, unknown. We do not know how many males actually mate and this number might be highly skewed. A third type of adjustment, for non-Poisson survival of offspring, is

$$N_e = (4N - 2)/(V_k + 2) \qquad (3)$$

where V_k is the variance in numbers of offspring per parental pair. When survival of each egg has a Poisson distribution with mean 2, then $V_k = 2$. The true variance is probably much greater. Consider a common larval site for *Anopheles,* puddles. Survival of all the eggs laid there is hit or miss, in large part dependent on whether the puddle dries up or is washed away. Assuming that each female lays several eggs when she oviposits, then some few females who lay in fortunate sites will produce more offspring than others, thereby increasing V_k and decreasing N_e.

Note that these equations all refer to discrete generations, an idealization that is not really appropriate for *An. gambiae*. It is possible to make adjustments when a stable age distribution can be inferred, but a rough approximation – that N_e is roughly the number of mosquitoes born during some time interval that make it to the age of

reproduction times the number of time intervals during which they reproduce – is more practical. In general multiple considerations to correct N so it approaches N_e can be combined with one another, simply by taking a composite function (Kimura and Ohta 1971).

Anopheles gambiae s.s. and metapopulations

Combining the information above we get the following picture of genetic structure on *Anopheles gambiae s.s*. at our focal research site, Banambani village, Mali (Table 3). The calculations and rationale are described in Taylor et al. (1993) and Lanzaro et al. (1998).

Table 3. Estimates of effective population size presented in relation to some of the structure in *An. gambiae* around Banambani, Mali. In this table the total effective population size is a bit larger than the sum of those for the forms it comprises. This is because of rounding error and because there are some unassigned or hybrid individuals.

Parameter	*Notation*	*Value (Banambani, Mali)*
Effective population size	$N_{e,tot}$	4,400
	$N_{e,Bam}$	900
	$N_{e,Sav}$	1,500
	$N_{e,Mop}$	1,900
Migration to adjacent population	m_{ss}	0.008
Gene flow among forms	$N_e m_{BS}$	16
	$N_e m_{SM}$	12
	$N_e m_{BM}$	2

The figures in Table 3 vary in their reliability, particularly the numbers and migration rates when population sizes are small during the dry season. Frankham, Ballou and Briscoe (2002) report that an average ratio of N_e/N across many species is approximately 0.1, not far off from the ratio estimated from Tables 1 and 3. There is little allowance here for year-to-year variation, which we know to be substantial in our study area (Table 1). In addition to temporal structure, the population of this species around Banambani is structured spatially (between villages) and non-dimensionally (non-random mating due to chromosomal forms), in the terminology of Taylor and Powell (1983). This degree of complexity is hard to capture and quantify even with multiple patterns of genetic variation.

A new development in ecological genetics – metapopulation analysis – holds much promise for furthering our understanding of structure in this species. A metapopulation is a set of local populations, many of which may be unable to sustain themselves, where local extinction may be frequent, but where migration and recolonization can retain a dynamic equilibrium.

Several types of metapopulations can be distinguished (see Hanski and Gilpin 1997). Three of these have particular significance for the introduction of genetically modified *An. gambiae*: (1) The 'classical' metapopulation of Levins (1969), a network of equivalent local populations that inhabit discrete patches. The probability of extinction is equal for all patches, as is the probability of recolonization and origin of migrants. (2) A mainland-island metapopulation, a system of habitat patches which are within dispersal distance of a very large (mainland) patch which never suffers extinction. (3) A source-sink metapopulation, where some patches have a negative growth rate (sinks) and are maintained by migration from patches with a positive

growth rate (sources). Which patches are sources or sinks may vary seasonally or otherwise.

Describing the metapopulation of *An. gambiae* around Banambani as one of these types or as an intermediate will require careful application of direct and indirect methods of assessing structure. Genetic data can be used to test particular hypotheses about structure and type as an extensive body of theory now exists that describes how the patterns of gene frequencies and molecular evolution are determined by metapopulation structure. A better description of the importance of drift for the population can be attained by examining this type of question. Detailed discussion of these is beyond the scope of this work, but the reader is referred to Pannell and Charlesworth (2000), Gavrilets, Acton and Gravner (2000), Whitlock (1999) and Slatkin (1977) for good examples of the utility of metapopulation theory.

We have used computer simulation to explore how a transposable element might move through a classical metapopulation under very simple conditions. The outcome of one such simulation is shown in Figure 4, for a transposable element released at a frequency of 0.1 in the Mopti population of one village in March. This hypothetical transposable element is associated with a single gene which induces 100% refractoriness to *Plasmodium*. We assume that there is no dissociation of the transposable element and the gene of interest. In addition, the transposable element has a fixed level of meiotic drive so that heterozygotes would contribute $2p(1-p)(1+i)$ of the modified gene rather than the $2p(1-p)$ of Mendelian segregation. We took $i=0.75$ for these simulations. We assumed that it had no negative fitness effect on the carrier. These conditions are oversimplified (Boëte and Koella 2002), but our purpose here is to explore the effect of structure on the movement of an ideal transposable element and gene for malaria refractoriness. The incorporation of more realistic transposable-element dynamics to this simulation is a future goal.

The status of the population at two-month intervals for two years is illustrated. The degree of red coloration in the figure shows what proportion of the population is carrying the transposable element. It is evident that incorporation of the transposable element is far from simple and depends critically on the population structure.

With the prospect of introducing a transposable element into the population, it is critical to understand more about the equivalency, permanence and demographic synchrony of patches. The results attained this far go part of the way to describing important characteristics of the metapopulation. The mathematical analyses and computer simulations made to date are still only rough. Much more detail, especially with regard to population properties during the dry season, contributions from different locations where immatures develop and more information on mating behaviour are desirable to attain a level of resolution that will make modelling of the effect of such a transposable element effective and any introduction successful.

Conclusion and recommendations

Estimates of effective population size are of little intrinsic interest by themselves. This is particularly evident when we consider genetic-modification schemes like the release of a transposable element into the metapopulation. In these circumstances neutrality cannot be assumed, so random genetic drift, to which N_e refers, is largely irrelevant to incorporation of the transforming genes. Rather, the importance of N_e lies in its utility to probe the structure of the population by allowing us to relate observed patterns of allele frequencies to particular aspects of population structure, which we

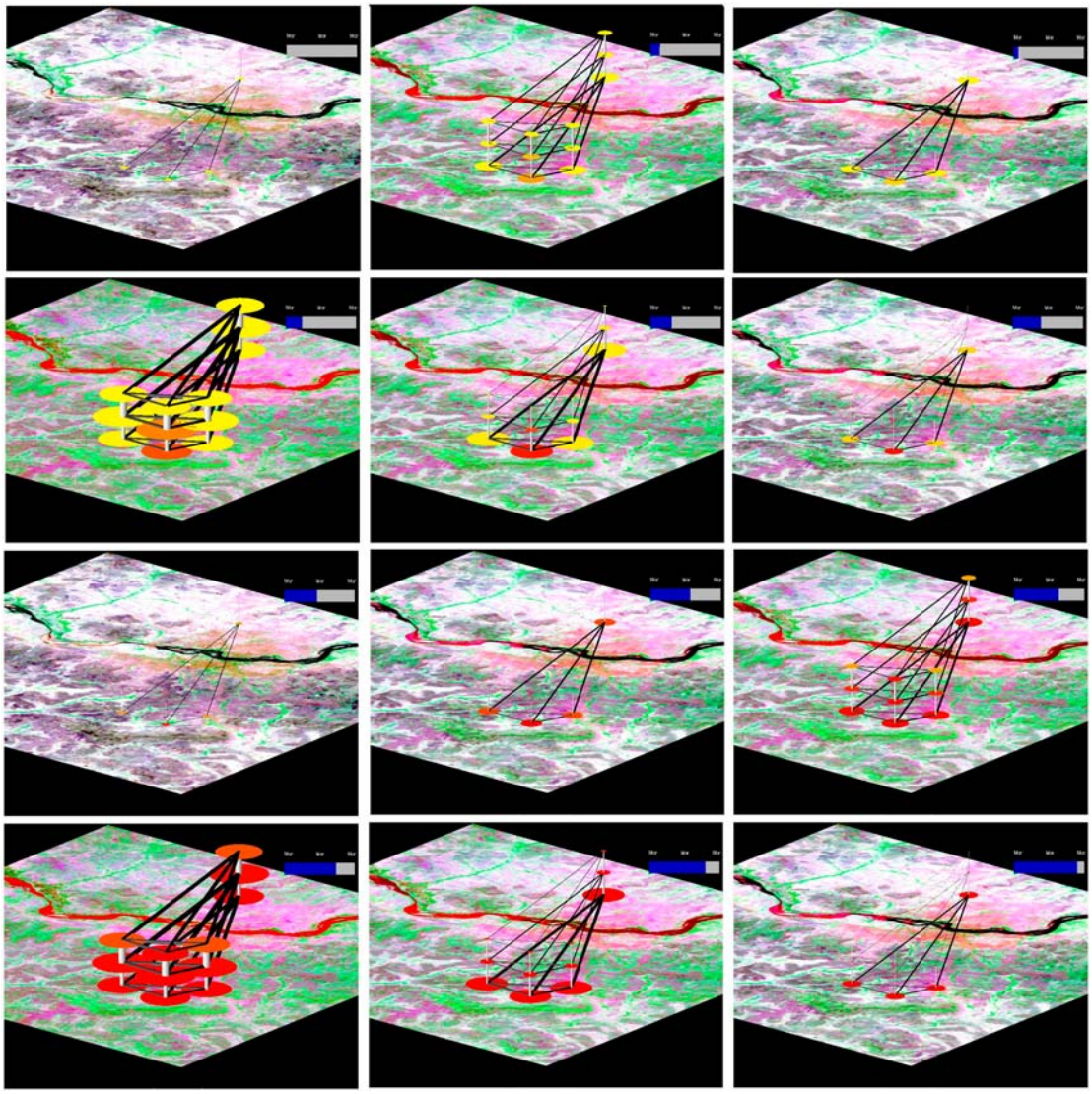

Figure 4. Simulated movement and frequency change of a transposable element through a metapopulation of *Anopheles gambiae s.s.* around Banambani, Mali. The release consists of 0.1 of the Mopti population in Banambani in March. Symbolic representations of the structured population are as in Figure 3, with the addition of red coloration to depict the proportion of the population which possesses the transposable element. The full movie from which these frames were excerpted is available at:
http://taylor0.biology.ucla.edu/~manoukis/Structure.

have seen is very relevant for determining the pattern of transformation. In this area, the estimation of N_e will remain critical and of enormous utility.

It is not known if genetic modification will proceed through the use of transposable elements or if malaria control will be accomplished through other methods. It may, for example, happen that sufficiently variable refractory mosquitoes can-not be made, that the genetic changes can-not be effected into natural populations or that ethical considerations will preclude usage of this method. Our research agenda should be constructed so that information acquired by our efforts will benefit malaria control by any means – such as bed nets, a vaccine, anti-malarial drugs or new insecticides.

Both direct and indirect measures of gene flow will be necessary for estimates of population structure. They each give somewhat different information because they are based on different assumptions (Slatkin 1987). Direct methods are necessarily bound

by space and time and have low sensitivity, which is to say they are not well suited to detecting rare events. On the other hand, recorded events are real and may give insight into yearly cycles or median conditions, given enough observations. Indirect methods typically assume complete neutrality and equilibrium conditions, and typically give results that are open to multiple interpretations. The use of both approaches helps to fill in the gaps left by the assumptions of one method or another, making estimates more robust. Especially in the case of rare events, which are of enormous importance to the prospects of genetic modification of *An. gambiae,* multiple sorts of evidence increase our confidence in the estimates.

In terms of population structure, the issues most in need of clarification are:

- Establish metapopulation structure of *An. gambiae*. In particular, determine if all patches are equivalent, if there are sources/sinks and especially what, if any, is the extirpation regime.
- Better estimates of gene flow and population size are needed so N_e can be used in estimating the importance of rare events.
- Examine in detail the nature and extent of gene flow between chromosomal forms in other parts of Africa.
- Extend research to other sites around the continent, especially east and central Africa. There is no reason to believe that what is true for the species in one area will also pertain to another.
- Identify and then survey in great depth island populations of *An. gambiae* that are serious candidates genetic-control trials.

One of the most urgent needs, in our opinion, is to establish the metapopulation structure of *An. gambiae* by directly measuring and indirectly inferring parameters that are of importance to determining metapopulation characteristics, such as synchrony, extirpation regime and patch equivalency. One important area here is to examine the extinction regime of the populations in the villages, which may be undertaken by both direct and indirect methods.

Improvement of existing estimates of hybridization, migration and population size during the wet season is also necessary for a more comprehensive characterization of the system. A promising avenue for improving some of these is revealed by the continuing development of coalescence theory, which may be described as an examination of genealogical patterns of genes in time. These powerful methods should enable researchers to efficiently probe fine-grained details and variation of the population.

This sort of investigation would be most useful if extended to other sites. In addition to the power of comparative studies that would be gained, some sense of the limits of variability in system dynamics could also be explored. Such an understanding will ultimately prove critical in predicting how a given system might change when faced with attempts at genetic manipulation or externally varying conditions, like changes in climate or increasing human development, for example.

Once we have a more complete quantification of the factors we already know are important to the structure of the species, realistic modelling of the effects of genetic modification of *An. gambiae* will be possible. The more complete our understanding of the system is as a metapopulation the more useful and accurate the modelling efforts can be.

Acknowledgments

This work has been supported by grants from NIH. We are indebted to Mario Coluzzi, Greg Lanzaro, Guimogo Dolo and Yeya T. Touré for help in formulating the ideas presented here. We also thank Joan Slottow, Maria Diuk-Wasser and Sigrid Rian for help preparing the illustrations presented in Figures 3 and 4, and Phil Lounibos for his many helpful suggestions on presentation.

References

Anderson, E.C., Williamson, E.G. and Thompson, E.A., 2000. Monte Carlo evaluation of the likelihood for N-e from temporally spaced samples. *Genetics,* 156 (4), 2109-2118.

Blower, J.G., Cook, L.M. and Bishop, J.A., 1981. *Estimating the size of animal populations*. Allen & Unwin, London.

Boëte, C. and Koella, J.C., 2002. A theoretical approach to predicting the success of genetic manipulation of malaria mosquitoes in malaria control. *Malaria Journal,* 1 (1), 1-7. [http://www.malariajournal.com/content/1/1/3]

Carnahan, J., Zheng, L., Taylor, C.E., et al., 2002. Genetic differentiation of *Anopheles gambiae s.s.* populations in Mali, West Africa, using microsatellite loci. *Journal of Heredity,* 93 (4), 249-253.

Costantini, C., Li, S.G., Della Torre, A., et al., 1996. Density, survival and dispersal of *Anopheles gambiae* complex mosquitoes in a West African Sudan savanna village. *Medical and Veterinary Entomology,* 10 (3), 203-219.

Crow, J.F., 1956. Breeding structure of populations. II. Effective population number. *In:* Kempthorne, O., Bancroft, T., Gowen, J., et al. eds. *Statistics and mathematics in biology*. University of Iowa Press, Ames, Iowa.

Crow, J.F. and Kimura, M., 1970. *introduction to population genetics theory*. Harper and Row, New York.

Della Torre, A., Costantini, C., Besansky, N.J., et al., 2002. Speciation within *Anopheles gambiae* - the glass is half full. *Science,* 298 (5591), 115-117.

Della Torre, A., Fanello, C., Akogbeto, M., et al., 2001. Molecular evidence of incipient speciation within *Anopheles gambiae s.s.* in West Africa. *Insect Molecular Biology,* 10 (1), 9-18.

Dolo, G., 2000. *Cytogenetique, genetique moleculaire et simulation par ordinateur de la dispersion du complexe Anopheles gambiae vecteur du paludisme a Banambani, Mali*, Universite du Mali.

Dolo, G., Dao, A., Traoré, S.F., et al., 1999. *Rapport de l'etude entomologique sur la transmission du paludisme dans six villages (Aout 1995 - Fevrier 1998). Technical Report*. Association pour le Développement de la Riziculture en Afrique de l'Ouest (ADRAO), Projet de Recherche du Consortium "Santé".

Favia, G., Della Torre, A., Bagayoko, M., et al., 1997. Molecular identification of sympatric chromosomal forms of *Anopheles gambiae* and further evidence of their reproductive isolation. *Insect Molecular Biology,* 6 (4), 377-383.

Favia, G., Lanfrancotti, A., Spanos, L., et al., 2001. Molecular characterization of ribosomal DNA polymorphisms discriminating among chromosomal forms of *Anopheles gambiae s.s. Insect Molecular Biology,* 10 (1), 19-23.

Frankham, R., Ballou, J.D. and Briscoe, D.A., 2002. *Introduction to conservation genetics*. University Press, Cambridge.

Fu, Y.X. and Li, W.H., 1999. Coalescing into the 21st century: An overview and

prospects of coalescent theory. *Theoretical Population Biology,* 56 (1), 1-10.
Futuyma, D., 1998. *Evolutionary biology*. 3rd edn. Sinauer, Sunderland.
Gavrilets, S., Acton, R. and Gravner, J., 2000. Dynamics of speciation and diversification in a metapopulation. *Evolution,* 54 (5), 1493-1501.
Hanski, I. and Gilpin, M.E., 1997. *Metapopulation biology : ecology, genetics, and evolution*. Academic Press, San Diego.
Hartl, D.L. and Clark, A.G., 1997. *Principles of population genetics*. 3rd edn. Sinauer, Sunderland.
Hudson, R.R., 1990. Gene genealogies and the coalescent process. *In:* Futuyama, D. and Antonovics, J. eds. *Oxford Survey of Evolutionary Biology. Vol. 7*. Oxford University Press, Oxford, 1-44.
Johnson, C.G., 1969. *Migration and dispersal of insects by flight*. Methuen, London.
Kimura, M. and Ohta, T., 1971. *Theoretical aspects of population genetics*. Princeton University Press, Princeton.
Lanzaro, G.C., Touré, Y.T., Carnahan, J., et al., 1998. Complexities in the genetic structure of *Anopheles gambiae* populations in west Africa as revealed by microsatellite DNA analysis. *Proceedings of the National Academy of Sciences of the United States of America,* 95 (24), 14260-14265.
Levins, R., 1969. Some demographic and genetic consequences of environmental hetrogeneity for biological control. *Bulletin of the Entomological Society of America,* 15, 237-240.
Pannell, J.R. and Charlesworth, B., 2000. Effects of metapopulation processes on measures of genetic diversity. *Philosophical Transactions of the Royal Society of London Series B Biological Sciences,* 355 (1404), 1851-1864.
Service, M.W., 1993. *Mosquito ecology : field sampling methods*. 2nd edn. Elsevier Applied Science, London.
Slatkin, M., 1977. Gene flow and genetic drift in a species subject to frequent local extinctions. *Theoretical Population Biology,* 12, 253-262.
Slatkin, M., 1985. Gene flow in natural populations. *Annual Review of Ecology and Systematics,* 16, 393-430.
Slatkin, M., 1987. Gene flow and the geographic structure of natural populations. *Science,* 236 (4803), 787-792.
Taylor, C., Touré, Y.T., Carnahan, J., et al., 2001. Gene flow among populations of the malaria vector, *Anopheles gambiae*, in Mali, west Africa. *Genetics,* 157 (2), 743-750.
Taylor, C.E. and Powel, J.R., 1983. Population structure of *Drosophila*: genetics and ecology. *In:* Ashburner, M., Carson, H. L. and Thompson, J. R. eds. *The genetics and biology of Drosophila. Vol. 3D*, Vol. Vol. 3D. Academic Press, London, 29-60.
Taylor, C.E., Touré, Y., Coluzzi, M., et al., 1993. Effective population size and persistence of *Anopheles arabiensis* during the dry season of West Africa. *Medical and Veterinary Entomology,* 7 (4), 351-357.
Touré, Y.T., Dolo, G., Petrarca, V., et al., 1998a. Mark-release-recapture experiments with *Anopheles gambiae s.l.* in Banambani Village, Mali, to determine population size and structure. *Medical and Veterinary Entomology,* 12 (1), 74-83.
Touré, Y.T., Petrarca, V., Traoré, S.F., et al., 1998b. The distribution and inversion polymorphism of chromosomally recognized taxa of the *Anopheles gambiae* complex in Mali, West Africa. *Parassitologia,* 40 (4), 477-511.
Tripet, F., Touré, Y.T., Taylor, C.E., et al., 2001. DNA analysis of transferred sperm

reveals significant levels of gene flow between molecular forms of *Anopheles gambiae*. *Molecular Ecology,* 10 (7), 1725-1732.

Turchin, P., 1998. *Quantitative analysis of movement : measuring and modeling population redistribution in animals and plants*. Sinauer, Sunderland.

Waples, R., 1989. A generalized approach for estimating effective population size from temporal changes in allele frequency. *Genetics,* 121 (2), 379-391.

Watterson, G.A., 1975. On the number of segregating sites in genetical models without recombination. *Theoretical Population Biology,* 7, 256-276.

Whitlock, M.C., 1999. Neutral additive genetic variance in a metapopulation. *Genetical Research,* 74 (3), 215-221.

Wright, S., 1969. *Evolution and the genetics of populations. Vol. II. The theory of gene frequencies*. University of Chicago Press, Chicago.