Chapter 5
Genomics of *Phaseolus* Beans, a Major Source of Dietary Protein and Micronutrients in the Tropics

Paul Gepts, Francisco J.L. Aragão, Everaldo de Barros, Matthew W. Blair, Rosana Brondani, William Broughton, Incoronata Galasso, Gina Hernández, James Kami, Patricia Lariguet, Phillip McClean, Maeli Melotto, Phillip Miklas, Peter Pauls, Andrea Pedrosa-Harand, Timothy Porch, Federico Sánchez, Francesca Sparvoli, and Kangfu Yu

Abstract Common bean is grown and consumed principally in developing countries in Latin America, Africa, and Asia. It is largely a subsistence crop eaten by its producers and, hence, is underestimated in production and commerce statistics. Common bean is a major source of dietary protein, which complements carbohydrate-rich sources such as rice, maize, and cassava. It is also a rich source of minerals, such as iron and zinc, and certain vitamins. Several large germplasm collections have been established, which contain large amounts of genetic diversity, including the five domesticated *Phaseolus* species and wild species, as well as an incipient stock collection. The genealogy and genetic diversity of *P. vulgaris* are among the best known in crop species through the systematic use of molecular markers, from seed proteins and isozymes to simple sequence repeats, and DNA sequences. Common bean exhibits a high level of genetic diversity, compared with other selfing species. A hierarchical organization into gene pools and ecogeographic races has been established. There are over 15 mapping populations that have been established to study the inheritance of agronomic traits in different locations. Most linkage maps have been correlated with the core map established in the BAT93 x Jalo EEP558 cross, which includes several hundreds of markers, including Restriction Fragment Length Polymorphisms, Random Amplified Polymorphic DNA, Amplified Fragment Length Polymorphisms, Short Sequence Repeats, Sequence Tagged Sites, and Target Region Amplification Polymorphisms. Over 30 individual genes for disease resistance and some 30 Quantitative Trait Loci for a broad range of agronomic traits have been tagged. Eleven BAC libraries have been developed in

P. Gepts
University of California, Department of Plant Sciences / MS1, Section of Crop and Ecosystem Sciences, 1 Shields Avenue, Davis, CA 95616-8780, USA.
e-mail: plgepts@ucdavis.edu

P.H. Moore, R. Ming (eds.), *Genomics of Tropical Crop Plants*

genotypes that represent key steps in the evolution before and after domestication of common bean, a unique resource among crops. Fluorescence in situ hybridization provides the first links between chromosomal and genetic maps. A gene index based on some *P. vulgaris* 21,000 expressed sequence tags (ESTs) has been developed. ESTs were developed from different genotypes, organs, and physiological conditions. They resolve currently in some 6,500–6,800 singletons and 2,900 contigs. An additional 20,000 embryonic *P. coccineus* ESTs provides an additional resource. Some 1,500 M_2 Targeting Local Lesions In Genomes populations exist currently. Finally, transformation methods by biolistics and Agrobacterium have been developed, which can be applied for genetic engineering. Root transformation via *A. rhizogenes* is also possible. Thus, the Phaseomics community has laid a solid foundation towards its ultimate goal, namely the sequencing of the *Phaseolus* genome. These genomic resources are a much-needed source of additional markers of known map location for marker-assisted selection and the accelerated improvement of common bean cultivars.

5.1 Introduction

5.1.1 Economic, Agronomic, and Societal Importance of Phaseolus Beans

FAO statistics provide insight into the economic status of common bean (*Phaseolus vulgaris*) throughout the world (http://faostat.fao.org/site/408/default.aspx; verified December 5, 2006). Dry beans were grown on 27.7 million ha in 148 countries in 2004 and total production was 18.7 million metric tons (MT). *Phaseolus* beans are the most important grain legume for direct human consumption. Bean production is more than twice that of chickpea, which is the second most important grain legume. Importantly, eight of the top ten producers are considered to be developing countries. When developmental status is considered further, it is seen that developing countries produce 86% of worldwide production of beans. The main production areas of common bean are Latin America (with Brazil and Mexico as the most important producers) and eastern Africa (where per capita consumption is the highest in the world in countries such as Burundi, Rwanda, and Uganda) (Broughton et al. 2003). Bean production is increasing in some Asian countries such as China and Myanmar, primarily for export purposes (see below). Green bean production amounted to 6.4 million metric tons on 890,000 ha in 2004.

In many of the countries in the world, especially developing countries, common bean is consumed as an important part of the diet and not exported. An extreme example is Brazil, the third leading producer, where less than 0.1% of the 2 million MT produced is exported. The importance of bean to diets in the developing world is reflected in the fact that for developing countries only 13% of production is exported. This contrasts with developed countries, which export 31% of their production. From an economic perspective, dry bean exports generate US$812 million for developing countries, whereas developed countries receive US$460 million

for their exports. The two major exporters are China and Myanmar. For Myanmar, beans (whether *Phaseolus* or other dry bean species) are its most important agricultural export generating US$253 million in 2004. For this very poor country, the bean export income is significant when balanced against their total of only US$2.9 billion in exports in 2004–2005.

Common bean is an important dietary component especially in some developing countries. For some, such as Brazil and Mexico, it is a major source of protein. In some of the least developed countries, such as Burundi, Rwanda, and Uganda, bean provides 40%, 31%, and 15% of the daily intake of total protein, respectively. For other developing countries, Nicaragua, Cuba, and North Korea, the percent total protein derived from bean is 19%, 13%, and 11%, respectively. For a major producer, like Brazil, beans provide 9% of the protein. Qualitatively, bean seed proteins provide sulfur essential amino acids such as lysine, but are poor in methionine, thus complementing cereals in this respect (Bressani 1983; Gepts and Bliss 1985). Not surprisingly, beans are also a major source of calories for residents of these countries. In addition, common bean plays an important role as a source of minerals, especially iron and zinc (Broughton et al. 2003), for which it also complements cereals. Genetic variation for seed content of these minerals has been demonstrated (Islam et al. 2002; Beebe et al. 2000a) and breeding for enhanced mineral content (biofortification) can therefore reasonably be expected (Guzmán-Maldonado et al. 2003).

For those countries where beans are an important crop, yield varies significantly. As expected, on average the yield in developed countries, 1,944 kg/ha, is significantly higher than those in developing countries, 1,035 kg/ha. The situation is more drastic for those least developed countries that depend heavily on beans as a food source. In Burundi, Rwanda, and Uganda the yields are 918 kg/ha, 671 kg/ha, and 638 kg/ha, respectively. Identifying and minimizing yield limiting factors is an ongoing concern for many bean improvement programs. In addition, given the prevalence of bean in these diets, modifying the nutrient content in general to make it a more balanced and nutritious food source is also receiving emphasis.

5.1.2 Phaseolus as an Experimental Organism

The genus *Phaseolus* is a member of the tropical tribe Phaseoleae, which also includes cowpea, pigeon pea, and soybean. The Phaseoleae tribe is part of the Phaseoloid-Millettioid clade, which diverged some 45–50 million years ago from the Hologalegina clade, which contains most temperate crop legumes, such as pea, alfalfa (and *Medicago truncatula*), chickpea, and lentil (Lavin et al. 2005). Synteny between *Phaseolus* and other legumes is negatively correlated with phylogenetic distance. Thus, the highest synteny levels are observed with the genus *Vigna* (cowpea and mung bean), followed by soybean, and distantly, the Hologalegina clade (Boutin et al. 1995; Lee et al. 2001; Yan et al. 2004; Choi et al. 2004; Moffet and Weeden 2006). For example, the region marked by the *Bng122-D0140-Bng*

171-Bng173 markers on linkage group B1 of common bean is syntenic with a region on LG G of soybean. This region harbors a cluster of disease and nematode resistance genes (Freyre et al. 1998; Foster-Hartnett et al. 2002; Kelly et al. 2003).

Phaseolus is a diploid genus with most species having 2n = 2x = 22 chromosomes (some species have 2n = 2x = 20). The genome size of *P. vulgaris* (580 Mbp/haploid genome) is comparable to that of rice (490 Mbp/haploid genome; Bennett and Leitch 2005). In common bean, the levels of duplication and the amount of highly repeated sequences are generally low. Mapping experiments demonstrated that most loci are single copy (Vallejos et al. 1992; Freyre et al. 1998; McClean et al. 2002). Gene families tend to be small, and the traditionally large families such as resistance gene analogs (Rivkin et al. 1999) and protein kinases (Vallad et al. 2001) are of moderate size. Further experiments are needed, however, to confirm these conclusions and compare these results to those of other legumes.

Studies of the evolution of common bean have uncovered several noteworthy features that are of interest to other crop plants. Unique among crop plants, common bean consists of two geographically distinct, evolutionary lineages (Andean and Mesoamerican) that predate domestication and trace back to a common, still extant ancestor located in Ecuador and northern Peru (Debouck et al. 1993; Kami et al. 1995). Patterns of marker diversity and virulence in pathogens and *Rhizobium* parallel those in the bean host, suggesting host-microbe coevolution (Guzmán et al. 1995; Geffroy et al. 1999, 2000; Kelly and Vallejo 2004; Araya et al. 2004; Mkandawire et al. 2004; Aguilar et al. 2004). Geffroy et al. (1999, 2000) have shown that Andean and Mesoamerican resistance specificities appeared in the same, presumably ancestral gene cluster. The inheritance of the domestication syndrome in common bean was the second among all crop plants and the first one in the legumes to be investigated (Koinange et al. 1996). The traits involved, such as growth habit (e.g., determinacy), photoperiod sensitivity and phenology, pod and seed size, and seed color, were not only important in domestication but remain crucial agronomic traits determining farmer and consumer acceptability.

In common bean, sequence variation value based on four genotypes was θ_W (x 10^3) = 4.8 (P. Cregan, personal communication). This compares with θ_W (x 10^3) = 0.97 for *Glycine max* (Zhu et al. 2003), θ_W (x 10^3) = 7.1 for *Arabidopsis thaliana* (Schmid et al. 2005), and θ_W (x 10^3) = 9.6 for *Zea mays* (Tenaillon et al. 2001). The higher polymorphism in common bean relative to soybean is presumably related to its diversification into the two geographic gene pools in the Andes and Mesoamerica, mentioned earlier.

5.2 Bean Genetic Resources

There are a large number of collections of *Phaseolus* germplasm collections, which include the wild and domesticated genotypes of the five domesticated *Phaseolus*

species (*P. vulgaris*: common bean; *P. coccineus*: runner bean; *P. dumosus*: year bean; *P. acutifolius*: tepary bean; *P. lunatus*: lima bean) and other wild *Phaseolus* species. The largest and most diverse is the World Phaseolus collection at the Centro Internacional de Agricultura Tropical (CIAT) in Cali, Colombia (http://isa.ciat.cgiar.org/urg/main.do?language=en). Other large collections are those of the USDA in Pullman, WA, USA (http://www.ars.usda.gov/main/ site_main.htm?modecode=53481500), the Institut für Pflanzengenetik und Kulturpflanzenforschung (IPK) in Gatersleben, Germany (http://fox-serv.ipk-gatersleben.de/), and the Centro Nacional de Recursos Genéticos e Biotecnologia (CENARGEN/ EMBRAPA) in Brasilia (http://www.cenargen.embrapa.br/). There are also national collections as well as a taxonomic collection focused on wild members of the Phaseolinae subtribe at the National Botanic Garden of Belgium (http://www.br.fgov.be/research/collections/living/phaseolus/index.html). Recently, the USDA collection at Pullman established a *Phaseolus* stock collection (http://www.ars.usda.gov//Main/site_main.htm?docid=9065). Together, these collections represent a substantial wealth of genetic diversity that is generally freely available for plant genetic and breeding research.

Several types of populations have been used for linkage mapping in common beans. In terms of population structure, the first molecular maps in common bean were based on first backcross and F_2 generation populations (Vallejos et al. 1992; Nodari et al. 1993a). While easy to develop, the backcross and F_2 populations tended to have few total genotypes and seed supply was limiting, problems that were overcome with the development of recombinant inbred (RI) populations from the initial mapping populations (Freyre et al. 1998; Yu et al. 1998). The BAT93 x Jalo EEP558 RI population has become the core mapping population for common bean because markers from other linkage maps have been mapped in this population and linkage groups from different maps can therefore be correlated (Freyre et al. 1998; Vallejos et al. 2001; Blair et al. 2003). In addition, over 15 other recombinant inbred line populations have been created to map individual or multiple traits (reviewed in Broughton et al. 2003; Kelly et al. 2003). The majority of mapping populations have been produced by crosses across gene pool boundaries (i.e., Mesoamerican x Andean) usually with divergent parents showing contrasting phenotypic characteristics (such as morphology and disease resistance) and high genetic polymorphism (Nodari et al. 1992; Broughton et al. 2003). Relatively few populations have been created from within-gene pools crosses (Andean x Andean or Mesoamerican x Mesoamerican: e.g., Frei et al. 2005; Kolkman and Kelly 2003). Compared to other crop species, genetic mapping in *P. vulgaris* has not utilized many inter-specific crosses for the construction of linkage maps except for the introgression of some disease resistance (Bai et al. 1997). Similarly, very few crosses between wild and domesticated common bean have been used for genetic mapping except for one study of a domesticated x wild cross, which was used to determine the inheritance of the domestication syndrome (Koinange et al. 1996) and the recent use of advanced backcrossing to analyze wild beans for positive yield quantitative trait loci (QTL) (Mauro Herrera 2003; Blair et al. 2006b).

5.3 Marker and Sequence Diversity

As with all plant species, diversity levels and organization of genetic diversity ("structure") were initially estimated with a wide array of molecular marker types. The majority of these marker analyses were designed to understand the organization of diversity in the species, which is now one of the best known among crop species. Based on phaseolin seed storage protein variation and partial reproductive isolation, Gepts and colleagues developed the two-gene pool concept for *P. vulgaris* (Gepts and Bliss 1985, 1986; Gepts et al. 1986; Koenig et al. 1990; Gepts 1990). This result was confirmed by an extensive isozyme analysis (wild: Koenig and Gepts 1989b; domesticated: Singh et al. 1991b) and mtDNA restriction fragment length polymorphisms (RFLP) (Khairallah et al. 1990, 1992). Those analyses also found that the within-gene pool variation was less than that found between gene pools. Allozyme diversity was a key component utilized to define three Mesoamerican (Durango, Jalisco, and Mesoamerica) and three Andean (Nueva Granada, Peru, and Chile) domesticated races in common bean (Singh et al. 1991a). More recently, allozyme data provided important ancestry information regarding the origin of genetic materials from southwestern Europe by demonstrating that the patterns of variation in this region are similar to those found in the Americas (e.g., Santalla et al. 2002). Essentially the allozyme data provided an important hypothesis regarding the origin of common bean and the levels of genetic diversity within the species.

Other marker types have provided data for further analysis of the organization of diversity in common bean. Nuclear RFLP comparisons also supported the two-gene pool concept, but unlike isozyme analyses, the levels of diversity within each gene pool were similar (Becerra-Velásquez and Gepts 1994). Using random amplified polymorphic DNA (RAPD) analysis, Freyre et al. (1998) identified a clear separation between Andean and Mesoamerican gene pools but also detected geographic structure within the Mesoamerican gene pool. A RAPD analysis led Beebe et al. (2000b) to observe subsets of landraces within each Mesoamerican race and define a new Guatemalan race within the Mesoamerican gene pool. These results suggest a level of diversity large enough to distinguish races and subraces. This was not the case for an amplified fragment length polymorphism (AFLP) analysis of Andean genotypes where the genotypes formed essentially a single large pool (Beebe et al 2001). AFLP markers, in contrast, proved to be a very powerful fingerprinting tool to distinguish closely related genotypes belonging to the same commercial class, such as the yellow bean class (Pallottini et al. 2004). AFLP and Inter Short Sequence Repeatmarkers have also proven to be very useful to assess the level and direction of gene flow between wild and domesticated bean populations (Papa and Gepts 2003; Papa et al. 2005; Payró de la Cruz et al. 2005; Zizumbo-Villarreal et al. 2005).

A suite of SSR markers, recently developed for common bean (Yu et al. 1999; Gaitán-Solís et al. 2002; Blair et al. 2003), provide a new tool for diversity analyses. Blair et al. (2006a) measured diversity among 44 genotypes with 129 simple sequence repeats (SSRs), and as expected, they observed two gene pools corresponding to the Mesoamerican and Andean genotypes. It was somewhat surprising that

the Andean genotypes were more diverse than the Mesoamerican lines. Recently, in a study of 172 landraces that represent a broad cross-section of the phenotypic diversity found in the USDA core collection of common bean, greater SSR diversity among the Andean genotypes was also observed (McClean et al. 2006). Applying a model-based approach to population structure analysis (Pritchard et al. 2000), six Mesoamerican and three Andean subpopulations were observed. In addition, a subpopulation, most similar to other Andean genotypes, consisted of landraces collected from throughout the range of the species. These sequence and marker analyses are forming a foundation on which association mapping can now be applied to common bean. Genome sequencing and extensive genome-wide marker development provide new avenues for crop improvement. Principal among these is association mapping, as 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). Because this approach can uncover potential causative single nucleotide polymorphisms (SNPs; Thornsberry et al. 2001) or markers linked to a gene associated with a trait of interest (Hagenblad et al. 2004), understanding sequence and marker variation is important to applying this approach (Nordborg et al. 2005).

For common bean, 27,000 DNA sequences are deposited in GenBank (http:www.ncbi.nlm.nih.gov; verified July 18, 2006). Among these, the vast majority of these are expressed sequence tag (EST) sequences (see section 5.1). EST sequencing projects typically sample not only different tissues but different genotypes. By sampling genotypic differences, and applying stringent screening criteria, polymorphisms such as SNPs and insertion/deletions (indels) can be discovered that define a level of sequence diversity within a species. Ramírez et al. (2005) analyzed contigs derived from over 21,000 ESTs and compared contigs derived from single genotypes. By comparing sequences from Negro Jamapa (Mesoamerican) and G19833 (Andean), they discovered 421 polymorphisms from 196 contigs between these two genotypes. Most (94%) of the polymorphisms were SNPs.

McConnell et al. (2006) compared DNA sequence differences between BAT93 and Jalo EEP558, the parents of the community-wide mapping population (see section 5.2). For their analysis, they considered 322 genes and compared sequence data from the 3' UTR and exon and intron sequences from the 3' end of the coding region of the genes. Among these genes, 70% were polymorphic; the majority (86%) of the polymorphisms were SNPs. The distributions of the SNPs were similar in exons (44%) and introns (39%). Only 15% of the differences were located in the 3' UTR. In contrast, indels were most often discovered within introns. A SNP occurred every 151 nt, and on average each gene contained 2.7 SNPs. That value is very similar to the 2.8 SNPs per contig reported by Ramírez et al. (2005). This value will certainly increase as the size of the sequenced fragment increases and more genotypes are compared.

An extensive comparison of intron 1 of dihydroflavonol 4-reductase (DFR) measured the sequence variation among 92 genotypes including both landraces and varieties (McClean et al. 2004). Among these genotypes, 20 haplotypes were defined based on 69 polymorphisms. The level of diversity was similar among landraces

and varieties. Furthermore, the Middle American gene pool was more diverse than the Andean pool. Recent results with intron 3 of chalcone isomerase show a similar pattern with greater diversity in the Middle American gene pool (McClean and Lee, unpublished results). By contrast, the level of diversity among all genotypes is lower for this region of the genome than for intron 1 of DFR and only 10 haplotypes were observed.

5.4 Genetic Mapping and Tagging in *Phaseolus vulgaris*

5.4.1 Genetic Mapping

Linkage mapping in common beans as in other crops has benefited from a range of molecular technologies that have greatly supplemented the number of genetic markers used in genetic maps for the species. As a result, a large number of markers are in use today for common bean mapping compared to early linkage maps that were based almost entirely on a limited number of morphological markers such as those for flower or seed color or certain pod traits and growth habit characteristics (Bassett 1991). Isozymes and seed proteins, both analyzed based on biochemical assays, were among the first molecular markers to be used in genetic maps but suffered from the limitation of requiring multiple protocols and different source tissues (Gepts 1988; Arndt and Gepts 1989; Koenig and Gepts 1989a; Vallejos and Chase 1991). With the advent of DNA technologies, RFLP markers emerged as the first DNA-based markers in common beans to be used on a large scale (Vallejos et al. 1992; Nodari et al. 1993a; Adam-Blondon et al. 1994). New mapping populations were created at UC Davis and the University of Florida (UF) based on the crosses BAT93 x Jalo EEP558 (F2) and XR-235-1-1 x Calima (BC1) to determine the linkage relationships of large numbers of RFLP markers with totals of 224 Bng and 108 D series clones mapped in these studies. As single copy markers, RFLPs were especially useful for comparative mapping and map integration and led to the creation of the core linkage map of Freyre et al. (1998). In that study, RFLP probes from the initial UC Davis and UF maps plus those of Adam-Blondon et al. (1994) were analyzed to create a core map for the BAT93 x Jalo EEP558 (RIL) population that combined RFLPs with additional marker types including RAPDs, allozymes and known genes. RFLPs were also useful for comparative mapping across legume species as they hybridized well to genomic DNA of soybean and mung bean (Boutin et al. 1995). Sequencing of the Bng probes used to establish the UF map showed them to be rich in gene sequences (Murray et al. 2002).

Polymerase chain reaction (PCR)-based markers added greatly to the efficiency of genetic mapping in common beans by increasing marker throughput, initially with multi-locus marker systems such as those of the RAPD and then AFLP techniques and more recently with single copy marker systems such as STS/SCAR Sequence-Tagged Sites/Sequence-Characterized Amplified Regions and SSR primer pairs. RAPD and AFLP markers were especially useful for saturating RFLP-anchored genetic maps and for creating genetic maps for additional populations

(Freyre et al. 1998; Johnson and Gepts 2002; Tar'an et al. 2002). As genetic maps became more saturated and based on a wide range of molecular marker types, the number of linkage groups coalesced to equal the haploid complement of chromosomes for beans (n = 11). The cumulative genetic distance represented by linkage maps also increased as a greater number of markers was used. Comparative mapping based on RAPD bands was frequently useful for linking genetic maps by correlating the presence of fragments with the evolutionary origin of parental genotypes in either the Mesoamerican or Andean gene pools. The advent of single-copy PCR-based markers such as SSR and STS/SCAR markers has further aided in map comparison and integrating genetic maps. Over 30 STS or SCAR markers developed for many different disease or insect resistance genes and seed color traits have been placed on the core genetic map (McClean et al. 2002; Blair et al. 2006c; Miklas et al. 2006b.

Meanwhile, SSR markers are proving informative due to their multiallelic nature, codominant inheritance, and wide distribution in the genome. So far, SSR markers have been derived from Genbank sequences (Yu et al. 1999, 2000; Métais et al. 2002; Blair et al. 2003; Guerra-Sanz 2004), SSR-enriched genomic libraries (Gaitán-Solís et al. 2002, Yaish and de la Vega 2003; Buso et al. 2006) and bacterial artificial chromosome (BAC) sequences (Caixeta et al. 2005a). The first effort towards the integration of SSR markers into a common bean genetic map was performed by Yu et al. (2000), mapping 15 SSRs onto a framework map based on RAPD and RFLP markers, followed by Blair et al. (2003), integrating 100 SSRs into two linkage maps with AFLP, RAPD, and RFLP markers.

Additional integration of new SSR loci is ongoing at Empresa Brasileira de Pesquisas Agropecuarias Arroz e Feijão (Santo Antônio de Goiás, GO, Brazil) (for BAT93 x Jalo EEP558) and CIAT (for DOR364 x G19833) and will allow the identification of more polymorphic and transferable markers that can be mapped across populations. Furthermore, existing SSRs have been useful in two recent mapping studies by Ochoa et al. (2006) and Blair et al. (2006c) showing their potential for anchoring new genetic maps. Apart from SSR and STS markers, a new set of gene-based markers is being implemented for genetic mapping in common bean that includes TRAP and resistance-gene analogs (RGA)-based markers targeting disease resistance genes (López et al. 2003; Mutlu et al. 2006; Miklas et al. 2006a). Meanwhile, gene-based SNP markers being mapped at CIAT, and mapping of EST sequences at North Dakota State University and at the University of Saskatchewan holds promise for the development of a transcriptional map for beans that could help in the establishment of correlations between candidate genes and specific QTLs.

For the future, use of reference populations, anchor markers, and comparative mapping between populations will continue to assist in the placement of new markers onto genetic maps and the comparison of genetic distances between markers. Comparative mapping has become important for correlating locations of QTL for biotic or abiotic stress resistance, nitrogen fixation, seed characteristics, and root traits (Nodari et al. 1993b; Tar'an et al. 2002; Beebe et al. 2006; Miklas et al. 2006b; Ochoa et al. 2006). In addition, the use of anchor markers, especially those that are highly polymorphic such as SSRs, will allow the genetic mapping of more narrow crosses that are often of interest to plant breeders.

5.4.2 Gene Tagging

More than 30 individual genes for disease resistance and a similar number of genes for QTL underlying major traits with significant impact to common bean agriculture in the tropics have been successfully linked with markers. Such genes tightly linked with markers are referred to as "tagged genes." The primary goal for gene tagging in common bean has been to identify markers tightly linked with disease resistance traits for the purpose of marker-assisted selection. The first linked marker (RAPD A14.1100) was identified for the *Ur-4* rust resistance gene (Miklas et al. 1993) and was used for gene pyramiding and retention of a less effective gene (*Ur-4*) in the presence of an epistatic gene, *Ur-11*, with broad effect against the hypervariable rust pathogen (Stavely et al. 1994). Since then, many resistance genes have been tagged (see recent reviews Miklas et al. 2006b; Kelly et al. 2003; Ragagnin et al. 2005) including *Ur-3*, *Ur-5*, *Ur-6*, *Ur-7*, *Ur-9*, *Ur-11*, and *Ur-13* for resistance to rust; *Co-1*, *Co-1*^{2}, *Co-2*, *Co-4*, *Co-4*^{2}, *Co-5*, *Co-6*, *Co-9*, and *Co-10*, for anthracnose resistance; five *Phg* genes (Caixeta et al. 2005b) for resistance to angular leaf spot; *I*, *bc-1*^{2}, and *bc-3* for bean common mosaic virus; *Bct* for beet curly top virus; *bgm-1* for bean golden yellow mosaic virus (BGYMV); and *Pse-1*, *Pse-2*, and *Pse-3* for resistance to halo bacterial blight. Other yet unnamed R genes for resistance to rust, anthracnose, and angular leaf spot have also been tagged. In addition to R genes, markers tightly linked to QTL with major effect against pathogens causing BGYMV, common bacterial blight, Fusarium root rot, and white mold diseases have been tagged, localized on core linkage maps, and in certain cases used effectively in marker-assisted breeding.

Genetic population structure used to tag genes include: near-isogenic lines (NILs) developed by conventional backcrossing as used to tag *Ur-4* (Miklas et al. 1993) and *Ur-11* (Johnson et al. 1995) genes; NILs developed from heterogeneous inbred lines as used for *Ur-3* (Haley et al. 1994) and *bc-1*^{2} (Miklas et al. 2000); bulked segregant analysis using F_2 or inbred populations including recombinant inbred lines (RILs) as for *bgm-1* (Urrea et al. 1996) and *Bct* (Larsen and Miklas 2004); and a whole map approach as with *Co-2* (Adam-Blondon et al. 1994) and *Ur-9* (Jung et al. 1996) genes. Selective (Miklas et al. 1996, 2003a; Kolkman and Kelly 2003) and whole genome mapping (Jung et al. 1996; Miklas et al. 2001) approaches have been widely used to identify and generate markers for QTL. NILs combined with bulked segregant analysis was used to reduce identification of false positive markers in tagging *Ur-4*, *Ur-11*, and other genes.

Most genes to date have been tagged with RAPD markers which have been converted to SCAR markers (see list by Miklas 2005). Exceptions include the *Ur-13* gene tagged with an AFLP marker (Mienie et al. 2005) and a QTL conferring resistance to white mold linked with the phaseolin seed protein locus (*Phs*) (Miklas et al. 2001). A few codominant RAPD markers have been generated as for *bgm-1* (Urrea et al. 1996), *bc-3* (Johnson et al. 1997), and *Co-4*^{2} (Awale and Kelly 2001), but most markers identified have been dominant and in coupling phase linkage (*cis*) with the R genes and QTL. The predominance of dominant markers reduces the efficiency of marker-assisted selection. Use of dominant markers in repulsion phase

(*trans*) linkage alone or in combination with markers in coupling can be used to improve selection efficiency (reviewed by Kelly and Miklas 1998). Codominant markers like SSRs accelerate the identification of homozygous individuals without the necessity for progeny tests, but still too few exist to be effective for gene tagging studies. Ideally, selection of markers flanking the resistance gene would be used to improve marker assisted selection (MAS) efficiency, but such flanking markers are generally not available for the genes and QTL that have been tagged thus far in common bean. For marker-assisted backcrossing, lack of flanking markers can be partially overcome by also selecting for the recurrent parent genomic background outside the target locus using random markers (Hagiwara et al. 2001). Other gene-sequence based marker systems like RGA (López et al. 2003), Resistance Gene Analog Polymorphism (Mutlu et al. 2006), and TRAP (Miklas et al. 2006a) show promise for tagging targeted traits but have not yet lead to MAS. As EST databases expand for common bean, concurrently STS and SNP markers will be generated, leading to additional and more direct marker systems for tagging genes.

The utility of tightly linked markers for MAS has been evaluated by surveying an array of accessions with and without the gene for presence and absence of the marker. These surveys have revealed the utility for many markers to be limited to specific gene pools (Miklas et al. 1993). Once limitations of a marker are understood, MAS can be implemented more effectively as a breeding tool. The utility of QTL for MAS is determined by expression of the QTL in multiple environments and in different populations, and by the effectiveness of MAS itself.

MAS has been implemented in different breeding strategies for common bean improvement. Gene pyramiding, as in the example for marker-assisted detection of *Ur-4* gene in the presence of *Ur-11,* led to release of Beldakmi-RR-7 navy bean (Stavely et al. 1994). Kidney and cranberry bean germplasm releases (Miklas et al. 2002; Miklas and Kelly 2002) benefited from MAS of the hypostatic *I* gene in the presence of *bc-3* to facilitate combining genes for more durable resistance to bean common mosaic virus (BCMV) and bean common mosaic necrosis virus (BCMNV). Marker-assisted backcrossing was used to rapidly deploy resistance genes to combat emerging disease epidemics: *bgm-1* in snap bean RGMR lines (Stavely et al. 2001) to combat BGYMV in Florida and *Co-4*^{2} in USPT-ANT-1 pinto (Miklas et al. 2003b) to combat anthracnose in Minnesota and North Dakota. The former exemplifies the additional benefits of marker-assisted backcrossing of a recessive gene and a gene that would otherwise be difficult to detect by direct pathogen screening. Resistance to BGYMV in the GMR lines (Singh et al. 2000a, 2000b) was selected and characterized by presence of markers for an R gene *(bgm-1)* and a major QTL conditioning resistance (linked with the SW12 marker). MAS for multiple QTL in early generations followed by phenotypic selection with the pathogen in later generations resulted in the release of pinto and kidney bean germplasm lines with improved resistance to common bacterial blight (Miklas et al. 2006c, 2006d). ABCP-8 pinto with enhanced resistance to common bacterial blight benefited from MAS for a single QTL (Mutlu et al. 2005a, 2005b).

The applications above pertain to breeding for resistance to individual pathogens, whereas, an even greater utility for MAS in the future will be simultaneous

introgression of R genes conferring resistance to different pathogens in the same genetic background. The concept of multiple disease resistance is relatively straightforward; however, it is not easily accomplished by conventional means due to the inherent difficulty associated with the simultaneous monitoring of several different R genes by direct pathogen screening. Defining reaction symptoms after multiple inoculations with different pathogens is an arduous task as is discerning epistatic interactions between the R genes. The availability of molecular markers tightly linked to the R genes can bypass these difficulties because the markers are not affected by the environment or by epistatic interactions among the genes to which they are linked. Multiple-disease-resistant Mesoamerican bean lines with "carioca-type" grains were developed by MAS for R genes: *Co-4*, *Co-6*, *Co-10*, *Ur-OuroNegro*, and *Phg-1* (Ragagnin et al. 2005). Subsequently, *Co-4* was replaced by the more effective allele, *Co-4*^{2} (Alzate-Marin et al. 2005). Similarly, MAS was used to develop black and red bean cultivars with multiple disease resistance (Costa et al. 2006). Breeding for multiple disease resistance using MAS is a dynamic process, and new resistance genes can be incorporated at any moment as long as the gene of interest is 'tagged' with tightly linked markers and introduced in the appropriate genetic background.

5.5 BAC Cloning and Utilization

5.5.1 Development of BAC Libraries

In the genus *Phaseolus*, eleven BAC libraries have been constructed (Table 5.1). Although most of these libraries have relatively high genome equivalents, it must be noted that this, as with all BAC libraries, represents a statistical assessment based upon the assumption that the restriction sites are completely random throughout the genome. Because it has been demonstrated that this is not the case, (e.g., Nodari et al. 1992), the genome coverage must be viewed only as a probability of a particular region being represented, not as an absolute representation of the number of times a region will appear in a library. For a more complete coverage of a region, several libraries constructed with different enzymes should be constructed and assayed.

The BAC libraries listed here represent a unique phylogenetically ordered set of libraries for use in evolutionary studies as well (Fig. 5.1). In addition to a library in *P. lunatus* cv. Henderson, chosen as an outgroup, all the other libraries were developed in *P. vulgaris*. These *P. vulgaris* libraries represent key steps in the intra-specific evolution and domestication of common bean. DGD1962 is a wild bean from northern Peru, representing the presumed ancestral gene pool of the species (Debouck et al. 1993; Kami et al. 1995). The remainder of the libraries is distributed in the two evolutionary lineages that were domesticated. In the Mesoamerican lineage G02771 and G12946 are wild Mexican beans that contain the three subfamilies of the Arcelin-Phytohemagglutinin-Alpha-amylase inhibitor (APA) seed proteins,

Table 5.1 Some characteristics of BAC libraries developed in *Phaseolus* spp.

Library Species[1]	Genotype	Number of Clones	Restriction Site	Average Clone Size (kB)	Genomes	Empty Clones (%)	Chloroplast	Reference[2]
Pv	Sprite	33,792	*Eco*RI	90	4.79			a
Pv	BAT93	110,592	*Hin*dIII	125	20.8	< 0.5	0.05	b
Pv	DGD1962	55,296	*Hin*dIII	105	8.7	< 0.5	0.4	b
Pv	G02771	55,296	*Hin*dIII	139	12.1	< 0.5	0.08	b
Pv	G19833	55,296	*Hin*dIII	145	12			c
Pv	G12949	30,720	*Hin*dIII	135	6.5	-		d
Pv	OAC H45	33,024	*Hin*dIII	100	5	~ 0		e
Pv	HR67	22,560	*Bam*HI	300	8.1	< 0.5		f
Pv	OAC Rex	31,776	*Bam*HI	150	5.6	< 0.5		f
Pv	G02333	24,960	*Hin*dIII	125	6			g
Pl	Henderson	55,296	*Hin*dIII	110	9.5	< 0.5	0.03	b

[1] Pv: *Phaseolus vulgaris*; *Pl*: *Phaseolus lunatus*

[2] a: Vanhouten and Mackenzie 1999; b: Kami et al. 2006; c: M. Blair, pers. comm..; d: Galasso et al. 2005; e: Yu et al. 2006: f: Perry et al. 2006; g: Melotto et al. 2003

including the arcelin subfamily (see section 5.4.2). G02333 is a Mexican landrace highly resistant to anthracnose.

BAT93 and OAC-HR45 and OAC-HR67 are breeding lines and OAC-Rex is a cultivar from the Mesoamerican gene pool. Such an array of BAC libraries can help in describing the structural modifications that have accompanied phenotypic changes occurring during domestication, not only in the genes themselves, but also in adjacent, regulatory regions. Furthermore, they will allow an analysis of

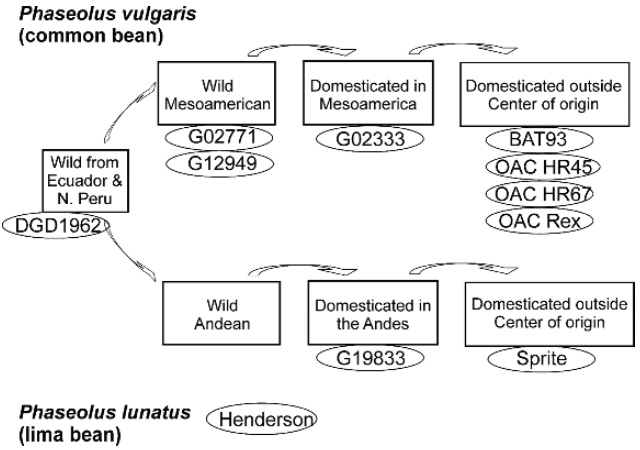

Fig. 5.1 Phylogenetic and genealogical distribution of BAC libraries in *Phaseolus* spp. Boxes represent different segments of the *P. vulgaris* gene pools and the general direction of their evolution. Names surrounded by ellipses are the genotypes in which BAC libraries have been established

the effects, if any, of selection exerted during and after domestication on genetic diversity in chromosome regions harboring domestication genes.

BAC libraries of the common bacterial blight (CBB) resistant variety OAC-Rex (Michaels et al. 2005), derived from OAC95-4 (Tar'an et al. 2001), and the CBB-resistant breeding line HR67 were constructed with a BiBAC2 vector that is designed to be used directly for plant transformation with Agrobacterium (Chang et al. 2003). The OAC-Rex library is being screened with a probe derived from the pvCTT001 microsatellite marker associated with a major CBB resistance locus on the B5 linkage group (Tar'an et al. 2001).

In the Andean gene pool, G19833 is a landrace from Peru, whereas Sprite is a bred variety. Thus, using this array of BAC libraries, it is possible to study overall structural evolution of the genome in *Phaseolus* both prior and after domestication. It is also possible to analyze phenotypic changes resulting from specific structural modification at the genome level. Some of the first information obtained towards this goal is shown in the next section for two disease resistance loci (*Co-4* and *I*) and one pest resistance locus (*APA*).

5.5.2 Whole-BAC Mapping and Sequencing

With the recent availability of BAC libraries in *Phaseolus* sp., a number of BACs have been completely sequenced for analysis. These results demonstrate both the strength and some of the weakness of this technology.

Using BAC clones identified from the Sprite library, Melotto et al. (2004) obtained the complete sequence of a 106,000 bp clone containing part of the *Co-4* locus for resistance to anthracnose. Analysis of the sequence demonstrated five copies of the *COK-4* gene, which is the putative gene responsible for the fungal resistance. Sequence alignment and phylogenetic analysis indicated that this gene has undergone several rounds of duplication and divergence. In addition, 19 other putative genes were identified by a BLAST homology search that had not been isolated from *Phaseolus* prior to this analysis. At least four putative retrotransposon elements were also identified. At the time of publication, this was the largest contiguous DNA sequence available for *P. vulgaris*.

Using BAC-end sequencing and hybridization, Vallejos et al. (2006) were able to construct a 425 kb contig around the *I* gene that conveys resistance to BCMV. This contig contained a large cluster of TIR-NBS-LRR sequences. Susceptible cultivars harboring the recessive *i* allele displayed simpler, more variable haplotypes compared to the resistant cultivars harboring the dominant *I* allele, which showed a single haplotype. Further data suggested that this locus may have expanded during evolution and domestication (Vallejos et al. 2006).

Kami et al. (2006) published the complete sequence for a BAC clone from the G02771 library containing the arcelin, phytohaemagglutinin, alpha-amylase inhibitor (APA) locus that is responsible for resistance to bruchid insects. G02771 is considered to be one of the more recently evolved *Phaseolus vulgaris* genotypes

based on the presence of the most recently identified variant of the APA family, the arcelins. As in the previous examples, this gene family has multiple members, two putative arcelins, three phytohaemagglutinins, and one alpha-amylase inhibitor, and has also undergone several duplication and divergence events that correlate with a change in function as well. At least six putative retrotransposons were identified in this clone along with 10 other putative coding regions unrelated to the APA gene family. A separate region was characterized in the clone that has high (>95%) identity to a portion of the chloroplast genome. Although these nuclear encoded chloroplast "splinters" have been documented in the Arabidopsis and rice genomes, this is the first time they have been observed in *Phaseolus* (Shahmuradov et al. 2003). Additional APA-containing clones from different BAC libraries (Table 5.1) will be sequenced in the near future to further understand the evolution of the APA locus. Already, Sparvoli and coworkers have undertaken a similar analysis in the G12949 genotype (Galasso et al. 2005). The BAC library screening resulted in three overlapping clones, containing several arcelin variants, the phytohemaglutinin-L and phytohemagglutinin-E genes and many retrotrotransposons. A fourth BAC clone, containing a second phytohemagglutinin-L gene and the alpha-amylase inhibitor gene, was part of the APA locus, but it was not possible to link it to the other three BACs (Galasso et al. unpublished results). Preliminary comparison with data obtained by Kami et al. (2006) shows a different degree of complexity of the APA locus between the two wild genotypes, demonstrating the utility of different libraries to describe the evolutionary pathway followed by a single complex locus and the need of further BAC sequencing to propose an evolutionary model for the APA gene family.

As these studies demonstrate, in addition to yielding information about the genes of primary interest, whole BAC sequencing can yield a lot of unanticipated information about genomic structure and genetic information. However, much of this information is limited in that there is currently a lack of sequences from other varieties and species from which to derive a useful hypothesis. As the cost of whole BAC sequencing declines, it is hoped more supportive sequence information will become available.

5.6 Molecular Cytogenetics

Common bean is a diploid species with 22 chromosomes (Sarbhoy 1978; Maréchal et al. 1978; and references therein). The chromosomes are small in size and similar in morphology. Therefore, it had not been possible for a long time to recognize all chromosome pairs, even when banding techniques were applied (Zheng et al. 1991). Although common bean, as other species in the genus, develops giant, polytene chromosomes in the suspensor cells of the immature proembryo (Nagl 1969), these chromosomes are not well-suited for detailed cytogenetic studies, because the endoreduplicated sister chromatids do not pair along their entire lengths, giving these chromosomes a loose appearance (Pedrosa 2003).

Major advances in the chromosome analysis in the species came with the application of the fluorescence in situ hybridization technique (FISH). This technique enables the localization of different DNA sequences along the chromosomes, making the identification of particular chromosome pairs possible. The first sequences to be localized on common bean mitotic chromosomes were the 5S and the 45S rDNA gene clusters (Moscone et al. 1999). Combined with chromosome morphology and the distribution of constitutive heterochromatin, these sequences could be used as chromosome markers to identify almost all chromosome pairs of the species. Interestingly, the two cultivars that were analyzed in this work had different number and size of 45S rDNA sites, which was sufficient to change the size of a few chromosome pairs. This analysis was therefore expanded to 37 accessions of common bean, representing the known genetic diversity of the species, and it became clear that this gene family has been amplified in the Andean gene pool, whereas the Mesoamerican gene pool seems to have retained the ancestral number of loci (3) in most of its representatives (Pedrosa-Harand et al. 2006).

By the time Moscone and co-workers published their work, several genetic linkage maps were already available for the species, including a unified core map (Freyre et al. 1998). The correlation between linkage groups and chromosomes was, however, not elucidated. Using pooled RFLP markers from each linkage group of a genetic map, this correlation could be established (Pedrosa et al. 2003). Each pool of markers could be used to identify a chromosome pair and an improved ideogram was constructed. This analysis showed no correspondence between linkage group and chromosome sizes. To understand these differences in recombination among chromosomes, a detailed chromosome map is now being built for common bean, similar to the existing maps of both model legumes (Pedrosa et al. 2002; Kulikova et al. 2001). Genetic markers from each linkage group have been used to select BACs, now available from different libraries. The selected clones are individually mapped to the chromosomes of the species, revealing the localization of the corresponding marker. A BAC clone corresponding to an anthracnose resistance gene locus has been mapped to linkage group B8 according to the common bean core linkage map (Melotto et al. 2004). BACs corresponding to several other genomic regions have been selected and have been used to build up the chromosome map. This map will enable the comparison of genetic and physical distances in several genomic regions and will anchor a future physical map in relation to the major chromosome landmarks: centromere, telomeres, and heterochromatin. The same BACs will be available for comparative studies with closely related species, giving insights into the chromosome evolution mechanisms in this group.

5.7 Functional Genomics

5.7.1 EST Development

Partial sequencing of cDNA inserts or expressed ESTs, obtained from many tissues and organs, has been used as an effective method for gene discovery, molecular

marker generation, and transcription pattern characterization. It is an efficient approach for identifying large number of plant genes expressed during different developmental stages and in response to a variety of environmental conditions.

In the last years, the number of *P. vulgaris* EST sequences has increased considerably. As of June 2006, 21,377 *P. vulgaris* EST sequences are available at the GenBank, National Center for Biotechnology Information (NCBI) (http://www.ncbi.nlm.nih.gov/dbEST/). Table 5.2 summarizes the information on publicly available cDNA libraries that have been used to generate EST sequences. The libraries are derived from different *P. vulgaris* genotypes including Mesoamerican and Andean gene pools, different plant organs and biological conditions. Contributions to the development of ESTs come from collaborative projects performed within the frame of Phaseomics among scientists from academic institutions in Brazil, Colombia, Mexico, and the United States.

The complete EST data set, generated by Ramírez et al. (2005) and Melotto et al. (2005), has been assembled into contigs to provide a single *P. vulgaris* gene index containing 20,578 ESTs (Graham et al. 2006). Of these, 6,787 were classified as singletons and the remaining assembled into 2,883 contigs (http://www.ccg.unam.mx/phaseolusest/) (Graham et al. 2006). In addition, in July 2005, The Institute for Genomic Research (TIGR) released version 1.0 of the TIGR Common_bean Gene Index (http://www.tigr.org). The input sequences are 21,290 and the output sequences are classified as 2,906 tentative consensus (TC) sequences and 6,578 singletons ESTs, which are values similar to those reported by Graham et al. (2006).

Table 5.2 Currently available *Phaseolus vulgaris* cDNA libraries and EST sequences

Genotype	Tissue/condition	Total ESTs	Attribution	Reference
Mesoamerican Negro Jamapa	Nodules elicited by *Rhizobium tropici* CIAT899	4,636	UNAM/UM[1]	Ramírez et al. 2005
Mesoamerican Negro Jamapa	Pods	3,667	UNAM/UM[1]	Ramírez et al. 2005
Mesoamerican Negro Jamapa	Roots	4,329	UNAM/UM[1]	Ramírez et al. 2005
Mesoamerican Negro Jamapa	Leaves	3,456	UNAM/UM[1]	Ramírez et al. 2005
Andean G19833	Leaves	4,938	CIAT[2]	Ramírez et al. 2005
Mesoamerican SEL 1308	Leaves	449	USP[3]	Melotto et al. 2005
Mesoamerican SEL 1308	Seedling shoots	2,474	USP[3]	Melotto et al. 2005
Mesoamerican SEL 1308	Seedling shoots inoculated with *Colletotrichum lindemuthianum*	2,332	USP[3]	Melotto et al. 2005

[1] Universidad Nacional Autónoma de México/University of Minnesota
[2] Centro Internacional de Agricultura Tropical, Cali, Colombia
[3] Universidade de São Paulo

The *P. vulgaris* EST sequence resource developed has provided tools for projects oriented to characterize the transcript profile of different bean organs and/or growth conditions. For instance, macroarray technology was used for transcriptome analysis of bean mature nodules elicited by the nitrogen fixing bacteria: *Rhizobium tropici* (Ramírez et al. 2005). This study led to the identification of genes over-expressed in nodules as compared to other plant organs such as roots, leaves, stems, and pods. Nodule transcript profile showed that genes related to nitrogen and carbon metabolism are integrated for ureide production (Ramírez et al. 2005).

In *P. coccineus*, some 20,000 ESTs were isolated from globular stage embryos of developing seed and deposited in GenBank by R. Goldberg (http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=Nucleotide&cmd=search&term=phaseolus+and+coccineus+and+embryo+and+est). Because of the close relationship between *P. vulgaris* and *P. coccineus*, these sequences can be used to isolate homologous sequences in common bean (Nanni et al. 2005).

Because genetic and genomic resources of common bean are limited, a comparative genomic approach using bean ESTs and those of other plant species is important to determine the usefulness of sequence information from other plants to study common bean. Bean unigenes share more orthologous sequences with soybean and *Arabidopsis thaliana* unigenes than with *Orzya sativa* and *Zea mays* unigenes (Melotto et al. 2005). However the degree of conservation among genes of bean, soybean, and *A. thaliana* varies between functional categories (Melotto et al. 2005). Single linkage clustering analysis of homologous sequences among five different plant species (*P. vulgaris, Lupinus albus*, soybean, *M. truncatula*, and *A. thaliana*) has been used to identify candidate genes that are relevant for adaptation to phosphorus deficiency in bean and are shared with other species (Graham et al. 2006).

Ongoing projects within the Phaseomics community will contribute to increase the number of *P. vulgaris* ESTs sequences in the near future. Kirstin Bett's group (U. Saskatchewan, Canada) is sequencing 20,000 ESTs from subtraction suppression cDNA libraries prepared from above-ground plant tissue of *P. vulgaris* cv. ICA Pijao and *P. angustissimus* PI535272 subjected to cold treatment (K. Bett, personal communication). Gary Stacey's group (U. of Missouri, USA) is sequencing a total of 20,000 ESTs from normalized subtracted cDNA libraries prepared from leaves of *P. vulgaris* cv. Early Gallatin treated with the fungal pathogen *Uromyces* (G. Stacey, personal communication). Mario Aguilar's group (U. Nacional de La Plata, Argentina) is sequencing 6,000 ESTs from a cDNA library prepared from *P. vulgaris* cv. Negro Jamapa hairy roots harvested after a short time of inoculation with *Rhizobium etli* SC15 (M. Aguilar, personal communication). Comprehensive information on EST sequences will provide relevant genetic and genomic resources that may be used to achieve the ultimate goal of improving bean crop quality and production.

5.7.2 TILLING

Targeted induced local lesions in genomes (TILLING) is a powerful reverse genetic approach that uses gene-specific primers for the identification of mutants of

a gene of interest from a large mutagenesis population (McCallum et al. 2000). Theoretically, given a sufficient population size, genome saturation can be achieved and mutants for any gene can be identified. Tagged mutagenesis, such as T-DNA or transposon-based systems, generally requires an effective transformation system. TILLING, however, does not rely on transformation and thus allows for reverse genetic approaches in transformation recalcitrant species, such as common bean (for transformation, however, see section 5.7.3).

Significant advances have been made in the development of a TILLING platform in common bean. The common bean genotype BAT93, from the Mesoamerican gene pool, was selected for TILLING by the common bean research community because it is one of the parents of the core genetic map (Freyre et al. 1998), it has been used in the generation of a large BAC library (Kami et al. 2006; see also section 5.5.1), has broad adaptation, and has desirable characteristics, such as disease resistance (Broughton et al. 2003). Based on genome saturation in other species, such as *Lotus japonicus* (Perry et al. 2003), it was estimated that about 5,000 mutagenesis lines would be required for genome saturation in BAT93. The magnitude of the project required collaboration and a TILLING consortium for tool development was created, including the University of Geneva (Geneva, Switzerland), USDA/ARS/TARS (Mayagüez, Puerto Rico), and CIAT (Cali, Colombia). Studies optimizing mutagen concentration determined that 35–40 mM ethylmethane sulfonate (EMS) resulted in the desired lethality rate of 60–70% in BAT93. Mutagenesis was confirmed using a screen for nodulation mutants and found that 35 mM EMS resulted in 10% putative nodulation-deficient mutants in a population of 348 M2 lines (C. Pankhurst, P. Lariguet, and W. Broughton, unpublished results). Thus, the initial BAT93 mutagenesis population showed a high mutation frequency and proved effective for forward genetic screening.

The TILLING consortium has currently produced about 1,500 M2 families and has advanced about 900 families to the M3 generation. DNA has been extracted from all M2 families and these DNAs will be used to confirm appropriate mutagen concentration and for TILLING analysis. To manage the large number of lines in the mutant population, the template DNA will be multiplexed for TILLING analysis and a database will be created for the organization of phenotypic and genotypic data. A critical component of TILLING is access to a genomic sequence for primer development. Common bean genomic and cDNA sequences of the selected loci will be searched using databases from other species such as *Medicago*, *Lotus*, soybean, pea, and *Arabidopsis* and from the common bean EST collections. Once the sequences of *P. vulgaris* loci of interest are available, gene specific primers will be designed using codons optimized to detect deleterious lesions (CODDLE) , developed for the discovery of deleterious lesions within coding sequences. The consortium expects to identify allelic series, which will allow for the study of gene function, as has been the case in previous TILLING efforts (Henikoff et al. 2004; McCallum et al. 2000).

These initial studies have shown that the common bean EMS mutagenesis protocol is effective at generating mutants and that mutants can be identified using appropriate screening protocols. To expand access to this TILLING platform, the consortium will establish seed banks for screening and distribution of the BAT93

mutagenesis population at its three locations and will provide the mechanisms for reverse genetic screening of the population through TILLING analysis. TILLING in common bean will have wide applications for both basic and applied research, as requests for mutants have already been received.

5.7.3 Transformation

Common bean has been transformed using two main approaches: particle bombardment or biolistics and Agrobacterium transformation. The applicability of the particle bombardment to introduce and transiently express genes into dry bean was demonstrated in the beginning of 1990s (Genga et al. 1992; Aragão et al. 1992, 1993). The bombardment of meristematic cells of embryonic axes revealed that foreign genes could efficiently reach the superficial cell layers, demonstrating the feasibility of producing transgenic bean using this method (Aragão et al. 1993). Russell et al. (1993) reported the production of transgenic navy bean plants using an electrical particle acceleration device. However, the protocol was time-consuming and the frequency of transformation was low (0.03%) and variety-limited. Later, Kim and Minamikawa (1996) reported the recovery of transgenic bean plants (cv. Goldstar) by bombarding embryonic axes. A transformation system was developed for regeneration of transgenic bean plants based on the bombardment of apical meristematic regions associated with an efficient tissue culture protocol for multiple shoot induction, elongation, and rooting (Aragão et al. 1996; Aragão and Rech 1997). The frequency of transformation ranged from 0.2 to 0.9% using linear or circular plasmid vectors, respectively (Aragão et al. 1996; Vianna et al. 2004). Morphology of the explants used during bombardment may influence the generation of transgenic plants because some varieties present the central region of the meristematic region covered by the primordial leaves (Aragão and Rech 1997). As shoot differentiation occurred in the peripheral layers of the meristematic ring, the number of cells that could be reached by the microparticles coated-DNA will be drastically reduced. Consequently, the efficiency of transformation could also be reduced.

One important constraint in the transformation system based on the bombardment of meristematic tissue of embryonic axes is the difficulty of having efficient selection of transformed cells. Using the selective agent imazapyr, an herbicidal molecule of the imidazolinone class capable of systemically translocating and concentrating in the apical meristematic region of the plant, we have been able to increase the recovery of fertile soybean (Aragão et al. 2000) and, more recently, dry bean plants (Bonfirm et al. 2007).

Since the first report on *P. vulgaris* transformation in 1993, a few groups have transformed bean to introduce agronomic traits. Our group has introduced useful genes, such as the *be2s1* gene to improve the methionine content of the seeds. In two of the five transgenic lines, the methionine content was significantly increased by 14 and 23% compared to that in non-transformed plants (Aragão et al. 1999). Russell et al. (1993) introduced into bean the *bar* gene, which encodes the phosphinothricin

acetyl transferase (PAT) and confers resistance to both phosphinothricin and the herbicide glufosinate ammonium, and the coat protein gene from the bean golden mosaic geminivirus (BGMV) in an attempt to produce virus-resistant plants. The introduced *bar* gene was shown to confer resistance to the herbicide under glasshouse conditions. However, transgenic bean plants expressing the BGMV coat protein gene did not exhibit virus resistance (D. Maxwell, unpublished results). Recently, we have generated highly BGMV-resistant plants by expression of a mutated *AC1* viral gene (Faria et al. 2006) and using interfering RNA (RNAi) (F.J.L Aragão and J. C. Faria, unpublished results). The transgenic bean lines were introduced into the breeding program to evaluate gene expression in different genetic backgrounds under glasshouse and field conditions. Transgenic bean lines containing the *bar* gene and resistant to the herbicide glufosinate ammonium were tested in field (Aragão et al. 2002). This was the first field release of a transgenic line of *P. vulgaris*.

Agrobacterium tumefaciens-assisted transformation is a standard protocol successfully applied to generate transgenic legume plants of some *Phaseolus* species (Zambre et al. 2005), although the frequency of transformation for *Phaseolus vulgaris* (common bean) has been very low (Zhang et al. 1997). The only report where *A. tumefaciens* was used to achieve expression of a lea gene that conferred abiotic tolerance in *P. vulgaris* is a very recent protocol based on sonication and vacuum assistance (Liu et al. 2005). Although this novel and cheap transformation method looks quite promising, the transformation efficiency must be improved.

Agrobacterium rhizogenes is a soil bacterium responsible for development of hairy root disease on a range of dicotyledonous plants (Tepfer 1990). Infection at the wounding sites by *A. rhizogenes* results in the transfer, integration, and expression of T-DNA from the root-inducing plasmid into the plant cells (Grant et al. 1991). Although *A. rhizogenes*-mediated root transformation has been described for a number of legumes, the induction of composite plants has not been previously reported for common bean (Díaz et al. 2000; Lee et al. 1993; Cheon et al. 1993; Stiller et al. 1997; Boisson-Dernier et al. 2001; and Van de Velde et al. 2003). Provided that root transformation is efficient, such composite plants are a simple, fast and reproducible strategic alternative to stable transformed lines, for functional genomics programs. A protocol originally developed for soybean (Bond and Gresshoff 1993) has been successfully modified for application in bean. After trying several *A. rhizogenes* strains (5), strain K599 (cucumopine-type) was the most efficient to induce hairy roots on several wild accessions (3), landraces (6), and cultivars (2) of *P. vulgaris* and three other species within the genus *Phaseolus* [*P. coccineus* (2), *P. acutifolius* (2), and *P. lunatus* (2)]. High transformation efficiency rates (75–90% frequency) were reproducibly found in three independent experiments with different *P. vulgaris* wild accessions, landraces and cultivars.

Briefly, for hairy root induction, beans seedlings are infected with a fresh culture of *A. rhizogenes* K599 by wounding the cotyledonary nodes with a syringe, and then covered with a plastic lid to increase humidity, because hairy roots are very sensitive to desiccation. Over 5–8 days a globular tumor develops on the stem at the wound site. Visible hairy roots emerged from the tumor as early as one week after infection. At three weeks post-infection, between 85–100% plants had abundant hairy roots.

Stem with primary root was removed from the plant by cutting about 1 cm below the cotyledon nodes. Plants with induced roots were repotted in fresh sterile vermiculite (Estrada-Navarrete et al. 2006). Transgenic roots in the genus *Phaseolus* offers a new strategy for over-expressing or suppressing endogenous genes. This method can be readily scaled up to perform functional genomics focused on root biology and root-microbe interactions.

5.8 Perspective

The ultimate goal of the Phaseomics (Phaseolus Genomics) Initiative is to sequence the common bean genome as a platform for the more efficient improvement of common bean (Broughton et al. 2003). With relatively limited resources, the Phaseolus community has laid the foundation to achieve this goal by developing genomic resources such as an impressive collection of germplasm and genetic stocks, mapping populations, BAC libraries, an incipient EST and TILLING assembly, and the development of tools such as FISH and transformation. Further work is necessary, however, including the development of a full physical map (including BAC-end sequencing) and correlation of this map with the genetic map, further assembly of EST and TILLING collections, establishment of a transcript map, and increasing the efficiency of transformation, among others. Already bean breeders have been very active in translating genomic information into mapping and tagging genes for agronomic interest for marker-assisted selection. Several cultivars and improved germplasm accessions have been released, which resulted from MAS. Currently, the limiting factor for MAS is the number of markers of known map location, primarily because of the lack of DNA sequence information available. A coordinated genomic effort in *Phaseolus* closely matched with breeders is a necessary condition for the full use of MAS and an accelerated development of improved bean cultivars and, hence, a more diverse and nutritious diet.

References

Adam-Blondon A, Sévignac M, Bannerot H, Dron M (1994) SCAR, RAPD and RFLP markers tightly linked to a dominant gene (*Are*) conferring resistance to anthracnose in common bean. Theor Appl Genet 88:865–870

Aguilar OM, Riva O, Peltzer E (2004) Analysis of *Rhizobium etli* and of its symbiosis with wild *Phaseolus vulgaris* supports coevolution in centers of host diversification. Proc Natl Acad Sci USA 101:13548–13553

Alzate-Marin AL, Arruda KM, Souza KA, Barros EG, Moreira MA (2005) Introgression of *Co-4*^{2} and *Co-5* anthracnose resistance genes into "Carioca" common bean cultivars with the aid of MAS. Ann Rep Bean Improv Coop 48:70–71

Aragão FJL, Rech EL (1997) Morphological factors influencing recovery of transgenic bean plants (*Phaseolus vulgaris* L) of a Carioca cultivar. Int J Plant Sci 158:157–163

Aragão FJL, Desa MFG, Almeida ER, Gander ES, Rech EL (1992) Particle bombardment-mediated transient expression of a Brazil nut methionine-rich albumin in bean (*Phaseolus vulgaris* L). Plant Mol Biol 20:357–359

Aragão FJL, Desa MFG, Davey MR, Brasileiro ACM, Faria JC, Rech EL (1993) Factors influencing transient gene-expression in bean (*Phaseolus vulgaris* L) using an electrical particle-acceleration device. Plant Cell Rep 12:483–490

Aragão F, Barros L, Brasileiro A, Ribeiro S, Smith F, Sanford J, Faria J, Rech E (1996) Inheritance of foreign genes in transgenic bean (*Phaseolus vulgaris* L) co-transformed via particle bombardment. Theor Appl Genet 93:142–150

Aragão FJL, Barros LMG, de Sousa MV, Grossi de Sa MF, Almeida ERP, Gander ES, Rech EL (1999) Expression of a methionine-rich storage albumin from the Brazil nut (*Bertholletia excelsa* H.B.K., Lecythidaceae) in transgenic bean plants (*Phaseolus vulgaris* L., Fabaceae). Genet Mol Biol 22:445–449

Aragão FJL, Sarokin L, Vianna GR, Rech EL (2000) Selection of transgenic meristematic cells utilizing a herbicidal molecule results in the recovery of fertile transgenic soybean [*Glycine max* (L.) Merril] plants at a high frequency. Theor Appl Genet 101:1–6

Aragão FJL, Vianna GR, Albino MMC, Rech EL (2002) Transgenic dry bean tolerant to the herbicide glufosinate ammonium. Crop Sci 42:1298–1302

Araya CM, Alleyne AT, Steadman JR, Eskridge KM, Coyne AP (2004) Phenotypic and genotypic characterization of *Uromyces appendiculatus* from *Phaseolus vulgaris* in the Americas. Plant Dis 88:830–836

Arndt GC, Gepts P (1989) Segregation and linkage for morphological and biochemical markers in a wide cross in common bean (*Phaseolus vulgaris*). Ann Rep Bean Improv Coop 32:68–69

Awale HE, Kelly JD (2001) Development of SCAR markers linked to *Co-4*^{2} gene in common bean. Ann Rpt Bean Improv Coop 44:119–120

Bai YH, Michaels TE, Pauls KP (1997) Identification of RAPD markers linked to common bacterial blight resistance genes in *Phaseolus vulgaris* L. Genome 40:544–551

Bassett MJ (1991) A revised linkage map of common bean. HortSci 26:834–836

Becerra-Velásquez VL, Gepts P (1994) RFLP diversity in common bean (*Phaseolus vulgaris* L). Genome 37:256–263

Beebe S, Gonzalez AM, Rengifo J (2000a) Research on trace minerals in the common bean. Food and Nutrition Bull 21:387–391

Beebe S, Skroch PW, Tohme J, Duque MC, Pedraza F, et al. (2000b) Structure of genetic diversity among common bean landraces of Middle American origin based on correspondence analysis of RAPD. Crop Sci 40:264–273

Beebe S, Rengifo J, Gaitan E, Duque MC, Tohme J (2001) Diversity and origin of Andean landraces of common bean. Crop Sci 41:854–862

Beebe SE, Rojas-Pierce M, Yan X, Blair MW, Pedraza F, et al. (2006) Quantitative trait loci for root architecture traits correlated with phosphorus acquisition in common bean. Crop Sci 46:413–423

Bennett M, Leitch I (2005) Angiosperm DNA C-Values database. Release 4.0 http://www.rbgkew.org.uk/cval/database1.html

Blair MW, Pedraza F, Buendia HF, Gaitán-Solís E, Beebe SE, et al. (2003) Development of a genome-wide anchored microsatellite map for common bean (*Phaseolus vulgaris* L). Theor Appl Genet 107, 1362–1374

Blair MW, Giraldo MC, Buendia HF, Tovar E, Duque MC, et al. (2006a) Microsatellite marker diversity in common bean (*Phaseolus vulgaris* L). Theor Appl Genet 113:100–109

Blair MW, Iriarte G, Beebe S (2006b) QTL analysis of yield traits in an advanced backcross population derived from a cultivated Andean x wild common bean (*Phaseolus vulgaris* L) cross. Theor Appl Genet 112:1149–1163

Blair MW, Muñoz C, Garza R, Cardona C (2006c) Molecular mapping of genes for resistance to the bean pod weevil (*Apion godmani* Wagner) in common bean. Theor Appl Genet 112: 913–923

Boisson-Dernier A, Chabaud M, Garcia F, Becard G, Rosenberg C, et al. (2001) Agrobacterium rhizogenes-transformed roots of *Medicago truncatula* for the study of nitrogen-fixing and endomycorrhizal symbiotic associations. Molec Plant-Microbe Inter 14:695–700

Bond JE, Gresshoff PM (1993) Soybean transformation to study molecular physiology. In: Gresshoff PM (ed) Plant responses to the environment. CRC Press, Boca Raton, FL, USA, pp 25–44

Bonfim K, Faria JC, Nogueira EOPL, Mendes EA, Aragão FJL (2007) RNAi-Mediated Resistance to Bean Golden Mosaic Virus in genetically engineered common bean (*Phaseolus vulgaris*) Molec Plant-Microbe Inter 20:717–726

Boutin S, Young N, Olson T, Yu Z, Shoemaker R, et al. (1995) Genome conservation among three legume genera detected with DNA markers. Genetics 38:928–937

Bressani R (1983) Research needs to upgrade the nutritional quality of common beans (*Phaseolus vulgaris*). Qualitas Plantarum Plant Foods Human Nutrition 32:101–110

Broughton WJ, Hernandez G, Blair M, Beebe S, Gepts P, et al. (2003) Beans (*Phaseolus* spp.) – model food legumes. Plant Soil 252:55–128

Buso GSC, Amaral ZPS, Brondani RPV, Ferreira ME (2006) Microsatellite markers for the common bean – *Phaseolus vulgaris*. Mol Ecol Notes 6:252–254

Caixeta ET, Borém A, Kelly JD (2005a) Development of microsatellite markers based on BAC common bean clones. Crop Breed Appl Biotech 5:125–133

Caixeta ET, Borem A, Azate-Marin AL, Fagundes SA, Silva MGM, et al. (2005b) Allelic relationships for genes that confer resistance to angular leaf spot in common bean. Euphytica 145:237–245

Chang YL, Henriquez X, Preuss D, Copenhaver GP, Zhang HB (2003) A plant-transformation-competent BIBAC library from the *Arabidopsis thaliana* Landsberg ecotype for functional and comparative genomics. Theor Appl Genet 106:269–276

Cheon CI, Lee NG, Siddique ABM, Bal AK, Verma DPS (1993) Roles of plant homologs of Rab1p and Rab7p in the biogenesis of the peribacteroid membrane, a subcellular compartment formed de-novo during root-nodule symbiosis. EMBO J 12:4125–4135

Choi H-K, Mun J-H, Kim D-J, Zhu H, Baek J-M, et al. (2004) Estimating genome conservation between crop and model legume species. Proc Natl Acad Sci USA 101:15289–15294

Costa MR, Tanure JPM, Arruda KMA, Carneiro JES, Moreira MA, et al. (2006) Pyramiding of anthracnose, angular leaf spot and rust resistance genes in black and red bean cultivars. Ann Rep Bean Improv Coop 49:187–188

Debouck DG, Toro O, Paredes OM, Johnson WC, Gepts P (1993) Genetic diversity and ecological distribution of *Phaseolus vulgaris* in northwestern South America. Econ Bot 47:408–423

Díaz CL, Spaink HP, Kijne JW (2000) Heterologous rhizobial lipochitin oligosaccharides and chitin oligomers induce cortical cell divisions in red clover roots, transformed with the pea lectin gene. Molec Plant-Microbe Inter 13:268–276

Estrada-Navarrete G, Alvarado-Affantranger X, Olivares J-E, Díaz-Camino C, Santana O, Murillo E, Guillén G, Sánchez-Guevara N, Acosta J, Quinto C, Li D, Gresshoff PM, Sánchez F (2006) *Agrobacterium rhizogenes*-transformation of the *Phaseolus* spp.: A tool for functional genomics. Molec Plant-Micr Inter 19:1385–1393

Faria JC, Albino MMC, Dias BBA, Cancado LJ, da Cunha NB, Silva LD, Vianna GR, Aragão FJL (2006) Partial resistance to bean golden mosaic virus in a transgenic common bean (*Phaseolus vulgaris* L.) line expressing a mutated rep gene. Plant Sci 171:565–571

Foster-Hartnett D, Mudge J, Larsen D, Danesh D, Yan HH, et al. (2002) Comparative genomic analysis of sequences sampled from a small region on soybean (*Glycine max*) molecular linkage group G. Genome 45:634–645

Frei A, Blair MW, Cardona C, Beebe SE, Gu H, et al. (2005) QTL mapping of resistance to *Thrips palmi* Karny in common bean. Crop Sci 45:379–387

Freyre, R, Skroch P, Geffroy V, Adam-Blondon A-F, Shirmohamadali A, et al. (1998) Towards an integrated linkage map of common bean. 4. Development of a core map and alignment of RFLP maps. Theor Appl Genet 97:847–856

Gaitán-Solís E, Duque MC, Edwards KJ, Tohme J (2002) Microsatellite repeats in common bean (*Phaseolus vulgaris*): Isolation, characterization, and cross-species amplification in *Phaseolus* ssp. Crop Sci 42:2128–2136

Galasso I, Lioi L, Lanave C, Campion B, Bollini R, et al. (2005) Identification and sequencing of a BAC clone belonging to the *Phaseolus vulgaris* (L) insecticidal *Arc4* lectin locus. Ann Rep Bean Improv Coop 48:40–40

Geffroy V, Sicard D, de Oliveira J, Sévignac M, Cohen S, et al. (1999) Identification of an ancestral resistance gene cluster involved in the coevolution process between *Phaseolus vulgaris* and its fungal pathogen *Colletotrichum lindemuthianum*. Molec Plant-Microbe Inter 12:774–784

Geffroy V, Sévignac M, De Oliveira J, Fouilloux G, Skroch P, et al. (2000) Inheritance of partial resistance against *Colletotrichum lindemuthianum* in *Phaseolus vulgaris* and co-localization of QTL with genes involved in specific resistance. Molec Plant-Microbe Inter 13:287–296

Genga AM, Allavena A, Ceriotti A, Bollini R (1992) Genetic transformation in *Phaseolus* by high-velocity particle microprojectiles. Acta Hort. 300:309–313

Gepts P (1988) Provisional linkage map of common bean. Annu Rep Bean Improv Coop 31:20–25

Gepts P (1990) Biochemical evidence bearing on the domestication of *Phaseolus* beans. Econ Bot 44(3S):28–38

Gepts P, Bliss FA (1984) Enhanced available methionine concentration associated with higher phaseolin levels in common bean seeds. Theor Appl Genet 69:47–53

Gepts P, Bliss FA (1985) F_1 hybrid weakness in the common bean: differential geographic origin suggests two gene pools in cultivated bean germplasm. J Hered 76:447–450

Gepts P, Bliss FA (1986) Phaseolin variability among wild and cultivated common beans (*Phaseolus vulgaris*) from Colombia. Econ Bot 40:469–478

Gepts P, Osborn TC, Rashka K, Bliss FA (1986) Phaseolin-protein variability in wild forms and landraces of the common bean (*Phaseolus vulgaris*): evidence for multiple centers of domestication. Econ Bot 40:451–468

Graham MA, Ramírez M, Valdés-López O, Lara M, Tesfaye M, et al. (2006) Identification of phosphorus stress induced genes in *Phaseolus vulgaris* L. through clustering analysis across several species. Funct Plant Biol 33:789–797

Grant JE, Dommiss EM, Conner AJ (1991) Gene transfer to plants using Agrobacterium. In: Murray DR (ed) Advanced methods in plant breeding and biotechnology. CAB International, Wallingford, pp 50–73

Guerra-Sanz JM (2004) New SSR markers of *Phaseolus vulgaris* from sequence databases. Plant Breed 123:87–89

Guzmán P, Gilbertson RL, Nodari R, Johnson WC, Temple SR, et al. (1995) Characterization of variability in the fungus *Phaeoisariopsis griseola* suggests coevolution with the common bean (*Phaseolus vulgaris*). Phytopathology 85:600–607

Guzmán-Maldonado SH, Martínez O, Acosta-Gallegos JA, Guevara-Lara F, Paredes-López O (2003) Putative quantitative trait loci for physical and chemical components of common bean. Crop Sci 43:1029–1035

Hagenblad J, Tang CL, Molitor J, Werner J, Zhao K, et al. (2004) Haplotype structure and phenotypic associations in the chromosomal regions surrounding two *Arabidopsis thaliana* flowering time loci. Genetics 168:1627–1638

Hagiwara WE, dos Santos JB, do Carmo LM (2001) Use of RAPD to aid selection in common bean backcross breeding programs. Crop Breed Appl Biotech 1:355–362

Haley S, Afanador L, Miklas P, Stavely J, Kelly J (1994) Heterogeneous inbred populations are useful as sources of near-isogenic lines for RAPD marker localization. Theor Appl Genet 88:337–342

Henikoff S, Till BJ, Comai L (2004) TILLING. Traditional mutagenesis meets functional genomics. Plant Physiol 135:630–636

Islam FMA, Basford KE, Jara C, Redden RJ, Beebe S (2002) Seed compositional and disease resistance differences among gene pools in cultivated common bean. Genetic Res Crop Evol 49:285–293

Johnson E, Miklas P, Stavely J, Martínez-Cruzado J (1995) Coupling- and repulsion-RAPDs for marker-assisted selection of PI 181996 rust resistance in common bean. Theor Appl Genet 90:659–664

Johnson WC, Guzmán P, Mandala D, Mkandawire A, Temple S, et al. (1997) Molecular tagging of the *bc-3* gene for introgression into Andean common bean. Crop Sci 37: 248–254

Johnson WC, Gepts P (2002) The role of epistasis in controlling seed yield and other agronomic traits in an Andean x Mesoamerican cross of common bean (*Phaseolus vulgaris* L.). Euphytica 125:69–79

Jorde LB (2000) Linkage disequilibrium and the search for complex disease genes. Genome Res 10:1435–1444

Jung G, Coyne D, Skroch P, Nienhuis J, Arnaud-Santana E, et al. (1996) Molecular markers associated with plant architecture and resistance to common blight, web blight, and rust in common beans. J Am Soc Hort Sci 121:794–803

Kami J, Becerra Velásquez B, Debouck DG, Gepts P (1995) Identification of presumed ancestral DNA sequences of phaseolin in *Phaseolus vulgaris*. Proc Natl Acad Sci USA 92:1101–1104

Kami J, Poncet V, Geffroy V, Gepts P (2006) Development of four phylogenetically-arrayed BAC libraries and sequence of the APA locus in *Phaseolus vulgaris*. Theor Appl Genet 112:987–998

Kelly J, Miklas PN (1998) The role of RAPD markers in breeding for disease resistance in common bean. Mol Breed 4:1–11

Kelly JD, Vallejo VA (2004) A comprehensive review of the major genes conditioning resistance to anthracnose in common bean. Hortscience 39:1196–1207

Kelly JD, Gepts P, Miklas PN, Coyne DP (2003) Tagging and mapping of genes and QTL and molecular marker-assisted selection for traits of economic importance in bean and cowpea. Field Crop Res 82:135–154

Khairallah MM, Adams MW, Sears BB (1990) Mitochondrial DNA polymorphisms of Malawian bean lines: further evidence for two major gene pools. Theor Appl Genet 80:753–761

Khairallah MM, Sears BB, Adams MW (1992) Mitochondrial restriction fragment polymorphisms in wild *Phaseolus vulgaris* – insights in the domestication of common bean. Theor Appl Genet 84:915–922

Kim JW, Minamikawa T (1996) Transformation and regeneration of French bean plants by the particle bombardment process. Plant Sci 117:131–138

Koenig R, Gepts P (1989a) Segregation and linkage of genes for seed proteins, isozymes, and morphological traits in common bean (*Phaseolus vulgaris*). J Hered 80:455–459

Koenig R, Gepts P (1989b) Allozyme diversity in wild *Phaseolus vulgaris*: further evidence for two major centers of diversity. Theor Appl Genet 78:809–817

Koenig R, Singh SP, Gepts P (1990) Novel phaseolin types in wild and cultivated common bean (*Phaseolus vulgaris*, Fabaceae). Econ Bot 44:50–60

Koinange EMK, Singh SP, Gepts P (1996) Genetic control of the domestication syndrome in common-bean. Crop Sci 36:1037–1045

Kolkman JM, Kelly JD (2003) QTL conferring resistance and avoidance to white mold in common bean. Crop Sci 43:539–548

Kulikova O, Gualtieri G, Geurts R, Kim DJ, Cook D, et al. (2001) Integration of the FISH pachytene and genetic maps of Medicago truncatula. Plant J 27:49–58

Larsen RC, Miklas PN (2004) Generation and molecular mapping of a sequence characterized amplified region marker linked with the *Bct* gene for resistance to *Beet curly top virus* in common bean. Phytopathology 94:320–325

Lavin M, Herendeen PS, Wojciechowski MF (2005) Evolutionary rates analysis of Leguminosae implicates a rapid diversification of the major family lineages immediately following an Early Tertiary emergence. Syst Biol 54:575–594

Lee JM, Grant D, Vallejos CE, Shoemaker RC (2001) Genome organization in dicots. II. Arabidopsis as a 'bridging species' to resolve genome evolution events among legumes. Theor Appl Genet 103:765–773

Lee, NG, Stein B, Suzuki H, Verma DPS (1993) Expression of antisense nodulin-35 RNA in *Vigna aconitifolia* transgenic root nodules retards peroxisome development and affects nitrogen availability to the plant. Plant J 3:599–606

Liu ZC, Park BJ, Kanno A, Kameya T (2005) The novel use of a combination of sonication and vacuum infiltration in Agrobacterium-mediated transformation of kidney bean (*Phaseolus vulgaris* L) with lea gene. Mol Breed 16:189–197

López CE, Acosta IF, Jara C, Pedraza F, Gaitán-Solís E, et al. (2003) Identifying resistance gene analogs associated with resistances to different pathogens in common bean. Phytopathology 93:88–95

Mackay TFC (2001) The genetic architecture of quantitative traits. Ann Rev Genet 35:303–339

Maréchal R, Mascherpa J-M, Stainier F (1978) Etude taxonomique d'un groupe complexe d'espèces des genres *Phaseolus* et *Vigna* (Papilionaceae) sur la base de données morphologiques et polliniques, traitées par l'analyse informatique. Boissiera 28:1–273

Mauro Herrera M (2003) Wild bean populations as source of genes to improve the yield of cultivated *Phaseolus vulgaris* L. Ph.D. thesis, University of California, Davis

McCallum CM, Comai L, Greene EA, Henikoff S (2000) Targeting induced local lesions in genomes (TILLING) for plant functional genomics. Plant Physiol 123:439–442

McClean PE, Lee RK, Otto C, Gepts P, , Bassett MJ (2002) Molecular and phenotypic mapping of genes controlling seed coat pattern and color in common bean (*Phaseolus vulgaris* L.) J Hered 93:148–152

McClean PE, Lee RK, Miklas PN (2004) Sequence diversity analysis of dihydroflavonol 4-reductase intron 1 in common bean. Genome 47:266–280

McClean PE, Lee RD, McConnell MD, Mamidi S, White CP (2006) Sequence and marker-based diversity in common bean. Plant Animal Genome 14:W144, p 40

McConnell M, Mamidi S, Lee R, McClean P (2006) DNA sequence polymorphism among common bean genes. Ann Rept Bean Improv Coop 49:143–144

Melotto M, Fransisco C, Camargo LEA (2003) Towards cloning the *Co-4*^{2} locus using a bean BAC library. Ann Rep Bean Improv Coop 46:51–52

Melotto M, Coelho MF, Pedrosa-Harand A, Kelly JD, Camargo LEA (2004) The anthracnose resistance locus *Co-4* of common bean is located on chromosome 3 and contains putative disease resistance-related genes. Theor Appl Genet 109:690–699

Melotto M, Monteiro-Vitorello CB, Bruschi AG, Camargo LEA (2005) Comparative bioinformatic analysis of genes expressed in common bean (*Phaseolus vulgaris* L.) seedlings. Genome 48:562–570

Métais I, Hamon B, Jalouzot R, Peltier D (2002) Structure and level of genetic diversity in various bean types evidenced with microsatellite markers isolated from a genomic enriched library. Theor Appl Genet 104:1346–1352

Michaels TE, Smith TW, Larsen J, Beattie AD, Pauls KP (2005) OAC Rex common bean. Can J Plant Sci 86:733–736

Mienie C, Liebenberg M, Pretorius Z, Miklas P (2005) SCAR markers linked to the common bean rust resistance gene *Ur-13*. Theor Appl Genet 111:972–979

Miklas PN (2005) List of SCAR marker for disease resistance: Aug 2005 update http://www.ars.usda.gov/SP2UserFiles/Place/53540000/miklas/SCARtable.pdf

Miklas PN, Kelly JD (2002) Registration of two cranberry bean germplasm lines resistant to Bean Common Mosaic and Necrosis Potyviruses: USCR-7 and USCR-9. Crop Sci 42:673–674

Miklas PN, Stavely JR, Kelly JD (1993) Identification and potential use of a molecular marker for rust resistance in common bean. Theor Appl Genet 85:745–749

Miklas PN, Johnson E, Stone V, Beaver JS, Montoya C, et al. (1996) Selective mapping of QTL conditioning disease resistance in common bean. Crop Sci 36:1344–1351

Miklas PN, Larsen RC, Riley R, Kelly JD (2000) Potential marker-assisted selection for *bc-1*^{2} resistance to bean common mosaic potyvirus in common bean. Euphytica 116:211–219

Miklas PN, Johnson WC, Delorme R, Gepts P (2001) QTL conditioning physiological resistance and avoidance to white mold in dry bean. Crop Sci 41:309–315

Miklas PN, Hang AN, Kelly JD, Strausbaugh CA, Forster RL (2002) Registration of three kidney bean germplasm lines resistant to Bean Common Mosaic and Necrosis Potyviruses: USLK-2 Light Red Kidney, USDK-4 Dark Red Kidney, and USWK-6 White Kidney. Crop Sci 42:674–675

Miklas PN, Delorme R, Riley R (2003a) Identification of QTL conditioning resistance to white mold in snap bean. J Am Soc Hort Sci 128:564–570

Miklas PN, Kelly JD, Singh SP (2003b) Registration of anthracnose resistant pinto bean germplasm line USPT ANT 1. Crop Sci 43:1889–1890

Miklas PN, Hu J, Grünwald NJ, Larsen KM (2006a) Potential application of TRAP (Targeted Region Amplified Polymorphism) markers for mapping and tagging disease resistance traits in common bean. Crop Sci 46:910–916

Miklas PN, Kelly JD, Beebe SE, Blair MW (2006b) Common bean breeding for resistance against biotic and abiotic stresses: from classical to MAS breeding. Euphytica 147:106–131

Miklas PN, Smith JR, Singh SP (2006c) Registration of common bacterial blight resistant dark red kidney bean germplasm line USDK-CBB-15 10.2135/cropsci2005.06-0110. Crop Sci 46:1005–1007

Miklas PN, Smith JR, Singh SP (2006d) Release of common bacterial blight resistant pinto bean germplasm lines USPT-CBB-5 and USPT-CBB-6. Ann Rep Bean Improv Coop 49:283–284

Mkandawire ABC, Mabagala RB, Guzman P, Gepts P, Gilbertson RL (2004) Genetic diversity and pathogenic variation of common blight bacteria (*Xanthomonas campestris* pv. *phaseoli* and *X. campestris* pv. *phaseoli* var. *fuscans*) suggests pathogen coevolution with the common bean. Phytopathology 94:593–603

Moffett MD, Weeden NF (2006) Investigation of synteny conservation between *Pisum* and *Phaseolus*. Abstract, Plant & Animal Genome XIV, 2006: http://www.intl-pag.org/14/abstracts/PAG14_P442.html Moscone EA, Klein F, Lambrou M, Fuchs J, Schweizer D (1999) Quantitative karyotyping and dual-color FISH mapping of 5S and 18S-25S rDNA probes in the cultivated *Phaseolus* species (Leguminosae). Genome 42:1224–1233

Murray J, Larsen J, Michaels TE, Schaafsma A, Vallejos CE, et al. (2002) Identification of putative genes in bean (*Phaseolus vulgaris*) genomic (Bng) RFLP clones and their conversion to STSs. Genome 45:1013–1024

Mutlu N, Miklas PN, Coyne DP (2006) Resistance gene analog polymorphism (RGAP) markers co-localize with disease resistance genes and QTL in common bean. Mol Breed 17: 127–135

Mutlu N, Miklas P, Reiser J, Coyne D (2005a) Backcross breeding for improved resistance to common bacterial blight in pinto bean (*Phaseolus vulgaris* L.). Plant Breed 124:282–287

Mutlu N, Miklas PN, Steadman JR, Vidaver AV, Lindgren D, et al. (2005b) Registration of pinto bean germplasm line ABCP-8 with resistance to common bacterial blight. Crop Sci 45:806

Nagl W (1969) Banded polytene chromosomes in the legume *Phaseolus vulgaris*. Nature 221:70–71

Nanni L, Losa A, Bellucci E, Kater M, Gepts P, et al. (2005) Identification and molecular diversity of a genomic sequence similar to SHATTERPROOF (*SHP1*) in *Phaseolus vulgaris* L. Plant Animal Genome XIII:P470, p 188

Nodari RO, Koinange EMK, Kelly JD, Gepts P (1992) Towards an integrated linkage map of common bean. I. Development of genomic DNA probes and levels of restriction fragment length polymorphism. Theor Appl Genet 84:186–192

Nodari RO, Tsai SM, Gilbertson RL, Gepts P (1993a) Towards an integrated linkage map of common bean. II. Development of an RFLP-based linkage map. Theor Appl Genet 85:513–520

Nodari RO, Tsai SM, Guzmán P, Gilbertson RL, Gepts P (1993b) Towards an integrated linkage map of common bean. 3. Mapping genetic factors controlling host-bacteria interactions. Genetics 134:341–350

Nordborg M, Hu TT, Ishino Y, Jhaveri J, Toomajian C, et al. (2005) The pattern of polymorphism in *Arabidopsis thaliana*. PLoS Biology 3:e196

Ochoa IE, Blair MW, Lynch JP (2006) QTL analysis of adventitious root formation in common bean under contrasting phosphorus availability 10.2135/cropsci2005.12-0446. Crop Sci 46:1609–1621

Pallottini L, Garcia E, Kami J, Barcaccia G, Gepts P (2004) The genetic anatomy of a patented yellow bean. Crop Sci 44:968–977

Papa R, Gepts P (2003) Asymmetry of gene flow and differential geographical structure of molecular diversity in wild and domesticated common bean (*Phaseolus vulgaris* L.) from Mesoamerica. Theor Appl Genet 106:239–250

Papa R, Acosta J, Delgado-Salinas A, Gepts P (2005) A genome-wide analysis of differentiation between wild and domesticated *Phaseolus vulgaris* from Mesoamerica. Theor Appl Genet 111:1147–1158

Payró de la Cruz E, Gepts P, Colunga GarciaMarín P, Zizumbo Villareal D (2005) Spatial distribution of genetic diversity in wild populations of *Phaseolus vulgaris* L. from Guanajuato and Michoacán, México. Genet Res Crop Evol 52:589–599

Pedrosa A (2003) Chromosomal organisation and physical mapping in legumes. Ph.D. thesis, University of Vienna, 143 p

Pedrosa A, Sandal N, Stougaard J, Schweizer D, Bachmair A (2002) Chromosomal map of the model legume *Lotus japonicus*. Genetics 161:1661–1672

Pedrosa A, Vallejos C, Bachmair A, Schweizer D (2003) Integration of common bean (*Phaseolus vulgaris* L.) linkage and chromosomal maps. Theor Appl Genet 106:205–212

Pedrosa-Harand A, Almeida CCS, Mosiolek M, Blair MW, Schweizer D, et al. (2006) Extensive ribosomal DNA amplification during Andean common bean (*Phaseolus vulgaris* L.) evolution. Theor Appl Genet 112:924–933

Perry JA, Wang TL, Welham TJ, Gardner S, Pike JM, et al. (2003) A TILLING reverse genetics tool and a web-accessible collection of mutants of the legume *Lotus japonicus*. Plant Physiol 131:866–871

Perry G, Reinprecht Y, Pauls KP (2006) Identification of common bacterial blight resistance genes in *Phaseolus vulgaris*. Abstract P16051, Plant Biology 2006, Joint Annual Meeting of the American Society of Plant Biologists and the Canadian Society of Plant Physiologists, Boston

Pritchard J, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155:945–959

Ragagnin VA, Alzate-Marin AL, Souza TLPO, Sanglard DA, Moreira MA, et al. (2005) Use of molecular markers to pyramiding multiple genes for resistance to rust, anthracnose and angular leaf spot in the common bean. Ann Rep Bean Improv Coop 48:94–95

Ramírez M, Graham MA, Blanco-López L, Silvente S, Medrano-Soto A, et al. (2005) Sequencing and analysis of common bean ESTs. Building a foundation for functional genomics. Plant Physiol 137:1211–1227

Rivkin M, Vallejos C, McClean P (1999) Disease-related sequences in common bean. Genome 42:41–47

Russell DR, Wallace KM, Bathe JH, Martinell BJ (1993) Stable transformation of *Phaseolus vulgaris* via electric discharge mediated particle acceleration. Plant Cell Rep 12:165–169

Santalla M, Rodino AP, de Ron AM (2002) Allozyme evidence supporting southwestern Europe as a secondary center of genetic diversity for the common bean. Theor Appl Genet 104:934–944

Sarbhoy RK (1978) Cytogenetical studies in genus *Phaseolus* Linn. I and II. Somatic and meiotic studies in fifteen species of *Phaseolus*. Cytologia 43:161–180

Schmid K, Ramos-Onsins S, Ringys-Beckstein H, Weisshaar B, Mitchell-Olds T (2005) A multilocus sequence survey in *Arabidopsis thaliana* reveals a genome-wide departure from a neutral model of DNA sequence polymorphism. Genetics 169:1601–1615

Shahmuradov IA, Akbarova YY, Solovyev VV, Aliyev JA (2003) Abundance of plastid DNA insertions in nuclear genomes of rice and Arabidopsis. Plant Mol Biol 52:923–934

Singh SP, Gepts P, Debouck DG (1991a) Races of common bean (*Phaseolus vulgaris* L., Fabaceae). Econ Bot 45:379–396

Singh SP, Nodari R, Gepts P (1991b) Genetic diversity in cultivated common bean. I. Allozymes. Crop Sci 31:19–23

Singh SP, Morales FJ, Miklas PN, Teran H (2000a) Selection for bean golden mosaic resistance in intra- and interracial bean populations. Crop Sci 40:1565–1572

Singh SP, Morales FJ, Terán H (2000b) Registration of bean golden mosaic resistant dry bean germplasm GMR 1 and GMR 5. Crop Sci 40:1836

Stavely JR, Kelly JD, Grafton KF (1994) BelMiDak-rust-resistant navy dry beans germplasm lines. HortScience 29:709–710

Stavely JR, McMillan RT, Beaver JS, Miklas PN (2001) Release of three McCaslan type, indeterminate, rust and golden mosaic resistant snap bean germplasm lines BelDade RGMR 4, 5 and 6. Ann Rep Bean Improv Coop 44:197–198

Stiller J, Martirani L, Tuppale S, Chian RJ, Chiurazzi M, et al. (1997) High frequency transformation and regeneration of transgenic plants in the model legume *Lotus japonicus*. J Exp Bot 48:1357–1365

Tar'an B, Michaels TE, Pauls KP (2001) Mapping genetic factors affecting the reaction to *Xanthomonas axonopodis* pv. *phaseoli* in *Phaseolus vulgaris* L. under field conditions. Genome 44:1046–1056

Tar'an B, Michaels TE, Pauls KP (2002) Genetic mapping of agronomic traits in common bean. Crop Sci 42:544–556

Tenaillon MI, Sawkins MC, Long AD, Gaut RL, Doebley JF, et al. (2001) Patterns of DNA sequence polymorphism along chromosome 1 of maize (*Zea mays* ssp *mays* L.). Proc Natl Acad Sci USA 98:9161–9166

Tepfer D (1990) Genetic transformation using *Agrobacterium rhizogenes*. Physiol Plant 79:140–146

Thornsberry JM, Goodman MM, Doebley J, Kresovich S, Nielsen D, et al. (2001) *Dwarf8* polymorphisms associate with variation in flowering time. Nat Genet 28:286–289

Urrea C, Miklas P, Beaver J, Riley R (1996) A codominant randomly amplified polymorphic DNA (RAPD) marker useful for indirect selection of bean golden mosaic virus resistance in common bean. J Am Soc Hort Sci 121:1035–1039

Vallad G, Rivkin M, Vallejos C, McClean P (2001) Cloning and homology modelling of a *Pto*-like protein kinase family of common bean (*Phaseolus vulgaris* L.). Theor Appl Genet 103:1046–1058

Vallejos C, Skroch P, Nienhuis J (2001) *Phaseolus vulgaris* – The common bean. Integration of RFLP and RAPD-based linkage maps. In: Phillips R, Vasil I (eds), DNA-based markers in plants. Kluwer, Dordrecht, the Netherlands, pp. 301–317

Vallejos CE, Astua-Monge G, Jones V, Plyler TR, Sakiyama NS, et al. (2006) Genetic and molecular characterization of the *I* Locus of *Phaseolus vulgaris*. Genetics 172:1229–1242 Vallejos CE, Chase CD (1991) Linkage between isozyme markers and a locus affecting seed size in *Phaseolus vulgaris* L. Theor Appl Genet 81:413–419

Vallejos CE, Sakiyama NS, Chase CD (1992) A molecular marker-based linkage map of *Phaseolus vulgaris* L. Genetics 131:733–740

Van de Velde W, Mergeay J, Holsters M, Goormachtig S (2003) *Agrobacterium rhizogenes*-mediated transformation of *Sesbania rostrata*. Plant Sci 165:1281–1288

Vanhouten W, Mackenzie S (1999) Construction and characterization of a common bean bacterial artificial chromosome library. Plant Mol Biol 40:977–983

Vianna GR, Albino MMC, Dias BBA, Silva LdM, Rech EL, Aragão FJL (2004) Fragment DNA as vector for genetic transformation of bean (*Phaseolus vulgaris* L.). Scientia Horticulturae (Amsterdam) 99:371–378

Yaish MWF, de la Vega MP (2003) Isolation of (GA)(n) microsatellite sequences and description of a predicted MADS-box sequence isolated from common bean (*Phaseolus vulgaris* L.). Genet Mol Biol 26:337–342

Yan HH, Mudge J, Kim DJ, Shoemaker RC, Cook DR, et al. (2004) Comparative physical mapping reveals features of microsynteny between *Glycine max*, *Medicago truncatula*, and *Arabidopsis thaliana*. Genome 47:141–155

Yu K, Park S, Poysa V, Gepts P (2000) Integration of simple sequence repeat (SSR) markers into a molecular linkage map of common bean (*Phaseolus vulgaris* L.). J Hered 91:429–434

Yu K, Haffne M, Park SJ (2006) Construction and characterization of a common bean BAC library. Ann Rep Bean Improv Coop 49:61–63

Yu KF, Park SJ, Poysa V (1999) Abundance and variation of microsatellite DNA sequences in beans (*Phaseolus* and *Vigna*). Genome 42:27–34

Yu Z, Stall R, Vallejos C (1998) Detection of genes for resistance to common bacterial blight of beans. Crop Sci 38:1290–1296

Zambre M, Goossens A, Cardona C, Van Montagu M, Terryn N, et al. (2005) A reproducible genetic transformation system for cultivated *Phaseolus acutifolius* (tepary bean) and its use to assess the role of arcelins in resistance to the Mexican bean weevil. Theor Appl Genet 110:914–924

Zhang ZY, Coyne DP, Mitra A (1997) Factors affecting *Agrobacterium*-mediated transformation of common bean. J Am Soc Hort Sci 122:300–305

Zheng J, Nakata M, Uchiyama H, Morikawa H, Tanaka R (1991) Giemsa C-banding patterns in several species of *Phaseolus* L. and *Vigna* Savi, Fabaceae. Cytologia 56:459–466

Zhu YL, Song QJ, Hyten DL, Van Tassell CP, Matukumalli LK, et al. (2003) Single-nucleotide polymorphisms in soybean. Genetics 163:1123–1134

Zizumbo-Villarreal D, Colunga-GarcíaMarín P, Payró de la Cruz E, Delgado-Valerio P, Gepts P (2005) Population structure and evolutionary dynamics of wild–weedy–domesticated complexes of common bean in a Mesoamerican region. Crop Sci 35:1073–1083