30

Overview of the identification of Africanized honey bees

Howell V. Daly

INTRODUCTION

This is the second opportunity that I have had to review progress and problems in the identification of Africanized honey bees. In 1984 at the National Conference of the Entomological Society of America, I reported that our 1978 morphometric analysis had proven reliable and the only method that had been extensively tested. For allozyme analysis, additional loci were needed because the fast allele of MDH was being found in Europe and could no longer be considered diagnostic for Africanized bees (Nunamaker & Wilson 1981, Badino *et al.* 1983). The paper by Carlson & Bolten (1984) had just appeared on the use of cuticular hydrocarbons. This approach seemed well suited to identification if problems with age-dependent variation and contamination could be overcome. The most exciting prospect was the use of DNA restriction fragment polymorphism. Avise *et al.* (1979) had shown the technique to be successful in natural populations of mice.

Today we have an array of new techniques for the identification of honey bees. Morphometric analysis now includes the simplified FABIS methods of Rinderer *et al.* (1986) and automatic image analysis described by S. Batra in this volume. Analysis of cuticular hydrocarbons has attracted much recent attention. We are fortunate to have reports by all the key investigators: D. A. Carlson, C. A. McDaniel, and R. K. Smith. The call for research on DNA restriction fragment polymorphism has been answered by several scientists. Credit goes to G. Hall (1986) for the first demonstration that genetic differences among honey bees could be detected by this method. He is joined in this volume by W. S. Sheppard, D. R. Smith and D. W. Severson. In this area we also have all the leading investigators represented. I have reserved mention of M. Spivak's contribution until last because it represents what I hope will be the future direction of our efforts to identify honey bees; namely, a combination of methods that are based on independent sources of information.

Why do we need so many ways to identify honey bees? The remainder of this paper is devoted to the lessons to be learned from systematic biology and a discussion of some obvious and not so obvious problems associated with our current methods.

CONCLUSIVE AND PROBABLE IDENTIFICATIONS

Identifications of organisms are made with various degrees of assurance. When specimens of a species have unique and clearly defined structural or other characters, then the identifications are irrefutable within the context of the current classification. Such identifications are *conclusive*. For example, *Apis mellifera* is easily distinguished from its congeners *A. dorsata* and *A. florea*. We expect to find structural, behavioral, biochemical, and genetic 'gaps' between these species because their evolutionary lineages are independent and have diverged. Even the nearest relative of *A. mellifera, A. cerana*, has an easily visible structural character: two veinlets extend distad from the large basal cell of the hind wing rather than one as in *A. mellifera*. This and other characters have provided consistent differences between the two species. The characters are also indirect evidence that the species do not interbreed in nature. Thanks to Ruttner & Maul (1983), we now have experimental evidence of reproductive isolation between the species. Therefore, specimens of *Apis mellifera* can be conclusively identified as members of a single biological species.

Geographic races, subspecies, genetic hybrid swarms, and biotypes within a species, however, usually cannot be conclusively identified. Rarely do sharp boundaries exist that permit us to easily recognize all members of distinctive populations. Differences tend to be reduced by the exchange of genetically determined traits among adjacent populations. Furthermore, intervening populations or hybrid zones usually show intermediate character states because most features of the phenotype are inherited as quantitative characters (Falconer 1981). Comparison of samples known to be typical of two or more populations allows one to estimate how often an identification based on certain characters is likely to be correct. Identifications under these circumstances are *probable* rather than conclusive.

ANALYSIS OF GEOGRAPHIC VARIATION AND THE FATE OF SUBSPECIES

The analysis, interpretation, and taxonomic recognition of geographic variation within species has been of continuing concern to systematists. Early attention focused on the description of subspecies, often based on a few characters of external anatomy. The assumption was that the variation of other unstudied characters would coincide, i.e. vary concordantly. Difficulties in delimiting subspecies led to the '75-percent rule' which, in effect, meant that to be valid about 90% of each of two subspecies in question should be distinguishable (Mayr 1969). When geographic variation in many characters was studied within species, numerous cases were found where variation was discordant. In 1953, Wilson & Brown argued that the naming of subspecies was inadequate to deal with complex variation. Even with quantitative methods the number of names was almost limitless. Subsequently, the practice of

naming subspecies, as well as theoretical interest in subdividing species into identifiable groups, markedly declined (Selander 1967). Analysis of geographic variation is still of central interest in speciation theory (Futuyma 1979) and as the basis for taxonomic decisions about species as a whole (Liebherr 1986).

From this brief historical review, the lesson for us is that the separation of Africanized honey bees from European honey bees is likely to be much more complicated than would appear on the first examination of a few characters. We need many sources of information to minimize the number of misclassified samples. Differences between Africanized and European bees in South America were doubtless maximal when the former were introduced in 1956. The ancestral populations in Europe and Africa had been separated by a distance of more than 70° in latitude, and the Sahara Desert had reduced gene flow for at least 2000 years. As the Africanized bees approach the United States, they have had 30 years to accumulate the effects of natural selection in new environments, mutation, and genetic drift. These processes usually promote evolutionary divergence, but Africanized bees also have had the opportunity to interbreed to an unknown extent with European bees. The resulting hybrids and backcrosses promote gene flow and serve to reduce the differences between the ancestral populations. Thus our task in identifying Africanized bees is similar to the problem of identifying subspecies; a problem that systematists have essentially abandoned.

BASE LINE DATA FOR IDENTIFICATION PROCEDURES

The accuracy of probable identifications depends entirely on how representative the initial samples are with respect to the total populations to be identified. At the outset of a project there is often the illusion of success when analysis of the first samples shows strong differences. However, a reliable procedure requires a large data base and repeated blind testing. Both Africanized and European bees in the Western Hemisphere are genetically heterogeneous. Any procedure for making probable identifications should be based on as broad a sampling of this heterogeneity as possible. Populations of Africanized bees also have been found to vary in time: in the Panama Canal Zone they became progressively smaller or morphometrically more African-like from 1982 to 1985 (Boreham & Roubik 1987). European bees are a mixture of races, including minor introductions from Africa even before the advent of Africanized bees (Morse *et al.* 1973). Samples must include both managed and feral colonies. The latter may be reservoirs of genetic variation not represented in managed colonies.

The sample unit is usually a collection of bees from a colony, and identification is based on pooled extracts or averages of characters of the collection. Some procedures can identify individual bees. A complication for all procedures is the possible mixture of Africanized and European workers in a single colony. This could occur by drift between colonies, or when an Africanized colony is in the process of taking over a European colony, or the queen may produce a mixture of daughters because she was inseminated by both kinds of drones. Inclusion of such mixed colonies in base line data will confuse the evaluation of differences between relatively pure Africanized and European colonies.

INTERPRETATION AND RISK OF PROBABLE IDENTIFICATIONS

Probability statements of identification must be interpreted within the context of the statistical procedure and the data base on which the analysis is made. For example, the morphometric analysis of Daly & Balling (1978) was made on collections of worker bees that were determined to be European or Africanized according to the geographic location of the collection and the judgement of the collector. The discriminant analysis of colonies was based on the means of measurements of the collections from colonies. The statement that a new colony is Africanized at 0.7 or 70% probability also indicates the sample is European at 0.3 or 30% probability. The sample could be of normal Africanized bees or normal European bees, but it is more likely to be the former based on the previous analysis of known Africanized and European bees. The statement about the colony does not mean that the colony is composed of 70% Africanized bees and 30% European bees or that workers have 70% Africanized genes and 30% European genes. To make such statements, the procedures must be able to distinguish individuals or be based on genetic analyses, respectively. Furthermore, the statement that a new sample is Africanized at 1.0 or 100% probability is not a conclusive identification; it is still a probable identification based on the initial analysis of known Africanized and European bees. Conflicting identifications of new samples by different methods are not uncommon. The new sample could be from a distinctive population not represented in the base line data of the original analyses.

All probable identifications carry the risk of actual misidentification. Any method (morphometric, biochemical, behavioural, or genetic) that yields a probable rather than conclusive identification carries this risk. When large numbers of samples are being identified, even a small risk becomes an important consideration in terms of the number of samples that may be misidentified. Special care must be exercised when one or a few Africanized samples are suspected in the midst of a large population of European bees. Because identifications are based on probability statements, the suspected Africanized bees may, in theory, be indistinguishable statistically from the 'tail' of a very large distribution of European samples.

An option available in discriminant analysis is to assign a 'prior probability' estimate (Norusis 1985) that takes into account the expected frequency of Africanized bees in a given area. The usual procedure is to set the prior probabilities equal. By suitably changing this estimate, a more extreme score is required to classify a sample as Africanized. The alternative that I prefer is to leave the prior probabilities equal but require a high level of probability before accepting a sample as identified. This practice was introduced by Rinderer *et al.* (1986) who required greater than 0.90 probability of membership in one of the groups before accepting the sample as identified. Those less than the criterion were considered unidentified and subject to other kinds of analysis.

CONCLUDING REMARKS

The problem of identifying Africanized bees can be considered at two extremes: (1) 'new introductions', or detection of the first Africanized bees to arrive in areas previously occupied by European bees; and (2) 'hybrids', or detection of genetic

crosses and backcrosses between Africanized and European bees in areas where they have interbred. In the first situation, Africanized and European bees are relatively distinct, and phenotypic methods are fast and inexpensive. For surveys, the most efficient approach is to use several steps in the identification such as those described by Sylvester & Rinderer (1986). However, rather than depend on characters from one aspect of the phenotype such as morphometrics, I recommend that different, uncorrelated aspects of the phenotype such as morphometrics, cuticular hydrocarbons, and allozymes be combined in joint probability statements. When less expensive genetic methods are developed, they could be added to the process.

In the second situation a spectrum of genotypes may exist in an area together with one or both parental types. Genetically based methods will be essential because phenotypic methods are not likely to provide the necessary discrimination of degrees of relationships.

ACKNOWLEDGEMENTS

I am indebted to Kim Hoelmer for his careful reading of the manuscript.

REFERENCES

Avise, J. C., Lansman, R. A. & Shade, R. O. (1979) *Genetics* **92** 279–295.
Badino, G., Celebrano, G. & Manino, A. *Heredity* **74** 443–446.
Boreham, M. M. & Roubik, D. W. (1987) *Bull. Entomol. Soc. Amer.* **33** 34–39.
Carlson, D. A. & Bolten, A. B. (1984) *Bull. Entomol. Soc. Amer.* **30** 32–35.
Daly, H. V. & Balling, S. S. (1978) *J. Kansas Entomol. Soc.* **51** 857–869.
Falconer, D. S. (1981) *Introduction to quantitative genetics* 2nd ed, Longman, New York.
Futuyma, D. J. (1979) *Evolutionary biology* Sinauer, Sunderland, Mass.
Hall, H. G. (1986) *Proc. Natl. Acad. Sci. USA* **83** 4874–4877.
Liebherr, J. K. (1986) *Univ. California Publ. Entomology* **106** 198 pp.
Mayr, E. (1969) *Principles of systematic zoology* McGraw-Hill, New York.
Morse, R. A., Burgett, D. M. & Ambrose, J. T. (1973) *Bee World* **54** 57–60.
Norusis, M. J. (1985) *Advanced statistics guide SPSSX* McGraw-Hill, New York.
Nunamaker, R. A. & Wilson, W. T. (1981) *J. Kansas Entomol. Soc.* **54** 704–710.
Rinderer, T. E., Sylvester, H. A., Brown, M. A., Villa, J. D., Pesante, D. & Collins, A. M. (1986) *Apidologie* **17** 33–48.
Ruttner, F. & Maul, V. (1983) *Apidologie* **14** 309–327.
Selander, R. K. (1967) *Systematic Zoology* **16** 286–287.
Sylvester, H. A. & Rinderer, T. E. (1986) *Amer. Bee J.* **126** 330–333.
Wilson, E. O. & Brown, W. L. (1953) *Systematic Zoology* **2** 97–111.