In situ *conservation of wild populations of* Phaseolus lunatus *in the Central Valley of Costa Rica*

Alain Maquet and Jean-Pierre Baudoin
Gembloux Agricultural University – FUSAGx, Gembloux, Belgium

In the conservation of plant genetic resources, particular attention is given to wild ancestral populations of cultigens. Such material represents a possible source of much genetic variation, extremely useful for a plant breeding programme.

Wild materials are not easy to maintain *ex situ* for various reasons, such as a lack of information on their ecogeographic distribution, a lack of knowledge of floral biology and mating system, variation in seed ripening and occurrence of seed dormancy. To circumvent these constraints, *in situ* conservation is considered as a dynamic conservation system enabling the continuing evolution of crop genepools, maintaining genetic integrity and adaptation potential of each population.

In an *in situ* conservation programme, the main goal is to preserve as many variants as possible of the target species, which means conserving locally common and rare alleles at as many sites as possible.

Several operations should be carried out when developing *in situ* conservation programmes. Such operations concern not only the target species but also the management of nature reserves where the wild population is scattered.

A key objective is to monitor population dynamics and this will require a full understanding of some major factors, mainly gene flow, breeding system, demography (describing and explaining changes in a number of individuals within a population) and phenology.

The model selected in the project

The plant model is a food legume, *Phaseolus lunatus* or Lima bean, and particularly the wild populations of the cultigen located in one of the major centres of diversity of the species: the Central Valley of Costa Rica.

Lima bean

Lima bean wild populations are made up of short-lived perennial self-compatible plants, with a mixed mating system. In fact the species is regarded as predominantly self-pollinating. Each plant has a climbing indeterminate growth habit; some are very vigorous, bear numerous pods and are characterized not only by perennial or multiannual behaviour but also by a significant soil seed bank.

The Central Valley

In Costa Rica, wild bean populations are mainly distributed in the Central Valley, particularly in zones of premontane and lower montane humid forests, at altitudes varying from 500 to 1800 m. The region is characterized by a mean annual rainfall varying between 1200 and 3000 mm, with wet and dry seasons.

In the Central Valley, Lima beans are grouped into populations of very varying size. More than 70% of these populations contain less than 5 plants and only 15% contain more than 50 plants. In the valley, Lima bean populations are mainly found in habitats disturbed by man: along the roads, in old traditional coffee plantations, in disturbed areas with grasses and scattered trees or bush thickets, or in more woody habitats.

Why an *in situ* conservation programme on Lima bean in the Central Valley?

Wild Lima bean populations in the Central Valley are seriously threatened by several forms of land use including growing urbanization, severe grazing, seasonal fires in pasture lands and sugarcane plantations, and the replacement of traditional small-scale coffee plantations interplanted with leguminous trees by modern high-input plantations. It is therefore urgent to set up an *in situ* conservation programme and in doing so investigate the major components/factors influencing the population dynamics of species having a similar behaviour or mating system.

Lima bean is regarded as a self-pollinated species with facultative allogamy. Pollination is by insects, mainly social and solitary bees visiting flowers for both pollen and nectar.

We conducted some experiments with a view to study gene flow. The movement of genes within and among populations is one of the most important factors that determine the pattern of genetic variation in wild populations. It is also responsible for the genetic structure within populations and its force constitutes an essential parameter for defining optimal sampling strategies for conservation purposes.

Gene flow depends on the efficiency of three types of gene dispersal: dispersal of seeds, pollen and flowers (through vegetative growth).

What are the results of these investigations?

- First we identified a major pollinator among all the insects visiting Lima bean flowers in the valley: this pollinator is the common bee *Apis mellifera*, known to forage on a rather restricted area.
- In all populations investigated, pollen dispersal was limited: most pollen transfers occurred across distances of less than 1 m. Frequencies of pollen transfer falls quickly beyond one metre. The corresponding dispersal variance $\sigma^2_p = 1.7\ m^2$, a value probably underestimated because we could not measure pollen transfer over distances greater than 6.9 m.
- Measurements of flower dispersal show great variability, mainly explained by the surrounding vegetation used as support. Mean dispersal variance of flowers was 2.7 m^2 (0.35 to 10.96 m^2).
- Mean dispersal variance of seeds varies between 1.34 and 2.46 with estimated mean value $\sigma^2_p = 1.68\ m^2$.
- Considering a mean outcrossing rate $t = 0.1$ (data from electrophoresis) and the mean adult plant density $d = 0.235\ pl/m^2 \Rightarrow NA = 96\ m^2$ and $Nb = 7.23$ individuals (where NA is the genetic neighbourhood area, an area from which the parents of central individuals may be treated as if drawn at random (Wright 1946), and Nb is the effective neighbourhood size, an estimate of the random genetic drift occurring in a single neighbourhood. For more details and references see Hardy *et al.* 1997).

Therefore random local genetic differentiation is expected. As the studied populations were spread over areas ranging from 100 m^2 to more than 1000 m^2, they could include several to many neighbourhood areas. The conclusion is therefore straightforward: allelic distribution within a population is expected to be highly structured.

The evaluation of genetic variability is mainly aimed at evaluating the whole genetic diversity represented by the wild Lima bean populations present in the Central Valley. Such investigations are necessary to identify the genotypes to be conserved but another important objective will be to understand the genetic organization of the wild material at two levels: intra- and interpopulation, and by doing so suggest guidelines for optimum conservation.

The genetic structure was evaluated using several parameters of population genetics:

- We calculated 5 indices characterizing the intrapopulation polymorphism: P (percentage of polymorphic loci), A (average allele number per locus), A_e (distribution of frequency of the different alleles in one locus), H_o (proportion of observed heterozygotes) and H_e (expected proportion of heterozygotes in panmixy conditions). Results showed a low percentage of polymorphic loci (10%), few alleles per locus, less than 1% heterozygotes and a large heterogeneity in the index values. At the intrapopulation level we also calculated the Wright consanguinity coefficient in order to evaluate the possible divergence of the breeding system from the panmixy or Hardy-Weinberg equilibrium. Data showed that in general the populations were not in Hardy-Weinberg equilibrium.
- We also studied the genetic diversity at the interpopulation level using other parameters, in particular the Nei diversity indices and F statistic analysis. If we consider a reference population subdivided into subpopulations, the total genetic diversity (H_t) is composed of two items: the intrapopulation genetic diversity (H_s) and the interpopulation genetic diversity (D_{st}). The genetic differentiation coefficient among the subpopulations (G_{st}) can be deduced from the ratio D_{st}/H_t. H_t and H_s are 0.193 and 0.082 respectively. D_{st} is estimated at 0.111. G_{st} value is around 0.52, as H_s and D_{st} are very similar. We can see therefore that 52% of the total heterozygosity came from genetic diversity between populations while 48% came from genetic diversity within population. Values as high as 52 % for G_{st} are indicative of a high genetic differentiation between populations. In our case several factors could explain such a high value: for example a limited interpopulation gene flow, a genetic drift responsible for marked fluctuations in allelic frequency or a natural selection on some populations (perhaps in relation with environmental data). If we now consider the F statistics, it is interesting to point out the high value of F_{it}, a parameter measuring the divergence of the genotypic frequencies of a population from the Hardy-Weinberg equilibrium. This high value is due to both F_{st} (measuring between population genetic differentiation) and F_{is} (consanguinity coefficient measuring the effect of non-panmixic crosses within subpopulations). Such deviation from the Hardy-Weinberg equilibrium is therefore the consequence of the breeding system, predominantly self-pollinating.

Demography is a discipline which describes and explains changes in the number of individuals within a population over time. Its study is essential for an *in situ* conservation programme because it will allow the evolution of a population to be predicted, together with its persistence and levels of survival or destruction; it will also suggest how to implement field operations or management designed to preserve these populations.

In plant population demography, the most commonly used method is based upon the projection matrix theory. Such a matrix provides a lot of informatic statistics which are useful for monitoring population dynamics.

The first step was to construct a life cycle graph for each population. In such a graph, individuals are classified into developmental stages or age classes. In our case, in the presence of a substantial soil seed bank, we were able to differentiate seed classes according to their age. All other individuals, juvenile or adult, were grouped into classes according to their developmental stage. This model can therefore be defined as a "mixed demographic model".

In the generalized life cycle graph, each node is associated with a particular age class for the seeds or with a developmental stage for the growing plants. In the model, the projection interval is one year, which means that an arrow joining two nodes expresses the possibility for an individual to move from one to the other class or to contribute by its reproduction to increase the individuals of the other class within one year (G_0 G_1 J L_1 L_2 L_3).

Without going into details, the analysis of such population projection matrices provides a range of measures of the population structure and behaviour that are very useful for conservation.

For example the dominant eigen value of matrix A is the asymptotic growth rate of the population when it reaches its stable structure, i.e. stable distribution of stages/ages. The asymptotic growth rate may be used as a measure of fitness for the population in its particular environment.

From the projection matrix we can also define a sensitivity matrix S. The higher the sensitivity value of a matrix element, the greater the importance of the corresponding phase of the life cycle to the population growth rate. Sensitivity analysis may then be used (a) to determine which phase of the life cycle of the individuals is the most critical for population survival, (b) to quantify the contribution of each vital rate to population growth, (c) to evaluate the effects of environmental perturbations on population dynamics, and (d) to propose management strategies.

The projection matrix models were interpreted with the help of the "Unified Life Models" software.

The information gathered by demography allowed the identification of the major determinants of the population dynamics and the response of populations to contrasting environments. Our investigations also demonstrated differences in the demographic patterns of the studied populations. This is well illustrated by the sensitivity values.

In the sensitivity analysis, values of restricted sensitivity are particularly high in the most natural and protected sites. This indicates a high sensitivity of the populations located in this environment to modification of their life cycle due to disturbances, in particular those resulting from human activities.

The information collected from demography was also very useful to develop conservation and management strategies.

For example in disturbed sites, e.g. along trails or at the edges of coffee plantations, conservation of wild Lima beans should favour the growth of young lignified individuals. As a management practice, we suggest maintaining mulch on the soil surface at the end of the dry season or installing a vegetative cover to guarantee both high air and soil humidity. This practice would favour growth of L_1 individuals. We can also promote recruitment of new individuals from seeds by weeding the soil surface just after seed dispersal at the end of the dry season. This practice will allow rapid germination by exposing the seeds to high temperature in order to break the induced seed-coat dormancy.

In the case of decreasing populations, germination delay caused by dormancy and presence of a soil seed bank are particularly important. Weeding will then favour early germination and allow the emergent individuals to reach the L_1 stage within one growing season.

Total seed production in the population located in the most disturbed sites could be favoured by selecting the adult plants yielding a large quantity of seeds. Seed production could also be favoured by weeding during the rainy season. This ensures sufficient regrowth of the cut individuals.

As human pressure is growing in the Central Valley, even the populations located in the most natural sites are endangered. Our studies have shown that management of populations grown in these sites must favour the survival of L_3 individuals that appear to be a key factor in population dynamics. Selective clearing could also be carried out in these sites to maintain a continuous recruitment of new plants that could potentially serve as a reserve of lignified adult plants. Management should also avoid the adverse effects of high plant densities, by removing some individuals. As fecundity does not play any role in the dynamics of populations located in natural sites subjected to low levels of disturbance, an individual plant's seed production must not be taken into account when selective cuttings of *Phaseolus lunatus* are made in these populations.

Results of the demographic study are now being integrated in the *in situ* management of both natural and synthetic populations in the Central Valley of Costa Rica.

The ultimate objective of this project is of course to integrate the various factors studied (phenology, demography, gene flow, genetic structure, ecological distribution of wild lLima bean population) with a view to implement an *in situ* conservation programme in the Central Valley. To do this, we should preserve as many variants as possible, taking into consideration the different populations, distinct ecological habitats, relations between genetic structure, life zones (according to Holdridges's classification), and soil types of the target zone. Another important aspect of the implementation of the *in situ* consideration programme was to decide between two field options to maintain the selected populations: either preserving all wild populations in their natural sites by preserving a core collection in protected sites not yet colonized by *Phaseolus lunatus*, or using the natural resources already present in the Central Valley.

The first investigations on this component were undertaken in 1997 by integrating the Geographic Information System (GIS) as a tool for mapping the Central Valley as well as selecting and managing conservation reserves.

All environmental data available for the Central Valley were fingerprinted to develop different maps, e.g. climates, life zones, soil types, land use, communication, water course network and position of all the Lima bean populations in the target area. This work was carried out with the close collaboration of the CIEDES (Centro de Investigacion en Desarrollo Sostenible) using the GIS/ARCINFO software.

By doing so we first combined life zones and soil maps to identify all possible environmental combinations in which genetic variation could occur. For example 40 out of 72 possible combinations were derived in the Central Valley, with different distributions of wild Lima bean populations. We could identify two major combinations, representing half of the Central Valley area and containing most of the populations (75%).

Each combination was visited with the aim to locate sites likely to be used as conservation areas. The choice of the sites was guided by the need to retain only sites not likely to be disturbed in the coming years. We decided, on the basis of the actual situation of the Central Valley, to implement small conservation areas ("microreserves") not yet colonized by wild Lima bean, relatively protected from any human disturbance, representing the various ecological regions of the Central Valley and having a microenvironment well suited for wild Lima bean populations. We selected some isolated woody sites, remote from cultivated land or human settlements, mainly located along streams and deep slopes, legally protected from cutting and weeding. In these microreserves, the woody species are very useful to support the lignified and climbing *Phaseolus lunatus* individuals.

From this first analysis and from the identification of microreserves, we suggested planning the *in situ* conservation and management as follows:

- the minimal area of patches was calculated according to neighbourhood area determination; the area varied from 56 m^2 to 150 m^2;
- the number of seeds to introduce was estimated to provide a mean density of 0.35 adult individuals per m^2, assuming the mortality and germination rates observed in the demographic study;
- each patch was cleared to facilitate germination;
- supports were provided for the climbing wild beans, either by preserving those already present or by introducing some useful species;
- 2-3 wide corridors were built up and cleared like the patches;
- a buffer zone of about 5 m is present to protect microreserves;

- in each microreserve, management of the protected populations was applied according to the results of both phenological and demographic studies. For example optimum management operations consisted in favouring L_1 stage individuals, that are the most important for population growth rate. Clearing and selective cutting of juvenile individuals were also practised at the beginning of the dry season, to avoid competition for resources.

References

Hardy, O., S. Dubois, I. Zoro Bi and J.P. Baudoin. 1997. Gene dispersal and its consequences on the genetic structure of wild populations of lima bean (*Phaseolus lunatus*) in Costa Rica. Plant Genetic Resources Newsletter 109:1-6.

Wright, S. 1946. Isolation by distance under diverse systems of mating. Genetics 31:39-59.