Chapter 10
Molecular Markers for Characterizing and Conserving Crop Plant Germplasm

G. Barcaccia

Abstract Molecular markers have revolutionized and modernized our ability to characterize genetic variation and to rationalize genetic selection, being effective and reliable tools for the analysis of genome architectures and gene polymorphisms in crop plants. The area of plant genomics that has shown the greatest development with respect to the use of molecular marker technology is that of population genetics. All DNA polymorphism assays have proven to be powerful tools for characterizing and investigating germplasm resources, genetic variation and differentiation of populations, on the basis of gene diversity and gene flow estimates. In the last decade, RFLP and PCR-derived molecular markers have also been extensively applied in plant genetics and breeding for Mendelian gene tagging and QTL mapping. As a matter of fact, the number of loci for which DNA-based assays have been generated has increased dramatically, the majority using PCR as methodology platform. The information acquired is now being exploited to transfer different traits, including biotic stress resistances and improved quality traits, to important varieties by means of marker-assisted selection (MAS) programs. Although the potential for take-up is now much wider than in the past, the progress seems nevertheless to be slow, albeit measurable. The most important challenges in the near future are certainly the molecular characterization of germplasm collections for preserving them from genetic erosion and the identification of phenotypic variants potentially useful for breeding new varieties. Knowing the presence of useful traits, genes and alleles would help in making decisions on the multiplication of plant accessions and the maintenance of seed stocks. There are no doubts that the use of molecular markers for characterization and conservation of genetic resources should be implemented so that potentially useful genes and genotypes can be added to core collections to make them exploitable by breeders. A new concept that might be successful is that of building crop plant collections primarily based on the knowledge of the presence of valuable genes and traits.

G. Barcaccia
Department of Environmental Agronomy and Crop Science, Faculty of Agriculture,
University of Padova, Campus of Agripolis, Viale dell'Università 16, 35020, Legnaro, (Italy)
e-mail: gianni.barcaccia@unipd.it

S.M. Jain and D.S. Brar (eds.), *Molecular Techniques in Crop Improvement*,
DOI 10.1007/978-90-481-2967-6_10, © Springer Science+Business Media B.V. 2009

G. Barcaccia

10.1 Introduction

Molecular markers have proven to be powerful tools for analyzing germplasm resources and assessing genetic variation within as well as genetic differentiation among populations. In fact, the area of plant genomics that has shown the greatest development with respect to the use of DNA marker technology is that of population genetics. However, both RFLP and PCR-derived markers have also been extensively applied in plant genetics and breeding for mapping Mendelian genes and QTLs.

The use of molecular markers for investigating and managing genetic resources should be implemented so that useful information on genes and traits can be added to core collections to make them exploitable by breeders.

This chapter deals with the use of molecular markers for characterization of crop germplasm and for decision-making in conservation programs. Case studies related to important self- and cross-pollinated leguminous and cereal crops for exploring germplasm resources and mapping agronomic traits in landraces and elite stocks are presented. Theoretical considerations and experimental observations are critically discussed.

10.2 Genetic Characterization and Its Use in Decision-Making for the Conservation of Crop Germplasm: Basic Concepts

The characterization, conservation and exploitation of crop plant germplasm maintained in gene banks propound a number of challenges to the researchers dedicated to the investigation of plant genetic resources. Common problems include the development of strategies for sampling representative individuals in natural and experimental populations, the improvement of tools and technologies for long-term conservation and for high-throughput characterization of large numbers of stored accessions. The knowledge of the genetic diversity present in a gene bank is crucial for developing sustainable conservation strategies and it is also essential for the profitable exploitation of a gene bank by specific breeding programs. As a matter of fact, germplasm characterization of plant accessions deposited in gene banks has been limited and this likely represents a major cause for the limited adoption of conserved accessions in crop breeding programs (Ferreira 2006). Consequently, the genetic characterization of accessions belonging to a given collection and the examination of genetic relationships among them should be strengthened and perpetrated not only for maintaining but also for exploiting crop genetic resources.

Conservation of the genetic resources in the agro-ecosystem in which they have evolved (in situ conservation) is now being more widely considered, as complementary to strategies based on gene banks (ex situ conservation), for limiting genetic erosion and so preserving genetic diversity. If it is true that in situ conservation has been proposed essentially for wild relatives of cultivated plants,

BookID 104577_ChapID 10_Proof# 1 - 11/07/2009

it is also true that when considered for major crops this alternative can very often be unfeasible from a socio-economic perspective (Negri et al. 2000; Lucchin et al. 2003). Moreover, on-farm conservation of landraces is seen as a dynamic system that could help maintaining intact the genetic adaptation to changing conditions and the technical, social, cultural and environmental context in which they have occurred and evolved. In view of this, the whole of morpho-phenological and agronomic traits together with molecular markers could be the basis for the recognition of marks of specificity and typicality of landraces which would further enhance not only their on-farm conservation, but also allow their on-market valorization (Table 10.1).

Conservation of genetic resources entails several activities, many of which can greatly benefit from knowledge generated through the use of molecular marker technologies. The same applies to activities related to the acquisition and collection of germplasm (i.e., accurate description of accessions and application of effective procedures) and its evaluation for useful agronomic traits. The availability of a robust genetic characterization ensures that decisions made on conservation strategies will be supported by this information and results in an improved germplasm management. Of the experimental activities associated to plant genetic resources, the morpho-phenological and molecular evaluation of germplasm adds value to plant genetic resources and it is particularly important because helps the identification of genes and traits, and thus provides the basic knowledge for the exploitation of collections in programs aimed at breeding new crop varieties.

Classical attempts to directly use plant accessions stored in germplasm banks in breeding programs have been mainly focused to the identification of sources of genes of interest, such as resistance to plant pathogens or pests, and their transfer to cultivated materials. Linkage drag has very often restrained breeders from the initiative of using accessions from germplasm banks mainly because the improved material of advanced breeding programs is far more attractive than any germplasm resource of unknown genetic origin and phenotypic adaptation or performance (Ferreira 2006). When such a risk is taken, the accessions are usually screened to reveal the presence of a gene of interest and typically a backcross program is then initiated to transfer the gene to an elite line or cultivar. This procedure, however, is usually limited to Mendelian traits under simple genetic control, whereas complex traits require more elaborated methods, such as the mapping of QTLs in order to

Table 10.1 Main genetic, cultural and socio-economic features of on farm and ex situ conservation schemes for crop plant diversity

	Conservation strategy	
	On farm	Ex situ
Genetic drift and inbreeding	Operating	Not operating
Genetic adaptation to changing conditions	Happening	Not happening
Cultural and socio-economic role of crops	Maintained	Eroded
Cost	Moderate to high	Low to moderate

integrates the classical backcrossing program with linkage information based on the use of molecular markers.

The term "characterization" indicates the description of a single character or the evaluation of a set of genetic traits in individuals and populations. Moreover, this term is also used by means of distinguishing genotypes or gene pools. Thus, characterization of plant genetic resources refers to the process and tool by which accessions can be evaluated, differentiated and identified. In broad terms, this identification can refer to any morphological descriptor or molecular polymorphism of an accession. In the agreed terminology of gene banks and management of germplasm collections, characterization usually stands for the description of qualitative traits or quantitative traits that are highly heritable, easily scored by the eye and equally expressed in all environments (International Plant Genetic Resources Institute, IPGRI). Under a molecular point of view, characterization refers more specifically to the detection of DNA polymorphisms as a result of differences in random sequences or specific genes by using molecular marker techniques.

Standard characterization and evaluation of accessions can be routinely carried out by using different methods, including traditional practices such as the use of descriptor lists of morphological characters. They can also involve the adoption of agronomic trials under various environmental conditions. Genetic analysis and, in particular, molecular characterization of accessions refer to the visualization and description of DNA markers that follow Mendelian inheritance patterns and that involve specific or random sequences in the genome. In this context, the application of RFLP or PCR-derived markers for assaying polymorphisms at single or multiple loci all qualify as genetic characterization methods. Because of its nature, molecular characterization clearly offers an enhanced power for detecting diversity through fingerprinting and genotyping individual genomes and haplotyping individual genes. In addition, characterization with molecular markers offers a greater power of detection compared to phenotypic traits because they reveal differences at the genotypic level and are not influenced by the environment.

10.3 Use of Molecular Markers for the Characterization and Conservation of Plant Genetic Resources

Genomic DNA-based marker assays have revolutionized and modernized our ability to characterize genetic variation and to rationalize genetic selection (Lanteri and Barcaccia 2006). Molecular markers are known as particularly effective and reliable tools for the characterization of genome architectures and the investigation of gene polymorphisms in crop plants.

Besides linkage mapping, gene targeting and assisted breeding, the plant DNA polymorphism assays are powerful tools for characterizing and investigating germplasm resources and genetic relatedness. These techniques include restriction fragment length polymorphism (RFLP) markers and PCR-based molecular markers,

BookID 104577_ChapID 10_Proof# 1 - 11/07/2009

such as simple sequence repeat (SSR) or microsatellite markers (Morgante and Olivieri 1993), amplified fragment length polymorphism (AFLP) markers (Vos et al. 1995). SSR and AFLP markers are the most widely exploited techniques for the characterization of crop plant genetic resources: both types of markers offer an almost unlimited supply of molecular traits for distinctive fingerprinting and genotyping of plant materials, respectively, with and without a prior knowledge of target DNA sequences.

In general, molecular markers detect polymorphism by assaying subsets of the total amount of the DNA sequence variation in a genome. Polymorphisms detected by the RFLP assay reflect the variation of restriction fragment sites. PCR-based polymorphisms result from DNA sequence variation at primer binding sites and from DNA length differences between primer binding sites. The SSR assay utilizes pairs of primers flanking each simple sequence repeat and polymorphisms differ for the number of repetitive di-, tri- or tetranucleotide units present at one locus. The AFLP assay detects polymorphisms at multiple loci and involves the use of combination of primers specific for two distinct four-base and six-base long restriction sites flanking the target sequence unit. Although RFLP markers have provided useful estimates of the genetic diversity and relatedness in crop plants, there is some concern about their discriminatory power. Increasing the number of probe-enzyme combinations may improve the number of RFLP marker loci detectable, but the level of polymorphism that can be revealed by PCR-based markers still remains higher. In fact, owing to their own genetic nature, SSR markers usually detect multiple alleles at a given locus while AFLP assays mainly detect single alleles at multiple loci randomly distributed in the genome. A more recently introduced method is represented by SNP markers based on the detection of single-nucleotide polymorphisms by direct DNA sequencing of target gene regions (Fig. 10.1).

As a matter of fact, until now AFLP markers have provided the most widespread and robust technique with the highest polymorphism information content that combines the reliability of the RFLP technique with the potentiality of the PCR technique.

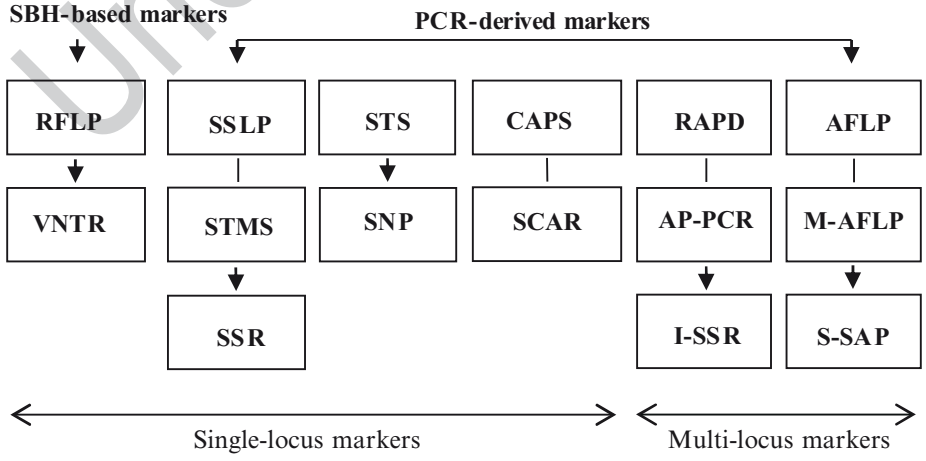

Fig. 10.1 Classification of the most commonly used molecular marker systems

Either AFLP and AFLP-derived markers, such as Microsatellite-AFLP and sequence-specific amplified polymorphism (S-SAP) markers, have been mostly used for fingerprinting and scanning whole genomes, characterizing single chromosomes, and tagging specific genes. However, microsatellite (SSR) markers along with SNP markers are now considered the most powerful and robust molecular marker systems for the analysis of whole genomes and single genes, and hence for the molecular characterization of single lines and population groups by means of genotyping or haplotyping.

Multi-locus marker systems, such as arbitrarily primed-PCR markers, AFLP and AFLP-derived markers visualize simultaneously many marker alleles: they supply an nearly unlimited number of polymorphisms and are exploitable over all species with no pre-existing genome or gene sequence knowledge, but show dominance (*i.e.*, only one allele identified, no possibility to discriminate between homozygous and heterozygous individuals). They exploit fingerprints, which are typically analyzed as pair-wise comparisons and whose results are to a limited extent reproducible and comparable among laboratories. Single-locus marker systems, such as RFLP, SSR and SNP markers, are usually characterized by co-dominance (*i.e.*, both alleles identified in heterozygous individuals) and thus are very informative in terms of polymorphisms, and supply reliable and comparable data. They allow to reconstruct genotypes and define haplotypes by multiplex analyses, however requires preliminary sequence information and thus are not always easily applicable.

Molecular markers are an irreplaceable tool to study biodiversity at the genetic level. Independently from the marker system considered, all types of plant DNA polymorphisms have proven to be powerful marker assays for assessing genetic variation and differentiation of populations, on the basis of gene diversity and gene flow estimates. In particular, the use of DNA markers allows to measure the genetic variation within single populations and to evaluate the genetic relatedness among populations, so that the formulation and implementation of germplasm maintenance and use programs can be optimized.

With the development of the PCR and DNA sequencing technologies associated with high-throughput screening systems, marker polymorphisms are now the choice for molecular-based surveys of genetic variation. Importantly, molecular markers showing different patterns of inheritance can now be investigated in nearly all of major crop species. The most widespread use of molecular markers in this context is the assessment of genetic variation within and genetic relationship among populations. Although in principle all types of molecular markers would be suitable for this purpose, microsatellite markers have been in the recent past and still are the most used in all the diversity studies. A standard set of markers for the major crop species should be selected and recommended to investigate the neutral genetic variability in the genome. In addition, one might also consider markers associated with important Mendelian genes and QTLs, reflecting the adaptive genetic potentials of individuals for a given qualitative or quantitative trait. For the most agriculturally important species, breeding research programs focus on the mapping of genes and polygenes so that an increasing number of markers for assisted characterization and selection of plant genetic resources will be available in the future.

Molecular markers are an indispensable tool to understand the genetic structures of populations. For the sampling of germplasm to create a plant gene bank, they are necessary but in no way sufficient to make adequate decisions. In addition to diversity information derived from molecular marker data sets, there are needs for understanding plant resource characteristics and specific knowledge on breeding values. It is therefore strongly recommended to concentrate co-ordinated phenotyping and genotyping efforts to fill the global maps of crop species diversity and to develop a better understanding of the rational decision-making process.

Molecular polymorphisms linked to quantitative traits or qualitative genes as well as nucleotide variants of causative genes can find practical application in marker-assisted selection (MAS) breeding programs.

Detailed linkage maps of specific chromosomes or chromosome blocks have been developed by analyzing the segregation of selected molecular markers of maternal and paternal origin in experimental populations. Moreover, wide-genome scanning approaches provide a method for rapidly identifying molecular markers linked to a specific genetic trait and accurately locating genes in a saturated genetic map. For instance, the identification of AFLP markers associated with specific genes based on the use of markers randomly distributed in the genome coupled with the use of near isogenic lines (NIL) or bulked segregant analysis (BSA) was applied not only in diploids but also in remarkably complex auto- and allo-polyploids. Both strategies are based on pooling the DNAs from individuals sharing the same genetic background but showing extreme classes of a given trait, and then screening for differences between genetic classes using molecular markers. Molecular differential screening of plants with contrasting characteristics is still considered one of the most powerful tools for identifying, isolating and using the genes underlying the expression of Mendelian traits and QTLs (quantitative trait loci). Similarly, natural populations can be exploited for discovering molecular markers linked to agronomically important genes by means of association mapping through linkage disequilibrium (LD) analysis. Interest in the study of LD, *i.e.* non-random association of alleles, in crop plants has increased dramatically in recent years because of two main factors (reviewed by Rafalski and Morgante 2004). First, genomic technologies enables rapid identification of haplotypes at many Mendelian loci, either by DNA sequencing or by detecting SNP markers. Second, in the presence of significant LD, it can be possible to identify genetic regions that are associated with a particular trait of interest (e.g., disease susceptibility) by genome scanning of individuals from an existing population. By contrast, if LD declines rapidly around the causative gene, the identification of genetic factors responsible for the trait of interest is possible by screening a limited set of candidate genes. Individual SNP markers or SNP haplotypes within a candidate gene are systematically tested for association with the phenotype of interest. Whole-genome scan and candidate gene approach are fundamentally similar methodologies, and differ primarily in the scale at which the analysis is performed.

An important point to be considered when using molecular markers in genetic characterization studies is the nature of the genomic DNA polymorphisms, since molecular markers can assay either neutral or adaptive variation. The genomics revolution of the last ten years has improved our understanding of the genetic

make-up of living organisms (Vendramin and Morgante 2006). Together with the achievements represented by complete genomic sequences for an increasing number of species, high throughput and parallel approaches are available for the analysis of transcripts, proteins, chemically-induced and transgenic mutants. All this information facilitates the understanding of the function of genes in terms of their relationship to the phenotype. Despite its great relevance, such an understanding could be of little value to population and conservation genetics because it will not elucidate the relationship between genetic variation in gene sequences and phenotypic variation in traits, but rather only that between a given gene and a mutant phenotype (Vendramin and Morgante 2006). The relationships between the phenotypic variation of complex traits and the molecular polymorphism of genes can be studied on the basis of a genomic approach. Work in model plant species such as Arabidopsis and rice has started to unveil a large number of genes involved in the determination of traits of adaptive significance, such as phenology and environmental stress tolerance or resistance. This progress will finally allow conservation genetics to directly analyze variation in genes involved in adaptive processes rather than in neutral markers. However, neutral markers will remain important to make inferences about stochastic processes affecting natural population evolution.

It is known that most molecular marker systems exploited for population genetics target genomic regions which are selectively neutral, even though molecular marker technologies which target specific genes do exist. The neutrality of markers is suitable for most uses in germplasm conservation and management. However, when the interest of conservation lies specifically in the diversity of traits of agronomic importance, some questions remain on the representativeness of markers. In such cases, the markers able to detect functional diversity are more suitable for the characterization and management of germplasm collections.

As a matter of fact, technology is rapidly evolving in molecular marker systems, moving from anonymous markers towards markers associated to a specific gene and/or a known chromosome position. One of the roles of biotechnology is that of supplying low-cost and high-throughput molecular tools for developing markers and making decisions in the processes of characterization and conservation of agricultural genetic resources.

10.4 Genetic Diversity and Similarity Statistics for Characterizing Plant Germplasm at the Population Level

Genetic diversity and similarity measurements are very useful for describing the genetic structure of populations. The genetic structure of natural populations of a crop plant species is strongly influenced by the reproductive system of their individuals and the union types occurring within populations. Breeding schemes that can be adopted as well as variety types that can be constituted depend on the reproductive barriers and mating systems of plants.

Natural populations of species that reproduce by apomixis or that propagate vegetatively are polyclonal, being composed by several genetically distinct clones and usually dominated by a few well-adapted genotypes. Therefore, genetic variation within populations is distributed among clones and most populations are characterized by different levels of differentiation among genotypes.

Landraces of self-pollinated species (e.g., bean, lentil, wheat and barley) are composed of a mixture of pure lines, genetically related but reproductively independent each other. Thus, genetic as well as phenotypic variation is mainly detectable among lines due to the presence within natural populations of fixed genotypes mainly homozygous for different alleles. Spontaneous hybridization is however possible to some extent depending on the species, environmental factors and germplasm stocks. Cultivated varieties of selfing species are usually represented by pure lines obtained by repeated self-pollination of a number of hybrid individuals stemmed from two parental lines chosen for complementary morphological and commercial traits.

Maize is one of the most commercially important cross-pollinated species. In many countries, existing landraces are selected by farmers for their own use and eventually sale to neighbors. Traditionally, landraces are developed by mass selection in order to obtain relatively uniform populations characterized by valuable production locally. Synthetics are also produced by intercrossing a number of phenotypically superior plants, selected on the basis of morpho-phenological and commercial traits. More rarely, plants are also evaluated genotypically by means of progeny tests. Compared to landraces, synthetics have a narrower genetic base but are equivalently represented by a heterogeneous mixture of highly heterozygous genotypes sharing a common gene pool. However, newly released varieties are exclusively represented by F_1 hybrids developed by private breeders and seed companies using inbred lines belonging to distinct heterotic groups.

Genetic characterization is providing new information to guide and prioritize conservation decisions for crop plants. The most urgently required action is the effective protection of all remaining wild ancestral populations and closely related species of crop plants, most of them now endangered. They are the only remaining sources of putative alleles of economic values that might have been lost during domestication events. It is equally important to ensure that the plant genetic resources selected for conservation include populations from the geographic areas representing the different domestication centres where high estimates of genetic diversity within and differentiation among populations are expected.

The assessment of the optimal plant and molecular marker sample size is a key step for the characterization of populations by computing genetic diversity and/or similarity statistics (Table 10.2). In particular, the choice of the most appropriate system (i.e., dominant vs codominant) and type (i.e., neutral vs functional) of markers and of the optimal number of markers and plants required to describe the genetic structure of a given population has to be carefully addressed.

For conservation purposes, individual genotypes of collected populations need to be identified to become part of the conservation scheme. Some general criteria can be defined concerning the desirable genetic properties of the sample: (1) it should represent the largest gene pool possible of the species; (2) it should

t2.1 **Table 10.2** Sizes of the plant sample for landraces in relation to the reproductive system of
t2.2 the species

t2.3 t2.4 Prevalent reproductive system of the species	Level and type of genetic variation in the population	Range of minimum sample size (seeds or plants)
t2.5 Apomixis	Low (clones)	20–30
t2.6 Amphimixis		
t2.7 Selfing	Intermediate (pure lines)	40–60
t2.8 Crossing	High (HW equilibrium)	80–100

have a maximum effective population size; (3) special genetic traits should be conserved. It is worth noting that maximizing the effective population size suggests collecting extreme genotypes, which may not be representative for the population. Using parameters from population genetics, the group of plants chosen should have minimum inbreeding and minimum relatedness among each other.

Two criteria should be followed for genetic marker-based characterization and conservation decision-making: priority populations for conservation should be those with the largest within-populations diversity and/or should maximize the conservation of between-populations diversity. Both within and between population diversity parameters are classically measured using molecular markers. In both cases, soundly based priority decisions for conservation at the global level will require the availability of large datasets.

Standard genetic diversity (H) and genetic differentiation (D) statistics of Nei (1973) and the inbreeding (F) coefficients of Wright (1965) are widely used to summarize the genetic structure of populations and the distribution of genetic variation using codominant molecular markers. The average (n_o) and the effective (n_e) number of alleles per locus well as the observed (H_o) and the expected (H_e) heterozygosity are the most commonly calculated population genetic parameters for the characterization of the within-subpopulation diversity. Let p_i denote the frequency of the i^{th} marker allele at a given locus, the genetic diversity computed as $H=1 - p_i^2$ is equivalent to the expected heterozygosity. This parameter can be computed for single subpopulations and the population as a whole. A hierarchical analysis of variance with estimation of F-statistics is then usually performed in order to measure the deficiency or excess of heterozygosity within subpopulation (F_{IS}) and between subpopulations (F_{IT}) as well the fixation index (F_{ST}) of the total population. In particular, F_{ST} measures the effect of total population subdivision in different subpopulations and it corresponds to the reduction of heterozygosity of subpopulations compared to the population as a whole (Fig. 10.2).

Genetic similarity (GS) estimates between populations are also very often computed using dominant molecular markers. These estimates are traditionally calculated in all possible pair-wise comparisons between individuals either within or between populations using different genetic similarity parameters. For instance, genetic similarity estimates between individuals can be calculated using the following formula of Dice (1945): $GS = 2M_{ij}/(2M_{ij} + M_i + M_j)$, where M_{ij} represents the number of shared markers scored between the pair of individual fingerprints (i and j) considered, M_i is the number of markers present in i but absent in j and M_j

BookID 104577_ChapID 10_Proof# 1 - 11/07/2009

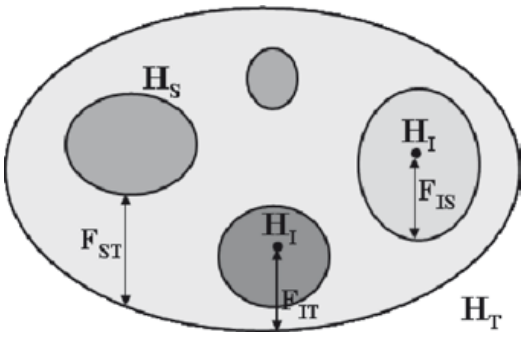

Fig. 10.2 Inbreeding (F) coefficients of Wright (1965) for assessing the genetic structure of populations

is the number of markers present in *j* but absent in *i*. Such estimate of genetic similarity corresponds to that calculated by the formula of Nei and Li (1979). The Simple Matching (SM) coefficient is also very popular for computing genetic similarity.

Any set of genetic diversities or similarities can be analyzed in terms of within and between population genetic variation, and more particularly, in terms of individual plant contributions to the total genetic variability. The most common approach used involves calculating a matrix of genetic distances or similarities on the basis of marker frequencies and marker fingerprint, respectively, generating UPGMA dendrograms and PCA centroids. Priority populations and individuals for conservation would therefore be the ones contributing most to the variation and differentiation of the germplasm set.

10.5 Using Molecular Marker-assisted Characterization and Conservation of Crop Plant Germplasm: Case Studies

Landraces are populations with high genetic variability and fitness to the natural and anthropological environments where they have originated. They represent not only a valuable source of useful traits, but also an irreplaceable bank of highly co-adapted genotypes. Knowledge of genetic variation within local populations and genetic differentiation with breeding stocks is expected to have a significant impact on the preservation and exploitation of plant germplasm resources. The case-studies reported in this paper refer to Italian maize and lentil landraces, durum and bread wheat pure lines, and Mesoamerican bean varieties.

In maize (*Zea mays* L.), unless in the last few decades conservation of local populations in gene-banks has taken place worldwide, the gradual replacement in cultivation of landraces by hybrids has resulted in genetic erosion. Many maize breeders are now concerned that genetic diversity within this species has been decreasing at an alarming rate as a consequence of modern hybrids and other agricultural changes. For instance, most inbreds have been developed from a limited number of elite lines and synthetics, a practice that heightens the risk of genetic uniformity in commercial maize production fields. Thus, maize breeders have

recently become more aware of the need for both assessing and maintaining genetic diversity among hybrid varieties and improving the management of genetic resources through the conservation of landraces. Information about genetic diversity and differentiation is known to be useful in planning crosses for hybrid development, assigning lines to some heterotic groups, maintaining genetic variability of landraces, and protecting inbreds and varieties. It can be obtained by surveying both qualitative and quantitative morphological traits or using molecular markers for investigating polymorphisms at the DNA sequence level.

Lentil (*Lens culinaris* Medik.) has been part of the human diet since a long time, being one of the first crops domesticated in the Near East. A great part of the landraces came from temperate areas of the Mediterranean basin, where lentils exist with colours that range from yellow to red-orange to green, brown and black. Two clear morphological groups based on seed characters could be established in this species, coinciding with the taxonomic description for the *macrosperma* and *microsperma* types. As in other Mediterranean countries, Italian landraces seem to be migrated following criteria related with seed characters preferred by farmers. Some of the landraces show an increasing importance in the food market of plant derivatives because lentil is one of the crops with the highest level of protein, being so a very important part of the diet.

Modern European wheat varieties are thought to display quite low levels of gene pool variation because of the high selective pressure applied in breeding programs. The genetic diversity of durum wheat (*Triticum turgidum* L. var. *durum*) and bread wheat (*Triticum aestivum* L. var. *aestivum*) elite germplasm has been traditionally estimated on the basis of morphological and quantitative traits, disease resistances, gliadin proteins and only recently by molecular markers. The adoption of molecular markers for the genetic characterization of wheat accessions is expected to play a key role not only for marker-assisted conservation but also for marker-assisted identification and selection of varieties in this species.

Regarding beans (*Phaseolus vulgaris* L.), yellow-coloured beans are among the traditional bean varieties grown principally in Mexico and Peru under several names, such as Azufrado and Canario. Originally, varieties from these countries represented two evolutionarily distinct bean groups as they originated from two different domestication centres, one in Mexico and the other in the southern Andes. Beans with yellow-coloured seeds are grown and consumed mainly in the north-western part of Mexico, but immigration from this country has created a market for yellow-seeded beans mainly in the USA and more recently also in Europe.

10.5.1 Genetic Anatomy of a Patented Yellow Bean (Phaseolus Vulgaris L.) Variety

Molecular markers find application in plant science and crop improvement to overcome limitations due to the absence of an appropriate and uniform legal protection of varieties and germplasm resources. In some countries, like USA, it is

BookID 104577_ChapID 10_Proof# 1 - 11/07/2009

possible to obtain a utility patent for varieties of crop species and it is also possible to obtain a Plant Variety Protection (PVP) certificate. Some of these awards can be controversial because of the perceived lack of novelty of distinctive morphological traits and the varieties themselves.

Genomic DNA fingerprinting was exploited by Pallottini et al. (2004) for assessing the genetic anatomy of a patented yellow bean (*Phaseolus vulgaris* L.), specifically the variety Enola. In order to check its origin, an AFLP-based genomic DNA fingerprinting study was carried out using a representative sample of 56 domesticated common bean accessions, including a subsample of 24 varieties with yellow coloured seeds morphologically similar to those of Enola. AFLP markers were detected using several primer combinations which revealed a total of 133 polymorphic markers. Most yellow-seeded beans, including Enola, were clustered in a tightly knit subgroup of the Andean gene pool. Enola was most closely related to the pre-existing Mexican cultivars: a sample of 16 individuals of Enola displayed full genetic identity with yellow-seeded beans from Mexico, namely Mayocoba and Azufrado Peruano (Fig. 10.3). Probability calculations of matching the specific Enola DNA fingerprint showed that the most likely origin of Enola is by direct selection within pre-existing yellow-bean cultivars from Mexico (Pallottini et al. 2004).

In particular, the least probable scenario was the one in which the Enola marker combination resulted from a cross between Andean and Mesoamerican genotypes, regardless of their seed colour represented in our sample (probability of 1×10^{-18}). The scenario with the highest probability represented selection without hybridization within cultivar Azufrado Peruano 87 (probability of 3×10^{-1}).

The assessment of the genetic anatomy of Enola as well as its genetic identity with pre-existing Central America local varieties raise questions about the rationales for the award of a utility patent and a PVP certificate.

10.5.2 Genetic Variation and Differentiation of Landraces of Lentil (Lens Culinaris Var. Microsperma L.) and Maize (Zea Mays Var. Indurata L.)

The Common Agricultural Policy of the EU concerning the quality of agricultural products and the preservation of landraces have driven more attention towards the Protected Geographical Indication (PGI) and Protected Designation of Origin (PDO) marks. Lentil (*Lens culinaris* var. *microsperma* L.) is extensively cultivated in the Mediterranean basin where it includes different landraces and modern varieties.

Barcaccia et al. (1998) demonstrated that the set up of molecular marker-based reference systems is feasible for the precise identification of single farmer's populations and for the preservation of the landrace as a whole. The mean Dice's genetic similarity estimates among lentil entries was shown to be 93%, ranging from 91% to 100%, when populations were analyzed as bulked DNA samples of 12 plants each. Although multi-locus DNA fingerprints were identical among most but not

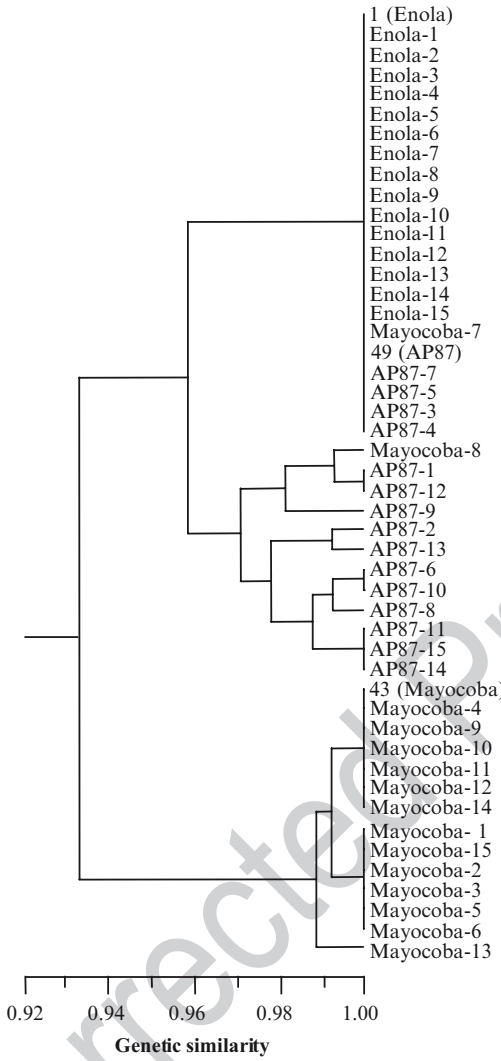

Fig. 10.3 UPGMA dendrogram showing the genetic relationships among Peruano-type bean cultivars: Enola, Azufrado Peruano 87 (AP87), and Mayocoba as assessed using AFLP markers (for additional information see Pallottini et al. 2004)

all the 26 chosen farmer's populations, indicating an homogeneous gene pool of the lentil landrace locally known as "Castelluccio of Norcia", some reliable polymorphic markers were visualized and three off-type farmer's populations were finally discovered. A number of polymorphisms were unambiguously shared with two Canadian (Laird) and Turkish (Kislik) commercial varieties used as genetic controls since their seeds were considered morphologically indistinguishable from the seeds of the Italian landrace under study. This finding cannot be explained by gene flow occurring among farmers but most likely by taking into account seed exchange with farmers who do not reproduce their own seed stocks. On the whole, the molecular approach proved to be essential for clearly identifying lentil local varieties and also for discriminating phenotypically similar imported lentil varieties.

Several types of molecular markers were also used for the characterization of the gene pool of Italian landraces of maize (*Zea mays* var. *indurata* L.). In particular, a comparative characterization of 10 field populations of the landrace "Nostrano of Storo" was carried out using different types of PCR-based markers (Barcaccia et al. 2003). The inbred line B73 and three synthetics (VA143,VA154 and VA157) selected from as many landraces were also used. Genetic diversity and relatedness were evaluated over 84 SSR and 53 Inter-SSR marker alleles using a total of 253 plants. Up to 23 alleles per SSR locus were scored while the average effective number of alleles per population was 6.99. Nei's total genetic diversity as assessed with SSR markers was $H_T = 0.851$ while the average diversity within populations was $H_S = 0.795$. The overall Wright's fixation index F_{ST} was as low as 0.066. Thus, more than 93% of the total variation was within population. Unique alleles over all SSR loci were found for six populations. An average of 17.7 marker alleles per Inter-SSR primer were scored with an effective number of marker alleles per locus of 1.34. The Shannon's diversity information index over all populations and I-SSR loci was 0.332, varying from 0.286 to 0.391. The extent of differentiation between populations was as low as $G_{ST} = 0.091$. Dice's genetic similarity matrices were estimated for both SSR and Inter-SSR markers. The mean genetic similarity coefficients within and between populations were respectively 0.269 and 0.217, for SSR markers, and 0.591 and 0.564, for Inter-SSR markers. UPGMA dendrograms displayed all field populations but one clustered into a distinct group, in which the syntheticVA154, selected from the "Marano Vicentino" landrace, was also included. One field population and the other two synthetics were clustered separately as well B73. The matrix correlation assayed by the Mantel's correspondence test was as high as 0.908. Findings suggest that, although a high variability can be found among plants, most plant genotypes belong to the same landrace locally called "Nostrano of Storo" (Barcaccia et al. 2003). Although gene flow from commercial hybrids might have occurred, the large number of polymorphisms and the presence of both unique alleles and alleles unshared with B73 and synthetics are the main factors underlying the value of this flint maize landrace as a source of genetic variation and peculiar germplasm traits. Because of its exclusive utilization for human consumption, such a molecular marker characterization will be a key step for obtaining a quality mark and so promoting the conservation and protection of the landrace.

The construction of a linkage map was then undertaken using several molecular marker systems, including AFLP and SSR markers (Fig. 10.4).

A set of mapped markers was adopted to study the effects of different conservation strategies (on farm, in situ and ex situ) on the population genetic structure of landraces (Pallottini 2002). Marker alleles showing significant frequency changes were distributed throughout the genome, but chromosome blocks carrying marker loci with differentiated genetic diversity values were detected. Some molecular markers revealed to be suitable also for the evaluation of the extent of gene flow that can occur between landraces and modern varieties.

The possibility of identifying the local populations of lentil and maize through their molecular characterization can be an essential element not only for conservation and breeding purposes, but also for certifying typical local products in order to

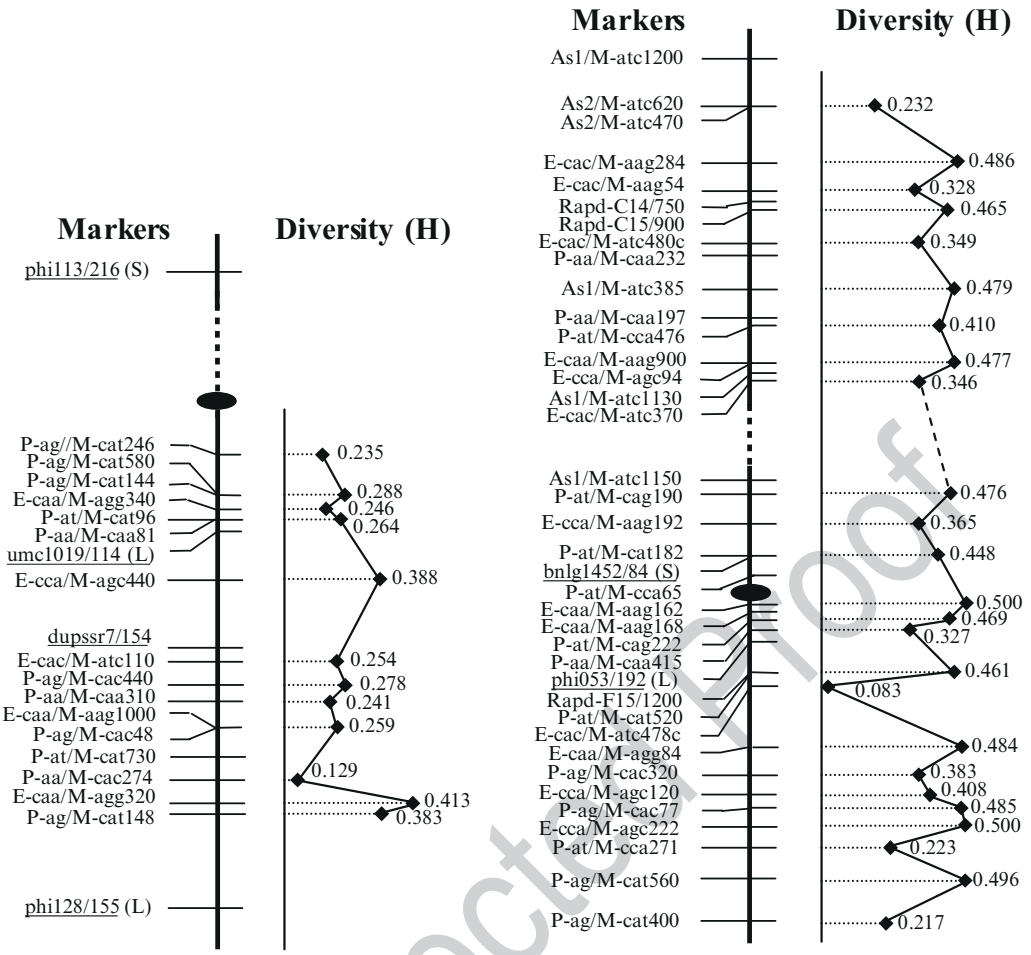

Fig. 10.4 Linkage groups of two maize chromosomes (5 and 3) carrying marker loci with differentiated genetic diversity (H) estimates (0.281 vs 0.409, on average)

avoid market frauds and to safeguard their gene pools. In the near future it could represent a basic requisite for their use in a serious and consumer-oriented production and marketing context.

10.5.3 Genetic Fingerprinting Durum Wheat (Triticum Durum l.) and Bread Wheat (Triticum Aestivum l.) Elite Germplasm Stocks for Multiple Breeding Purposes

The genetic diversity of durum wheat (*Triticum turgidum* L. var. *durum*) and bread wheat (*Triticum aestivum* L. var. *aestivum*) modern variety collections was estimated on the basis of molecular markers using AFLP technology for genomic DNA fingerprinting.

A total of 38 durum wheat and 26 bread wheat genomic DNA samples isolated from commercial varieties and experimental lines were investigated by fluorescent

BookID 104577_ChapID 10_Proof# 1 - 11/07/2009

AFLP markers using five different *Pst*I/*Mse*I and *Eco*RI/*Mse*I primer combinations previously selected on the basis of their ability of detecting polymorphisms. As many as 267 clearly detectable markers were scored, of which 59 (41.6%) and 73 (51.4%) proved to be polymorphic among varieties within and between species, respectively. Dice's genetic similarity (GS) estimates among the 64 pure lines were calculated in all possible pair-wise comparisons and the correspondent similarity matrix was used for the construction of UPGMA dendrograms and for the definition of centroids according to PCA analysis. Mean genetic similarity estimates within durum wheat and within bread wheat were 92% and 89%, respectively. In each species, a few multi-locus genotypes showing almost full identity were found. Several species-specific and variety-specific DNA markers were also scored: the latter types will be cloned, sequenced and converted into easily detectable single-locus markers.

On the whole, more than 68% of the total genetic variation found in wheat materials was explained by the first two principal coordinates. The observed number (n_o) and the effective number (n_e) of alleles were equal to 1.416 and 1.163 in durum wheat and to 1.514 and 1.184 in bread wheat, respectively. Nei's genetic diversity (H) estimates over all genomic loci were also comparable for the two species (0.102 and 0.119, respectively). Linkage disequilibrium (LD) tests were performed for all pair-wise comparisons of marker alleles. The number of significant LD was 78 over 142 loci (0.78%) in durum wheat and 139 over 142 loci (1.42%) in bread wheat. Preliminary data suggest the finding of a few AFLP markers displaying highly significant linkage disequilibrium ($P < 0.01$) with a number of wheat resistance genes, including yellow (stripe) and brown (leaf) rust, powdery mildew, Fusarium head blight and Septoria leaf spot diseases (Table 10.3).

Table 10.3 Estimates of linkage disequilibrium (LD) between molecular markers and resistance factors in bread and durum wheat varieties

Marker	Pathogen	LD	*r*	χ^2	*P*
A22	BR	–0.0831	–0.3508	4.68	0.0306
D51	YR	–0.0893	–0.3613	4.96	0.0259
A22	FUS	–0.1177	–0.5026	9.60	0.0019
D36	SEP	0.0693	0.3885	5.74	0.0166
E45	PM	–0.0436	–0.3254	4.02	0.0049
B27	BR	–0.0917	–0.0167	12.91	0.0003
D67	BR	–0.0651	0.0108	11.92	0.0006
C23	YR	0.0799	0.7826	7.63	0.0057
D15	YR	–0.0799	–0.4793	7.63	0.0057
C32	FUS	0.0888	0.8696	11.30	0.0008
E41	FUS	0.1154	0.6500	7.80	0.0057
C37	SEP	0.1154	0.6500	7.80	0.0052
D62	SEP	–0.1006	–0.3353	8.35	0.0052
C37	PM	0.1346	0.5948	8.33	0.0039
A06	PM	–0.1228	-0.0737	7.10	0.0077

YR: yellow (stripe) rust; BR: brown (leaf) rust; PM: powdery mildew; FUS: Fusarium head blight; SEP Septoria leaf spot.

This work is currently in progress. As future perspectives, additional *Pst*I/*Mse*I primer combinations will be assayed in the same set of wheat materials. It is known that *Eco*RI and *Pst*I enzymes sample different regions of the wheat genome: *Pst*I is methylation-sensitive and cuts principally in unmethylated DNA regions of the genome, containing expressed and mainly single-copy genes, whereas *Eco*RI is methylation-insensitive and cuts DNA throughout the genome. The final aim is that of assembling a database of DNA polymorphisms for the durum and bread wheat germplasm. This information is potentially useful for tracing single pure lines through genetic fingerprints and also for planning experimental crosses between pure lines on the basis of their genetic distances. Moreover, selected AFLP markers potentially linked to resistance genes will be converted into sequence-tagged site markers, such as sequence characterized amplified region (SCAR) or cleaved amplified polymorphic sequence (CAPS) markers, and then assayed in experimental populations segregating for the resistance traits in order to test their suitability and reliability for use in wheat MAS programs.

10.5.4 Effects of Different Conservation Strategies on the Population Genetic Structure of Maize Landraces as Assessed with Molecular Markers

The on farm, in situ and ex situ conservation methods may exert a different influence on the genetic structure of populations grown by farmers. This influence should be accurately evaluated to avoid genetic erosion and conservation programme failure. In fact, the loss of genetic diversity could be due to inbreeding that can result from drift and migration to natural and human selection and gene flow. Each of these factors has a different relative importance on the types of conservation methods. Molecular markers were used to investigate the influence of the conservation strategy on the genetic structure of farmer populations grown for 2 years with three different methods: (1) on farm conservation by farmers, using own seeds and traditional agronomic practices; (2) in situ conservation in the original area but taking into account the spatial isolation from other fields cultivated with hybrid varieties; and (3) ex situ conservation far from the original area with no gene flow due to the total absence of fields grown with the same crop (Pallottini 2002). Statistical tests failed to reveal any significant difference in terms of diversity/similarity absolute values among the populations conserved according to the three distinct strategies (Fig. 10.5).

Nevertheless, about 10% of the comparisons performed for the marker allele frequency parameter at the total assayed loci showed significant differences. Even the differences between genetic variation parameters computed for mapped and random marker loci were significant. In particular, some marker loci were more affected than others by changes of the marker allele frequency depending on the conservation method. These markers, distributed throughout the genome, could be related to important genes involved in the adaptation to environmental conditions or responsible for traits evaluated in the selection by farmers.

BookID 104577_ChapID 10_Proof# 1 - 11/07/2009

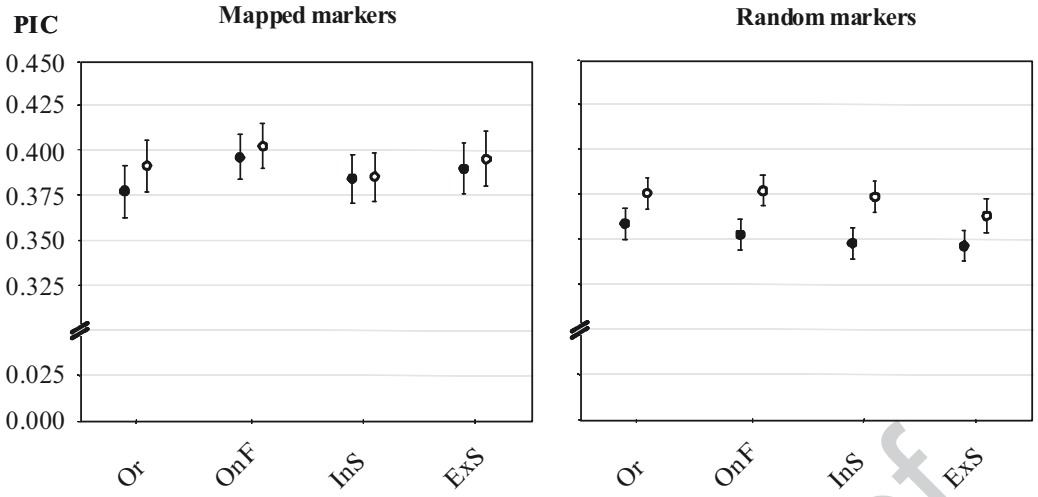

Fig. 10.5 PIC values and standard errors computed in the original population and in the populations obtained from on farm, in situ and ex situ conservation strategies, using two sets of mapped and random markers

In sum, although all conservation methods studied have determined the significant changes to the genetic structure of the farmer populations, the genetic variation and diversification that occurred with ex situ conservation was much stronger than that observed for in situ and on-farm conservation. It is worth mentioning that to monitor these changes, the level at which the investigation is performed is essential. When the mean values of the more common genetic diversity and/or similarity indexes are taken into account, no significant differences are highlighted because of the large set of molecular data and the occurrence of bidirectional changes of marker allele frequencies over all marker loci. Consequently, variation of the marker allele frequency has to be computed and interpreted at each single marker locus or between pairs of marker loci, but not on the whole molecular marker data set. We identified chromosome blocks in linkage groups 2, 3, 9 and 10 carrying marker loci affected by strong LD that could be associated to genes influenced by selection and that could have played a role in the adaptation to the different environmental situations (Barcaccia et al. unpublished results).

10.6 Using Molecular Characterization to Make Informed Decisions on the Conservation of Crop Genetic Resources

Information about the genetic make-up of gene bank accessions contributes towards decision-making for conservation activities, which range from collecting and managing genetic resources to identifying genes to add them value for breeding purposes.

Well-informed sampling strategies for germplasm material destined for ex situ conservation and designation of priority sites for in situ conservation are both

crucial for successful conservation efforts. In turn, defining strategies depends on knowledge of location, distribution and extent of genetic diversity. Molecular marker-based characterization, by itself or in combination with other data, such as morpho-phenological traits, provides reliable information for assessing, among other factors, the amount of genetic diversity/similarity, the structure of genetic variation in samples and populations, rates of genetic differentiation among populations and the distribution of biodiversity in populations from different locations.

Molecular characterization helps to determine the breeding behaviour of species, adaptive and reproductive success of individuals and the existence of gene flow among individuals, that is, the movement of alleles within and between populations of the same or related species, and its consequences (Papa and Gepts 2003). Molecular data improve or even allow the elucidation of phylogeny, and provide the basic knowledge for understanding taxonomy, domestication and evolution. As a result, information from DNA markers or DNA sequences offers a good basis for better conservation approaches.

Selected molecular technologies render cost-effective and comprehensive genotype profiles and gene haplotypes of accessions generated through DNA fingerprinting and DNA sequencing that may be used to establish the identity of the material under study. Simultaneously, in addition to the presence of redundant materials or duplicated accessions, these technologies can detect contaminants, and in the case of mixtures, contamination with introgressed genes from other accessions or commercial varieties as well. Moreover, molecular marker data sets provide the baseline for monitoring natural changes in the genetic structure of the accessions as well as changes occurring as a result of human intervention (e.g., seed regeneration or sampling for replanting in the field). Whatever the case, analysis of molecular information allows the design of strategies for either purging the consequences of inappropriate procedures or amending them to prevent future inconveniences (De Vicente et al. 2006).

In the next years, the area of crop plant genomics that might show the greatest development with respect to the use of molecular marker technology is likely that of applied breeding programs.

An increasingly important role of genetic characterization is identifying useful genes in germplasm, that is, maximizing conservation efforts. Because the major justification for the existence of germplasm collections is for the use of the conserved accessions, it is important to identify the valuable genes that can help to breed new varieties able to meet the challenges of current and future agriculture (De Vicente et al. 2006). Characterization has benefited from several approaches resulting from advances in molecular genetics such as Mendelian gene tagging and QTL mapping. Research in this field has led to the acknowledgement of the value of wild relatives, in which modern techniques have discovered useful variation that could contribute to varietal improvement. Knowledge of molecular information in major crops and species, and of the synteny of genomes and colinearity of genes, has also opened up perspectives for identifying important agronomic genes or functional variants in other crop types, particularly those receiving little attention from formal research.

BookID 104577_ChapID 10_Proof# 1 - 11/07/2009

Managing biodiversity means not only genetic characterization through DNA polymorphism detection, as it requires information used to address key issues of both ex situ and in situ plant germplasm management and to assist in the process of decision making (Barcaccia and Lanteri 2006). For ex situ crop germplasm maintenance, molecular tools may contribute to the sampling, management and development of core collections as well as the utilization of genetic diversity. For the in situ and on farm preservation strategies of genetic resources, molecular markers might help in the recognition of the most representative populations within the gene pool of a landrace and the identification of the most suitable strategies for their managing and use.

Analysis of genomic DNA samples based on the detection of molecular markers should be capable of identifying plant varieties unambiguously and definitive and also be effective for calculating the genetic distance between plant germplasm accessions. The ability to identify new varieties and determine their diversity with respect to previously registered varieties promises to be the prime requirement for a valuable market of plant varieties, as well as a guaranteed genetic value of plant materials.

New molecular approaches have been recently developed for adapting the current PCR-based techniques to target functional diversity. New findings from genomics research indicate that there is a tremendous genetic potential locked up in germplasm collections that can be released only by shifting the paradigm from searching for phenotypes to searching for superior genes with the aid of molecular linkage maps (Barcaccia and Lanteri 2006). At present the increasing information available from genome scanning and gene mapping means that molecular markers known to be tightly linked to traits of agronomic interest can be better addressed for characterizing genetic diversity and help in identifying genetic variation of use to breeders. Furthermore, the identification of Mendelian genes and QTLs controlling a given qualitative or quantitative trait and the availability of their DNA sequences may facilitate the classification of variation in germplasm pools. High resolution genetic maps and linkage groups enable markers closely linked to agronomically important traits to be used and the increasing numbers of SNPs and ESTs for genes and transcripts, respectively; provide routes for more targeted sequence-based approaches. Classification of the sequence variants at a target locus would substantially reduce the amount of work needed to assess their potential for breeding and lead to the identification of superior haplotypes, alleles and/or genotypes. The information acquired is now being exploited to transfer different traits, including biotic stress resistances and improved quality traits, to important varieties by means of MAS programs.

Germplasm in collections can undergo structural molecular characterization (i.e., based on the investigation of anonymous DNA sequences) and functional molecular characterization (i.e., based on the identification of genes and their functions). The information gathered from structural characterization not only provides increased clarity on existing genetic diversity and its organization in individuals, but it also useful to determine the structure of populations providing the basis for functional characterization.

The increasing number of genome and transcriptome sequencing projects has opened the opportunity to design functional molecular markers on expressed sequences of known chromosome position for characterizing and exploring genetic resources. Moreover, this information enables the compilation of large amounts of sequence data that can be used to develop markers linked to specific genes and to discover novel functional variations. In addition, the development of technologies continues and this means, on one side, increased markers and, on the other, decreased costs, so to allow their application in the tasks of characterizing and preserving plant genetic resources, which usually involve large numbers of samples.

New developments are also taking place in designing better approaches to access new and useful genetic variation in collections, such as allele mining and association mapping studies. Allele mining focuses on the detection of allelic variation in important genes and/or traits within a germplasm collection. If the targeted DNA is available, either a gene of known function or a given sequence of unknown origin, then the allelic variation in a collection, usually due to point mutations, can be successfully identified. Association mapping studies of natural populations are an alternative to segregation analysis in experimental populations for identifying useful genes by correlation of molecular markers to a specific phenotype. These studies can be performed on a germplasm collection and also on other breeding materials as long as significant linkage disequilibrium (LD) exists.

The importance of the variation captured in genetic resources in allowing evolution and/or facilitating plant breeding has been long recognized. However, appreciating the variation held in collections is not sufficient. Conservation of genetic resources needs to be combined with an enhanced use of conserved materials. Worldwide germplasm collections of crop plant species maintained ex situ in gene banks together with that held in situ and on farm situations harbour abundant quantities of hidden allelic variants. The challenge is to unravel the mysteries of this variation so that it can be used for the benefit of humankind. More and more, technologies have increased throughputs, which generally means the generation of progressively larger amounts of genetic data. Genotyping individuals to identify the available allelic variation that makes up the phenotypes provide the groundwork on which genetic resources can be used in plant breeding. Phenotyping is very much linked to the usefulness of good molecular characterization, together forming the basis of progress in modern genomics research in crop plants (De Vicente et al. 2006).

10.7 Conclusion

In conclusion, the most important challenges in the near future are certainly the molecular characterization of germplasm collections for preserving them from genetic erosion and the identification of phenotypic variants potentially useful for breeding new varieties. Knowing the presence of useful traits, genes and alleles would help in making decisions on the multiplication of accessions and the maintenance of seed stocks for responding to an expected higher demand of materials.

Use of molecular markers for characterization and conservation of genetic resources should be implemented so that genotypes with known and useful genes and alleles can be added to core collections to make them exploitable by breeders. This will facilitate the use of, and add value to, crop plant germplasm resources. A new concept that might be successful is that of building collections primarily based on the knowledge of the presence of valuable genes and traits.

[AU2] ## References

Barcaccia G, Albertini E, Torricelli R et al. (1998) Molecular characterization of lentil and emmer landraces. National Conference of Biodiversity, Alghero, Sassari, Italy, 8–11 September, pp 841–846

[AU3] Barcaccia G, Lucchin M, Parrini P (2003) Characterization of a flint maize (*Zea mays* var. *indurata*) Italian landrace: II. Genetic diversity and relatedness assessed by PCR-based markers. Genet Resour Crop Evol 50:253–271

Barcaccia G, Pallottini L, Soattin M et al. (2003) Genomic DNA fingerprints as a tool for identifying cultivated types of red chicory (*Cichorium intybus* L.) from Veneto, Italy. Plant Breed 122:178–183

Barcaccia G, Albertini E, Pallottini L et al. (2005) Fingerprinting wheat varieties for multiple breeding purposes. Proceedings of the XLIX national conference of agriculture genetics, Potenza, 12–15 September, C.06

Barcaccia G, 0Pallottini L, Parrini P et al. (2006) A genetic linkage map of a flint maize (*Zea mays* var. *indurata* L.) Italian landrace using a one-way pseudo-testcross strategy and multilocus PCR-based markers. Maydica 51:469–480

De Vicente MC, Guzmán FA, Engels J et al. (2006) Genetic characterization and its use in decision-making forthe conservation of crop germplasm. In: Ruane J Sonnino A (eds) The role of biotechnology in exploring and protecting agricultural genetic resources, FAO, Rome. pp 129–138

Dice LR (1945) Measures of the amount of ecological association between species. Ecology, 26:297–302. Proceedings of the international workshop. The role of biotechnology, Villa Gualino, Turin, Italy, 5–7 March 2005. pp 129–138

Ferreira ME (2006) Molecular analysis of gene banks for sustainable conservation and increased use of crop genetic resources. In: Ruane J, Sonnino A (eds) The role of biotechnology in exploring and protecting agricultural genetic resources, FAO, Rome. pp 121–127

Lanteri S, Barcaccia G (2006) Molecular markers based analysis for crop germplasm preservation. In: Ruane J, Sonnino A (eds) The role of biotechnology in exploring and protecting agricultural genetic resources, FAO, Rome. pp 55–66

Morgante M, Olivieri AM (1993) PCR-amplified microsatellites as markers in plant genetics. Plant J 3:175–182

Negri V, Tosti N, Falcinelli M et al. (2000) Characterization of thirteen cowpea landraces from Umbria (Italy). Strategy for their conservation and promotion. Genet Resour Crop Evol 47:141–146

Nei M (1973) Analysis of gene diversity in subdivided populations. Proc Natl Acad Sci U S A 70:3321–3323

Nei M (1978) Estimation of average heterozigosity and genetic distance from a small number of individuals. Genetics, 89:583–590

Nei M, Li WH (1979) Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc Natl Acad Sci U S A 76:5269–5273

Pallottini L, Garcia E, Kami J et al. (2004) The genetic identity of a patented yellow bean. Crop Sci 44:968–977

835 Pallottini L (2002) Variation of the genetic structure of populations in relation to different
836 strategies of germplasm conservation (*on farm*, *in situ* and *ex situ*) in local varieties of maize:
837 the case study of the "Nostrano di Storo" (*Zea mays* var. *indurate* L.).Ph.D. thesis, Padua, Italy,
838 University of Padua, pp 316
839 Papa R, Gepts P (2003) Asymmetry of gene flow and differential geographical structure of
840 molecular diversity in wild and domesticated common bean (*Phaseolus vulgaris* L.) from
841 Mesoamerica. Theor Appl Genet 106:239–250
842 Rafalski A, Morgante M (2004) Corn and humans: recombination and linkage disequilibrium in
843 two genomes of similar size. Trends Genet 20:103–111
844 Vendramin G, Morgante M (2006) Genetic diversity in forest tree populations and conservation:
845 analysis of neutral and adaptive variation. In: Ruane J, Sonnino A (eds) The role of biotechnology
846 in exploring and protecting agricultural genetic resources, FAO, Rome, pp 145–147
847 Vos P, Hogers R, Bleeker M et al (1995) AFLP: a new technique for DNA fingerprinting. Nucleic
848 Acids Res 23:4407–4414
849 Wright S (1965) The interpretation of population structure by F-statistics with special regard to
850 systems of mating. Evolution 19:395–420
851 Wright S (1978) Evolution and the genetics of populations. Variability within and among natural
852 populations. University of Chicago Press, USA, p 4