Host range, specificity and recruitment: synthesis of session 2

A.S. McClay

Alberta Environmental Centre, Bag 4000, Vegreville, Alberta, T9C 1T4, Canada

Abstract. The papers in the session on 'Host Range, Specificity and Recruitment' centred around three questions: how to determine the host range of a biological control agent; whether we can understand the origins of that host range; and whether we can use that understanding to increase the effectiveness or safety of biological control of weeds. Many methods have been used for determining insect host-ranges in the field; however, it is suggested that the possible use of gut-content analysis has been neglected. Questions of sample size and replication in screening tests have never been analysed statistically. These questions cannot be answered without an analysis of acceptable levels of risk; it is suggested that screening should be viewed as a risk-assessment process. The major trend in recent studies of insect host-ranges has been the emergence of an evolutionary perspective, both on the macroevolutionary or phylogenetic level, and on the microevolutionary or genetic level. A phylogenetic approach may help to assess the stability of a particular host-plant association over evolutionary time, and hence contribute to the risk-assessment process. Phylogeny reconstruction may also be useful in determining the areas of origin or centres of diversification of target-weed species. Finally, it is suggested that it would be feasible to develop a comprehensive global database of insect host-plant records, and that this would be a useful tool in biological control of weeds as well as in other areas.

Introduction

The papers presented in this session are linked by a natural set of questions, which have always been central concerns in biological control of weeds:

- how can we determine which plant species will be attacked by a given insect or plant pathogen?
- can we understand how insects and plant pathogens come to attack the particular plant species which form their host range (and not others)?
- can we use this understanding to increase the effectiveness or safety of biological control of weeds?

I do not propose to undertake a detailed review of the methods or concepts of host-specificity testing here, as these issues have been extensively discussed in previous Symposia and elsewhere (e.g. Cullen 1990; Weidemann and TeBeest 1990). I will comment on two aspects of host-range testing which have not received much attention in the past, one relating to technique and one relating to design. The second of these leads into a consideration of the aims of host-specificity testing. I will then discuss the relevance of phylogenetic and evolutionary approaches to host-range studies, and finally forward a proposal for making more effective use of the existing body of knowledge on insect-host-plant associations.

Host-range testing

Gut content analysis - a possible tool?

Many experimental methods have been used in the assessment of the host range of candidate biocontrol agents for weeds. One approach which is widely used in feeding studies with other groups of organisms, however, seems to have been surprisingly neglected in weed biocontrol. Pomerinke and Thompson (1995) used gut content analysis to show that the weevil Cleomidius trivittatus (Say) feeds only on Astragalus mollissimus Torr. (Fabaceae). I have not found any other record of the use of this technique for determining the host range of a candidate biocontrol agent. A variety of techniques could be used to identify plant material in the gut of an insect. If the potential food-plants can be distinguished morphologically (by characteristic trichomes, epidermal cell shapes, etc.) they could be identified by direct microscopic examination (Pomerinke and Thompson 1995). Specific polyclonal antibodies for discriminating among a group of potential host-plant species could be
developed by preparing immune sera to extracts of each of the plants and cross-absorbing with extracts of the other plants to remove potential cross-reactive antigens. Plant proteins in the gut contents could then be detected by Western blot analysis or by ELISA (Hagler and Naranjo 1994; Hagler et al. 1994; N. Kedersha personal communication.). Methods could probably also be developed to detect characteristic plant secondary-compounds or specific DNA sequences.

Gut content analysis of field-collected insects could be a useful supplement to other methods for determining the actual field host-range of insects which accept a range of related plant species in laboratory feeding tests. It would be primarily useful for species with mobile, externally-feeding adults or larvae, which may have fed on plants other than those on which they were collected or observed feeding.

**Statistical aspects - sample size and replication**

One aspect of host-range testing which has received little attention is the question of sample sizes, replication and statistical analysis. How many individual insects or test plants should be used in order to get reliable information on the host-specificity of a candidate agent? How many times should a test be replicated? How many eggs should be laid in an oviposition test to give reliable results? What kind of statistical analysis is appropriate for host-range tests? Sample sizes in screening tests appear to have been determined entirely by the judgment of investigators in the past. Harley and Forno (1992), for instance, recommend that at least five pairs of adults should be used, and that each test should be repeated at least three times, but give no reasoning to support this. Although this is not necessarily a serious flaw in previous testing, it is always desirable to be able to justify our procedures to other scientists and to regulators. Questioning the rationale for sample sizes can also lead to a clearer appreciation of exactly what types of conclusions can or cannot be drawn from screening tests.

The standard approaches for estimating required sample sizes do not seem to fit the ease of screening tests. Normally these depend on calculating the sample size required to detect a given level of difference between two treatments, for a known level of variability in the response to be measured. The null hypothesis would thus be that attack on a test plant species is the same as on the target weed. However, rejecting this null hypothesis is not an interesting result - we are not interested in whether attack on the test plants is somewhat less than on the target weed, but in showing that attack on the test plants is low enough that they will not be at risk from the agent in the field. This is not solely a statistical question, but requires a risk assessment (Chesson 1990; Bourchier and McCarty 1995; Harris and McEvoy 1995) and a decision about acceptable levels of risk.

**Risk assessment**

The fact that agent screening is a risk-assessment process can be seen clearly if we ask what we are trying to prove. The traditional answer to this question, and one often given to the general public, is that the aim of screening is to prove that agents are safe for release, or that they will cause no damage to non-target plants. Logically this is impossible. With a finite sample we can never prove a negative statement about an entire population, and we know from the evolutionary studies referred to below that host shifts do occur. A more realistic aim for the screening process is to assess, as accurately as possible, the risks associated with introducing the agent into a new area.

Several models of the risk-assessment or risk-management process exist; a typical structure is shown in Fig. 1. The terminology of these models has been mainly derived from environmental toxicology, and sounds a little strained when applied to introductions of living organisms. However, the models have the benefit of requiring us to think about the consequences of introducing biocontrol agents in a broader context than simply assessing the specificity of a particular agent-host association. "Receptor characterization" involves...

---

![Fig. 1. General risk-management framework (from Bourchier and McCarty 1995).](image-url)
identifying the organisms likely to be affected by the introduction of the agent. These need not be limited to host plants; Jayanth et al. (1993) found that adult Zyogramma bicolorata Pallister (Coleoptera: Chrysomelidae) would feed on foliage of a non-host plant, sunflower (Helianthus annuus L.) in the field if pollen of the host species, Parthenium hysterophorus L., had been deposited on the foliage. 'Exposure characterization' includes how the agent gets into the environment, where it goes and how long it persists. In biological control terms, this requires us to look at the possible dispersal mechanisms and pathways of the agent, its habitat preferences, climatic requirements, rates of population increase, and its potential for evolutionary change. Human dispersal of the agent should not be overlooked in this context, as illustrated by Pemberton (1995).

Risk-assessment concepts have been incorporated into screening tests in several ways. Wan and Harris (this Volume) attempted to predict the relative levels of attack by a proposed biological control agent for Cirsium arvense (L.) Scop. on various native non-target Cirsium spp., by comparing rates of field host-finding, adult feeding, oviposition, larval survival and cold-hardiness on the different non-target species. The analysis included scenarios where the non-target species occur by themselves or together with C. arvense, using data from both choice and no-choice tests. Another approach is to compare the characteristics of a proposed agent with those of other species already present in the proposed area of introduction (Bruckart et al. this Volume). If it can be shown that the risks from an agent are less severe than those of other species already present and considered acceptable in the system, there is a good case for releasing the agent.

When viewed as a risk assessment, screening is not an isolated process which attempts to determine an inherent, abstract property of the agent, but an attempt to assess the real-world consequences of a specific, proposed action. Tests should therefore be designed with a view to predicting the behaviour and effects of the candidate agent when released into a specific, ecological context. This context includes the climatic, floristic, faunistic, geographical, economic and political characteristics of the proposed release area. As part of the risk-assessment process, risks associated with alternative courses of action should also be considered. What would be the consequences of using alternate means of control for the weed, or of leaving it uncontrolled? Such a comparison would, in many cases, show that the perceived or real risks of biological control are much less than those of the alternatives.

This approach would set clearer standards for the acceptance of candidate biocontrol agents, as acceptable levels of risk would need to be defined. In turn, these standards should help to determine the levels of precision, and hence of replication, needed in screening tests. For example, if we know that the rate of larval survival on a critical non-target species must be less than 1% of that on the target weed, experiments can be designed to provide data of the required level of precision.

The final stage of a risk assessment is to test predictions of risk against observed effects in the field. The most testable predictions are likely to be those concerning attack on non-target species related to the target weed. In many cases, permission has been granted to release biocontrol agents despite the fact that pre-release studies have shown that they have some tendency to attack related non-target species. These releases may have been approved because attack on non-targets was considered to be an artefact of test procedures and predicted to be insignificant under field conditions, because the non-target species were not considered important, or because of the urgent need for control of the target species. In all these cases, it is important to compare the outcome of releases with the predictions from pre-release studies.

I expect to have the opportunity to test one such case. The chrysomelid beetle Lema cyanella (L.) was approved for release in Canada as a biocontrol agent for C. arvense on the basis of tests by Peschken (1984). He predicted that although some native North American Cirsium spp. were accepted in laboratory and field-cage tests, these species would be relatively free from attack in the field because of their lower density, different habitats, and, or, their utilization by native phytophagous insects. Lema cyanella has now overwintered and bred successfully at two sites in Alberta (McCay this Volume). At one of these sites, a few patches of the native thistle Cirsium flodmannii (Ryd.) Arthur occur, interspersed with larger areas of C. arvense. If L. cyanella becomes established at this site, it should be possible to estimate relative attack-rates on the two species and compare these with the predictions of Peschken (1984).

Similar predictions have been made by Hasan and Delfosse (1995) for attack by a rust on native
Australian *Heliotropium* spp., by Hill and Halley (1995) and Ockers et al. (1995) for attack by biocontrol agents for introduced *Solanum* spp. on *eggplant* (*Solanum melongena* L.) in South Africa, by Kok et al. (1992) and Blossey et al. (1994) for attack by biocontrol agents for *Lythrum salicaria* L. on native Lythraceae in North America, and by Willis and Ash (this Volume) for attack by an eriophyid mite on native *Hypericum* species in Australia. Careful monitoring of the results of these and similar releases will be a valuable test of our ability to predict the extent of recruitment by introduced biocontrol agents onto non-target hosts.

**Evolutionary approaches to host-range**

Perhaps the most significant recent development in studies of insect-host-plant associations has been the increasing emergence of an evolutionary approach. This includes both macroevolutionary studies on the phylogeny of host-plant associations (e.g. Farrell et al. 1992; Radke and Singer 1995; Briese this Volume; Kovalev and Zaïsev this Volume) and microevolutionary studies on the genetics of host-plant selection (Futuyma et al. 1994; Carrière and Rothberg 1995). This trend has been encouraged by the wider acceptance of cladistic methods, the availability of computer methods for reconstruction of phylogenies, and the increasing availability of molecular tools for looking at DNA and protein sequences. The interest of weed biocontrol practitioners in this area is obvious: if we understand the evolutionary development of host-plant associations, we are in a better position to assess their stability and hence the risks of host-plant shifts.

If insects are primarily 'carried along' by their host plants, speciating only in response to the appearance of new-host lineages, the phylogenies of insects and host plants should be highly congruent. A lack of congruence indicates that host-shifting has taken place during the evolution of the insect clade (Brooks 1988). Despite the limited number of studies which have been carried out in this area, a consensus is emerging (Anderson 1993). Few cases are known where the phylogeny of an insect group corresponds strictly to that of the host plants (Farrell and Mitter 1990). Most studies have shown some degree of incongruence between the insect and host phylogenies, suggesting that host shifts have been relatively frequent over evolutionary time (e.g. Weintraub et al. 1995). These shifts are, however, subject to chemical, taxonomic, or ecological constraints, so that insects are more likely to shift hosts to species which are chemically similar, closely related or grow in the same habitats as their present hosts. This leads to the observed pattern in which many insect taxa have speciated in association with particular plant groups, while occasional closely-related species are associated with unrelated plants. The genus *Aphthona* (Coleoptera: Chrysomelidae), for example, has diversified on *Euphorbia* spp. (Euphorbiaceae) which were presumably the ancestral hosts. Within the genus, occasional host transfers have occurred to species of other families including Rosaceae, Cistaceae, Iridaceae and Geraniaceae (Maw 1981). A phylogenetic analysis of cases such as this should help to indicate the frequency of such host shifts over evolutionary time, and perhaps help to point out groups in which host associations are particularly stable or unstable. In principle, it should be possible to use a phylogenetic reconstruction to date the age of a host association. An estimate that a particular host association has been stable for (say) four million years should be a powerful argument in a risk assessment.

A few caveats should be noted in connection with the application of phylogenetic methods in biological control. Phylogenies reconstructed from different types of data do not always agree. The phylogeny of the subtribe Ambrosiinae (Asteraceae), a group containing several important weed genera such as *Ambrosia, Xanthium, Iva, and Parthenium*, was reconstructed by Miao et al. (1995) from chloroplast DNA data and by Karis (1995) from morphological data. While the two reconstructions agree on some broad points, such as the basal position of the genus *Parthenium* within the subtribe, they differ widely on many other details. It is not yet clear which approach should be considered more reliable. The frequent occurrence of hybridization may also complicate efforts to reconstruct plant phylogenies (Rieseberg 1995), and hybrids between plant species may play an important role in insect host-range expansion or shifting (Flaate and Whitham 1993).

An interesting possible application of evolutionary approaches in biological control concerns determining the centre of origin of a target weed. It is usually recommended that surveys for biological control agents should be carried out in the weed’s centre of origin, on the basis that this is where the maximal diversity of potential biocontrol agents will occur. For some weeds the centre of origin is difficult to determine, either because the species and the genus to
which it belongs now have a cosmopolitan distribution, as with *Chenopodium* spp. (Holm et al. 1979), or because of the difficulty of identifying forms occurring in the area of introduction with those within the native range, as with leafy spurge (*Euphorbia* spp.) (Harvey et al. 1988). Qin et al. (1994) used a phylogeny based on morphological characters to identify the sister groups and hence the probable area of origin of the scale insect *Ceroplastes sinensis* Del Guercio (Hemiptera: Coccidae). This species was previously supposed to be native to Asia, but the distribution of its sister groups strongly suggests an origin in the neotropics. A similar approach should be feasible with weeds. Nissen et al. (1995) discussed the use of DNA-based marker systems to determine the genetic diversity of weedy species; the application of phylogenetic analysis to such data should make these methods even more powerful.

**Databases**

The final area I wish to discuss is the possibility of improving our ability to search and use the vast amount of information which has already been gathered on insect-host-plant associations. I propose that the weed biocontrol community should consider the development of a comprehensive, global, electronic database of insect-host-plant records. I believe that recent advances in information technology, particularly the ability to link computers through global networks, have brought this goal within reach. The ultimate aim would be to allow a user to retrieve quickly a comprehensive list of known host-plant associations for a given insect or plant taxon; included in each record would be information on the geographical area of occurrence of the association, and a reference to the original publication or other data source from which it was derived.

The primary literature on insect-host-plant associations is scattered through an enormous number of journals, books and monographs in many languages and formats; a bibliography of such sources up to 1979 was produced by Kingsolver et al. (1984). These publications vary from original reports of individual associations, e.g. Pomeranke and Thompson (1995), through listings of species associated with particular plant (Palmer and Pullen 1994) or insect (Scott 1977) taxa, to extensive compilations of records for large groups (Scherf 1964; Richards 1976; Stone 1991). While some publications include host-plant lists or indices, in others host-plant information can only be located by a page-by-page search.

Electronic databases, including BIOSIS, CAB ABSTRACTS, AGRICOLA and Zoological Record can be helpful in searching this literature. However, they are databases of literature references rather than direct sources of host-plant records, and may miss many publications if the relevant names are not included in the title, keyword or abstract fields. They are also limited in their time coverage, most going back only to the 1970s. The system I propose here would index host-plant records directly, and would have complete retrospective coverage.

The availability of this resource would facilitate weed biocontrol projects in many ways. A list of potential candidate agents could be obtained quickly for any given target weed. Searching for the alternate recorded hosts of these or related species would eliminate clearly polyphagous species, and identify plant species which should be included in screening tests. Analysis of the geographical areas of occurrence would help to identify centres of diversity which would be promising for further field surveys, as well as pointing out any species already recorded from the areas of introduction. The duplication of effort currently involved in searching through the same literature sources at the start of each new project would be eliminated.

The database would also have applications outside the field of biological control of weeds, including identification of potential pests of new or newly-introduced crop or ornamental plants in a given area, rapid assessment of the risks of newly-introduced insects, quarantine interceptions, conservation and biodiversity studies, and research on the evolution of host-plant associations.

This database may seem an ambitious goal, however, a consideration of the efforts which have already been undertaken, the available technology, and the resources required, show that it is not as unrealistic as it might appear. A number of insect-host-plant databases have already been constructed or are in preparation for particular taxa or geographical areas. The Phytophagous Insects Data Bank (PIDB) includes published and unpublished records covering the majority of British phytophagous insects (Ward 1988). The HOSTS database at the Natural History Museum (London) includes about 42500 Lepidoptera host-plant records from published and unpublished sources, with strength in North America and the Oriental region.
(G. Robinson personal communication). Palmer (1995) describes a database used for field and literature records obtained during foreign exploration. Databases are currently being constructed for phytophagous insects of Argentina (Braun personal communication) and insects associated with Proteaceae in South Africa (Wright personal communication). These projects would provide a useful basis for a global database: the World-Wide Web may provide a means whereby physically-separate databases can be linked into a single functional system.

I have made a rough estimate of the storage space and input time required for a global insect-host-plant database. Bernays and Chapman (1994) estimate that there are about 310000 described species of phytophagous insects. It is probably optimistic to assume that host-plants are known for 10% (31000) of these. Most species are oligophagous. However, a small number of species have many known host plants, and duplicate records will often occur. The PIDB (Ward 1988) currently contains about 50000 host-plant associations for 6935 insect species (L. Ward personal communication), giving an average of 7.2 associations per species. As the British fauna is better known than most areas of the world, I assume an average of five records per insect species, giving a total of 155000 records. Basic information in each record (plant and insect names, authors, higher taxa, locality, data source) could be stored in about 300 bytes. This would give a total database size of about 50 MB. Even allowing for a wide margin of error on this estimate, and for the need for supplementary storage space for index files, bibliographies, and so on, this would be well within the capacity of a standard SQL database system running on a Windows NT or similar workstation. A data entry rate of 100 records per person-day should be achievable, even allowing time for trouble-shooting and for obtaining print literature sources. Total data entry requirements would thus be 5-6 person-years. The funding required would therefore be substantial, but quite small in relation to the total funding currently devoted to biological control of weeds worldwide.

A number of issues would need to be resolved in order to ensure the usefulness of the database. These include questions such as: the data fields which should be included in each record; the criteria for recognition of a host association (Ward 1988); the treatment of ambiguous, duplicate, vague, or negative records; the inclusion of laboratory or ‘unnatural’ feeding records; the tracking of name changes, synonymy and re-identifications; questions of ownership, authorship and access to data; and, of course, funding and cost-recovery issues. I suggest that a working group of biocontrol researchers, taxonomists, and information specialists should be formed to make recommendations on these issues.

The project of a global host-plant database would fit well within the context of a number of current initiatives which apply modern information technology to taxonomic and biodiversity studies. The ‘Species 2000’ program, for instance, is an initiative of the International Union of Biological Sciences, developed in response to the United Nations Convention on Biodiversity, which aims to develop an indexing system for all groups of organisms, with an ultimate goal of listing all known species on Earth. Coordination with these initiatives should help to ensure that taxonomic information in the host-plant database is accurate and can be updated.

Acknowledgements

I thank Neal Spencer and Kevin Floate for their comments on the manuscript, Andrew Keddie and Nancy Kedersha for discussion on methods of gut-content analysis, and Lena Ward, Jack Coulson, Gaden Robinson, Larry Speers, Mark Wright, Karen Braun, and Robin McDaniel, for information and discussion on host-plant databases.

References


Session 2 synthesis


