Inferring Modes of Colonization for Pest Species Using Heterozygosity Comparisons and a Shared-Allele Test

J. A. Sved,* H. Yu,* B. Dominiak† and A. S. Gilchrist*,[1]

**Fruit Fly Research Centre, School of Biological Sciences, University of Sydney, New South Wales 2006, Australia and* †*New South Wales Agriculture, Orange, New South Wales 2800, Australia*

Manuscript received April 9, 2002
Accepted for publication November 22, 2002

ABSTRACT

Long-range dispersal of a species may involve either a single long-distance movement from a core population or spreading via unobserved intermediate populations. Where the new populations originate as small propagules, genetic drift may be extreme and gene frequency or assignment methods may not prove useful in determining the relation between the core population and outbreak samples. We describe computationally simple resampling methods for use in this situation to distinguish between the different modes of dispersal. First, estimates of heterozygosity can be used to test for direct sampling from the core population and to estimate the effective size of intermediate populations. Second, a test of sharing of alleles, particularly rare alleles, can show whether outbreaks are related to each other rather than arriving as independent samples from the core population. The shared-allele statistic also serves as a genetic distance measure that is appropriate for small samples. These methods were applied to data on a fruit fly pest species, *Bactrocera tryoni*, which is quarantined from some horticultural areas in Australia. We concluded that the outbreaks in the quarantine zone came from a heterogeneous set of genetically differentiated populations, possibly ones that overwinter in the vicinity of the quarantine zone.

AN extreme version of source-sink population models is that where each new sink population becomes extinct within one or few generations. Since they do not contribute to the migrant pool they have been termed "black-hole sinks" (Holt and Gaines 1992). As the evolutionary significance of such populations is minor or nonexistent, the scenario has received little or no attention from evolutionary biologists. However, the situation where small propagules repeatedly arise (despite soon disappearing) is of interest because it describes the situation often confronted in quarantine programs (*e.g.*, medfly; Davies *et al.* 1999). After repeated outbreaks of a pest species within a quarantine area, authorities seek information about the mode of introduction of the outbreaks to prevent future infestations. In many cases the mode of introduction will be obvious (*e.g.*, ships, their ballast, or importations of infested foodstuffs). Limited quarantine resources can then be specifically allocated to deal with the incursions.

However, in other situations, the route of introduction will not be apparent, especially if there are no obvious physical barriers to the dispersal or migration of the pest. Where isolated outbreak populations continually arise within a quarantine area, but the likely source population is distant (relative to the unaided dispersal range of the organism), two basic alternatives exist. In the first, the introduction is human assisted, with pests being carried into the quarantine area, *e.g.*, on vehicles or in infested produce. Second, the pest could be dispersing naturally, establishing unobserved populations closer to the quarantine area. If these cryptic populations are sufficiently close to the quarantine area, they may be the source of outbreaks in the adjacent quarantine area. The human-assisted mode involves a single long step while the second mode involves two or more steps via one or more intermediate populations. If the intermediate populations remain cryptic (typically due to a lack of monitoring in nonquarantine areas), forensic data such as microsatellites will be available only for the core and the outbreak populations.

Where data from only the core and the outbreak populations are available, the task of inferring the origin of the outbreaks can be approached by simulating the various sampling processes that may have led to the outbreaks. The simplest model assumes each outbreak, or founder event, in the quarantine area arises from a single introduction of individuals directly from the core population. The effect of adding various intermediate (cryptic) populations can then be investigated. The likelihood of the observed data can then be assessed under each scenario. Since each outbreak may involve only a few founding individuals (A. Meats, A. D. Clift and M. Robson, unpublished results), genetic distance measures such as F_{st} may be greatly affected by genetic drift and consequently be uninformative as to the origins of the outbreaks. Resampling simulations have also been used by Noor *et al.* (2000) and Pascual *et al.* (2001)

[1]*Corresponding author:* School of Biological Sciences, A12, The University of Sydney, NSW 2006, Australia.
E-mail: stuartg@bio.usyd.edu.au

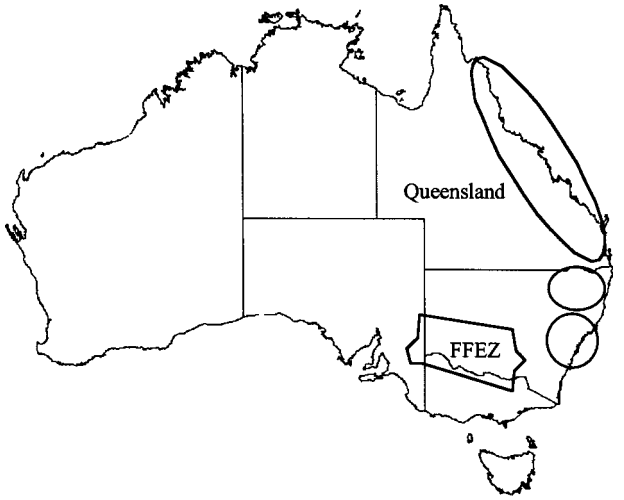

FIGURE 1.—Simplified map of Australia with three circles representing the major population groups of *B. tryoni* along the east coast, as identified in YU *et al.* (2001). The site of the fruit fly exclusion zone (FFEZ) in a part of the continent that is more marginal for *B. tryoni* is also indicated.

to study colonization by *Drosophila subobscura*. However, those studies dealt with only a single colonization event and were therefore unable to test for multistep introductions.

In this study, we applied resampling methods to microsatellite data from repeated but isolated outbreaks (colonizations) within and near a quarantine zone. The aim of this article is to demonstrate computationally straightforward procedures for distinguishing between different modes by which propagules from a distant source reach a quarantine area. We compared both the heterozygosity and the occurrence of sharing of alleles between the simulated and observed populations. The methods are useful in situations where large amounts of drift obscure patterns of isolation by distance.

We apply these methods to the problem of Queensland fruit fly, *Bactrocera tryoni*, in southeastern Australia. A previous study (YU *et al.* 2001) examined samples from areas of Australia where the fly is present in substantial numbers each year. The analysis showed that stable differences exist along the east coast of Australia, despite the high mobility of the species. The three main population groupings found are shown in Figure 1. The largest, most northerly grouping, which covers all of coastal Queensland, forms the core range of the species. The interest of this study is confined to southern fruit-growing regions in the fruit fly exclusion zone (FFEZ), where small outbreaks have been detected in the past few years. The core population of flies in Queensland must almost certainly be the ultimate source of outbreak flies. The critical question is one of whether the outbreaks come directly from this core population (*i.e.*, by human-assisted transport, a scenario assumed most likely by regulatory authorities) or from unknown intermediate populations.

TABLE 1

The origins of the 26 samples and sample sizes of the outbreak populations examined in this study

No.	Location	Year	Sample size	Propagule size[a]
1	Gilmore	1997	19	3
2	Tumut	1997	25	20
3	Gundagai	1997	20	4
4	Yenda	1997	8	4
5	Wagga Wagga	1997	6	2
6	Cootamundra	1997	20	8
7	Temora	1997	19	13
8	Young	1997	20	5
9	Grenfell	1997	19	7
10	Cowra	1997	15	9
11	Gilmore	1998	20	11
12	Batlow	1998	19	3
13	Tumut	1998	22	16
14	Gundagai	1998	20	8
15	Yenda	1998	7	5
16	Griffith	1998	5	3
17	Hillston	1998	7	8
18	Peak Hill	1998	20	7
19	Dunedoo	1998	20	7
20	Narromine	1998	20	7
21	Coonamble	1998	20	10
22	Hillston	1999	7	25
23	Wagga Wagga	1999	19	4
24	West Wyalong	1999	15	7
25	Corbie Hill	1999	13	3
26	Cudgell	1999	6	2
	Total		411	7

[a] The founder size needed to explain sample variance (see MATERIALS AND METHODS).

MATERIALS AND METHODS

Samples: The data set consists of 26 samples (Table 1), coming from regions of Australia in and around the FFEZ (Figure 2). The size of samples varied from 5 to 25 (Table 1, column 3), with a mean size of 15.9. All outbreak samples were trapped on the permanent trapping grid maintained by New South Wales (NSW) Agriculture, while the core population (Queensland) samples were collected during 1994–1998 as part of an annual collection program. In selecting this data set, several samples of four or less flies were excluded. In some of the larger collections, not all flies sampled were used for molecular analysis.

Microsatellite screening: We amplified and scored 6 *B. tryoni* microsatellites following the methods detailed in YU *et al.* (2001). Complete classification was not possible in all cases. The overall data set consists of 4838 classifications, 97.4% of the 414×12 possible observations. The 6 microsatellites were chosen from ~20 isolated microsatellites on the basis of high heterozygosity. Heterozygosities in the 6 ranged from 61 to 90%, with a mean of 72%.

ANALYSIS AND RESULTS

Genetic distances: Before embarking on the analysis outlined in the Introduction, we used a conventional

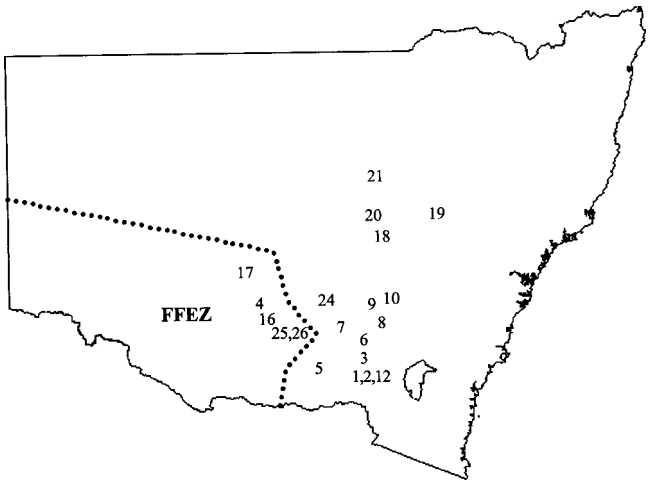

FIGURE 2.—Map showing sample sites, numbered following Table 1. Where locations were sampled in >1 year, only the number for the first year is shown. The dotted line indicates the border of the fruit fly exclusion zone (FFEZ).

distance analysis to try to detect patterning in the outbreak samples. Our previous analysis of endemic populations of *B. tryoni* on which Figure 1 is based (YU *et al.* 2001) relied heavily on chi-square tests for distinguishing populations. The validity of this analysis depended on the fact that chi-square tests showed that most samples of flies within regions were homogeneous. This homogeneity extended throughout the entire Queensland coastal region. It also extended for the 5 years of the sample. By contrast, although the genetic distances between Queensland, Northern NSW, and Sydney (*i.e.*, the three population groupings identified in Figure 1) were not large, they were statistically significant as judged by the heterogeneity chi-square test. This significance, in comparison to the homogeneity within regions, gave confidence in the reality of the divergence.

By contrast with the results from endemic populations, heterogeneity was found between many outbreak samples in neighboring regions and between years for the same sampling locations. Of the pairwise comparisons between samples, >50% (223 out of 325) gave significant differences. However, there were no clear patterns among the significance tests. An unrooted neighbor-joining tree was constructed to summarize the relationships between the samples (Figure 3). Since a drift-only model was suitable for the data, the tree was drawn using F_{st} genetic distances, calculated using the Gendist program in the Phylip group of programs (FELSENSTEIN 1993). Bootstrap support for the tree is very poor, which precluded any inference based on the pattern of the genetic distances.

A Mantel test (SOKAL and ROHLF 1995) was carried out and failed to show any correlation between physical distance and genetic distance ($r = 0.0665$, $P = 0.15$). Given the likely peripheral structure of any outbreak populations, it is not surprising that large fluctuations in frequency occur even in related populations. The small sample size adds to any difficulty in establishing relatedness using methods that depend on allele frequencies in the samples.

Models of colonization: Figure 4 shows three possible scenarios for the ancestry of outbreak samples. As previously mentioned, the analysis of endemic distributions carried out by YU *et al.* (2001) revealed a large core population of *B. tryoni* in Queensland that is the likely ultimate source for all outbreak samples. The simplest hypothesis (hypothesis 1) proposes that the outbreak samples come directly from the core population. Despite the distances involved, the possibility of human-assisted travel via the carriage of fruit makes this a realistic possibility. Alternative hypotheses are that the outbreaks come from a population or populations that are descended by one or more generations from the core population. Extreme forms of this hypothesis are that the descendant populations make up one single source population (hypothesis 2) or that they are independent of each other (hypothesis 3).

It should be emphasized that hypothesis 1 postulates that the sample individuals are themselves members of the core population. A more likely scenario may be one in which a small number of outbreak flies (a propagule) gives rise to a population, and the sample flies come from this population (hypothesis 3a). The distinguishing feature of hypothesis 3 is the fact that each sample is independently derived from the core population. This is true whether the sampled flies come from a population that is descended from the propagule by one generation (hypothesis 3a) or many generations (hypothesis 3b). These two possibilities are, in practice, difficult to separate.

Hypothesis 2 can also be formally subdivided into cases where the intermediate population is descended from the core by either a single generation (hypothesis 2a) or multiple generations (hypothesis 2b). Hypothesis 2a is an unlikely one in the case of the Queensland fruit fly, since the outbreak samples are geographically widespread and therefore unlikely to come from a population founded by a single generation previously. For the purposes of calculation, however, the hypothesis is a convenient one.

The three tests used in our study are set out below. Each uses a resampling strategy and the tests examine (a) the significance of the observed heterozygosities, (b) the variance of these heterozygosities, and (c) the significance of the extent of sharing of alleles.

Test a looks for a reduction of heterozygosity from the main population, *i.e.*, it tests hypothesis 1. Test b looks at whether the samples are consistent with coming from a single source population (hypotheses 1 and 2). Test c looks at whether the source populations could all be independent of each other (hypothesis 3).

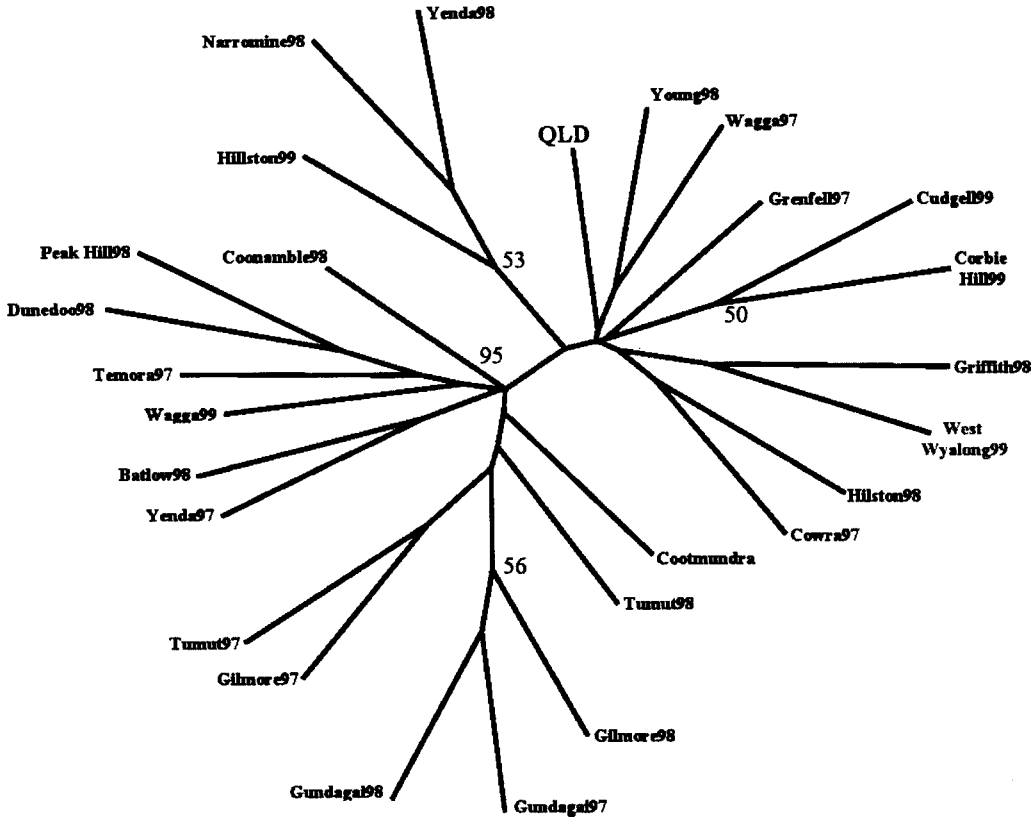

FIGURE 3.—Neighbor-joining tree constructed using F_{st} estimates as genetic distances. Bootstrap support values are shown only where the support exceeded 50% (from 500 resamplings). All other values are <50%.

Test a—heterozygosity tests: Specific information on hypothesis 1 is given by the analysis of heterozygosity, calculated as follows. Because different loci had differing numbers of observations, an overall weighted estimate of heterozygosity was used, calculated for each sample i ($i = 1, 2, \ldots, I$), as

$$H_i = 1 - \sum_{j=1}^{l}\left(\sum_{h=1}^{k_j} n_{ihj}^2 / n_{i.j}\right) / n_{i..},$$

where n_{ihj} is the number of observations in the ith outbreak sample of the hth allele ($h = 1, 2, \ldots, k_j$) at the jth microsatellite locus ($j = 1, 2, \ldots, l$), where the total number of observations in the ith sample at the jth microsatellite locus is

$$n_{i.j} = \sum_{h=1}^{k_j} n_{ihj},$$

and where the total number of observations in the ith sample is

$$n_{i..} = \sum_{j=1}^{l} n_{i.j} = \sum_{j=1}^{l}\sum_{h=1}^{k_j} n_{ihj}.$$

The first summation is over l microsatellite loci, and the second is over the k_j alleles at the jth locus.

The expected overall heterozygosity in the ith sample is reduced by a factor $(1 - 1/n_{i..})$ compared to the population from which it comes (*cf.* NEI 1987, Equation 7.39). A corrected heterozygosity, $H_i/(1 - 1/n_{i..})$, was calculated for each sample. All but one of the 26 values lie below the heterozygosity in the Queensland core population (Figure 5).

For each population, a test of whether heterogeneity was significantly less than expected was carried out, based on computer resampling from the core popula-

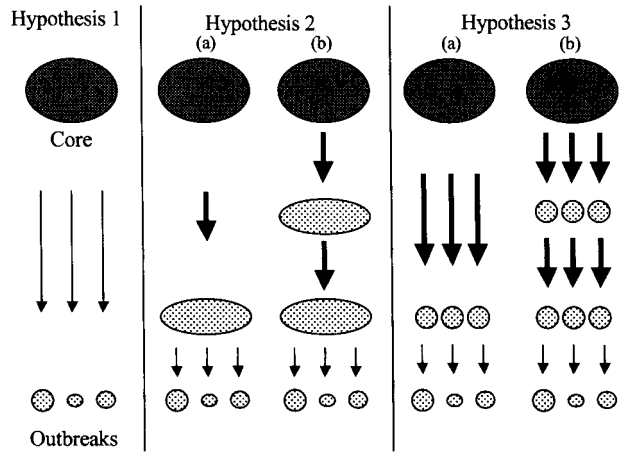

FIGURE 4.—Three hypotheses for the mode by which outbreak samples are derived from the main core population in northern Australia. They differ in the presence or absence of intermediate populations and the number (one or many) of any intermediate populations.

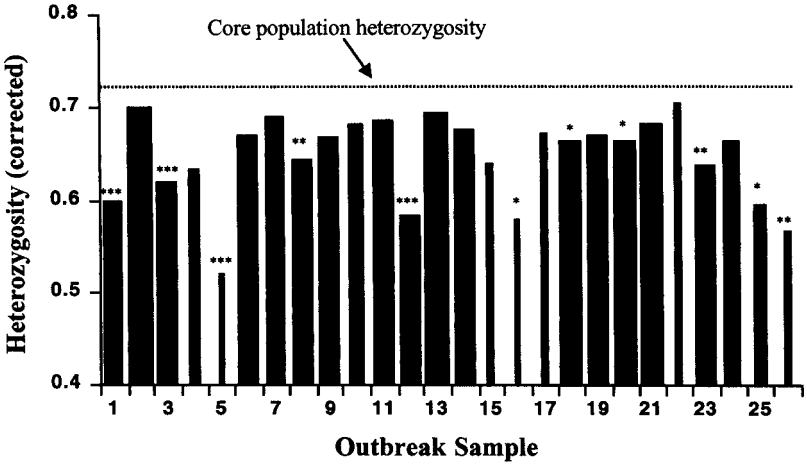

FIGURE 5.—Heterozygosity values in 26 samples. Thickness of the bars represents the sample size. Significantly reduced heterozygosity values are indicated (*$P < 0.05$, **$P < 0.01$, ***$P < 0.001$).

tion, calculating the uncorrected H_i, and recording the fraction of cases that were less than the observed uncorrected heterozygosity. Figure 5 shows the significance levels of each sample. Nearly half of the heterozygosity levels are significantly reduced, and the reduction of heterozygosity of the mean of all samples is highly significant.

Although hypothesis 1 is ruled out by these heterozygosity considerations, they do not rule out the possibility that each sample comes from a population started by a propagule one generation previously (hypothesis 3a). Estimates of the effective sizes of a propagule needed to produce the observed levels of heterozygosity are obtained from the formula $H_i = H_c\,(1 - 1/2n_i)$ (*cf.* NEI 1987, Equation 13.12), where n_i is the effective size of the propagule giving rise to sample i, H_c is the level of heterozygosity in the core population, and H_i is the level of heterozygosity in sample i, corrected for sample size. This can be rearranged to give $n_i = H_c/[2(H_c - H_i)]$. The values of n_i are shown in Table 1, column 5. In all cases except for sample 22, the samples would need to come from very small propagules, as low as a single pair. Whether these values are unrealistically low for fly outbreaks is difficult to tell.

Test b—variance of heterozygosity tests: One feature of Figure 5 is the considerable amount of variation in the heterozygosities of different populations. It is of interest to test the significance of this "heterogeneity of heterozygosities," to see whether the heterozygosities are consistent with samples taken from a single population (hypotheses 1 and 2). Significance tests were conducted by simulating overall data sets of 26 samples and counting those cases where the variance was above the observed variance. Four different assumptions were used as the basis of these simulations:

i. Direct sampling from a core population (hypothesis 1).
ii. Sampling from a single population founded from a propagule of size 7 (haploid size = 15) one generation previously (hypothesis 2a). This propagule size was chosen to give the observed average level of heterozygosities in the samples, using the formula $n = H_c/[2(H_c - \overline{H})]$, where $\overline{H}$ is the average heterozygosity and $2n$ is the haploid size.
iii. Sampling from a population founded from a large propagule (hypothesis 2b). A diploid size of 500 (haploid size = 1000) was chosen. For a population maintained at this size, 69 generations are needed to produce the observed average level of heterozygosity.
iv. An artificial population obtained by pooling all of the outbreak samples.

All tests agreed in showing, at around the 1% significance level, that the different samples have reduced heterozygosity to different extents. Figure 6 gives the results for tests i and ii, showing that there is little difference between the expected distributions for these two cases.

Test c—a test for sharing of alleles: In addition to the loss of heterozygosity, it was noted that several outbreak samples contained alleles that exist only at low fre-

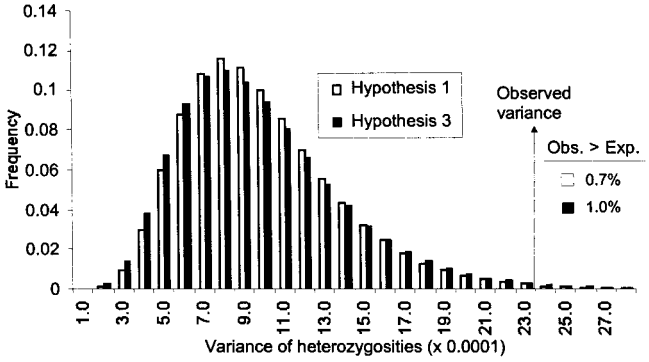

FIGURE 6.—The distribution of variances of heterozygosities for 10^6 simulations of hypotheses 1 and 3a.

quency in the Queensland and other endemic populations. This suggested the possibility of defining relatedness of populations on the basis of sharing of alleles, particularly rare alleles, and of developing a test of significance of such sharing of alleles. The null hypothesis in this case is that the samples come from populations that are independently derived from the core population (hypothesis 3). For ease of calculation, hypothesis 3a is used as the basis for calculation. The test focuses only on the presence or absence of a particular allele in each sample and does not take into account frequencies.

The population of alleles existing in the core population may be denoted as A_{hj}, where j is the locus number (1 . . . 6) and h is the allele number (1 . . . k_i). The frequencies, assumed known, are p_{hj}, where $\Sigma p_{hj} = 1$ for each locus j summed over alleles h. The probability that an allele will be present in sample i under hypothesis 3a is

$$Q_i = 1 - \sum_{g=1}^{m_i} C_g^{m_i} (1 - p_{hj})^{m_i - g} p_{hj}^{g} \left(1 - \frac{g}{m_i}\right)^{n_i},$$

which is calculated as follows. The haploid size of the propagule leading to sample i is equal to m_i (where $m_i/2$ is the diploid propagule size in Table 1). The calculation makes the simplifying assumption that m_i is a known constant. The probability that the propagule contains g copies of allele A_{hj} is the binomial probability $C_g^{m_i}(1 - p_{hj})^{m_i - g} p_{hj}^{g}$. The probability that this allele is not passed on in the sample is $(1 - (g/m_i))^{n_i}$, where n_i is the haploid sample size. The product of these two probabilities needs to be summed over all possible values of g to obtain the overall probability of absence $1 - Q_i$. The probability of presence of the allele is then the complement of the probability of absence.

The overall likelihood of an observed data set, in which the allele is present in some samples and absent in others, is obtained by multiplying the relevant probabilities for individual samples. In practice it is easier to take logarithms and to sum them to obtain an overall log likelihood.

Following this approach, we calculated likelihoods and tested for significance by resampling as outlined below. However, we found that the test was heavily influenced by likelihood values associated with cases where an allele was absent from, rather than present in, the sample. For this reason we calculated a "part likelihood," including probabilities where an allele was present in a sample and ignoring the contribution to the likelihood where an allele was absent. Although this statistic cannot be directly interpreted as a likelihood, it serves as the basis for a significance test and also provides a satisfactory distance statistic.

The test of significance: While the likelihood values in column 5 of Table 2 are very low, their absolute values cannot immediately be interpreted. Clearly the likelihood of any particular observed set of values diminishes as the number of observations (samples) increases. The significance test consists of calculating the sum of the part-likelihood values, simulating data sets under hypothesis 3a, and calculating the fraction of cases where the likelihood falls below the observed likelihood. We excluded cases where the allele was present in only one sample, on the grounds that such cases do not contribute to information on allele sharing. This exclusion had a small effect on the result.

TABLE 2

Lowest likelihood values for the outbreak data set

Allele (locus/allele)	Core frequency/1848	No. of samples with allele		Part likelihood
		Observed	Expected	
6/133	5	10	0.8	−35.17
2/141	19	14	2.9	−30.97
6/105	8	7	1.3	−21.27
2/150	53	13	7.2	−17.91
6/117	11	6	1.7	−16.93
3/152	1	3	0.2	−14.93
5/155	113	20	12.5	−14.77
4/126	50	12	6.8	−14.75
1/151	9	5	1.4	−14.52
1/152	40	9	5.7	−14.17

The first column identifies the locus (1–6) and the allele at that locus (size in base pairs). The next column shows its frequency in the core population, and the following columns show how many samples, out of 26, contain one or more copies of the allele, the expected number of populations containing the allele under hypothesis 3a, and the part likelihood with contributions from all samples containing the allele (see text).

One further modification to the test was necessary, referring to the assumption of known frequencies in the core population. Although the sample size for the core population was much larger than the outbreak sample size, 1848 for each microsatellite locus, the estimates from the core population are nevertheless subject to standard errors. Allowance was made for this source of variation, using an extra sampling stage for the core population, generating samples of size 1848 at each locus, on the basis of the observed frequencies from the core (Queensland) population. The sampled values were then used as the input frequencies for the outbreak sampling. The procedure of core sampling was replicated 1000 times, and each set was used for 1000 replicates of the outbreak samples.

The sampling process using observed frequencies assumes symmetry of the sampling process. In reality, the sampling should be aimed at generating a distribution of starting frequencies that lead to the observed frequencies, rather than vice versa. The expected variances are proportional to $p \cdot (1 - p)/n$, so that variation in the value of p should not lead to large sampling differ-

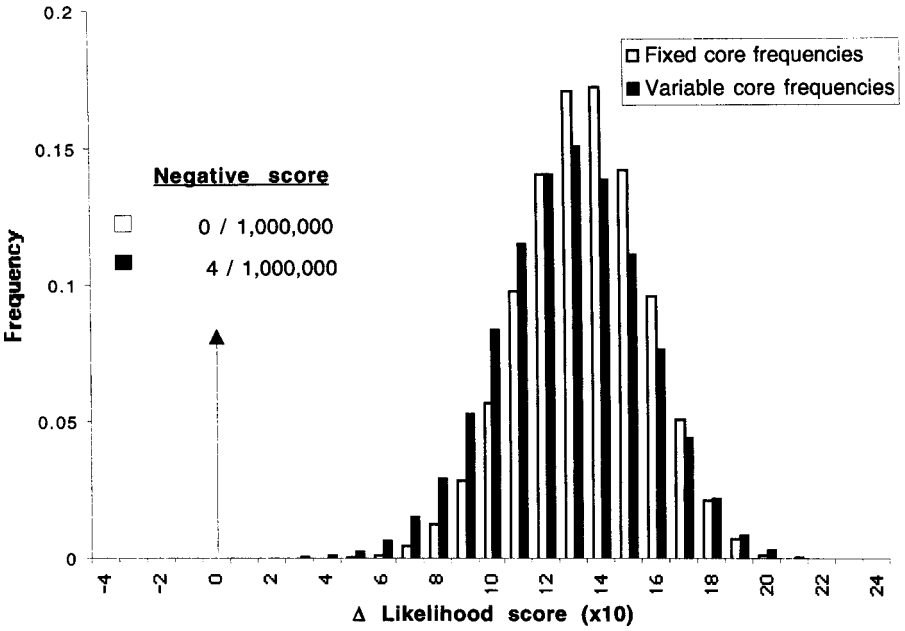

FIGURE 7.—Distribution of likelihoods of the extent of sharing of rare alleles for 10^6 simulated data sets. Calculations were made either using the observed core sample frequencies (fixed core frequencies) or from a resampling of the observed core sample (variable core frequencies). Values on the *x*-axis are the difference between the log likelihood of the observed sample and the log likelihood of the simulated samples. The likelihood of the observed data set (zero point) is indicated by the arrow.

ences. A different approach to a related problem, the assignment of individuals on the basis of unknown population frequencies, was taken by RANNALA and MOUNTAIN (1997). Their approach makes the Bayesian assumption of an equal *a priori* probability density for the frequencies of all alleles.

Results from the significance tests for hypothesis 3 are given in Figure 7. Allowance for the finite size of the core sample is clearly an important aspect of the test. However, the results are highly significant after allowance for this effect.

Genetic distance: The rare-allele test has been used in the above analysis as an overall measure of relatedness of the complete data set. It can also be used to provide a measure of relatedness, or genetic distance, between pairs of samples. For each pair of samples, a likelihood is calculated. Cases contributing to the calculation are those where pairs of samples share an allele, and the part likelihoods for each of the two samples are summed over all such cases. A high negative log likelihood, corresponding to a low probability, indicates a closer than expected relatedness, so that the positive log likelihood can be used as a measure of genetic distance.

A Mantel test was carried out using the positive log-likelihood values as the measure of genetic distance. Physical distance was measured using the log of the distance in kilometers and zero for cases of repeated sampling in different years at the same location. The correlation was equal to 0.11, which was significant with a probability value of ~3%.

DISCUSSION

The allele-sharing test: Interpretation of the results from the 26 outbreak samples was complicated by the large amount of heterogeneity between samples. However, we were struck by the fact that some alleles that are rare in the core population appeared in an unexpectedly large number of the outbreak samples. For example, an allele occurring at a frequency of <0.5% in the core population was present in 10 of the 26 samples of average size 15 diploid individuals. We used this as the basis of a statistical test to show that alleles, particularly rare ones, are shared between the samples significantly more often than expected by chance.

A key assumption in the allele-sharing test is that different samples are independent of each other. This assumption is, however, consistent with alternative possibilities that all individuals in the samples are independently drawn from the core population (hypothesis 1) or that each sample comes from a population that is itself established independently from the core population (hypothesis 3). An extreme form of the latter hypothesis is the case where the independent populations are established for just a single generation and then sampled to give the observed values (hypothesis 3a). Less extreme variants of this hypothesis would assume that the outbreak populations are larger and established over many generations. It seems, in practice, difficult to distinguish between these possibilities, and for simplicity we examined only the single generation model in detail. However, it seems clear, *e.g.*, from the variance heterogeneity test, that the consequences of hypotheses 3a and 3b are very similar.

The underlying argument of the allele-sharing test is that the high level of significance is due to nonindependence of the samples. The test assumes that the gene frequencies are known from the core population, although it takes into account the fact that they are not

known exactly. One alternative to nonindependence is the possibility that certain alleles have a selective advantage in the outbreak region, causing them to become systematically and independently increased in frequency. It is unlikely that microsatellite alleles would be affected in this way, although the formal possibility remains that they are tightly associated with alleles whose frequencies change systematically. A more likely possibility is that the samples are not independent of each other, but are samples from intermediate populations where initially rare alleles have, by chance or through selection via hitchhiking, increased in frequency.

Some comment needs to be made on the use of the partial likelihood as the statistic for the significance test. The results of Table 2 show that the high negative likelihoods are those associated with rare alleles. However, this does not necessarily mean that the test is optimized for such alleles. Any statistic can be chosen for such a test, and it is possible that assigning a greater weight to rare alleles might improve the power of the test. For our data, no obviously superior general weighting method was found.

We also attempted to examine the power of the test by computer simulation of data sets. We found that chance played a considerable role in the process. In many cases, rare alleles were absent in most generated samples, and the test had little power. Six loci, even of high heterozygosity, are insufficient to ensure that the test will always have power in detecting relatedness.

The focus on rare alleles in the shared-allele test suggests that there may be some similarity with the way in which Slatkin (1985) uses such alleles as a measure of gene flow. There are, however, substantial differences between the two analyses. Our analysis makes no assumptions of equilibrium whereas Slatkin's analysis assumes that an equilibrium has been reached. We excluded alleles present in only one sample, and these alleles form the basis for Slatkin's test. While there may be a common underlying reason for the utility of rare alleles in the two cases, there are substantial differences between the two uses.

Genetic distance: We used the part-likelihood statistic as the basis of a genetic distance measure. The genetic distance calculated in this manner was only loosely connected to the measure of genetic drift Φ^* (Latter 1973; equivalent to the coancestry coefficient θ for a pure drift model). The overall correlation between the two measures was 5%, as opposed, for example, to a correlation of 94% for Φ^* and D (Nei 1972).

The measure was able to detect a significant correlation between genetic distance and geographical distance (Mantel test). For our data, therefore, the distance measure appears to be a useful one. The combination of small population sizes and small sample sizes presumably leads to such large fluctuation in gene frequencies that these are not useful measures of relatedness. It must be remembered, however, that the basis of the likelihood distance measure is the existence of a well-characterized set of gene frequencies from a core population, which will usually not be available.

Descent hypotheses: The three hypotheses put forward (Figure 4) are extreme examples, all of which can be excluded in their simplest form. Hypothesis 1 can be excluded on the grounds of the lower heterozygosity in the samples compared to the core population. However, as pointed out previously, a close variant of hypothesis 1, hypothesis 3a, in which each sample is taken from a one-generation-old propagule rather than directly from core flies, cannot be excluded on the grounds of reduced heterozygosity.

While hypothesis 3 is consistent with heterozygosity considerations, it can be ruled out by the allele-sharing test. This shows that the samples cannot be derived independently of each other by any series of generational and sampling events.

Hypothesis 2, likewise, can be ruled out by two different lines of evidence. Substantial heterogeneity was found between most samples, showing that the outbreaks cannot come from a single uniform source. Similarly the differing heterozygosity values (Figure 5) argue that the samples come from populations that are inbred to differing extents.

Hypotheses 2 and 3 are extremes of a continuum. In one case the derivative populations are indistinguishable portions of the same overall population. In the other, they are entirely independent. It seems clear that the real situation must lie somewhere between the two. The area in question from which the samples are taken is large, and the overall density of flies is low. Therefore it is not surprising that there should be considerable regional variability.

A key question, from a control point of view, is whether residual overwintering flies are in some or in all regions from which the samples are taken. The existence of residual populations outside of the core population seems to be a necessary consequence of the rejection of hypothesis 3. However, without a complete sampling of the entire region, it is clearly impossible to state whether samples come from long-term populations in their immediate region or whether noncore source populations are somewhere in the general vicinity. Formally, we cannot rule out even the possibility that the source of samples is unrelated to the core population, although this seems an unlikely scenario in our case. A more complete general picture must be dependent on a greater density of samples and possibly on a larger selection of genetic markers.

We are grateful to Marianne Frommer, Sasha Curthoys, and Merryl Robson for their assistance in this study. Comments from two anonymous referees helped in the writing of the article. This work was supported by grants from Woolworths Supermarkets, the Australian Research Council, and Horticulture Australia.

LITERATURE CITED

Davies, N., F. X. Villablanca and G. K. Roderick, 1999 Bioinvasions of the medfly *Ceratitis capitata*: source estimation using DNA sequences at multiple intron loci. Genetics **153:** 351–360.

Felsenstein, J., 1993 PHYLIP (Phylogeny Inference Package), Version 3.5c (http://evolution.genetics.washington.edu/phylip.html).

Holt, R. D., and M. S. Gaines, 1992 Analysis of adaptation in heterogeneous landscapes: implications for the evolution of fundamental niches. Evol. Ecol. **6:** 433–447.

Latter, B. D. H., 1973 The island model of population differentiation: a general solution. Genetics **73:** 147–157.

Nei, M., 1972 Genetic distance between populations. Am. Nat. **106:** 283–292.

Nei, M., 1987 *Molecular Evolutionary Genetics*. Columbia University Press, New York.

Noor, M. A. F., M. Pascual and K. R. Smith, 2000 Genetic variation in the spread of *Drosophila subobscura* from a nonequilibrium population. Evolution **54:** 696–703.

Pascual, M., C. F. Aquadro, V. Soto and L. Serra, 2001 Microsatellite variation in colonizing and palearctic populations of *Drosophila subobscura*. Mol. Biol. Evol. **18:** 731–740.

Rannala, B., and J. L. Mountain, 1997 Detecting immigration by using multilocus genotypes. Proc. Natl. Acad. Sci. USA **94:** 9197–9201.

Slatkin, M., 1985 Rare alleles as indicators of gene flow. Evolution **39:** 53–65.

Sokal, R. R., and F. J. Rohlf, 1995 *Biometry*. W. H. Freeman, New York.

Yu, H., M. Frommer, M. Robson, A. Meats, D. C. A. Shearman *et al.*, 2001 Microsatellite analysis of the Queensland fruit fly *Bactrocera tryoni* (Diptera: Tephritidae) indicates spatial structuring: implications for population control. Bull. Entomol. Res. **91:** 139.

Communicating editor: M. W. Feldman