

Considerations for Design of Insecticide Resistance Monitoring Programs

R. T. ROUSH AND G. L. MILLER¹

Department of Entomology, Mississippi State University,
Mississippi State, Mississippi 39762

FORUM: J. Econ. Entomol. 79: 293-298 (1986)

ABSTRACT Monitoring is critical to resistance management, but there has been very little discussion in the literature about the statistical design of monitoring programs. Some general considerations show that the LD_{50} , a standard measure for resistance monitoring, is very inefficient compared with diagnostic tests that accurately distinguish between resistant and susceptible individuals. Even with diagnostic doses, sample sizes at any given location must often be very large (on the order of hundreds of individuals per population) to reliably detect resistance when it is present at frequencies of <10%. For those species where it is difficult to collect large numbers of individuals, resistance detection may not be a practical component of resistance management.

RESISTANCE MONITORING is often considered essential to insecticide and acaricide resistance management (Dennehy and Granett 1984, Staetz 1985), but surprisingly little has been published on the necessary characteristics of an effective monitoring program. Sample size requirements, in particular, have not been discussed in detail (e.g., Anonymous 1968, 1970, 1972). Two important papers on insecticide resistance detection, for example, simply state that "at least 100" insects should be tested from each population (Anonymous 1974, Miyata 1983). This report is an effort to stimulate improvements in the theory and practice of resistance monitoring. We will consider the interpretation of dose/mortality data, test methodology, and sample size requirements.

The precision required in a monitoring program depends on its purpose. Some monitoring programs are designed only to test whether a pesticide control failure was due to resistance rather than to poor application or other factors. A similar type of program may be used to determine if resistance is a problem in a specific location or time to improve pesticide choice (e.g., Dennehy and Granett 1984). In such cases, the primary objective is to document resistance problems; we will refer to these processes as resistance documentation.

A program to detect resistance before control failures occur requires greater precision in estimating the frequency of resistant individuals than does a documentation program. As we adopt resistance management tactics, such resistance detection will become more important for two reasons. First, resistance frequencies must be very low for some resistance management strategies to be effective (Tabashnik and Croft 1982, Curtis 1985). Thus, we should measure resistance frequencies

before such strategies are implemented. Second, once any type of resistance management approach has been implemented, we need to monitor for changes in resistance frequency to determine if the program is effective.

If resistance monitoring programs are to be used in a practical way for resistance management, they should be designed to detect resistant individuals at a frequency (i.e., phenotypic frequency) of 1%. After resistance frequencies reach this level, control can theoretically be lost in as little as one to six generations, depending on the circumstances (Georghiou and Taylor 1977, Tabashnik and Croft 1982). Although it would be desirable to detect resistance at even lower frequencies to allow more time to develop a resistance management program, 1% may be a practical limit in most cases, as discussed later in this paper.

Dose/Response Versus Diagnostic Dose. Resistance monitoring programs generally involve comparisons of LD_{50} 's, LD_{90} 's (or LC or LT values, which we will use interchangeably), and slopes between field-collected populations or laboratory strains or both (e.g., Twine and Reynolds 1980, Staetz 1985). This can be an adequate way to document resistance that has reached high levels, but it is very inefficient for detecting an incipient resistance outbreak, as illustrated by Fig. 1.

Fig. 1 was produced by using data from Beeman's (1983) study on the inheritance of malathion resistance in *Tribolium castaneum* (Herbst) to generate an expected dose/response curve for a partially resistant strain. Beeman found that this resistance was due to a single locus with two alleles, which would produce RR, RS, and SS genotypes in a heterogeneous population. A discriminatory exposure of 1-3 h killed essentially all SS but no RS individuals. In a heterogeneous population where the R allele frequency (p) was 0.1, the frequency of resistant heterozygotes (RS) would

¹ Current address: Dept. of Zoology, Weber State College, Ogden, UT 84408.

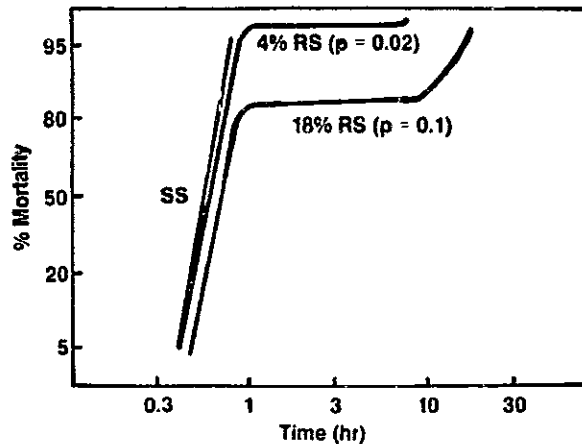


Fig. 1. Effect of segregation for a major resistance gene on the dose/response curve of a predominately susceptible strain, where p = frequency of the R allele. Drawn from Beeman's (1983) data on dose/responses of SS, RS, and RR genotypes of malathion-resistant *Tribolium*.

be 18%, assuming Hardy-Weinberg proportions (i.e., $2pq$ of $p^2 + 2pq + q^2$; the p^2 RR homozygotes [1%] will be ignored here for brevity). In such a population, a dose that killed 80% of the SS individuals would kill 80% (mortality of SS genotype) \times 82% (proportion of population that is SS), or 66% of all individuals in the population. Similar calculations were made at other doses.

As shown in Fig. 1, the frequency of resistant individuals must be fairly high (say at least 20%) before the LD_{85} is appreciably changed. The slope of the dose/mortality line or the LD_{85} might be better indicators of resistance, but both require that data be collected at doses causing high mortality levels to give any indication of a problem. In Fig. 1, for example, data would have to be taken at greater than a 1-h exposure before changes in the slope or LD_{85} could be noted. Neither the LD_{85} nor the slope would change much at the low but important resistance frequencies of 1-4%.

Consideration of Fig. 1 indicates that a diagnostic test would be more efficient than a dose/response regression in monitoring for resistance. In the *T. castaneum* example, the discriminatory 3-h malathion exposure would be diagnostic. Testing all of the sample at this dose would make better use of the individuals available since none would be wasted at the lower mortality levels where percentage mortality is not particularly revealing.

Test Methodology. The choice of a diagnostic test for a particular type of resistance is relatively easy when resistance has already been characterized for the species under consideration. A discriminatory dose may be readily established, as for *T. castaneum*. In other cases, biochemical tests such as electrophoresis are available to diagnose resistance in individual insects and mites (e.g., Miyata 1983). As we will show below, such tests are particularly useful because they reduce the num-

ber of individuals that must be sampled to detect resistance at any given level of certainty.

However, the situation is more difficult when resistance has not yet developed or where a standard test is not perfectly diagnostic. It is tempting simply to estimate an LD_{85} with an existing standard technique and multiply this dose by some factor (say 2 or 3) and use this as a diagnostic dose, but one problem with this approach was illustrated by Dennehy et al. (1983). The recognized standard technique for resistance in spider mites has been a slide dip assay, but Dennehy et al. (1983) showed that the slide dip assay discriminates poorly between dicofol-resistant and susceptible mites. In slide dip assays, the LD_{85} for the S strain killed 70% of the resistant mites. A dose 2- or 3-fold the susceptible LD_{85} would have killed >98% of the R strain. Thus, with some standard techniques, arbitrary increases above an LD_{85} may greatly reduce the ability to detect resistance rather than enhance it.

A partial solution to this problem was illustrated by Dennehy et al. (1983), who found that a leaf residue assay broadly discriminated between the R and S strains. Although it will not always be obvious before resistance is characterized that one test technique is more diagnostic than another, a suggestion from the Dennehy et al. (1983) report is that different test techniques should be tried. It may be particularly useful to try tests that are ecologically realistic. In this example, mites in the field are exposed to pesticides on leaf surfaces where they can continue to move and feed. Mites in the field are not stuck to slides and dipped into pesticide solutions, as called for in the standard technique. Many other standard topical tests (e.g., Twine and Reynolds 1980, Gunning et al. 1984) also do not closely simulate field exposure. Tests that more closely simulate field exposure may not always improve the detection of resistance, but should (perhaps more importantly) help to establish the relationship between laboratory assays and field failures, as called for by Ball (1981). For example, if a laboratory assay closely matches field conditions, as does a leaf residue test, and the tested strain shows high survival, concern about resistance in the field would be justified. If the mites died in a residue (or leaf spray) assay, one might be inclined to disregard the slide dip assay no matter what its results were.

A second problem with the LD_{85} approach is, in many cases, the choice of a susceptible population. A great deal of natural variation in LD values can be expected between relatively unexposed field populations (e.g., Twine and Reynolds 1980, Staetz 1985). Even laboratory populations can vary significantly from generation to generation (Wolfenbarger et al. 1982). There is no easy solution to this problem. We want to avoid failing to detect a resistance problem (and therefore want to avoid using an overly high LD_{85} estimate); at the same time, however, we want to avoid test results that

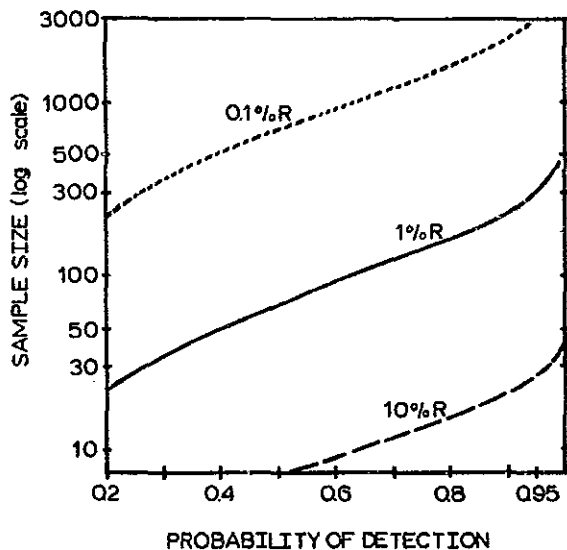


Fig. 2. Sample sizes necessary to detect at least one resistant individual with a diagnostic test when resistance is present at phenotypic frequencies of 0.1, 1, and 10%.

suggest a resistance problem when none exists (false positives). The strain(s) or population(s) used as the susceptibility-type(s) will simply have to be agreed upon by the monitoring community. However, in cases where the costs of resistance are very high relative to checking out false positives, it may be prudent to use a more susceptible strain to minimize the chances that a resistance outbreak will be overlooked.

The LD₉₅ will generally have a large estimation error (i.e., 95% CL) in most analyses. This might suggest that an LD₉₅, with a somewhat narrower confidence interval, would be a better test, but sample size considerations outweigh this benefit, as discussed below. Robertson et al. (1984) have recently shown how to properly select doses to improve the statistical accuracy of the LD₉₅, and similar procedures could be used for the LD₉₀. It will generally be useful for statistical tests (discussed below) to check the LD₉₀ empirically by testing several thousand individuals at the estimated dose.

Sample Sizes. Where perfectly diagnostic doses are available, the sample size required to detect resistance at any given frequency can be found by assuming that resistance is distributed as a binomial random variable. Here we only consider the frequency of resistant individuals, where resistance is defined by the assay technique, rather than resistance allele frequency. Thus, this discussion makes no assumptions about the genetics of resistance.

The probability of detecting at least one resistant individual, $P(x \geq 1)$, is simply one minus the probability of not detecting a resistant individual, $1 - P(x = 0)$. For the binomial distribution, $P(x = 0) = (1 - f)^n$ (Snedecor and Cochran 1967) where n is the sample size and f is the frequency of the resistant phenotype. Therefore,

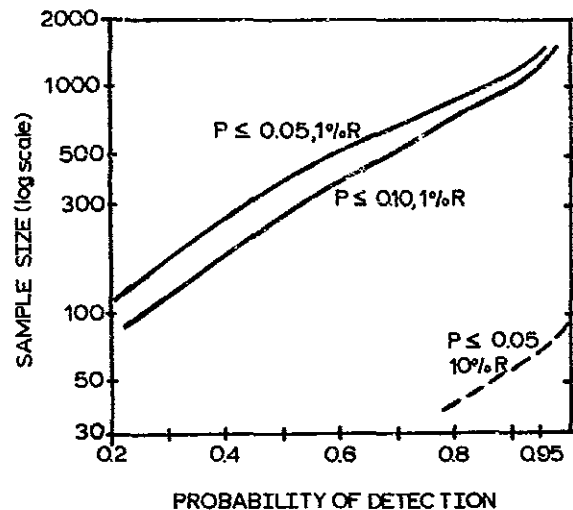


Fig. 3. Sample sizes necessary to detect resistance (i.e., to find statistical significance in a Z test) when an LD₉₅ is used and the frequency of resistant individuals is 1 or 10%. The P values indicate the level of α assumed for the statistical test.

$$1 - P(x \geq 1) = (1 - f)^n$$

$$n \log(1 - f) = \log[1 - P(x \geq 1)]$$

Solving for n gives

$$n = \frac{\log[1 - P(x \geq 1)]}{\log(1 - f)} \quad (1)$$

(for log to any base), which allows one to find the necessary sample size for a given probability, P , of detecting at least one resistant individual for a given resistance frequency, f . Sample results are shown in Fig. 2, which shows that large sample sizes are needed to detect low resistance frequencies. There is, for example, only a 63% chance of detecting resistance present at a frequency of 1% in a random sample of 100 individuals. To obtain a 95% probability of detecting resistance present at this frequency, one needs a random sample of 298 individuals. On the other hand, a sample of 50 individuals would be suitable for documenting resistance present at a frequency of 10% or higher.

Where the standard test technique is not perfectly diagnostic, one cannot simply rely on detecting at least one resistant individual. The use of an LD₉₀ (or similar lethal dose) as recommended above, assumes that 1% of the treated susceptible individuals will survive. Thus, a statistical test must be used to determine if the observed fraction of survivors is significantly greater than the 1% expected to avoid reacting to too many false positives.

A suitable test would be a one-sided Z test at the $100(1 - \alpha)\%$ confidence level with correction for continuity (Snedecor and Cochran 1967: 209):

$$Z = \frac{|s - ng| - \frac{1}{2}}{\sqrt{ng(1 - g)}} \quad (2)$$

where s is the observed number of survivors, n is the sample size, and g is the fraction of survivors expected in a susceptible strain. In the context of this statistical test, resistance detection means obtaining a statistically significant number of survivors when resistance is present at (or greater than) the assumed frequency. The approximate probabilities of detecting resistance with this test are shown in Fig. 3 for various sample sizes when resistant individuals are present at frequencies of 10 and 1%. These probabilities were calculated by adding the probabilities of all possible combinations for the numbers of resistant and susceptible survivors that would fail to give significance, and then subtracting this sum from 1. A probability estimate for resistance detection at a sample size of 100 at 1% R for $\alpha = 0.05$ is not shown in Fig. 3, but provides a concise example to illustrate the method. By rearranging formula 2, the number of survivors, s , necessary to achieve significance is

$$s = Z \sqrt{ng(1-g)} + ng + \frac{1}{2}$$

When $Z = 1.65$ (for $P \leq 0.05$ in a Z test), $n = 100$ and $g = 0.01$, $s = 3.14$. To achieve statistical significance at the 95% confidence level, therefore, one needs at least four survivors. The probability of observing four or more survivors is equivalent to $1 - P$ (observing three or fewer survivors). Since the survivors may include susceptible or resistant individuals (or some combination of those), for each number of survivors, say X , there are $X + 1$ combinations of susceptible and resistant animals that sum to X surviving individuals. In the current example, the possible combinations of resistant and susceptible survivors, each of which makes up 1% of the theoretical population (1% R frequency; tests at the LD_{95} for the susceptibles) are as follows:

$X = 3$	$X = 2$	$X = 1$	$X = 0$
3 0	0 2	1 0	0 0
2 1	1 1	0 1	
1 2	2 0		
0 3			

Each of these combinations has independent probability of occurring in a sample of 100. For example, the binomial probability of finding three resistant (or susceptible) survivors in a sample of 100 is ca. 0.061. Similarly, the probability for zero susceptible survivors in such a sample is ca. 0.366. The number of resistant individuals drawn in any given sample should be independent of the number of susceptible survivors. Therefore, the probability of obtaining three resistant and zero susceptible survivors is $0.061 \times 0.366 = 0.0223$. Similar calculations were made for the other nine pairs in the matrices shown above.

The sum of these probabilities, 0.858, was subtracted from 1.0 to give the probability, 0.142, of having four or more survivors (i.e., finding significance). Note that this procedure assumes that resistance was detected even when zero resistant sur-

vivors were found, as long as four or more susceptible survivors occurred. We considered this to be a legitimate result because the correct decision (resistance is present) would have been made, even if for the wrong reason. The curves shown in Fig. 3 are approximate because round-off errors occurred when only integers (whole individuals) were used for the number of survivors.

A little experimentation with the Z test will convince the reader that a small error in the estimation of the LD_{95} can cause considerable error in finding significance. That is why we recommend that a large empirical test of the LD_{95} be made. To follow up on another point mentioned earlier, further experimentation will convince the reader that an even larger sample size would be required to achieve significance if a lower lethal dose, say the LD_{99} , were used.

As shown in Fig. 3, sample sizes necessary to achieve a high probability of detecting resistance when a statistical test must be used can be very high when resistance is rare. A 95% probability of detecting resistance at a 1% frequency and $\alpha = 0.05$ requires a sample of 1,500 individuals, roughly 5-fold the number required for a perfectly diagnostic test (Fig. 2). Increasing α to 10% (i.e., placing a 90% confidence on the test) does not help very much (Fig. 3). The situation is not so bleak when resistant individuals are present at a 10% frequency, where a sample size of about 70 will provide a 95% probability of detection at $\alpha = 0.05$ (Fig. 3), only about twice the number required for a perfectly diagnostic test (Fig. 2). Nonetheless, as a general rule, reliance on an LD_{95} for detection requires larger to very much larger samples than if a perfectly diagnostic test is available.

The sample sizes given in Fig. 2 and 3 should be viewed as minimal to achieving the desired probability of detection. On the assumption that the LD_{95} may kill some R individuals, the sample size should perhaps be doubled. Field-collected individuals must often be reared in the laboratory for one or more generations before testing (e.g., Twine and Reynolds 1980). Many field-collected individuals will fail to reproduce so that the genetic number actually tested will often be much smaller than the number sampled from the field no matter how many are tested in the laboratory. In many cases, the field sampling scheme will have to consider the dispersive characteristics of the species. For example, spider mites at low densities do not disperse extensively. All of the mites on a given leaf may be closely related (Helle and Overmeer 1973). Therefore, a random sample of 30 leaves with an average of 10 mites per leaf probably constitutes a genetic sample of closer to 30 than 300 independent individuals. We have not tried to deal with what constitutes a field population here because that depends on species and habitat-specific dispersal, but sampling should often be geographically intensive since resistance is often localized (e.g., Dennehy and Granett 1984). That

is, one should consider samples of hundreds per community rather than per region to detect resistance foci before they spread.

Saving Survivors. A possible alternative to the large sample sizes called for in the previous section is to rear offspring from the survivors of an LD_{50} (or similarly high value) treatment and test them for resistance. This is desirable for a number of reasons, particularly since intense selection on such a colony may produce a strain that can be used to develop a diagnostic test for resistance. Unfortunately, however, this process may take several generations. The statistical approach outlined above will probably be necessary for quickly identifying resistance problems.

Practical Considerations and Conclusions. The use of LD_{50} , slope, and LD_{95} in resistance monitoring appears to be inefficient compared with a diagnostic test for resistance detection and documentation. In addition, even with diagnostic tests, the sample sizes necessary to detect resistance at the nominally low frequency of 1% can be quite large (on the order of several hundred individuals per location).

The inadequacy of most current monitoring programs may be illustrated by the recent development of pyrethroid resistance in *Helicoverpa* (= *Heliothis*) *armigera* (Hübner) in Australia. A monitoring program using the test procedure recommended by the Entomological Society of America (ESA) was in place as early as 1977 (Gunning et al. 1984). At least 50 larvae were collected from each site, which presumably exceeded the ESA standard of "at least 10 mated females" (Anonymous 1970, p. 147). At least 150 larvae were tested; LC_{50} 's and slopes were estimated. In spite of this effort, there was no hint that resistance was developing until control failures occurred in the field at Emerald in 1983. It is important to note that samples were tested from Emerald in 1979 and 1982, yet even these failed to show any changes in either LC_{50} 's or slopes.

The inadequacy of this standardized approach was recognized by Gunning and colleagues, who adopted a discriminating dose approach in 1983. This example suggests that most currently accepted procedures for monitoring resistance are inadequate for resistance management, including several that call for the development of dose/response curves and minimum samples at the highest doses of only 20-50 (e.g., Anonymous 1968, 1970, 1972) or even 100 individuals (Anonymous 1974).

For some species, such as spider mites, aphids, and mosquitoes, testing of the large sample sizes indicated in this paper may be economically feasible. On the other hand, high confidence in detection at a 1% resistance frequency may be practically impossible for many critical species (e.g., *Heliothis*) unless radically new collection and testing techniques are developed. One such approach for codling moths, *Cydia pomonella* (L.), has been suggested by Riedl et al. (1985). In cases where

such techniques are not available, it may be necessary to adopt prophylactic resistance management tactics (i.e., adopt practices on the assumption that resistance will develop otherwise, rather than rely on resistance monitoring to warn of impending problems).

There is much to be gained by resistance monitoring. However, for monitoring to achieve its full potential in resistance management, it may be necessary to modify standardized monitoring procedures that were first adopted over a decade ago. Not only will it be necessary to use increased sample sizes, it may also be desirable to modify the test techniques themselves.

Acknowledgment

We thank K. R. Hopper, S. B. Stark, W. J. Drapala, J. C. Schneider, J. R. Leeper, J. DeAngelis, and particularly B. H. Stanley and M. D. Morris for critical comments on the manuscript. This article is Miss. Agric. For. Exp. Stn. Paper No. 6209.

References Cited

- Anonymous. 1968. First conference on test methods for resistance in insects of agricultural importance. Method for the boll weevil and tentative method for spider mites. Bull. Entomol. Soc. Am. 14: 31-37.
1970. Second conference on test methods for resistance in insects of agricultural importance. Standard method for detection of insecticide resistance in *Heliothis zea* (Boddie) and *H. virescens* (F.). Tentative methods for detection in *Diabrotica* and *Hypera*. Bull. Entomol. Soc. Am. 16: 147-153.
1972. Standard methods for detection of insecticide resistance in *Diabrotica* and *Hypera* beetles. Bull. Entomol. Soc. Am. 18: 179-182.
1974. Recommended methods for the detection and measurement of resistance of agricultural pests to pesticides. FAO Plant Prot. Bull. 22: 127-137.
- Ball, H. J. 1981. Insecticide resistance—a practical assessment. Bull. Entomol. Soc. Am. 27: 261-262.
- Beeman, R. W. 1983. Inheritance and linkage of malathion resistance in the red flour beetle. J. Econ. Entomol. 76: 737-740.
- Curtis, C. F. 1985. Theoretical models of the use of insecticide mixtures for the management of resistance. Bull. Entomol. Res. 75: 259-265.
- Dennehy, T. J., and J. Granett. 1984. Monitoring dicofol-resistant spider mites (Acari: Tetranychidae) in California cotton. J. Econ. Entomol. 77: 1386-1392.
- Dennehy, T. J., J. Granett, and T. F. Leigh. 1983. Relevance of slide-dip and residual bioassay comparisons to detection of resistance in spider mites. J. Econ. Entomol. 76: 1225-1230.
- Georghiou, G. P., and C. E. Taylor. 1977. Operational influences in the evolution of insecticide resistance. J. Econ. Entomol. 70: 653-658.
- Gunning, R. V., C. S. Easton, L. R. Greenup, and V. E. Edge. 1984. Pyrethroid resistance in *Heliothis armigera* (Hübner) (Lepidoptera: Noctuidae) in Australia. J. Econ. Entomol. 77: 1283-1287.
- Helle, W., and W. P. J. Overmeer. 1973. Variability

- in tetranychid mites. *Annu. Rev. Entomol.* 18: 97-120.
- Miyata, T. 1983. Detection and monitoring methods for resistance in arthropods based on biochemical characteristics, pp. 99-116. In G. P. Georghiou and T. Saito [eds.], *Pest resistance to pesticides*. Plenum, New York.
- Riedl, H., A. Seeman, and F. Henrie. 1985. Monitoring susceptibility to azinphosmethyl in field populations of the codling moth (Lepidoptera: Tortricidae) with pheromone traps. *J. Econ. Entomol.* 78: 692-699.
- Robertson, J. L., K. C. Smith, N. E. Savin, and R. J. Lavigne. 1984. Effects of dose selection and sample size on precision of lethal dose estimates in dose-mortality regression. *J. Econ. Entomol.* 77: 833-837.
- Snedecor, G. W., and W. G. Cochran. 1967. *Statistical methods*. Iowa State University, Ames.
- Staetz, C. A. 1985. Susceptibility of *Heliothis virescens* (F.) (Lepidoptera: Noctuidae) to permethrin from across the cotton belt: a five-year study. *J. Econ. Entomol.* 78: 505-510.
- Tabashnik, B. E., and B. A. Croft. 1982. Managing pesticide resistance in crop-arthropod complexes: interactions between biological and operational factors. *Environ. Entomol.* 11: 1137-1144.
- Twine, P. H., and H. T. Reynolds. 1980. Relative susceptibility and resistance of the tobacco budworm to methyl parathion and synthetic pyrethroids in southern California. *J. Econ. Entomol.* 73: 239-242.
- Wolfenbarger, D. A., J. R. Raulston, A. C. Bartlett, G. E. Donaldson, and P. P. Lopez. 1982. Tobacco budworm: selection for resistance to methyl parathion from a field-collected strain. *J. Econ. Entomol.* 75: 211-215.

Received for publication 10 September 1985; accepted 12 November 1985.
