Ecogeographic surveys

14

N. Maxted[1], M.W. van Slageren[2] and J.R. Rihan[3]

[1]*School of Biological Sciences, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK:* [2]*Genetic Resources Unit, ICARDA, PO Box 5466, Aleppo, Syria:* [3]*Department of Biology, Biomedical Sciences Building, University of Southampton, Southampton SO9 3TU, UK.*

Introduction

To make the most efficient use of limited resources, plant germplasm collectors must have a clearly defined set of target taxa, and must know as much as possible about where (and when) to find these plants within their general target region. Much time and effort can be wasted if collectors do not know enough about the geographic distribution, ecology, phenology and diversity of the plants they are looking for before setting out into the field.

Conditions at the localities inhabited by a species will be characterized by more or less specific environmental constraints. The passport data associated with herbarium specimens, germplasm accessions and other plant records can be used to identify these constraints. For example, if passport data for a particular species or genotype indicate that in the past it has only been found on limestone scree slopes above 2000 metres in southwest Asia, then localities occurring within these parameters are clearly where one should initially look, if further material is being sought. A combination of ecological and geographic passport data from existing collections and the literature can be used to predict where plants may be found and when they are likely to be ready to collect.

Plant collectors are thus like detectives: they gather and analyse the clues in passport data in order to trace the plants that interest them. This is the essence of ecogeographic investigations, the subject of this chapter. IBPGR (1985) summarizes the three major components of ecogeography as the study of:

- distributions of particular species in particular regions and ecosystems;
- patterns of intraspecific diversity;

- relationships between ecological conditions and the survival or frequency of variants.

It concludes that: 'Field data provide a basis for determining how to maximize the sampling of genetic diversity. Ecogeographic information can be used to locate significant genetic material and representative populations can then be monitored.'

An ecogeographic investigation can focus on various levels of the taxonomic hierarchy. It can be concerned with informal categories such as landraces, cultivar groups or races within a crop (Chapter 7); or formal categories such as subspecies, varieties, and forms within a species or species within a series, genus or tribe. It can also deal with wild or cultivated taxa. The kinds of data will be largely the same, though in practice there will be a difference in the sources of the data. For example, among herbaria, only the few specialist institutes that concentrate on cultivated material (e.g. the N.I. Vavilov Institute for Plant Industry in St Petersburg, Russia) will have collections which do not under-represent the variation found within crops. Thus, gene banks will usually be better sources of ecogeographic data on cultivated material than herbaria.

An ecogeographic model

The steps involved in undertaking an ecogeographic study will be discussed below. First, a definition: an ecogeographic study is a process of gathering and synthesizing taxonomic, geographic and ecological data. The results are predictive and can be used to assist in the formulation of collecting and conservation priorities.

The difference between a 'study' and a 'survey' is one of degree. Ehrman and Cocks (1990) provide a good example of an ecogeographic study of the annual legumes of Syria. They present a detailed analysis of the climatic and soil characteristics that influence the distribution of the annual legume species they sampled throughout the country. They suggest that species diversity and seed production are related to annual rainfall and that populations in the drier areas face greater threat of genetic erosion. Based on their analysis of the ecogeographic data, they propose a detailed list of conservation priorities. However, this was only possible because of the very detailed ecogeographic data gathered over several years by the authors. Clearly, considerable time and resources are required to undertake such a study. If ecogeographic data are to be used as a routine part of collecting and conservation, then the quicker, less expensive, option of undertaking a survey is likely to be favoured. A survey will focus on collating data recorded by other plant collectors, rather than obtaining new data. It may be restricted to a literature search and gathering passport data from herbarium specimens and gene bank accessions.

In practice, all conservation activities, collecting not least among them, are necessarily preceded by some form of ecogeographic data collation and analysis. Though this may not follow exactly the methodology proposed here, most ecogeographic surveys or studies will be articulated in three phases, as follows:

- Phase I – Project design:

1. Project commission.
2. Identification of taxon expertise.
3. Selection of target taxon taxonomy.
4. Delimitation of the target region.
5. Identification of taxon collections.
6. Designing and building the ecogeographic database structure.

- Phase II – Data collection and analysis:

1. Listing of germplasm conserved.
2. Survey of taxonomic, ecological and geographic data sources.
3. Collection of ecogeographic data.
4. Data verification.
5. Analysis of taxonomic, ecological and geographic data.

- Phase III – Product generation:

1. Data synthesis.
2. Ecogeographic database, conspectus and report.
3. Identification of conservation priorities.

Phase I – Project design

Project commission

Ecogeographic projects may start in a variety of different ways. An individual collector may simply decide it is necessary to gather some background data prior to setting off for the field, or an international organization may commission a full ecogeographic study of a particular target gene pool as a preliminary to developing a comprehensive conservation strategy. Whatever the case, a taxon or taxa from a defined geographic region must be considered to be of sufficient interest to warrant time-consuming and possibly expensive background research to support subsequent collecting, conservation and use. The range of the study may vary from one species in a restricted area to a whole genus throughout the world, e.g. *Arachis* species in Brazil, *Aegilops* sect. *Sitopsis* in the Near East or *Hordeum* species worldwide. An example of a project commission is provided by Edmonds (1990) in the report of her herbarium survey of the genus *Corchorus* in Africa. She states that:

> A general survey of *Corchorus* L. species was commissioned by the International Jute Organization to provide the necessary background

> data on which future germplasm collecting expeditions could be based ... the survey was required to identify those wild species for potential use in the future genetic improvement of jute, in addition to identifying the countries and locations where collecting expeditions would be most profitable.

Care should be taken in identifying a specialist to undertake or supervise ecogeographic studies. The taxonomy of wild species is sometimes difficult, identification aids often lacking or of poor quality, and retrieving data from older herbarium specimens presents special problems. The ecogeographer need not be an expert in the target taxon, but should have some background knowledge of the group and be experienced in the use of identification aids. Misidentification of material will diminish the predictive value of the data collected. Employing a specialist to gather ecogeographic data might at first sight be considered extravagant, but many taxon specialists may be willing to undertake or supervise such a study, if they consequently had the opportunity to see material from the taxon's centre of diversity to which they might not otherwise have access. The specialist should also, if possible, have a good understanding of the geography of the region to be studied, especially in the case of local studies. This is illustrated by Sánchez and Ordaz (1987), who found local geographic and ecological expertise invaluable in their study of *Zea mexicana* (teosinte) in Mexico. Local expertise may also prove vital in trying to decipher locality details from specimen labels written by hand several decades ago.

Identification of taxon expertise

The acquisition of ecogeographic data will prove much easier if advice is sought from taxon experts at an early stage. They will be able to advise on the accepted taxonomy of the group, recommend (and perhaps provide) possibly obscure Floras, monographs and other literature, advise on any relevant databases, suggest which herbaria and/or gene banks should be visited and provide the ecogeographer with useful local connections. The authors of relevant scientific papers will probably be the first contact points. *Index Herbariorum* (Holmgren *et al.*, 1990) and its companion volume (Holmgren and Holmgren, 1992) list the researchers associated with different herbaria and what their specialities are. Increasingly, herbaria are acquiring electronic mail facilities, and there is a list of Plant Taxonomists Online (contact Jane Mygatt, jmygatt@bootes.unm.edu). A database of experts in botany and mycology worldwide is maintained at the University of Oulu, Finland (contact Anne Jäkäläniemi at the Department of Botany, anne.jakalaniemi@oulu.fi).

Selection of target taxon taxonomy

It is clearly important to have a good taxonomic understanding of the target group prior to undertaking an ecogeographic study. This can be obtained from various sources, in particular target taxon specialists, monographs, recent revisions of the group and, increasingly, taxonomic databases (see section on 'Survey of taxonomic, ecological and geographic data sources' below and Chapter 10). These will help the ecogeographer determine the generally accepted classification of the group, which will list the taxa currently considered members of the target group, their accepted names and the more common synonyms. The taxonomic limits to the study will thus be set.

There may be various alternative classifications of the target group. The ecogeographer must consider these and make a decision as to which one to adopt. Wild species are usually described using a combination of morphological characteristics. A classification using the biological species concept, where genetic data are given greater importance, may be more appropriate if the aim is to conserve maximum genetic variation in the target taxon. However, there are few such biologically based classifications available. They tend to be restricted to well-known crop plants and their allies, where the genetic relationships among the taxa have been extensively studied and the make-up of the gene pool is relatively well understood. Increasingly, however, genetic diversity studies are also being carried out on wild plants using biochemical and molecular markers.

Knowing the accepted classification of a group will provide leads to other literature: iconography, distribution maps, identification aids, autoecological studies, ethnobotanical investigations (including check-lists of vernacular names), bibliographies. More obscure groups may lack a recent revision or monograph, but the researcher must still collate whatever published taxonomic data are available to provide the backbone to the study.

Delimitation of the target region

The target region under study may be restricted by the terms of reference of the project commission, but if it is unspecified the taxon should be studied throughout its range. The commissioning agent may restrict the survey to a specific area (e.g. the Sahel, Vietnam, South America) if the area is fairly clearly defined and/or a complete study would be too costly in time and resources. However, restriction of the target region to save resources in the short term may ultimately prove to be a false economy. Multiple studies of the same taxon, possibly by different authors, are likely to form a less coherent whole. Rihan (1988), who undertook an ecogeographic survey of the forage *Medicago* species of the Mediterranean and adjacent arid/semiarid areas, found that her target region was floristically ill-defined. The natural distribution of the species she was studying did not coincide with the target region she was commissioned to cover. Both of these considerations may unnecessarily limit the predictive value of the ecogeographic study (see also Funk, 1993; Stressey, 1993).

Having established the limits of the target region, additional information on the target taxon can be obtained from local Floras and field guides. These can provide more detailed information on local geographic distribution and ecological preferences. Guides to which Floras cover which parts of the world are provided by Frodin (1984) and Davis *et al.* (1986). Information on sources of information on wild species, including more on these works, is provided in Chapter 10.

Identification of taxon collections

The researcher undertaking an ecogeographic study will need to visit the major herbarium collections of target taxon specimens from the target region. Travelling to herbaria may be expensive and so the selection of which ones are to be visited is crucial. Target taxon and region specialists will be able to suggest which herbaria and libraries the ecogeographer should concentrate on. Part 1 of *Index Herbariorum* (Holmgren *et al.*, 1990) records for each herbarium the historical plant collections conserved there. Part 2 (1–7), published by various authors between 1954 and 1988 in volumes 2 (A–D), 9 (E–H), 86 (I-L), 93 (M), 109 (N–R), 114 (S) and 117 (T–Z), of *Regnum Vegetabile*, an occasional series of the International Association for Plant Taxonomy, is an alphabetical index of the most important plant collectors, giving the present location of their specimens.

The important collections to be seen during the study fall into two categories: major international herbaria and local herbaria in the target region. The relative advantages and disadvantages of two categories of herbarium for the ecogeographer are given in Table 14.1. Because each kind of herbarium has its own strengths and weaknesses, it is important that both should be visited in the course of an ecogeographic survey. Another factor that should be considered when selecting which herbaria to visit is their age. This is especially important for the smaller, regional herbaria. Recently established herbaria are likely to contain a higher proportion of recently collected specimens. These specimens commonly have more comprehensive and more legible passport data than older collections, so newer herbaria are more likely to yield better-quality ecogeographic data.

Davis and Heywood (1973) stress in a similar context that it is important to sample material from as many herbaria as possible, so that a true estimate of within-taxon variation can be made. Likewise, the broader the sampling of ecogeographic data associated with herbarium specimens or germplasm, the more likely the data will prove ecologically and geographically predictive.

Some of the ecogeographic data included in the database compiled

Table 14.1. Relative advantages and disadvantages of major international herbaria and regional herbaria.

	Advantages	Disadvantages
Major international herbaria	**1.** Broad taxonomic coverage, possibly material used in the production of revisions and monographs **2.** Broad international geographic coverage, possibly material used in the production of local Floras **3.** Skilled researchers available to provide general advice **4.** Appropriate taxonomic and geographic specialists **5.** Type material of target taxa **6.** Good botanical library	**1.** Predominance of old collections, making extraction of passport data more difficult and likely predictive value lower **2.** Geographic names associated with older collections sites may have changed more recently.
Regional herbaria	**1.** Good local regional coverage of target region **2.** Better-documented material, as the herbarium is likely to have been more recently established **3.** Regional specialists present, who can assist in deciphering local geographic names	**1.** Limited resources for herbarium maintenance **2.** Lack of target taxon specialists **3.** Limited botanical library

by Maxted (1990) were taken direct from the author's own germplasm collection database. This illustrates the point that ecogeographic data can equally well be obtained from the passport data recorded by previous germplasm collectors. The importance generally given by plant genetic resources workers to such data means that the passport data associated with germplasm accessions will often be of a higher standard than that associated with herbarium specimens. However, systematic germplasm acquisition programmes have only been established relatively recently. As a result, for many species herbarium specimens may provide the only source of detailed ecogeographic data. The sources of information on existing germplasm collections are discussed further below and in Chapter 9.

Designing and building the ecogeographic database structure

The ecogeographic data for the target taxon can be recorded on paper, but, for those who have access to (and experience of) computers, it is much more efficient to collate data directly into a database. Hardware and software requirements need to be considered early on. For example, a portable computer will be necessary if most data collation is to be done away from the base institution. If data are likely to be numerous, faster (and probably more expensive) machines will be preferred.

It is not uncommon to use a word-processing package for data management, but this has serious limitations. For serious users, a database management system (DBMS) is recommended. The main advantages of using a DBMS are (Date, 1981; Painting *et al.*, 1993):

- data capture and editing are easier;
- there is less chance of introducing errors while copying data between formats;
- searching and retrieving data for reports are easier;
- complicated sorting and indexing on multiple fields are possible;
- identification of duplicate records is possible;
- the structure of a file can be altered in response to changing information needs.

Many DBMSs are available, differing in their flexibility, ease of use and capabilities. Examples include dBASE IV, FoxPro (or FoxBASE), Paradox, etc. (Tatian, 1993). ORACLE and INFORMIX[1] are more complicated packages. Users should determine the type of DBMS (and, indeed, computer hardware) that is being used by any collaborators they may have. This may constrain their own choice, though many DBMSs can export a dBASE-compatible database or a flat-file format ASCII file, allowing data exchange.

Specialized packages are available for the management of herbarium label and related information. Examples include TROPICOS (Pankhurst, 1991) and the Botanical Research and Herbarium Management System (BRAHMS) developed by D. Filer (Oxford Forestry Institute). However, these may not be entirely appropriate for the ecogeographer's purposes, who will probably therefore need to develop his or her own database structure before data collation can begin. Maxted (1991) and Pankhurst (1991) discuss the design and construction of databases in taxonomy and related fields. More general practical advice on database design is given by Painting *et al.* (1993).

In general, database structure should be kept as simple as possible. A database is composed of records, equivalent to the horizontal lines in

[1] dBASE IV is a trademark of Ashton-Tate Corp., FoxPro and FoxBASE are trademarks of Fox Software Inc; ORACLE is a trademark of Oracle Corp.; and INFORMIX is a trademark of Informix Software Inc.

a report, and fields, each equivalent to a single column in a report. Each record could describe a separate herbarium sheet, accession, country or species, for example. Each field would then provide specific details about a different aspect of the specimen, country or species.

Fields should be defined for all information that can be expressed in a limited number of words or numbers. It is not reasonable to enter complete narrative information for all records if the narrative is more than a couple of lines (80 characters) long. Data in this form will be difficult to search and retrieve. In general, if a field contains more than ten words or word-and-number combinations, it should be divided into separate fields. Indeed, if a field can be split, it should be. Combining different data into a single field should certainly be avoided. For example, latitude and longitude can be entered in a number of ways, but it is best to have separate numeric fields for degrees and minutes and a character field for hemisphere (E, W, N, S). Similarly, if measurements are given in the original data in a variety of units (e.g. distances in kilometres and miles), there should be a numeric field for the data and a separate character field in which the unit used is specified; this will allow later transformation of the data into a common unit. Data in numeric fields can be manipulated mathematically, which is not the case for character fields, but many software packages will turn missing data or blanks in numeric fields into zeros. Room for extra characters should be allowed when setting field widths. Field names should be unique, descriptive and simple (Painting *et al.*, 1993).

When it is possible to enter more than one piece of data for a given field and record, these should be accommodated in separate records or in separate fields. For example, if variation within an accession in a continuous numerical descriptor like plant height exists and needs to be shown, create separate fields for maximum and minimum value or for mean and standard error. In the case of discrete descriptors, if a variety of flower colours occur in a species, for example, and a particular population appears mixed from a herbarium or germplasm sample, the options are as follows: (i) open separate records for each state if the descriptor is particularly significant and the number of different states large; (ii) note the presence/absence or frequency of each state (colour form, in this case) in a separate field; or (iii) note the most common state in one descriptor, the second most common in another, and so on, to the extent desired. Instead of separate fields, a single field could be used, with the data separated by standard delimiters, but this may complicate retrieval and analysis (but see Hintum, 1993). When a particular system has been chosen for recording a descriptor, this should be rigidly adhered to in building up the database. The 'rules' to be followed for each descriptor should be written down before starting to enter data.

Codes and/or a standard wording should be adopted whenever possible in entering data. This promotes data consistency. It can also speed up data entry and verification and facilitate exchange. A listing must be maintained of what the codes mean. In many cases, it is possible to use accepted standard codes, but one system should be selected and used consistently. Examples include standard codes for herbaria (Holmgren *et al.*, 1990), for authors of plant names (Brummitt and Powell, 1992) and for political units (International Standards Organization, 1981). The International Union of Biological Sciences Commission on Taxonomic Databases (TDWG) was established to facilitate data standardization and data exchange between botanical databases. It is producing sets of standard codes for botanical data, for example for botanical recording units (Hollis and Brummitt, 1992). Standards in preparation include ones for: economic use; habitat, soil and landscape; life-form; and plant occurrence and status. Information can be obtained from the TDWG Secretariat, based at the Missouri Botanical Garden. Since 1963 the World Conservation Union (IUCN) has been working on a system for describing the conservation status of species (for history, see Fitter and Fitter, 1987). A quantitative system for determining categories of threat is being developed (Mace *et al.*, 1992; IUCN, 1994). The United Nations Environment Programme (UNEP)/GEM's Harmonization of Environmental Measurement (HEM) Programme aims to 'enhance the compatibility and quality of information on the state of the environment worldwide'. It collates and disseminates information on environmental data, including models and classification schemes. It is working with the World Conservation Monitoring Centre (WCMC) on a classification of vegetation. Published germplasm descriptor lists (which include passport, characterization and evaluation data) should be examined before beginning to develop a database (Chapter 8).

A problem that commonly occurs in database construction is that of data repetition. For example, in a database of information on the provenance of herbarium specimens, an entry in a PROVINCE field will always be associated with the same COUNTRY field: Chiapas is always in Mexico, Uttar Pradesh always in India, etc. One could clearly have a single file with both PROVINCE and COUNTRY fields, but this would be uneconomic in terms of both data-inputting time and computer storage and more prone to inputting errors, as there may be dozens of records from Uttar Pradesh for which the word 'India' would have to be entered, for example. A better alternative is to create a second database file containing only PROVINCE and COUNTRY fields, each province being listed only once. One could then have the main file of locality information containing the PROVINCE field but no COUNTRY field and link this to the second file through the common field PROVINCE to access the country information. A DBMS that allows the processing of data in linked files in this way is sometimes called relational, though the technical definition of the term is somewhat stricter (Pankhurst, 1991).

A simple linked file structure, in this case dBASE files, is illustrated in Fig. 14.1 (Maxted, 1990). Only linking fields are shown. The example illustrates the inclusion of the three basic kinds of ecogeographic data in the one database: ecological, geographic and taxonomic. Files GENNOME, SECTNOME and TAXONOMY contain the taxonomic data,

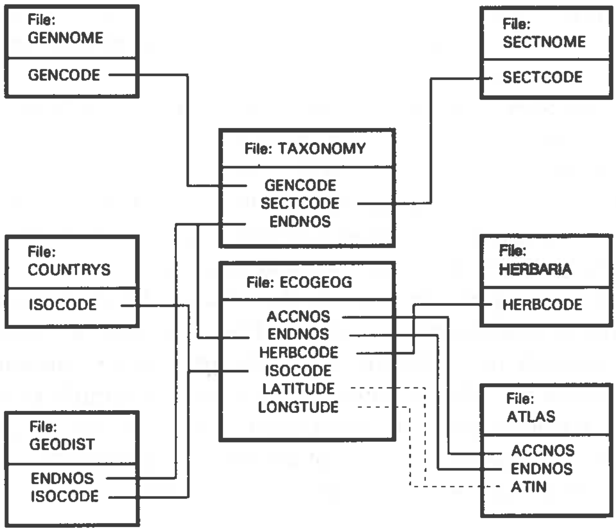

Fig. 14.1. Ecogeographic database file relations (Maxted 1990). Only fields within files that are linked to fields within separate files are shown. Solid lines connecting fields indicate direct links and hatched lines indicate partial links.

file ECOGEOG contains the bulk of the ecological data and files COUNTRYS, GEODIST and ATLAS contain the geographic data. One file, HERBARIA, contains curatorial data, the addresses of the herbaria visited. ATLAS contains the input data for a mapping program, which plotted dot-distribution maps using the latitude and longitude data held in the database. However, the program used by Maxted (1990) required the locality data in a specific input format, so the field ATIN is derived from the fields LATITUDE and LONGTUDE, but not identical to them. The field ENDNOS contains the taxon identification code and illustrates the role of the linking field. When building up the main ECOGEOG file, time was saved by entering the species name as a code. However, the full Latin name and generic section corresponding to each taxon code are found in the file TAXON.

As stressed by Painting *et al.* (1993), a register should be kept of field specifications (full descriptor name, field name, field width, field type, data entry rules, data validation rules, indexing, whether a linking field, file name).

Consistency is important when entering data into the database. It assists retrieval and report generation as well as sorting routines. Consistency will be better if data are entered all at once, rather than in fits and starts, as data entry conventions will be more easily remembered. Entries should be made as they will be required to look in a final report: the software should not be relied on to correct errors in data entry. For example, genus name should have the first character entered as upper-case. A hand-held scanner can be used to acquire optical images of herbarium labels and the like for later printing, examination and entry into the database.

Herbaria, especially young and rapidly developing institutions, are increasingly entering their specimen information in in-house databases, for example using the TROPICOS·or BRAHMS software already mentioned. For her ecogeographic survey of Mediterranean *Medicago*, Rihan (1988) obtained specimen passport data on diskette from the databases of the herbarium of the Madrid Botanical Garden, while Maxted (1990) used geographic data from the Vicieae Project Database (Bisby, 1984) in his ecogeographic study of *Vicia* subgenus *Vicia*. If a new data set is received on diskette, it needs only to be appended to the existing database file. Before doing so, however, it should be verified that the two files have compatible structures (fields of the same name, length and type), or data may be lost or appended incorrectly. Diskettes should also always be checked for computer 'viruses'.

Phase II – Data collection and analysis

Listing of germplasm conserved

Before embarking on the detailed data collation phase of the project, current conservation activities should be reviewed. If sufficient genetic variation of the target taxon from the target region is already safely conserved either *ex situ* or *in situ*, then there may be little justification for further collecting. Details of what material is currently being conserved can be obtained from the catalogues and databases of botanical gardens, gene banks and *in situ* conservation areas. Identifying these sources may prove time-consuming, but taxon experts may help guide the ecogeographer (Chapter 10). The International Plant Genetic Resources Institute (IPGRI) produces international directories of germplasm collections on a crop basis and also maintains a parallel database, which may be queried on demand (Chapter 9).

Care must be taken when interpreting information on current genebank or botanical garden holdings. The material held may be incorrectly identified, though it may be possible to check the identification by consulting voucher material or identifying living material. The actual quantities of germplasm available could also be misleading: gene banks and botanical gardens are encouraged to duplicate their holdings in other collections, so the total number of accessions held around the world can give a false impression of the genetic diversity conserved. The ecogeographer should also consider that, although accessions may be held in a gene bank, the material may for various reasons be unavailable to potential users and so create a false impression of a taxon's conservation status.

Survey of taxonomic, ecological and geographic data sources

Chapters 9 and 10 deal with sources of published information on the environment and on wild species, respectively, and Chapter 12 with sources of ethnographic, especially ethnobotanical, information. Increasingly, however, information is becoming available in media other than the conventional printed literature. Abstracts of publications (and, in some cases, the full text) may be available on microfiche, compact disk read-only memory (CD-ROM) or on-line bibliographic databases, for example (Chapter 13). There may have been other attempts to survey herbarium label information, the results of which may or may not be formally published. For example, some herbaria and other organizations are developing floristic and indigenous knowledge databases (Chapters 10 and 12). Herbaria may hold some label data in card catalogues. The example of the East Africa Herbarium in Nairobi may be instructive. It maintains card catalogues on plant uses (cited by Peters *et al.*, 1992) and on local names gleaned from herbarium labels and other sources. Databases are being developed from both catalogues.

Increasingly, computer networks, particularly the academic network known as the Internet, are being used as sources of information. The Internet links together some one million computers worldwide, which means that there are probably tens of millions of users. Many scientific interest groups have been set up on the Internet. Software, such as listserv and Usenet, supports electronic discussion groups and distributes electronic newsletters and scientific papers. In addition, many important information resources, such as university libraries and public domain software and databases, are being made available on the Internet. Compilations of listservs, Usenet news groups and information archives of relevance to biologists are provided by Dr Una Smith's 'A Biologist's Guide to Internet Resources' (smith-una@yale.edu). TAXACOM (taxacom@harvarda.harvard.edu) is perhaps the best-known mailing list on taxonomy and related subjects. How to gain access to Internet resources is described by Krol (1992).

The main categories of data on taxa that may be obtained from the literature and other information sources are listed in Box 14.1 (*indicates data that could be coded). The collation of much of this information may be undertaken while visiting major herbaria, which often have good botanical libraries attached and some of which have access to the Internet.

Collection of ecogeographic data

The kinds of information that the ecogeographer may obtain from the passport data associated with herbarium specimens and germplasm accessions are given in Box 14.2 (* indicates data that could be coded). Characterization and evaluation data could be added for germplasm accessions. This is an extensive (though not exhaustive) list, and it is

Box 14.1

- Accepted taxon name.*
- Locally used taxon name.*
- Botanical description.
- Taxonomic affinities.
- Distribution within the target region.*
- Timing and periodicity of local flowering and fruiting.*
- Habitat preference.*
- Topographic preference.*
- Soil preference.*
- Geological preference.*
- Climatic and microclimate preference.*
- Pollination and breeding system.*
- Germination requirements.
- Seed storage type.*
- Dispersal system.
- Genotypic and phenotypic variation, including karyotype.
- Biotic interactions, including seed predation (pests, pathogens, herbivores).
- Archaeological information (e.g. palynology).
- Ethnobotanical information (e.g. vernacular name, local uses).
- Conservation status* (e.g. IUCN Red List status).

unlikely that all items will be recorded during a given study. There are certain data items, however, that must be recorded for the study to yield predictive results, and these are shown in bold.

The herbaria of the world contain millions of specimens and the number of specimens of any one target taxon can be vast, but the scope of ecogeographic investigations will be limited by the availability of time and resources. During the course of an ecogeographic project several thousand specimens of the target taxon may be seen. Each of these specimens will require identification: the scientific names written on herbarium sheets should always be checked. However, only a proportion of the specimens seen and identified will be selected to have their passport data recorded in the database. The researcher must be discriminating. Specimens are more likely to be selected if they have detailed ecogeographic passport data or if they show features of particular taxonomic, ecological or geographic interest, i.e. they are odd or rare forms, come from unusual environments or are found on the edges of their natural range. Maxted (1990) found that data from about a third of the specimens seen during the study were finally included in the ecogeographic database.

It is important that the ecogeographer place particular emphasis on obtaining reliable specimen locality data for those specimens which are to be included in the database. Ideally, only specimens for which latitude

Box 14.2

- Herbarium, gene bank or botanical garden where specimen is deposited.*
- Name(s) of collector(s) and collecting number.
- **Collecting date* (to derive flower and fruiting time).**
- **Sample identification.***
- Phenological data* (presence of flowers and/or fruits).
- **Locality,* latitude and longitude or even greater detail if possible.**
- **Altitude.***
- Soil type.*
- **Habitat.***
- Vegetation type.*
- Slope and aspect.*
- Land use and/or farming system.*
- Phenotypic variation.
- Evidence of pests and pathogens.*
- Competitive ability.*
- Palatability.*
- Vernacular names.
- Local uses.*

and longitude data are recorded or for which these data can be established should be selected for inclusion in the database. In practice, it may be advisable to include specimens with two levels of detail, those for which full latitude and longitude details can be obtained and those with major country subunit detail (i.e. province or state) (Rhoades and Thompson, 1992). Specimens that lack even this lower level of geographic data should not be included unless they are particularly noteworthy.

Early plant collectors could not have predicted the detailed analysis that would subsequently be based on the information they recorded. This is therefore often very limited: locality data may be ambiguous and ecological details missing. Older specimen labels are almost invariably handwritten, which adds the problem of having to read the script, which may be in a foreign language. Herbarium staff may be able to help decipher semi-illegible labels. During the ecogeographic study of Vicieae and Cicereae from the southern republics of the ex-Soviet Union (N. Maxted, in prep.) assistance from workers in local herbaria proved invaluable not only in herbarium label interpretation, but also in the precise identification of specimen localities.

A gazetteer of local geographic names and localities will also help in this. There is no comprehensive world gazetteer yet available, but the *Atlas of the World* (Times Books, 1988) contains an extensive gazetteer and the *Official Standard Names Gazetteer* is being constructed and is available in country volumes from the US Board of Geographical Names.

Herbaria sometimes develop their own unpublished gazetteers, perhaps in a card catalogue or computer database. Details of the localities mentioned in standard Floras are sometimes published in separate volumes or in appendices. As pointed out by Forman and Bridson (1992), some problems may be encountered in finding localities in gazetteers: (i) the word on the specimen label may not be a locality; (ii) the name of the locality may have changed; (iii) political and administrative boundaries may have moved; (iv) a variant or incorrect spelling of the locality name may be being used; and (v) the name may be common and therefore have more than one entry in the gazetteer. Older maps, atlases and gazetteers, and even travel books, can be a useful source of localities if names or boundaries have changed (see also Room, 1979). If names have been transliterated from other scripts (e.g. Arabic, Chinese or Cyrillic), the various possible alternative renderings should be checked. In the case of common names, some effort will have to be made to at least partly reconstruct the collector's route.

This will also be necessary if specimen locality data are of poor quality. Fortunately, many collectors keep detailed collecting notebooks and/or diaries, which are usually to be found at the herbarium or gene bank conserving their collections. Almost invariably handwritten in field conditions, they may prove time-consuming to decipher. If notebooks are not available, it may be possible to tentatively estimate where a given specimen may have been collected by comparing its collecting date and/or collecting number with those of specimens whose localities can be recognized. Herbarium determination lists, arranged in collecting number order, are useful in reconstructing itineraries. There is an extensive secondary literature on early collectors, which can help in tracing itineraries. *Taxonomic Literature* (generally known as *TL-2*), published as volumes 94 (A–G), 98 (H–Le), 105 (Lh–O), 110 (P–Sak), 112 (Sal–Ste), 115 (Sti–Vuy) and 116 (W–Z)) of *Regnum Vegetabile* under the editorship of F.A. Stafleu and R.S. Cowan (1976–1988), has a 'Bibliography and biography' section for each author listing thousands of secondary literature references on itineraries. Volumes 125 (A–Ba) and 130 (Be–Bo) are supplements. The reports of germplasm collecting missions usually include a map showing the itinerary followed.

Forman and Bridson (1992) provide a list of biographies and gazetteers that may assist in the pin-pointing of important herbarium specimen collecting localities. It is arranged by regions (plus a general section), but does not include Europe and North America.

Specimens should be positively selected to represent the breadth of geographic and ecological conditions under which the target taxon is found. It is desirable to collect detailed passport data from a broad range of representative specimens. Duplicating data should be avoided.

Recently collected specimens often have higher-quality passport data which is easier to read, being often typewritten. These data are also more likely to have remained current. Extensive use of specimens collected in past decades may provide important details about changes in distributions, but is likely to yield less useful information about contemporary populations. This might present a special problem when trying to locate populations of threatened, rare and restricted taxa. For example, recent collecting activities (Slageren, 1990) have indicated that populations of *Aegilops uniaristata* seem unlikely to have survived in Turkey beyond 1900, so the widely available herbarium specimens of this taxon from that country would provide a false indication of its distribution and frequency.

Though fruit characters are important in some groups, flowers are required to identify most species, which means that herbarium specimens are usually of flowering material. Such material will help give an indication of the most appropriate time for germplasm collecting, but will not be as useful as material that is in fruit, especially mature fruit. The ecogeographer should be particularly on the lookout for fruiting material of the target species. Samples of fleshy fruits may be preserved in alcohol, rather than dried and pressed, and may be stored separately within the herbarium institute.

The database will inevitably contain many gaps. In general, it is much easier to extract curatorial or geographic than ecological data from herbarium specimens. This is illustrated by the percentages of fields containing data in the central file of the ecogeographic database compiled by Maxted (1990). Few of the specimens included in the database lacked collector's name, collecting number, collecting date and locality details, though the degree of detail recorded for the latter varied considerably. Ecological details were found to be much less commonly recorded on herbarium specimen labels: soil type was recorded in about 25% of cases, altitude in about 55% and habitat 65%.

It should be noted that the amount of ecological information gathered in a study such as Maxted's (1990) may be unusual. This is because a relatively high proportion of the specimen data included was taken directly from the author's own germplasm collection database and not from the herbarium specimens of other collectors. In the latter case, the data are likely to be much less complete. This point is illustrated by Rihan's (1988) ecogeographic study of *Medicago* species. The data she recorded were taken entirely from herbarium specimens. The percentage of entries containing soil type, altitude and habitat data declined to 15, 26 and 52% respectively, while the amount of locality information recorded was approximately equal in both projects.

The ecogeographer may be able to infer various features of collecting sites (e.g. latitude and longitude, geology, soil, altitude) from locality data by reference to appropriate maps or databases (Chapters 9 and 16). Whether this will be possible will depend on the precision of the locality data available, the topography of the collecting area and the precision of the environmental data required. For example, if the collecting site is situated on a gently undulating plain, then a crude estimate of altitude may be gained from the locality data, as the altitude is unlikely to vary significantly in the vicinity of the collecting site. However, if the site is situated in a mountainous area, then the altitude is likely to vary markedly within relatively short distances and so estimates of site altitude based on locality might be very wide of the mark unless this is extremely precisely specified. This kind of secondarily derived data should be flagged in the database to distinguish it from data derived directly from herbarium labels. Specimens should also be flagged, if possible, to denote the accuracy of locality data, whether available in the original record or derived from other locality information; in practice, this may mean giving a code for the scale of map used to pin-point a locality and read off coordinates (Rhoades and Thompson, 1992).

Ecogeographers will be faced with the question of how many specimens should be entered into the database before the amount of extra information gained from each additional specimen fails to significantly increase the predictive value of the data set. There is no specific answer. However, the compiler should be on the look out for the point when novel ecogeographic combinations no longer occur in the specimens being examined and the latitudinal/longitudinal extent of the distribution of the species has ceased to expand: the full range of geographic and ecological niches that the taxon inhabits will then probably have been recorded in the database.

After the database is set up and all the available data have been entered, the compiler will probably find that there are still a number of gaps in many of the fields for many records. A record should be marked (flagged) when all possible information has been collected. This will help to determine the completeness of the data set. If many of the records have significant amounts of missing data, certain kinds of mathematical analysis will not be possible or will give misleading results.

Data verification

Before the database can be deemed complete, errors must be spotted and corrected. A lot of errors can be avoided by appropriate time-of-entry features of the software, such as allowing only a limited number of valid responses (perhaps listed in a menu) or storing a default value. Once all the data have been entered, verification needs to be carried out. Tatian (1993) discusses the subject and suggests the following checks:

1. Range and rule checks: are some values outside the allowed range?
2. Inter-record checks: have some records been entered twice?
3. Visual comparison with original forms (or double entry of all data): have some records not been entered, or entered in incomplete fashion?
4. Interfile checks: are data consistent among files?

A useful way of carrying out range and rule checks involves indexing the database (i.e. rearranging the records in alphabetical or numerical order) on each field in turn. Records with typing errors, invalid and out-of-range entries, etc. in the indexed field can then be easily picked out by browsing through the file. Mapping latitude and longitude data may reveal errors if particular localities are shown up as obvious outliers in impossible places (e.g in the middle of a lake). Collectors often send duplicate sets of herbarium specimens to different international herbaria. Germplasm accessions are also commonly duplicated. The compiler should search the database for these duplicates and be aware of their possible effect on data analysis. The database software should be able to pick out records that match for a set of fields.

Analysis of taxonomic, ecological and geographic data

The raw ecological and geographic data included in the database can be analysed to help identify the habitats favoured by the target taxa and the geographic limits of its distribution. Useful aids to the interpretation of ecogeographic data are tables and bar charts indicating the number (or percentage) of specimens seen from different geographic or ecological units (e.g. climate type, soil type, aspect, shading characteristics, habitat). It should be possible to get the data for such charts directly from the database: a DBMS will usually be able to count the number of records which have particular entries in a specified field or fields. Data arranged in this fashion will help to characterize the ecological niche of the target taxon. For continuous ecological factors, such as altitude, latitude and soil pH, correlation with the frequency of occurrence of specimens along the gradient can be calculated. Correlation of morphological characters with environmental gradients will help to indicate clinal adaptation, in both wild and cultivated material.

One of the most thoroughly statistically tested ecogeographic data sets is that reported by Cocks and Ehrman (1987), Ehrman and Cocks (1990) and Ehrman and Maxted (1990) for the annual legumes in Syria. These authors undertook comprehensive fieldwork over several years, during which they gathered extensive ecogeographic data and were able to use these data to predict potential areas of conservation. They divided Syria into climatic regions and then recorded the percentage of sites for each region where each annual legume species was found. The authors also studied the percentage of sites of each soil type in which various taxa were found. The influence of both climatic factors and soil alkalinity on the distribution of various species was clearly demonstrated.

Such methods deal with one environmental factor at a time, or a single morphological variable. Ecogeographic data, however, are multivariate, in that two or more items of data are available for each record (e.g. each collecting site, germplasm accession or herbarium specimen). Ehrman and Cocks (1990) used various methods of cluster analysis on their environmental data to classify the collecting sites into groups or classes (clusters) the members of which had climates which were more similar overall (rather than as regards any one single variable) to one another than they were to members of any other class. A more detailed discussion of multivariate data analysis is provided in Chapter 15.

Another approach to the study of ecogeographic data involves mapping collecting sites. These distribution maps can be used in conjunction with topographical, climate, geological or soil maps. Stace (1989) stresses the importance of the means of visually displaying plant distribution. This can take two forms: (i) shading or enclosing an area with a single line, or (ii) using various kinds of dot-distribution maps.

The use of an enclosing line is ambiguous, as it provides no indication of the frequency of the taxon within the region. A single outlying specimen might erroneously suggest that the taxon is continuously present throughout an entire region. The occurrence of a species is often sparse at the periphery of its range and there is rarely a distinct cut-off line. Indicating presence in this manner also means that any variation due to local ecological and geomorphological factors within the individual provinces or countries cannot be shown. The problems associated with enclosed line maps can be illustrated with an example taken from Edmonds' (1990) ecogeographic survey of African *Corchorus*. Figure 14.2 shows the distribution of *C. aestuans*. A crude enclosed line map would shade the entire area from northwestern Uganda to southern Tanzania, which would not bring out the fact that in Uganda the species is represented by a single record and the majority of records are from south-eastern Kenya and central and eastern Tanzania. Enclosing line maps do have advantages, however. Westman (1991), for example, used such maps in conjunction with climatic isoline maps to calculate the

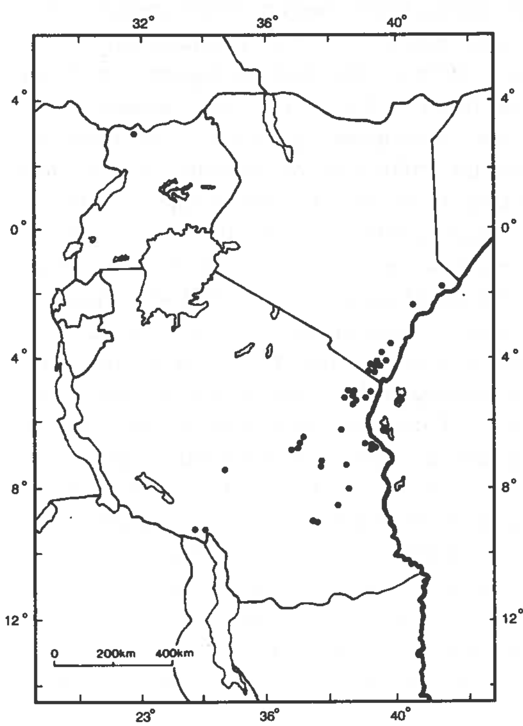

Fig. 14.2. Dot-distribution map for *Corchorus aestuans* in east Africa (taken from Edmonds, 1990).

percentage of the total areal distribution of various species in each climatic category in California.

Enclosing line maps can be used to indicate concentration of species. Such isoflor maps do not show actual species distributions: each line is a contour delimiting a greater or lesser concentration of species. Species distributions are superimposed on to a single map then contour lines are drawn around areas of the map with the same number of species. An example of an isoflor map for *Vicia* sect. *Narbonensis* is shown in Fig. 14.3 (Maxted, 1990). Of course, isoflor maps can be produced for infraspecific taxa within species (e.g. subspecies, crop landraces) as much as for species within sections or genera.

To represent distribution patterns in detail there is a general trend towards the use of dot-distribution maps (Stace, 1989). These may be of two types: (i) presence is indicated in subregions (e.g. grid squares) within the study region, or (ii) spots represent actual localities. The former option places the onus on the researcher to determine the presence or absence of the taxon in each square. This ensures evenness of coverage, but will be impractical if the target region is large and time allowed for the study limited. Species mapping is discussed in detail by Miller (1994).

Morphological or ecological information can also be superimposed

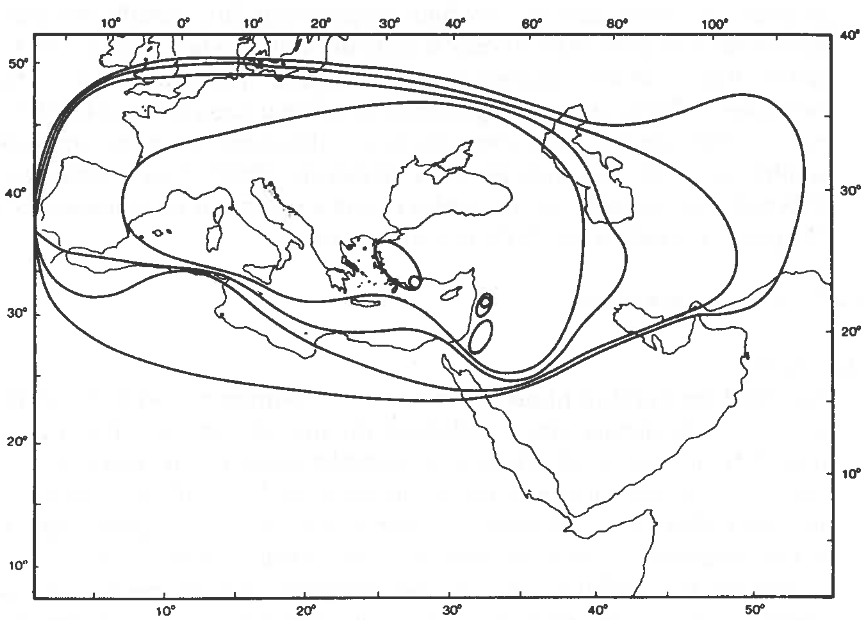

Fig. 14.3. Isoflor map for *Vicia* sect. *Narbonensis* in Europe, North Africa and the Middle East (taken from Maxted, 1990). Each contour line encloses one more species than the one immediately outside it.

on to a dot-distribution map, as Strid (1970) has done for the uppermost internode length of various populations of *Nigella arvensis* in the Aegean region. The position of a rectangle indicates the location of the population, while its height shows the relative length of the internode. Pie charts can be used to display the relative frequency of a character in different places. They are commonly used to show allelic frequencies in populations at different geographic locations, but have also been used to display morphological variation, e.g. stigma colour in *Crocus scepusiensis* (Rafinski, 1979). New (1958) used pie charts to demonstrate seed-coat variation in *Spergula arvensis*, the relative size of the circles indicating the size of the population sampled; different sized circles could also be used to indicate the number of database records from which the frequency has been estimated. Pie charts can also be used to compare the distribution of specimens with physical characteristics, e.g. altitude, temperature or soil type. Daday (quoted by Jones, 1973) used pie charts to show the relationship between the distribution of cyanogenic forms of *Lotus corniculatus* in Europe and January isotherms. Again, there is more on this in Chapter 15.

Mapping software can considerably simplify the production of distribution maps. Large scale maps of the world are available in digitized form, but if more detailed maps are required they may need to be customized or even digitized from scratch. Mapping programs will allow the user to import latitude and longitude coordinates from the ecogeographic database and plot them on to customized maps. The import facility allows the locality details to be transferred from the ecogeographic database directly, without re-entering data, while the ability to customize the base map allows a suitable scale to be used to display the distribution. More sophisticated mapping programs have built-in databases. These are called geographic information systems (GIS), computer hardware and software packages designed to store, analyse and display spatially referenced data (Haslett, 1990). The development of GIS will increasingly prove useful in the analysis of ecogeographic data. Chapter 16 deals with GIS in more detail.

Phase III – Product generation

Data synthesis

The final production phase of the project commences with the synthesis of all the disparate data collected during the study. The researcher should be aware of the degree of completeness of the database, or the collections on which it was based, in terms of how fully the target region has been effectively covered. If a particular habitat is under-represented in the database, it may be because the taxon is really not found there or because that habitat has not been sampled, or even because the target taxon has not been recognized in such a habitat. This problem must be considered if the results of the analysis and the inferences drawn from them are not to be misleading.

The database contents are summarized, together with the other data abstracted from the literature, into an ecogeographic conspectus. The pattern of the data included in both the database and conspectus can then be interpreted, the results of the analysis displayed and suggestions for appropriate target taxa and regions which warrant conservation discussed in the ecogeographic report.

The ecogeographic database, conspectus and report

The ecogeographic database, conspectus and report are the essential products of an ecogeographic study. The ecogeographic database contains the raw data. The conspectus summarizes the available taxonomic, geographic and ecological information for the target taxon through part or the whole of its range. The report interprets the data held in the other products and will help the ecogeographer identify conservation priorities.

The conspectus is arranged by plant names, which can be listed either alphabetically or systematically. In both cases it is helpful to provide an index to the taxa included in the study. The conspectus should summarize information from both the database and the literature survey. An abbreviated version of the ecogeographic conspectus, basically an annotated checklist containing, in coded form, the information shown in bold below, will be useful in the field when it comes to the actual germplasm collecting (Hammer, 1991). If possible, the following information should be included in the full conspectus:

- **accepted taxon name**, author(s), date of publication, place of publication;
- reference to published descriptions and iconography;
- short morphological descriptions or keys for important taxa or those that may be difficult to identify;
- **phenology, flowering season;**
- ethnobotanical notes, especially **vernacular name** and **local uses;**
- **geographic distribution**, i.e. countries, provinces or districts from which the taxon is recorded, including reliable records from the literature;
- distribution maps (preferably dot-distribution) produced directly from the latitude and longitude data held in the database;
- geographic notes, including an interpretation of the taxon's geographic distribution;
- **ecological notes**, including: altitude (minimum and maximum); habitat, topographic, soil, geological, climate and microclimatic preference; biotic interactions;
- **taxonomic notes**, including notes on any distinct genotypic and phenotypic variation within the taxon;
- conservation notes, containing an assessment of the variation currently conserved *ex situ*, the potential genetic erosion faced and the conservation status of the taxon in the field.

If the scope of the investigation is broad in the geographic or ecological sense, it may be necessary to provide a summary of the ecogeographic data for each geographic or ecological subunit. This can be illustrated with reference to Edmonds' (1990) survey of African *Corchorus*. The survey covered a vast geographic region and to increase the predictive value of the survey she lists the flowering time for each species in each country rather than providing one time range for the whole of Africa.

A listing of the specimens used during the study and a summary of the synonyms which have recently and frequently been used in the target region may also prove useful. Both these listings would significantly expand the size of the conspectus and so should perhaps be included as appendices.

The ecogeographic report discusses the contents of the database and conspectus and must draw general conclusions concerning the group's ecogeography, presenting a concise list of conservation priorities. If possible, the following points should be covered:

- the delimitation of the target taxon;
- the classification of the target taxon that has been used, and why;
- the mode of selection of representative specimens;
- the choice of hardware and software;
- the ecogeographic database file structures and inter-relationships;
- database contents;
- target taxon ecology;
- target taxon phytogeography and distribution patterns, with a summary of the distribution in tabular form;
- any interesting taxonomic variants encountered during the study;
- current and potential uses of the target taxon;
- the relationship between the cultivated species and their wild relatives;
- any particular identification problems associated with the group (identification aids to vegetative, floral and fruiting specimens should be provided);
- *in situ* and *ex situ* conservation activities associated with the target taxon, including the extent of diversity already conserved;
- genetic erosion threat faced by the group;
- priorities and suggested strategy for future conservation of the target taxon.

As discussed earlier, it is less easy to obtain ecological data from herbarium specimen passport data than it is to obtain geographic or taxonomic data. This may hamper drawing firm ecological conclusions from the study products. However, whatever information is available is a valuable asset and will aid in selecting conservation priorities. An example of the level of detail that might be included in an ecogeographic conspectus is given in Appendix 14.1.

The ecogeographic survey may yield information that is not of direct use in the identification of conservation priorities for the target taxon but which may, at a later date, be incorporated into purely ecological, taxonomic or other products. Maxted found that, during the ecogeographic survey of the Vicieae and Cicereae of the southern republics of the former Soviet Union, he gathered a large quantity of data on localities (place names, latitude, longitude and altitude) within the region. This will form the basis of a gazetteer for the region, which will prove useful to subsequent ecogeographic studies.

Identification of conservation priorities

The principal aim of the ecogeographic survey must be to provide a sound basis for the identification of conservation priorities and strategies, which includes collecting priorities and strategies. During the survey process, data from the literature, herbarium specimens and germplasm accessions are collated, summarized and synthesized into the three ecogeographic products. The pattern of variation within the target region and the target taxon is investigated and an estimate of potential genetic erosion and current conservation status made. On the basis of the various products of the ecogeographic survey or study, the ecogeographer can formulate future conservation priorities and strategies for the target taxon.

Within the target region, areas may be identified which are of particular interest either because of the plants found there or because of local conditions, e.g. areas with high concentrations of diverse or endemic taxa, low rainfall, high frequency of saline soils or extremes of exposure. If a taxon is found throughout a particular region, then the researcher can use the ecogeographic data to actively select a series of diverse habitats to sample. If a taxon has been found at one locality, but not at others with similar ecogeographic conditions, then a possible suggestion is that these localities be searched. Within the target group, specific taxa, populations and variants can be identified which warrant specific consideration, e.g. poorly known taxa, species whose potential as crops has not previously been noted, populations with particular adaptations or rare and endangered taxa.

Having ascertained the level of variation within the target taxa and the potential target region in the process of ecogeographic data compilation and analysis, this information must be assessed in the context of current conservation activities. Is sufficient genetic material of a particular taxon from a particular, interesting ecogeographic niche already safely conserved either *in situ* or *ex situ*? If not, should effort be expended to collect this material? Analysis of herbarium material may indicate that there is a rare relative of the winged bean (*Psophocarpus tetragonolobus*) growing on a small edaphic enclave in western Kenya, but, if a review of conservation activities indicates that large collections of this species are conserved in the Kenyan National Gene Bank and the material is also duplicated elsewhere, re-collecting would probably not be justified.

The ecogeographic survey or study should conclude with a clear, concise statement of the proposed conservation strategy for the target group and proposed conservation priorities. Questions should be considered such as whether population levels should be monitored to assess the threat of genetic erosion, whether a national or international collecting team should be directed to collect the priority target taxa, whether it is possible to conserve the taxa *in situ* and so on. Edmonds (1990), for example, proposes five specific missions to collect rare and endangered *Corchorus* species. If the ecogeographic data have been gathered solely from herbarium specimens, the ecological data obtained may be insufficient to draw detailed conclusions on the target taxon's habitat preferences. A survey mission to the target region may therefore be advisable, to obtain a clearer idea of the pattern of genetic variation and of the appropriate habitat types and clarify the conservation strategy to be proposed.

Once specific areas have been selected for collecting, a route that covers the maximum number of such areas in the minimum time can be suggested. With many species there is a narrow 'collecting window' during which collectors must find the target population. If they are too early the seeds will not be ripe, too late and the material may have shattered. The phenological data (in combination with both climate averages and weather data from the year of collecting) will indicate approximately when a collecting team should visit the target region, an estimate that will need to be refined on the basis of up-to-the-minute information on local weather conditions (Chapter 16).

Conclusion

Herbaria, gene banks, botanical gardens (and their associated libraries) are storehouses of botanical data as much as of plants, pressed and live. These data can be used to facilitate plant conservation. Analysis of a taxon's ecology, geographic distribution and taxonomy is a necessary prerequisite for assessing its conservation status and permits the prediction of which areas and habitats the taxon is likely to be found in. Once located, populations of the taxon can be monitored, sampled if necessary and effectively conserved.

Ecogeographic studies will always be limited by time and resources and it will be impossible to collate every piece of information available. However, if the study is planned carefully and undertaken efficiently, the data that are collated can be predictive. The results of an ecogeographic study will not always lead the collector to the exact localities of desired plant populations, although this is by no means impossible, but they can certainly identify the likely areas of current occurrence. The ecogeographic database and conspectus will also provide sufficient information to permit the conservationist to assess collecting and conservation priorities.

Acknowledgements

We are particularly grateful to T. Hodgkin, D.G. Debouck and L. Guarino for their comments on this manuscript and to R. Reid for discussion of the concept of ecogeographic studies. M. Perry and E. Bettencourt contributed substantially to the section on database design. K. Painting wrote the section on the Internet. D.M. Spooner added some important references.

References

Bisby, F.A. (1984) The Vicieae Database Project: products and services. *Webbia* 38:639–644.

Brummitt, R.K. and C.E. Powell (1992) *Authors of Plant Names*. Royal Botanic Gardens, Kew.

Cocks, P.S. and Ehrman, T.A.M. (1987) The geographic origin of frost tolerance in Syrian pasture legumes. *Journal of Applied Ecology* 24:678–683.

Date, C.J. (1981) *An Introduction to Database Systems*. Addison-Wesley, Reading.

Davis, P.H. and Heywood, V.H. (1973) *Principles of Angiosperm Taxonomy*. Krieger, New York.

Davis, S.D., Droop, S.J.M., Gregerson, P., Henson, L., Leon, C.J., Lamlein Villa-Lobos, J., Synge, H. and Zantovska, J. (1986) *Plants in Danger: What Do We Know?* IUCN, Gland.

Edmonds, J.M. (1990) *Herbarium Survey of African* Corchorus *L. Species*. Systematic and Ecogeographic Studies on Crop Genepools 4. IBPGR, Rome.

Ehrman, T. and Cocks, P.S. (1990) Ecogeography of annual legumes in Syria: distribution patterns. *Journal of Applied Ecology* 27:578–591.

Ehrman, T.A.M. and Maxted, N. (1990) Ecogeographic survey and collection of Syrian Vicieae and Cicereae (Leguminosae). *FAO/IBPGR Plant Genetic Resources Newsletter* 77:1–8.

Fitter, R. and M. Fitter (eds.) (1987) *The Road to Extinction. Problems of Categorizing the Status of Taxa Threatened with Extinction*. IUCN and UNEP, Gland.

Forman, L. and D. Bridson (1992) *The Herbarium Handbook*. Revised edition. Royal Botanic Gardens, Kew.

Frodin, D.G. (1984) *Guide to the Standard Floras of the World*. Cambridge University Press, Cambridge.

Funk, V.A (1993) Uses and misuses of floras. *Taxon* 42:761–772.

Hammer, K. (1991) Checklists and germplasm collecting. *FAO/IBPGR Plant Genetic Resources Newsletter* 85:15–17.

Haslett, J.R. (1990) Geographical Information Systems: a new approach to habitat definition and the study of distributions. *Tree* 5:214–218.

Hintum, Th.J.L. van (1993) A computer compatible system for scoring heterogeneous populations. *Genetic Resources and Crop Evolution* 40:133–136.

Hollis, S. and Brummitt, R.K. (1992) *World Geographical Scheme for Recording Plant*

Distributions. Hunt Institute for Botanical Documentation, Pittsburgh, for the International Working Group on Taxonomic Databases for Plant Sciences.

Holmgren, P.K. and N.H. Holmgren (1992) Plant specialists index. *Regnum Vegetabile* 124:1–394.

Holmgren, P.K., Holmgren, N.H. and Barnett, L.C. (1990) *Index Herbariorum I: The Herbaria of the World*. 8th edition. International Association for Plant Taxonomy and New York Botanical Garden, New York.

IBPGR (1985) *Ecogeographic Surveying and* In Situ *Conservation of Crop Relatives*. Report of an IBPGR task force meeting held at Washington DC. IBPGR, Rome.

International Standards Organization (1981) *Listing of ISO Codes*. ISO 3166. British Standards Institute, London.

IUCN (1994) *Draft IUCN Red List Categories*. Version 2.2. IUCN, Gland.

Jones, D.A. (1973) Co-evolution and cyanogenesis. In: Heywood, V.H. (ed.) *Taxonomy and Ecology*. pp. 213–242. Systematics Association Special Volume No. 5. Academic Press, London.

Krol, E. (1992) *The Whole Internet User's Guide and Catalog*. O'Reilly & Associates Inc., Sebastopol CA.

Mace, G., N. Collar, J. Cook *et al.* (1992) The development of the new criteria for listing species on the IUCN red list. *Species* 19:16–22.

Maxted, N. (1990) An herbarium based ecogeographic study of *Vicia* subgenus *Vicia*. Internal report. IBPGR, Rome.

Maxted, N. (1991) A revision of *Vicia* subgenus *Vicia* using database techniques. PhD thesis. University of Southampton, Southampton.

Maxted, N. (1993) An ecogeographic study of *Vicia* subgenus *Vicia*. Internal report. IBPGR, Rome.

Miller, R. (ed.) (1994) *Mapping the Diversity of Nature*. Chapman and Hall, London.

New, J.K. (1958) A population study of *Spergula arvensis* L. *Annals of Botany, New Series* 22:457–477.

Painting, K.A., M.C. Perry, R.A. Denning and W.G. Ayad (1993) *Guidebook for Genetic Resources Documentation*. IBPGR, Rome.

Pankhurst, R.J. (1991) *Practical Taxonomic Computing*. Cambridge University Press, Cambridge.

Peters, C.R., E.M. O'Brien and R.B. Drummond (1992) *Edible Wild Plants of Sub-Saharan Africa*. Royal Botanic Gardens, Kew.

Rafinski, J.N. (1979) Geographic variability of flower colour in *Crocus scepusiensis* (Iridaceae). *Plant Systematics and Evolution* 131:107–125.

Rhoades, A.F. and L. Thompson (1992) Integrating herbarium data into a geographic information system: requirements for spatial analysis. *Taxon* 41:43–49.

Rihan, J.R. (1988) An herbarium based ecogeographic survey of forages *Medicago* species in the Mediterranean and adjacent arid/semi-arid areas. Internal report. IBPGR, Rome.

Room, A. (1979) *Place Name Changes Since 1900*. Scarecrow Press, London.

Sánchez, J.J.G. and Ordaz, S. (1987) *El Teocintle en Mexico*. Systematic and Ecogeographic Studies on Crop Genepools 2. IBPGR, Rome.

Slageren, M.W. van (1990) *Genetic Resources Unit. 1989 Annual Report*. ICARDA, Aleppo.

Stace, C.A. (1989) *Plant Taxonomy and Biosystematics*. Edward Arnold, London.

Strid, A., (1970) Studies in the Aegean flora. XVI. Biosystematics of the *Nigella arvensis* complex with special reference to the problem of non-adaptive radiation. *Opera Botanica* 28:1–169.

Stressey, T.F. (1993) The role of creative monography in the biodiversity crisis. *Taxon* 42:313–321.
Tatian, P.A. (1993) The use of computers in field research. *Journal for Farming Systems Research-Extension* 3:97–118.
Times Books (1988) *Atlas of the World.* 7th edition. Times Books, London.
Westman, W.E. (1991) Measuring realized niche spaces: climatic response of chaparral and coastal sage scrub. *Ecology* 72:1678–1684.

APPENDIX 14.1
An example of ecogeographic conspectus construction (Maxted, 1993)

Vicia* L. sect. *Atossa (Alef.) Asch. & Graebner, Syn. Mitteleur. *Fl.* 6(2):949 (1909).

Ref. Pub. Description: Kupicha, *Notes Roy. Bot. Gard. Edinburgh* 34:320 (1976).

Perennial; erect or climbing; stem slender or stout. Stipules entire or semi-hastate; 2.5–9 x 1–5 mm; edge entire or with 1–6 teeth. Leaf 25–154 mm; apex tendrilous or mucronate; 2–28 leaflets per leaf; leaflet 10–85 x 3–38 mm; symmetric; margins entire. Peduncle 7–32 mm; with 1–8 flowers. Calyx mouth oblique; lower tooth longer than upper; base gibbous; pedicel 1–3 mm. Flowers 12–22 mm; all petals approximately equal length; standard yellow, blue or purple; shape platonychioid (limb and claw same width); claw bowing absent; upper standard surface glabrous. Wing marking absent; wing limb with slight basal folding. Legume 16–43 x 6–9 mm; oblong; laterally flattened; sutures straight or curved; valves glabrous (without hairs); septa absent; 2–7 seeds per legume. Seeds 2.5–4 x 3–5 mm; round or oblong; not laterally flattened; hilum over half seed circumference; lens positioned near hilum; testa surface smooth.

Number of taxa: 4

Chromosome numbers: 12, 14, 16, 18

Geographic distribution: Europe, Near East and Asia eastward to the Pacific.

Geographic notes: This section is relatively widely distributed from Iceland to Japan, but the extent of the distribution is largely due to a single species, *V. sepium*. The other three species of the section are much more restricted. There are two centres of diversity, one concentrated in northern Yugoslavia and the other in the Caucasus.

Ecological notes: All four species are encountered in similar habitats, stable semi-shaded woodland (coniferous, mixed or deciduous), woodland edges or hedgerows. They show no preference for a particular soil type, but are more commonly encountered at altitudes over 500 m (except *V. sepium*). The four species can be found in open or dense vegetation, under dry or moist conditions.

Taxonomic notes: The four species easily form three series. *V. balansae* and *V. abbreviata* are closely related.

Series *Truncatulae* (B. Fedtsch. ex Radzhi) Maxted, *Kew Bull.* 47(1):130 (1991).

Number of taxa: 2

Chromosome number: 14

Geographic distribution: southeast Europe and west Asia.

Geographic notes: The two species of this series are commonly found in northeast Turkey and the Caucasus, though *V. abbreviata* is also found further west in southeast Europe.

Taxonomic notes: Stankevich (1988) considers the two taxa of this series to be subspecies of *V. abbreviata*. After studying natural populations in the Caucasus she concluded that the two taxa intergrade from one to the other. This, she considered, was especially apparent in the subalpine zone between Karmadon and Chmi in North Ossetia, Russia. She argues that the two taxa have been able to remain morphologically distinct due to their preference for different ecological niches. While collecting in the Caucasus (spring–summer 1989), I located six populations of *V. balansae*. At five of these localities, *V. abbreviata* was equally abundant. Within the five sites, where both species were found, neither species showed a clear niche distinction and no putative hybrid forms were encountered. Therefore Maxted (1993) concluded the specific distinction should be retained.

V. balansae Boiss., *Fl. Orient.* 2:569 (1872)

Ref. Pub. Description: *Fl. Tur.* 3:304; *Fl. USSR* 13:457; *Illust. Fl. Iran* 32, Fig. 3.

Phenology: May–August

Chromosome number: 12, 14

Geographic distribution: SUN, TUR.[2]

Ecology: Alt. 550–2700 m; Hab. moist alpine pastures and forests.

V. abbreviata Fischer ex Sprengel, *Pl. Min. Cog. Pug. Prim.* 1(86):50 (1813).

Common synonym: *V. truncatula* Fischer ex Bieb., *Fl. Taur.-Cauc.* 3:473 (1819).

Ref. Pub. Description: *Fl. Iran.* 43–44; *Fl. Tur.* 3:303–304; *Fl. USSR* 13:456–457; *Illust. Fl. Iran.* 32, Fig. 3.

Phenology: May–August

Chromosome number: unknown

Geographic distribution: AUT, BGR, DEU, IRN, ROM, SUN, TUR, YUG.

Ecology: Alt. 100–2400 m; Hab. mountain forest and forest margins.

[2] These are ISO country codes.

Useful addresses

Harmonization of Environmental Measurement Programme (UNEP-HEM)
c/o GSF - Research Centre for Environment and Health
Ingolstädter Landstrasse 1
W-8042 Neuherberg b. München
Germany
Tel: +49 89 31875488
Fax: +49 89 31873325
E-mail: keune@gsf.de

International Union of Biological Sciences Commission on Taxonomic Databases (TDWG) Secretariat
Dr James L. Zarucchi
Missouri Botanical Garden
PO Box 299
St Louis
Missouri 63166
USA

Oxford Forestry Institute
University of Oxford
South Parks Road
Oxford
UK