Common Bean Germplasm Molecular Analysis: A Biotechnological Approach for Breeding

Marta Zulema Galván1* • Sebastián Alberto Stenglein2 • Pedro Alberto Balatti1-3

1 CONICET, Estación Experimental Agropecuaria, INTA, Ruta Nacional 68 Km 172, Cerrillos (4403) Salta, Argentina
2 Laboratorio de Biología Funcional y Biotecnología (Biolab)-CEBB, Facultad de Agronomía de Azul-UNCPBA, CONICET, CC47-7300, Azul, Argentina
3 Centro de Investigaciones de Fitopatología (CIDEFI), Facultad de Ciencias Agrarias y Forestales (UNLP) CC327, (1900) La Plata, Argentina

Corresponding author: * martazgalvan@gmail.com

ABSTRACT

Argentina, which is a major producer of common bean (Phaseolus vulgaris L.), represents the southern most limit of the Andean diversification center of the species. The diverse environmental conditions of these places and human selection favored the development of a great variability of wild beans and landraces, which is endangered due to the destruction of habitats by forest exploitation and agriculture. Information on the variability of these resources is essential to set conservation strategies and design breeding programs aimed at enlarging the genetic base of commercial beans. This work is an overview of the marker-based studies on landraces and wild bean genetic diversity, with special emphasis on Argentinean beans, as a first step for the optimal exploitation of the naturally available bean genetic resources, to generate new traits and improve crop performance. The identification of diversity and hybridization between populations is enhanced by the application of the new tools and the information generated by bean genomic research. Gene flow, which appears to occur fairly frequently in bean, has to be studied in more detail in this region in order to facilitate the transfer of useful alleles from the unexploited germplasm to improved lines, broadening the genetic diversity available for breeding. Some resistance gene analogs (RGAs) have been described within the Andean gene pool and only a few have been functionally characterized or linked to a phenotype. Therefore, a strategy for the exploitation of bean germplasm variability based on the detection of RGAs is also mentioned, though more work should be devoted at identifying these sequences in Andean landraces and wild beans.

Keywords: genetic variability, molecular markers, Phaseolus vulgaris, domestication

INTRODUCTION

Common bean (Phaseolus vulgaris L.) is one of the most important sources of protein for the world, considering that beans are approximately half of the grain legumes consumed worldwide (Broughton et al. 2003).

The center of origin of common bean has been located in the American continent, which is, as a consequence of this, the main source of beans diversity (Harlan 1971). The primary gene pool of P. vulgaris is composed of cultivars and wild populations, including the later the immediate ancestors of common bean cultivars (Burkart and Brücher 1953). The wild relative of cultivated common bean, Phaseolus vulgaris var. aborigineus (Burm.) Baudet, is an herbaceous climbing annual plant that is found in tropical and subtropical areas of America from northern Mexico (Chihuahua) to Argentina (Córdoba) (Gepts et al. 1980; Koemig et al. 1990; Toro et al. 1990; Menéndez Sevillano 2002). Common bean is a noncentric crop with multiple domestication sites throughout Middle and Andean South America (Harlan 1971; Gepts et al. 1986). Regarding this, the intraspecific organization of genetic variation in P. vulgaris suggests that from a nucleus of diversity located in Ecuador and northern Peru (Kami et al. 1995; Tohme et al. 1996) wild beans dispersed both northwards and southwards, consolidating two geographically distinct gene pools: one in Mesoamerica and the other one in the Andes (Gepts 1998) (Fig. 1).

Domestication of common bean from wild relatives took place about 10,000 years ago (Ladizinsky 1998). Domestication is the first genetic bottleneck imposed on wild germplasm by farmers and the early domesticated landraces carry only a subset of the genetic variation found in the wild ancestors (Ladizinsky 1998). In addition to this, considerable changes in the management and systems of bean production occurred threatening the vulnerable in situ
Moreover, the most promising progenies resulting from donations from 59 national germplasm banks and from germplasm bank at CIA T, which were obtained through the southernmost limit of the Andean diversification center of Northwestern Argentina is an area that also represents the ranking fourth among common bean exporting counties. It cultivates an area of approximately 270,000 ha, concentrating in three Northern provinces, Jujuy, Salta and Tucuman. It exports 98% of a total production of 328,000 t, ranking fourth among common bean exporting counties.

Conservation of germplasm that evolved through centuries of cultivation (Debouck and Tohme 1988; Debouck et al. 1993). Like in other parts of the world in Central and South America traditional landraces have been replaced by more profitable or alternative crops leading this to a further erosion of the genetic base of common bean. Due to the superior performance of elite crops over their related wild species, modern plant breeding programs are often based on repeated crossings of a limited number of closely related elite lines (Grandillo et al. 2007). In this way, the genetic variation contained in the unadapted germplasm remains unexploited, leading to varieties which might be more susceptible to pathogens and/or stresses.

In addition to this, the rate at which locally adapted landraces are being lost and at which natural habitats are being damaged is alarming, therefore the international community is investing resources in preserving plant collections in seed banks. Since 1977, CIAT (Centro Internacional de Agricultura Tropical; Cali, Colombia) has been the main center of collection and preservation of germplasm of Phaseolus. Currently, there are over 29,000 domesticated and more than 13,000 wild accessions of P. vulgaris in the germplasm bank at CIAT, which were obtained through donations from 59 national germplasm banks and from collections made at the centers of origin and diversification. Moreover, the most promising progenies resulting from P. vulgaris breeding efforts at CIAT are also included in the accessions bank, which is the main source of materials carrying genes with novel and desirable traits for the generation of improved genotypes in plant breeding programs. Therefore, wild relatives and landraces offer a vast genetic resource that can be used to broaden the genetic base of common bean modern varieties, and efforts should be diverted to collect, multiply and conserve them.

Argentina is a major producer of common beans and cultivates an area of approximately 270,000 ha, concentrated in three Northern provinces, Jujuy, Salta and Tucuman. It exports 98% of a total production of 328,000 t, ranking fourth among common bean exporting counties. Northwestern Argentina is an area that also represents the southernmost limit of the Andean diversification center of common bean and has an active germplasm bank with more than 500 domesticated and wild accessions located at the INTA (Instituto Nacional de Tecnologia Agropecuaria) Experimental Station in the province of Salta.

Here we give an overview of the marker-based studies on landraces and wild bean genetic diversity, with special emphasis on Argentinean beans, as a first step for the optimal exploitation of the naturally available bean genetic resources, to introduce new traits to commercial cultivars, improving their performance. Details about bean genomics and use of genetic diversity from wild germplasm have already been well documented and are therefore not covered here (Broughton et al. 2003; Acosta-Gallegos et al. 2007; Gepts et al. 2008; McClean et al. 2008).

GENETIC VARIABILITY IN THE MAJOR GENE POOLS OF COMMON BEAN

A wide array of molecular marker types has been used to estimate the levels of genetic diversity and its structure in the primary gene pool of common bean. Most of them were designed aimed at understanding the organization of diversity in the species, which is now one of the best known among crop species (Gepts et al. 2008). Molecular markers have been used to (i) identify and confirm the two major gene pools of common bean (Andean and Mesoamerican); (ii) analyze the effect of domestication on genetic diversity; (iii) identify and confirm ecogeographic races among domesticated beans in each of the two gene pools; (iv) identify a putative ancient gene pool in Ecuador and northern Peru; (v) assess the relative levels of genetic diversity in the Andean and Mesoamerican wild and domesticated gene pools; and (vi) document the importance of gene flow from domesticated to wild populations (Acosta-Gallegos et al. 2007) (for references see below).

Phaseolin, the major storage protein of common bean seeds (49% of the seed protein), has been used to trace the evolutionary origin of beans, showing that wild beans, whether belonging to the Mesoamerican or the Andean gene pool, hold a higher level of genetic diversity than cultivated ones (Gepts et al. 1986; Gepts and Bliss 1988; Gepts 1998). Isozyme and phaseolin have been used as tools to analyze common bean evolution and also the effect of domestication upon diversity (Gepts et al. 1986; Koenig and Gepts 1989; Koenig et al. 1990; Singh et al. 1991a; Becerra-Velasquez and Gepts 1994; Santalla et al. 2002). The finding that phaseolin type and isozyme pattern of cultivated beans were shared by the wild beans growing in the same area (Gepts et al. 1986; Koenig et al. 1990) suggested multiple domestication events. Distinct wild populations were domesticated along a vast geographical range, leading to the formation of two gene pools of cultivated beans, too (Koenig and Gepts 1989; Singh et al. 1991a). However, phaseolin is coded by a single complex locus (Emani and Hall 2008) and its lack of polymorphism in the domesticated gene pool prevented the detection of more subtle genetic differences between closely related landraces or cultivars (Gepts et al. 1997).

DNA markers revealed higher levels of polymorphism than phaseolin or isoenzymes and they have been major tools that contribute to the understanding of the structure of genetic diversity in common bean. Several molecular markers confirmed the existence of two major gene pools of origin in wild and domesticated beans, like RFLP (Restriction Fragment Length Polymorphism) (Becerra-Velasquez and Gepts 1994; Hammam et al. 1995), M13-related and chloroplast DNA sequences (Khairallah et al. 1990; Sonnante et al. 1994), RAPDs (Haley et al. 1994; Fofana et al. 1997; Johns et al. 1997; Duarte et al. 1999; Métais et al. 2000; Franco et al. 2001; Galván et al. 2001; Maciel et al. 2001), AFLP (Tohme et al. 1996; Caicedo et al. 1999; Maciel et al. 2003; Rossi et al. 2009), ISSR (González et al. 1998; Galván et al. 2003; Serna et al. 2003) and SSR (Blair et al. 2006a; Kwak and Gepts 2009; Burle et al. 2010; Galván et al. 2010). Differentiation between these two major gene pools of common bean can be assessed. The Americas Journal of Plant Science and Biotechnology
geographic gene pools has been well documented based on both phenotypic and molecular information and suggests that *P. vulgaris* may be undergoing incipient speciation (Kwak and Gepts 2009). This hypothesis was first supported on the observation of partial reproductive isolation between gene pools, including hybrid weakness in the F₁ and later generations (Gepts and Bliss 1985; Koinange and Gepts 1992). Studies with molecular markers supported the existence of a third gene pool of wild beans in Ecuador-Northern Peru (Debouck et al. 1993), which was considered ancestral of the Andean and the Mesoamerican pools mainly because wild populations were not involved in domestication and their phaseolin type has not been found in the domesticated gene pools. Analysis of the sequences of phaseolin genes and chloroplast genomic DNA supported that this unique bean germplasm is actually the primitive ancestor of the species (Kami et al. 1995; Chacón et al. 2005). Studies of wild accessions from northern Peru and Ecuador based on SSR and AFLP markers confirmed that they are positioned between Andean and Mesoamerican beans, though more closely related to Mesoamerican wild types (Kwak and Gepts 2009; Rossi et al. 2009). Considering that this population had a single phaseolin type and lower gene diversity than other wild populations, some authors suggested it may be a relic that represents a fraction of the genetic diversity of the ancestral population (Kwak and Gepts 2009). Alternatively, the reduced genetic diversity may reflect the narrow ecological amplitude of this group on the Pacific slope of the Andes (Debouck et al. 1993). Recent studies of linkage disequilibrium and population structure in wild and domesticated beans using AFLP markers indicate the occurrence of a strong bottleneck in the Andean wild population before domestication, suggesting a Mesoamerican origin of *P. vulgaris* (Rossi et al. 2009).

DNA markers like phaseolin confirmed the reduction of genetic diversity induced by the “domestication bottleneck” in common bean (Gepts et al. 1986; Sonnante et al. 1994; Islam et al. 2004; Papa et al. 2007). Wild beans show marked phenotypic differences compared with their corresponding domesticated forms (Gepts and Debouck 1991). These differences are the result of a constant process of selection performed along thousands of years for adaptation to cultivated environments and were named the domestication syndrome (Koinange et al. 1996).

Several studies suggested independent domestications in the Andean and Mesoamerican region (Gepts et al. 1986; Gepts 1998; Kwak and Gepts 2009; Rossi et al. 2009; Galván et al. 2010). The process of common bean domestication is described in detail elsewhere (Gepts and Debouck 1991). These differences are the result of a constant process of selection performed along thousands of years for adaptation to cultivated environments and were named the domestication syndrome (Koinange et al. 1996).

AFLP markers indicated that a large portion of the genome (16-18%, approx. 100 Mbp) was likely under the effects of selection, probably because of linkage to the loci selected during domestication (Papa et al. 2007). Farmers and breeders selecting for domesticated alleles also might have selected against many other tightly linked genes (Papa et al. 2007). As a result of this, some genome regions mostly affected by domestication harbor much higher levels of genetic variation in wild populations in comparison with the domesticated populations (Papa et al. 2005). Hence, tagging the domestication loci would be useful to identify chromosomal regions that may harbor historically less exploited diversity of the wild germplasm (Papa et al. 2007).

Moreover, AFLP (Amplified Fragment Length Polymorphism) markers showed a marked geographic structure of diversity in the Mesoamerican populations and suggested the occurrence of a single domestication at the state of Jalisco (Papa and Gepts 2003; Kwak and Gepts 2009; Kwak et al. 2009). On the other hand, the Andean gene pool was found to be less geographically structured than the Mesoamerican one (Papa and Gepts 2003; Blair et al. 2006a; Kwak and Gepts 2009; Kwak et al. 2009; Rossi et al. 2009). The number and location of the domestication events in the Andes are still unknown (Fig. 1). However, recent studies based on AFLP data support the occurrence of a single domestication event in Mesoamerica, and they also tend to support the same scenario in the Andes (Rossi et al. 2009).

Many marker types provided data for the existence of three Andean (Nueva Granada, Peru and Chile) and four Mesoamerican (Durango, Jalisco, Mesoamerica and Guatemala) races among domesticated beans (phaseolin; Gepts and Bliss 1986; Gepts et al. 1996; Koinange et al. 1990; allozymes: Koenig and Gepts 1989; Singh et al. 1991a; mitochondrial RFLPs: Khairallah et al. 1992; RAPD markers: Beebe et al. 2000; SSR: Blair et al. 2006a; Kwak and Gepts 2009) (Fig. 1). Most races were first identified based on morphological traits (Evans 1973; Singh 1991) breeding behavior (Nienhuis and Singh 1986) and reproductive isolation (Gepts and Bliss 1985; Singh and Gutiérrez 1984; Singh et al. 1991b). In addition, within Mesoamerican races, sub-races were defined on the basis of phenotype and RAPDs (Beebe et al. 2000). However, this is still under debate because SSR markers allowed to distinguish Mesoamerican races but two of them (Diaz and Blair 2006; Kwak and Gepts 2009) and AFLP, which cover a large portion of the genome, were unable to show the existence of races in the Andean pool (Rossi et al. 2009). Further studies should give light to the events responsible of race development in the Andean gene pool. The racial structure of landraces suggests the existence of a level of diversity large enough to be exploited in breeding programs.

One important characteristic of plant populations mostly related with breeding is their reproduction behavior. In spite of the fact that common bean is considered a self-pollinated species, several evidences support the existence of gene flow between wild and domesticated beans (Beebe et al. 1997; Menéndez Sevillaño 2002; Papa and Gepts 2003; Payró et al. 2005; Zizumbo-Villarreal et al. 2005). Weedy populations arise naturally from crosses between landraces and sympatric wild forms (Beebe et al. 1997). Papa and Gepts (2003) using AFLP markers identified Mesoamerican beans that were genetically intermediate between domesticated and wild types. In addition, the authors found that gene flow occurs predominantly from domesticated to wild types, event that might lead to a reduction in genetic diversity in wild Mesoamerican populations (Papa and Gepts 2003; Papa et al. 2005).

Gene flow between different bean populations suggests that interbreeding complexes of wild, weedy and cultivated types may be important mechanisms to generate variability in beans. Complexes might broaden the genetic base of some domesticated populations (Beebe et al. 1997). Most races were first identified based on morphological traits (Evans 1973; Singh 1991b) breeding behavior (Nienhuis and Singh 1986) and reproductive isolation (Gepts and Bliss 1985; Singh and Gutiérrez 1984; Singh et al. 1991b). In addition, within Mesoamerican races, sub-races were defined on the basis of phenotype and RAPDs (Beebe et al. 2000). However, this is still under debate because SSR markers allowed to distinguish Mesoamerican races but two of them (Diaz and Blair 2006; Kwak and Gepts 2009) and AFLP, which cover a large portion of the genome, were unable to show the existence of races in the Andean pool (Rossi et al. 2009). Further studies should give light to the events responsible of race development in the Andean gene pool. The racial structure of landraces suggests the existence of a level of diversity large enough to be exploited in breeding programs.

**INTEGRATION OF GENETIC AND GENOMIC TOOLS FOR BEAN GERMPLASM EXPLOITATION**

Genomic research is generating new tools, such as functional molecular markers, which supported with bioinformatics, and new knowledge about statistics and inheritance phenomena, could increase the efficiency and precision of crop improvement (Varshney et al. 2005). As a result, genome sequencing and extensive genome-wide marker development provide a platform for the more efficient improvement of common bean (Broughton et al. 2003). In the last few years, the Phaseomics community (Phaseolus Genomics, www.phaseolus.net) has laid a solid foundation towards sequencing the common bean genome by developing genomic resources such as an impressive collection of germplasm and genetic stocks, mapping populations, BAC libraries, EST, SNP and TILLING assembly, and the development of tools such as FISH and transformation (reviewed by Gepts et al. 2008; Fonseca et al. 2010; Hyten et al. 2010).

Additionally great efforts have been made and are under
way to use wild ancestors and to introgress some of the diversity that was lost during domestication, in order to improve bean agricultural yields under optimal as well as stress conditions (reviewed by Acosta-Gallegos et al. 2007). Here we describe some strategies designed to exploit bean germplasm variability based on the integration of genetic and genomic tools, which vary according to the objectives of the breeding program.

Advanced backcross QTL analysis

The limited use of exotic genetic resources for the improvement of quantitative traits can be explained by the fact that the transfer of traits from unadapted germplasm, that carries many undesirable genes, into elite lines is a laborious process, that requires an efficient selection procedure and many generations of backcrossing to the adapted parent in order to recover most of the desirable agronomic traits, which might not always be successful (Grandillo et al. 2007). Advanced backcross QTL analysis (AB-QTL) is a breeding method, proposed by Tanksley and Nelson (1996), that integrates the process of QTL discovery with variety development, by identifying and transferring useful QTL alleles from unadapted (e.g., landraces, wild species) to elite germplasm, thus broadening the genetic diversity available for breeding. The AB-QTL strategy differs from other QTL mapping methods in that the molecular marker and phenotypic analyses are delayed until advanced generations, like BC2 or BC3, where the frequency of the donor-parent genome is reduced and the segregating population resembles the recurrent parent of the cross (Grandillo et al. 2007).

The AB-QTL analysis was used to identify QTL loci for agronomic performance including seed yield in a BC2 F5 population from the cross of a large-seeded Andean cultivar, 'ICA Cerinza', and a wild common bean accession from Colombia G244404 (Blair et al. 2006b). One hundred fifty-seven lines were evaluated at three locations in Colombia and a genetic map based on microsatellites, sequence characterized amplified region (SCAR) and phaseolin markers, was used. As a result, 41 significant QTL for eight traits were identified, including five for seed weight, two for flowering time, and one for yield, which were consistent across environments. The wild accession contributed positive alleles for yield and other traits to the introgressed lines showing the advantages that advanced backcrossing has in common bean improvement (Blair et al. 2006b).

Association mapping based on linkage disequilibrium

Association mapping is an alternative to linkage analysis that uses the natural sequence diversity within a species to define the various loci controlling a complex trait (Jorde 2000; Mackay 2001). The primary goal of association mapping is to establish, based on linkage disequilibrium (LD), the correlations between genotypes and phenotypes in a sample of individuals. Biparental populations such as doubled haploid (DHs), F2 or recombinant inbred lines (RILs) have been widely used to construct molecular marker maps and to identify genes or QTLs for traits of interest. However, these mapping populations are the products of just one or a few cycles of meiotic recombination, limiting the resolution of genetic maps, and are often not representative of germplasm that is actively used in breeding programs. By contrast, the use of unrelated genotypes or natural populations in association mapping might improve the resolution for identifying genes contributing to a quantitative trait (Varshney et al. 2005). Details about this approach have been reviewed in Gupta et al. (2005).

Understanding sequence and marker variation is important to apply association mapping. For common bean, more than 150,000 DNA sequences are deposited in GenBank (http://www.ncbi.nlm.nih.gov). The vast majority are expressed sequence tag (EST) sequences from different tissues and genotypes (NCBI 2010). By sampling different genotypes polymorphisms such as SNPs and insertion/deletions (indels) can be identified and characterized, what might help to define the level of sequence diversity within the species (Galeano et al. 2009). Linkage disequilibrium and population structure studies based on these sequence and marker analyses are forming a foundation on which association mapping can now be applied to common bean (Gepts et al. 2008; Hanai et al. 2009; Kwak and Gepts 2009; Rossi et al. 2009).

CHARACTERIZATION OF ARGENTINEAN COMMON BEAN GENETIC RESOURCES

Landraces and wild beans diversity and population structure

Northwestern Argentina represents together with Bolivia the Southern limit of the Andean Center of distribution of wild common bean (Phaseolus vulgaris var. aborigineus) and probably an area of domestication of this species (Gepts et al. 1986; Koenig and Gepts 1989; Gepts and Debouch 1991; Singh et al. 1991b; Beebe et al. 2001; Islam et al. 2002).

Bean cultivation started in Argentina with the Spanish immigration. New settlers began to grow white bean types (e.g. ‘Alubia’), in the provinces of Salta and Tucumán in the northwestern region of the country, which were mainly exported to Spain and other European countries (Santalla et al. 2002). Argentinean domesticated germplasm include mostly bean populations of races Nueva Granada and Peru. However, race Chile and Mesoamerica have also been found (Singh et al. 1991b) which could be due to a germplasm exchange in pre-Columbian times (Kaplan and Kaplan 1988).

A bean-breeding program started, in 1984, at two INTA Experimental Stations, located in Cerrillos (Salta) and Leales (Tucumán). The project committed in the introduction of different types of bean cultivars, derived from international breeding programs carried on at CIAT (Centro Internacional de Agricultura Tropical) and EMBRAPA (Empresa Brasileira de Pesquisa Agropecuaria). Bean improvement was aimed at developing cultivars adapted to the agroecological characteristics of the area of bean production in Argentina. The cultivars included in the program were mostly black and white beans, although other types like Navy, Cranberry, Carioica and Light Red Kidney beans were also included. The analysis with RAPD markers of the most frequently cultivated beans developed in the northwestern region of the country revealed low levels of polymorphism around a narrow genetic base of the gene pool available for breeding (Galván et al. 2001). In spite of this, cultivars were grouped in direct relation with their pedigrees, gene pool of origin and phenotypic characteristics (e.g. growth habit, size color and shape of the seed), results that were confirmed by means of ISSR, which turn out to be better tools than RAPD markers to identify beans by gene pool of origin (Galván et al. 2003).

In Argentinean germplasm of wild beans grown in Argentina in a large area extended through the provinces of Jujuy, Salta, Tucumán, San Luis and Córdoba (Burrart and Brücher 1953) along the valleys of the Eastern Andean Mountain Range (altitude 700–3000 masl, latitude 22–27° S, longitude 63–66° W) (Menéndez Sevillano 2002; Santalla et al. 2004) (Fig. 2) The diverse environmental conditions of these places might have favored the development of landraces and wild populations of common bean suggesting the existence of great variability, which is endangered due to the destruction of natural habitats in order to increase the amount of land devoted to forest exploitation and agriculture.

Since the 1970s different collecting missions were carried out in this region (Menéndez Sevillano 2002), not only to collect, document, study and multiply wild beans and landraces but also to increase the availability of genetic diversity in gene banks. There are no major barriers between wild and domesticated beans, since intercrossings
generate fully fertile hybrids (Singh et al. 1995; De Ron et al. 2004). As a result, these resources remained important for the genetic improvement of domesticated types.

Many authors reported high levels of morphological diversity in Argentinean landraces and wild populations (Menéndez Sevillano 2002; De Ron et al. 2004; Santalla et al. 2004; Galván et al. 2006), suggesting that the Andean gene pool has a large genetic base in this region.

Andean wild beans produce small seeds with mottled seed coat patterns, which is considered a phylogenetically ancient feature with a marked selection advantage (Brücher 1988).

Domestication and subsequent evolution led to populations with a wide variety of flower and seed colors and patterns, some of which are not observed among wild beans such as the lilac and white flower colors (Santalla et al. 2004) (Fig. 3). Moreover, with the harvesting of bean seed under cultivation a greater range of colors became established and mottled seeds persist only in a modified form in the cranberry primitive landraces (Santalla et al. 2004).

Andean landraces evolved through long periods of traditional cultivation for human consumption in small isolated villages, been grown and marketed as mixtures, which can be separate into different types according to seed color and pattern (Kaplan 1981; Freyre et al. 1996). Ecological and human factors accounted for the existence and stability of bean mixtures that could guarantee the flexibility of the varieties and their capacity to adapt to environmental variation (De Ron et al. 2004). Andean farmers most probably have been avid plant selectors and maintained great levels of variability in the domesticated materials commonly used for consumption, as dry seed and fresh pod and for children toy and aesthetic use (Menéndez Sevillano 2002; De Ron et al. 2004). These uses explain the existence of large and extra-large seed size, variation in traits related to pod size, and diversity in seed color and pattern in this region (De Ron et al. 2004).

In addition to morphological variability, phaseolin and allozyme variation also revealed a large genetic base in Argentinean bean landraces which could suggest that domestication occurred within a diverse genetic wild structure (Cattan-Toupance et al. 1998; De Ron et al. 2004; Santalla et al. 2004; Galván 2007; Galván et al. 2010). Initially, biochemical markers suggested that gene flow between wild and domesticated beans appears to have been limited and has not appreciably modified the organization of the domesticated gene pool in northern Argentina (Santalla et al. 2004). However, different results were found using RAPD and ISSR markers (Fig. 4) (Galván et al. 2006; Galván 2007). Cluster analysis and principal coordinates analysis using molecular data grouped wild beans and landraces in two highly similar clusters, also showing a great variability within each cluster (Fig. 5). The fact that some of the wild beans analyzed were interspersed among landrace materials supports the idea of the existence of gene flow between wild and domesticated beans. This pattern has also been observed before in Mexico (Papa and Gepts 2003) and suggests that in these regions gene flow may occur predominantly from domesticated to wild populations. This asymmetric gene flow leads to populations that are phenotypically wild but include (presumably neutral) markers that originated from domesticated populations, explaining the inclusion of these wild accessions in a predominantly domesticated cluster (Galván 2007; Galván et al. 2010).

The topography of Northwestern Argentina contributes to the geographic isolation of wild populations, leading this to the maintenance of distinct ecotypes. Wild beans genetic variability, analyzed by means of molecular markers, was associated with population distribution, revealing a geographic structure of diversity in wild types. A certain degree of outcrossing and gene flow among neighboring populations might have helped to maintain high levels of intra-population variability (Galván et al. 2006; Galván 2007).

The presence of wild-weedy-cultivated complexes was reported in this region based on morphological data (Menéndez Sevillano 2002; De Ron et al. 2004). Indeed, there is evidence of variability in the reproductive system of domesticated and wild beans accessions, which, in certain areas, may have led to gene flow between sympatric populations (Ibarra-Pérez et al. 1977; Hoc and García 1999; Hoc et al. 2006). In Santa Victoria, in the province of Salta (22°15′S latitude and 24° 58′W longitude), wild beans and landraces showed high genetic similarity based on ISSR markers (Fig. 5) (Galván 2007). Furthermore, some wild characters, such as small seeds, dehiscent pods and mottled seed color patterns, were present in the offspring of one landrace suggesting the occurrence of hybridization events between landraces and wild beans (Galván 2007).

Selection appears to be a major evolutionary factor maintaining the identity of sympatric wild and domesticated populations (Papa and Gepts 2003). In Santa Victoria local farmers cultivate primitive landraces along with wild beans.
collected in inaccessible areas, such as the humid forests of Baritú National Park (Menéndez Sevillano 2002). Intermediate forms between wild and domesticated beans have been found in this area (Menéndez Sevillano 2002; Santalla et al. 2004; Hoc et al. 2006), suggesting that human intervention might enhance gene flow between them leading to an increase in the genetic variability of primitive landraces. This might allow adaptation to environmental variation and also generate new combinations in seed patterns.

Morphological and molecular characterization of the wild-weedy-cultivated complexes may allow them to function as pre-breeding trials, in which exotic wild germplasm is gradually introgressed into types with higher levels of domestication. As mentioned before, the introgression of traits from wild or weedy germ plasm into cultivated beans in modern breeding programs is difficult due to the prevalence of non-domesticated traits coded by dominant genes (Gepts and Debouck 1991; Singh 1991b). The resulting
introggression would complement modern breeding programs in the exploitation of genetic variability (Beebe et al. 2000).

All the findings described here might help to develop a plan for bean germplasm in situ and ex situ conservation in order to provide sources of diversity to breeding programs for the generation of improved bean cultivars. This is obviously important since it has been particularly difficult to improve the yield potential of common beans (Whitie et al. 1992; Kornegay et al. 1992). Further analysis of the variability and fertility of the hybrids generated between wild populations and landraces and a more detailed sampling should be planned to determine the extent of the hybridization process and its influence on the evolution of bean landraces from Northwestern Argentina.

**Disease resistance related sequences in Argentinean wild beans**

Wild beans are a potential source of novel alleles that can be exploited to improve yield and other quantitative traits of domesticated beans, since they must possess a wide array of adaptive traits to be able to colonize many diverse ecological niches including semiarid areas to moist tropical environments (Acosta-Gallegos et al. 2007). The current status of efforts to transfer useful traits from wild species into common bean has been reviewed by Acosta-Gallegos et al. (2007).

Diseases are, among several other traits, one of the most important factors that provoke reductions in yields in most bean producing areas (Beebe and Pastor-Corrales 1991). Although more than 200 pathogens have been reported for beans, only some of them such as angular leaf spot (causal organism *Pseudocercospora griseola*), anthracnose (*Colletotrichum lindemuthianum*), Bean common mosaic virus (BCMv), Bean golden mosaic virus (BGMv) and rust (*Uromyces appendiculatus*) have been reported to cause considerable economic losses.

Since wild beans have coexisted with pests and pathogens on an evolutionary time scale, they have developed resistances to different plant threats (Acosta-Gallegos et al. 2007) and should be a good source of resistance genes. For example, resistance to bruchids (*Zabrotes subfasciatus*) which is conferred by the seed protein arcelin, and was absent in thousands of domesticated common bean populations (Schoonhoven and Cardona 1982), was found only in a few highly resistant individuals in a wild bean population from Mexico (Acosta-Gallegos et al. 1998).

Plants may be tolerant and/or resistant to pathogens and this is related with the genes involved in the response. Tolerance is often a quantitative trait determined by several plant genes, in this case the pathogen infects and damages the plant but, disease severity remains at low levels, resulting in undetectable reductions in yield. On the other hand, plant resistance to pathogens is determined by the presence of resistance (*R*) genes, response described by Flor (1955) as the gene for gene resistance hypothesis. Avirulence, encoded in the genome of the invading pathogens, code for a product that is recognized by a receptor protein, which is encoded by an *R* gene of the host (Scofield et al. 1996). Several genes from different species have been cloned and their analysis revealed remarkable similarities in their structure. *R* genes belong to a large class of genes that code for large proteins that participate in protein-protein interactions and signal transduction (Skauskas et al. 1993) and their similarity with respect to conserved motifs such as a nucleotide binding site (NBS) near the N-terminus, and a leucine-rich repeat (LRR) residue near the C-terminus (Bent et al. 1997). NBS have been found to play a key role in activating a cascade of molecular events that lead to disease resistance. By means of the polymerase chain reaction and a pair of degenerate oligonucleotides synthesized based on several conserved motifs of resistance genes, homologues to *R* genes have been amplified and cloned and are known as resistance gene analogs (RGAs) (Kanazin et al. 1996; Leister et al. 1996; Yu et al. 1996).

Based on the sequence of *R* genes, RGAs encoding NBS and LRR motifs have been found in common bean (Geffroy et al. 1999; Rivkin et al. 1999; Ferrier-Cana et al. 2003; López et al. 2003). As a result of correlations made between the *R* gene evolution and the host-pathogen gene-for-gene coevolution processes at the population level, a gene cluster containing *R* genes against *C. lindemuthianum* and RGAs from different gene pools of common bean was identified (Geffroy et al. 1999; Ferrier-Cana et al. 2003). Furthermore, evidences were found supporting a partial resistance against anthracnose and QTL co-localization with “candidate genes” of resistance (Geffroy et al. 2000). RGAs also have been found to reside either close or in some of the *R* gene clusters for angular leaf spot, anthracnose, bean golden yellow mosaic virus, and rust (Rivkin et al. 1999; López et al. 2003).

Despite the abundance of RGA genes in many plant species only a small number of NBS-LRR sequences have been functionally characterized as resistance genes (Zhou et al. 2004; McHale et al. 2006). One of the main limitations to identify *R* genes is probably the lack of strains of a specific pathogen that can be recognized by a disease resistance gene co-located with a candidate NBS-LRR sequence (Geffroy et al. 2008). Moreover, the number of *R* genes that can be detected is determined by the number of different isolates of the pathogen that have been identified and isolated (Bennetzen and Hubert 1992).

In the past few years, angular leaf spot has become one of the most important causes of yield losses of common bean in Argentina (Fig. 6). In order to study the regional genetic diversity of *P. griseola*, the etiological agent of the disease, some collecting missions were carried out in different locations within the main area of common bean production in Argentina since 2001 (Stenglein and Balatti 2006). These studies provided basic information about the diversity of the fungus which also helped to develop cultivars with enhanced tolerance to the disease. Pathotypes and genetic variability were determined based on a set of bean differentials and molecular markers, revealing that *P. griseola* in Argentina displays high pathogenic and genomic diversity (Stenglein and Balatti 2006). The number of haplotypes found based on ISSR and RAPD markers among 45 isolates was 18 and they were unrelated with pathotypes. A recent analysis of 68 isolates from different producing areas in Argentina (27°-25°S latitude and 62°-67° W longitude) using a higher number of ISSR primers, revealed 34 haplotypes and a high level genetic variability not only between Andean and Mesoamerican isolates but also within both groups, suggesting the existence of great diversity of the pathogen in the region (Fermoselle et al. 2007). All these findings are helpful tools to develop breeding strategy to generate new resistant materials.

Also it was found that the wild bean species, *P. vulgaris* var. *aboriginus*, showed a high level of tolerance to most...
of the *P. griseola* pathotypes tested (Stenglein 2007). Previous observations revealed the presence of diversity among epidermal characters that might contribute to plant resistance to pathogens in wild beans from northwestern Argentina (Stenglein et al. 2005). Given the level of tolerance of *P. vulgaris var. aborigineus* to both Andean and Middle American pathotypes of *P. griseola*, and the high levels of tolerance/resistance to *C. lindemuthianum* previously reported (Cattan-Toupance et al. 1998), the presence of this species was considered a good candidate to look for resistance genes.

In the genome of wild common bean fifteen putative RGAs were detected using 4 degenerate PCR primers (Stenglein 2007). These sequences proved to be highly homologous to resistant genes and RGAs from other plant species and common bean as revealed by a BLAST analysis within sequences of the NCBI database. Among 22 putative RGA clones sequenced, 15 were highly homologous to previously isolated R genes falling most of them within sequences found in legumes, which is consistent with the idea that major sequences clades among R genes are family-specific (Meyers et al. 1999; Pan et al. 2000). The sequences had high levels of homology (≥ 80%) to RGAs of the *U*-6, a gene that confers resistance to bean rust, a QTL associated with resistance to anthracnose and angular leaf spot, and to an RGA that partially contributed to resistance against the *Bean golden yellow mosaic virus* (BGYMV) and anthracnose (Stenglein 2007). Further analyses are required in order to reveal a role in plant-microbe interaction for these sequences. However, this work is the first step aiming to the identification of *R* genes in *P. vulgaris* var. *aborigineus*, and demonstrates the potential importance of this wild common bean as a source of resistance genes (Stenglein 2007).

**CONCLUSIONS AND FUTURE PROSPECTS**

Common bean commercial cultivars have a narrow genetic base and it would be desirable to enlarge it by means of the unexploited germplasm that remains in many areas in Northern Argentina. Considering that the rate of germplasm loss is high, efforts should be diverted, as soon as possible at collecting, multiplying and conserving these bean resources. Wild relatives and landraces growing in the southernmost limit of the Andean gene pool offer a vast genetic resource to broaden the genetic base of common bean varieties.

The integration of classical breeding with the new tools and information continuously generated by genomic research will facilitate the transfer of useful alleles from landraces and wild beans to improved lines, broadening the genetic diversity available for breeding.

These new techniques will also allow the screening of larger areas of the bean genome enhancing the identification of diversity and gene flow between populations. This process appears to occur fairly frequently in common bean, but it has to be studied in more detail in this region. Researchers should work in identifying the intensity and direction of gene flow between sympatric populations, and the effect that human intervention has on it, for a more efficient use of this process in breeding programs aimed at obtaining new improved materials.

Nowadays breeders look for cultivars with high and stable yield and resistant to stresses and pathogens. The probability of commercial bean cultivation with these characteristics is high since beans seem to carry several disease resistance genes, so far only some of them and RGAs have been described within the Andean gene pool and only a few have been functionally characterized or linked to a phenotype. Therefore, more work should be devoted at identifying disease resistance genes in Andean landraces and wild beans.

**ACKNOWLEDGEMENTS**

The authors express their gratitude to Dr. A. Lanteri for her invaluable contributions for this research. MZ Galván and SA Stenglein are researchers of CONICET and PA Balatti is a member of CICBA. This work was supported by a research grant from CICBA, UNLP, and CONICET-CSIC.

**REFERENCES**


Duarte JM, Santos JB, Melo LC (1999) Genetic diversity among common bean cultivars from different races based on RAPD markers. *Crop Science* 39, 1497-1507

Duarte JM, Santos JB, Melo LC (1999) Genetic diversity among common bean cultivars from different races based on RAPD markers. *Crop Science* 39, 1497-1507

Evans A (1973) Commentary upon plant architecture and physiological efficiency in the field bean. In: Hall T (Ed) *Potential of Field Beans and Other Food Legumes in Latin America*, CIAT, Cali, Colombia, pp 279-286

The Americas Journal of Plant Science and Biotechnology 4 (Special Issue 1), 60-69 ©2010 Global Science Books


