

1 Introduction

*Mohar Singh¹, Hari D. Upadhyaya² and
Ishwari Singh Bisht¹*

¹National Bureau of Plant Genetic Resources, Pusa, New Delhi, India;

²International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru, Hyderabad, India

Major grain legumes, including common bean, pea, chickpea, faba bean, cowpea, lentil, pigeon pea, peanut, Asian *Vigna*, grass pea and horsegram, occupy considerable area under cultivation globally and form important constituents of global diets for both vegetarian and nonvegetarian peoples. These grain legumes have the ability to fix nitrogen, which reduces fertilizer use in agriculture, besides their high protein content. Despite this significant role, global production has increased only marginally in the past 50 years. The slow production growth, along with increasing human population and improved buying capacity, has substantially reduced per capita availability of grain legumes. Further, production can be enhanced more if the loss caused by several biotic and abiotic stresses is minimized. To overcome these major constraints, there is a need to identify stable donors in genetic resources for discovering useful genes and alleles and designing crops resilient to climate change. However, excellent performance has been achieved by applying new approaches for germplasm characterization and evaluation like development of core sets, mini-core sets, reference sets and trait-specific subsets, etc. In parallel, genomic resources such as molecular markers including simple sequence repeats (SSRs), single nucleotide polymorphism (SNPs), diversity arrays technology (DArT) and transcript sequences, e.g. expressed sequence tags (ESTs) and short-read transcript sequences, have been developed for important legume crops. It is anticipated that the use of genomic resources and specialized germplasm such as mini-core collection and reference sets will facilitate identification of trait-specific germplasm, trait mapping and allele mining for resistance to various biotic and abiotic stresses and also for useful agronomic traits. Furthermore, the advent of next-generation sequencing technologies coupled with advances in bioinformatics offers the possibility of undertaking large-scale sequencing of crop germplasm accessions, so that modern breeding approaches such as genomic selection and breeding by design can be realized in the coming future for legume genetic enhancement. Here we summarize brief details on the genetic and genomic resources research on important grain legumes.

1.1 Common Bean

The common, or kidney, bean (*Phaseolus vulgaris* L.) is the centrepiece of the daily diet of more than 300 million people. It is the most important food legume, far ahead of other legumes. Nutritionists characterize the common bean as a nearly perfect food, because of its high protein content and high amounts of fibre, complex carbohydrates and other dietary elements. The common bean was domesticated more than 7000 years ago in two centres of origin – Meso-America (Mexico and Central America) and the Andean region. Over the millennia, farmers grew complex mixtures of bean types across various production systems, resulting in a vast array of genetic diversity in common beans with a wide variety of colours, textures and sizes to meet the growing conditions and taste preferences of many different regions. Given current trends in population growth and bean consumption, demand for this crop in Latin America, sub-Saharan Africa and even in Europe and other parts of the world can be expected to grow in the future. International Centre for Tropical Agriculture (CIAT) scientists are convinced that new bean cultivars with higher yields, multiple disease resistance and greater tolerance to drought and low soil fertility will enable farmers to increase bean productivity and achieve greater yield stability. New production technology, together with the bean crop's wide adaptability, will help it remain an attractive option for small-farmer cropping systems. One potent source of solutions to problems in bean production is the great genetic diversity available for research and development in the world *Phaseolus* collection maintained at CIAT's Genetic Resources Unit (GRU) in trust for the Food and Agriculture Organization (FAO). The collection includes over 36,000 entries, of which 26,500 are cultivated *Phaseolus vulgaris*, about 1300 are wild types of common bean (<http://isa.ciat.cgiar.org/urg/main.do?language=en>), and the rest are distant relatives of the common bean. CIAT scientists have also created more manageable core collections. The core collection of domesticated common bean contains about 1400 accessions, while the collection of wild common bean consists of about 100 accessions. In recent years bean researchers at CIAT and in national programs of Latin America and sub-Saharan Africa have been evaluating the core collection for a wide range of useful traits, such as insect and disease resistance and tolerance to low phosphorus. Useful materials have been identified and incorporated into breeding programs at CIAT and elsewhere.

While focusing mainly on dry beans, CIAT scientists are also working to improve the green snap beans. Demand for fresh snap beans for domestic consumption or export is growing in Africa, Asia and Latin America, and sales are an excellent source of cash income for small farmers. Much of the CIAT's strategic research on dry beans, especially that dealing with diseases and pests, is readily applicable to snap beans. Classical breeding within the primary gene pool of common bean has given excellent results in the last two decades, with tangible benefits to the farming community. More recently, CIAT scientists have begun to integrate various biotechnology techniques into problem-solving research on the crop. CIAT scientists have succeeded in hybridizing common bean with the distantly related species *Phaseolus*

acutifolius, or tepary bean, which possesses genes for resistance to common bacterial blight (CBB), leafhoppers and drought. The resulting breeding lines have shown high levels of resistance to CBB. CIAT researchers have also developed a molecular marker-assisted approach to improving beans for resistance to bean golden mosaic virus (BGMV) that has cut breeding time and effort by about 60%. The results of recent molecular marking and selection work are highly encouraging, demonstrating not only the effectiveness of the strategy by the Standing Committee on Agricultural Research (SCAR) for selecting BGMV-resistant beans but also its efficiency.

1.2 Pea

Pea (*Pisum sativum* L.) is one of the world's oldest domesticated crops. Its area of origin and initial domestication lies in the Mediterranean, primarily in the Middle East. The range of wild representatives of *P. sativum* extends from Iran and Turkmenistan through Anterior Asia, northern Africa and southern Europe. The genus *Pisum* contains the wild species *P. fulvum* found in Jordan, Syria, Lebanon and Israel; the cultivated species *P. abyssinicum* from Yemen and Ethiopia, which was likely domesticated independently of *P. sativum*; and a large and loose aggregate of both wild (*P. sativum* subsp. *elatius*) and cultivated forms that comprise the species *P. sativum* in a broad sense.

Currently, no international organization conducts pea breeding and genetic resources conservation, and no single collection predominates in size and diversity. Important genetic diversity collections of *Pisum* with over 2000 accessions are found in national gene banks in at least 15 countries, with many other smaller collections worldwide. A high level of duplication exists between the collections, giving a misleading impression of the true level of diversity. However, the numbers of original pea landraces mainly from Europe, Asia, the Middle East and North Africa/Ethiopia have not been documented. The much smaller collections of wild relatives of pea are less widely distributed; there is more clarity when tracing these accessions to their origin. There are still important gaps in the collections, particularly of wild and locally adapted materials, that need to be addressed before these genetic resources are lost forever (Maxted, Shelagh, Ford-Lloyd, Dulloo, & Toledo, 2012). Many studies have been conducted on *Pisum* germplasm collections to investigate genetic and trait diversity. Several major world pea germplasm collections have been analysed by molecular methods and core collections were formed. The key priority is the collection and conservation of the historic landraces and varieties of each country in *ex situ* gene banks. The overall goal should be to ensure maintenance of variation for adaptation to the full range of agro-ecological environments, end uses and production systems. Wild peas have less than 3% representation in various national collections despite their wide genetic diversity. There is an urgent need to fully sample this variation, since natural habitats are being lost due to increased human population, increased grazing pressure, conversion of marginal areas to agriculture and ecological threats due to future climate change. It is urgent to implement

a comprehensive collection of wild relatives of peas representing the habitat range from the Mediterranean through the Middle East and Central Asia while these resources are still available, since these are likely to contain genetic diversity for abiotic stress tolerance. Genetic diversity available in wild *Pisum* species has been poorly exploited. The most attention has been given to *P. fulvum* as a donor of bruchid resistance and source of novel powdery mildew resistance (*Er3*). Relatively few genotypes with high degree of relatedness have been used as parents in modern pea breeding programs, leading to a narrow genetic base of cultivated germplasm. There are several current efforts to make either genome-wide introgression lines or at least simple crosses with the intent of broadening the genetic base. Further investigations, particularly in the wild *P. sativum* subsp. *elatius* gene pool, are of great practical interest. Molecular approaches will allow breeders to avoid the linkage drag from wild relatives and make wide crosses more successful and practical.

1.3 Chickpea

The genus *Cicer* comprises one cultivated and 43 wild species. Chickpea probably originated from southeastern Turkey. Four centres of diversity were identified in the Mediterranean, Central Asia, the Near East and India, as well as a secondary centre of origin in Ethiopia. Further, chickpeas spread with human migration toward the west and south via the Silk Route. It is grown and consumed in large quantities from Southeast Asia to India and in the Middle East and Mediterranean countries. It ranks second in area and third in production among the pulses worldwide. Most production and consumption of chickpea takes place in developing countries. It is a true diploid and predominantly self-pollinated legume, but cross-pollination by insects sometimes occurs. Thirty five of the chickpea wild relatives are perennials and the other nine (including the cultivated species) are annuals. Based on seed size and shape, two main kinds of chickpea are recognized: the desi type, closer to the putative progenitor (*C. reticulatum*), is found predominantly in India and Ethiopia and has small, angular, coloured seeds and a rough coat. They have a bushy growth habit and blue-violet flowers. The kabuli type, predominantly grown in the Mediterranean region, has large, beige-coloured and owl-head-shaped seeds with a smooth seed coat. Their plants have a more erect growth habit and white flowers. It is estimated that more than 80,000 accessions are conserved in more than 30 gene banks worldwide (http://apps3.fao.org/wIEWS/germplasm_query.htm?i_l=EN). The gene bank at International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), India, is one of the largest gene banks, holding greater than 20,000 accessions of chickpea from about 60 countries. Other major collections (more than 12,000 accessions) are held at the National Bureau of Plant Genetic Resources (NBPGR), New Delhi, India; International Center for Agricultural Research in the Dry Areas (ICARDA) in Aleppo, Syria; Australian Temperate Field Crops Collection, Victoria, Australia; the United States Department of Agriculture (USDA); and the Seed and Plant Improvement Institute, Iran. Currently there is a reasonable number of wild annual

Cicer species, but still limited availability of perennial species. Less than 1% of the *Cicer* accessions (conserved in about 10 gene banks worldwide) represent wild species. Priority should be given to the conservation of chickpea in primary and secondary centres of diversity. *Cicer* genetic resources could be much better utilized. A representative core collection (10% of the entire collection) and a mini-core collection (10% of the core or 1% of entire collection) are being developed at the ICRISAT and evaluated extensively for useful traits (Upadhyaya & Ortiz, 2001). However, recent advances in plant biotechnology have resulted in the development of a large number of molecular markers, genetic and physical maps, as well as the generation of expressed sequenced tags, genome sequencing and association studies showing marker–trait associations, which has facilitated the identification of quantitative trait loci (QTLs) and discovery of genes/alleles associated with resistance to several abiotic and biotic stresses, beside agronomic traits.

1.4 Faba Bean

Faba bean (*Vicia faba* L.) is a major food and feed legume, because of the high nutritional value of its seeds, which are rich in protein and starch. Seeds are consumed dry, fresh, frozen or canned. The main faba bean producer countries are China, some in Europe, Ethiopia, Egypt and Australia. Geographical distribution and objectives of the breeding programs developed for this species therefore reflect where consumption is highest. In relation to the size of the market and in comparison with soybean, the faba bean selection programs are few and small. The role of *ex situ* and on-farm collections is even stronger for this crop due to the absence of a natural reservoir of wild accessions and to the modernization of agriculture, which progressively phases out numerous landraces. Botanic and molecular data suggest that the wild ancestor of this species has not yet been discovered or has become extinct. At the world level, more than 38,000 accession entries are included in about 37 listed germplasm collections. A large genetic variability has already been identified in *V. faba* in terms of floral biology, seed size and composition, and also tolerance to major biotic and abiotic stresses. More knowledge is needed on the interactions of *V. faba* with parasitic and pollinator insects, on traits related to environmental adaptation and impacts on nitrogen fixation in interaction with soil rhizobia and on bioenergy potential, which strengthens the demand for new and large phenotyping actions. Diversity analysis through genotyping is just beginning. The use of amplified fragment length polymorphism (AFLP) or SSR markers has allowed genetic resources to be distinguished according to their geographical origin and structuring of germplasm collections. Conservation of gene sequences among legume species and the rapid discovery of genes offer new possibilities for the analysis of sequence diversity for *V. faba* genes and evaluation of their impact on phenotypic traits. Projects that combine genotyping and phenotyping must be continued on *V. faba*, so that core collections can be defined; these will help in the discovery of genes and alleles of interest for faba bean breeders (Rispaill et al., 2010).

1.5 Cowpea

Cowpea (*Vigna unguiculata* (L.) Walp.) is cultivated widely in the tropics and has multipurpose uses: as food for human beings, fodder for livestock and atmospheric nitrogen fixers. Cowpea grains rich in protein are consumed in different forms in several parts of the tropics. The average grain yield of cowpea in West Africa is approximately 492 kg ha^{-1} , which is much lower than its potential yields. This low productivity is due to a host of diseases, insects, pests, parasitic weeds, drought, poor soils and low plant population density in the farmer's field. Despite a large number of cowpea accessions (about 15,000) maintained at the International Institute of Tropical Agriculture (IITA), recent studies demonstrated that genetic diversity in cultivated cowpea is low. Researchers, however, found a high level of random amplified polymorphic DNA marker diversity in landraces from Malawi. However, *ex situ* collection of cowpea and wild *Vigna* germplasm from different parts of the world were assembled in the IITA gene bank. These genetic resources have been explored to identify new traits and to develop elite cowpea varieties. Many cowpea varieties with high yield potential have been developed and adopted by the farmers. Efforts are continuing to develop better-performing varieties using conventional breeding procedures, while molecular tools are being developed to facilitate progress in cowpea breeding (Agbicodo et al., 2010).

1.6 Lentil

Lentils have been part of the human diet since Neolithic times, being one of the first crops domesticated in the Near East. Archaeological evidence reveals that they were eaten 9500–13,000 years ago. Lentil colours range from yellow to red-orange to green, brown and black. Lentils also vary in size, and are sold in many forms, with or without the skins, whole or split. Lentils are relatively tolerant to drought and are grown throughout the world. The FAO has reported that the world production of lentils primarily comes from Canada, India, Turkey and the United States. About a quarter of the worldwide production of lentils is from India, most of which is consumed in the domestic market. Canada is the largest export producer of lentils in the world.

Extensive collections of lentil germplasm now exist in various gene banks around the world. This germplasm including wild *Lens* species has been used in plant introduction strategies and in efforts to widen the potential sources of increasing genetic diversity in the breeding programmes of lentil. Improved techniques are emerging to overcome hybridization barriers between species, and as a result interspecific hybrids have been successfully obtained between species. Several interspecific recombinant inbred line populations have been developed. Selected and backcrossed lentil lines are currently in advanced yield trial stages, and desirable traits such as yield, disease resistance and agronomic traits have been incorporated into cultivated lentil especially from *Lens ervoides*, generating a wider spectrum of variability. Secondly, further expansion of the overall pool of germplasm and examination of

allelic variation at the nucleotide level will benefit lentil-breeding programmes by augmenting phenotype-based variation to further advance cultivar development. Genomic resources for lentils are limited now, but this situation is changing rapidly as the cost of genotyping has declined. As a result, two successive EST projects were undertaken under the NAPGEN EST project initiative and an Agricultural Development Fund project initiative. It has been emphasized that creation of intraspecific and interspecific genetic populations, genetic maps, association maps, QTLs and marker-assisted selection technologies for implementation in the breeding programme will enhance deployment of genes responsible for traits of interest. The economical use of genomic technologies for use in germplasm resource management and genetic improvement is on the near horizon.

1.7 Pigeon Pea

Pigeon pea (*Cajanus cajan* (L.) Millspaugh) is an important grain legume of the Indian subcontinent, Southeast Asia and East Africa. More than 85% of the world pigeon pea is produced and consumed in India, where it is a key crop for food and nutritional security of the people. The centre of origin is the eastern part of peninsular India, including the state of Orissa, where the closest wild relatives occur. Though pigeon pea has a narrow genetic base, vast genetic resources are available for its genetic improvement. The ICRISAT gene bank maintains about 13,216 accessions, whereas the Indian NBPGR bank maintains a total of about 12,900 accessions. Evaluation of small-sized subsets such as core (10% of whole collection) and mini-core (about 1% of the entire collection), developed at the ICRISAT, has resulted in identification of promising diverse sources for agronomic and nutrition-related traits, as well as resistance to major biotic and abiotic stresses for use in pigeon pea improvement programs. Wild relatives of pigeon pea are the reservoir of several useful genes, including resistance to diseases, insect pests and drought, as well as good agronomic traits, and have contributed to the development of cytoplasmic male sterility systems for pigeon pea improvement. Availability of genomic resources, including the genome sequence, will facilitate greater use of germplasm to develop new cultivars with a wider genetic base.

1.8 Peanut

The domesticated peanut (*Arachis hypogaea* L.) is an amphidiploid or allotetraploid having two sets of chromosomes from two different species, thought to be *A. duranensis* and *A. ipaensis*. These likely combined in the wild to form the tetraploid species *A. monticola*, which gave rise to the domesticated peanut. This domestication may have taken place in Paraguay or Bolivia, where the wildest strains grow today. Certain cultivar groups are preferred for particular uses based on differences in flavour, oil content, size, shape and disease resistance. For many uses, the different

cultivars are interchangeable. Most peanuts marketed in the shell are of the Virginia type, along with some Valencias selected for large size and the attractive appearance of the shell. Spanish peanuts are used mostly for peanut candy, salted nuts and peanut butter. Most runners are used to make peanut butter. Although India and China are the world's largest producers of peanuts, they account for a small part of international trade, because most of their production is consumed domestically as peanut oil. Exports of peanuts from India and China are equivalent to less than 4% of world trade. The major producers/exporters of peanuts are the United States, Argentina, Sudan, Senegal and Brazil. These five countries account for 71% of total world exports. In recent years, the United States has been the leading exporter of peanuts in the world.

Further, the number of accessions in the ICRISAT gene bank are about 13,500. Most of them have been characterized and evaluated for their reaction to diseases, insect pests and other desirable agro-morphological characteristics, leading to identification of 506 useful genetic stocks. Most of the germplasm is conserved as pods or seeds in the gene bank, while rhizomatous *Arachis* species are conserved as whole plants. ICRISAT serves as the world's largest repository of peanut germplasm and has distributed about 60,000 peanut germplasm samples free of cost to the international scientific community.

Despite significant progress, peanut genetic resource activities still suffer from several limitations in assembly and characterization. The establishment of a peanut genetic resources network is proposed to overcome many such limitations. However, sufficient numbers of molecular markers that reveal polymorphism in cultivated peanut are available for diversity assessments. In a study, the amount and distribution of genetic variation within and among six peanut botanical varieties, as well as its partitioning among three continents of origin (South America, Asia and Africa) was assessed at 12 SSR loci by means of 10 sequence-tagged microsatellite site primers. Discriminant function analysis reveals a high degree of accordance between variety delimitation on the basis of morphological and molecular characters. Landraces from Africa and Asia were more closely related to each other than to those from South America. Nei's unbiased estimate of gene diversity revealed very similar levels of diversity within botanical varieties. Landraces from South America had the highest diversity and possessed 90% of alleles, compared with Africa (63%) and Asia (67%).

1.9 Asian *Vigna*

Asian *Vigna* species constitute an economically important group of cultivated and wild species, and a rich diversity occurs in India and other Asian countries. Taxonomically, cultigen and conspecific wild forms are recognized in all major cultivated Asiatic pulses, mung bean (*V. radiata*), urd bean (*V. mungo*), rice bean (*V. umbellata*) and azuki bean (*V. angularis*) except for moth bean (*V. aconitifolia*), which has retained a wild-type morphology. The cultivated species, *V. radiata* and *V. mungo*, are of Indian origin. The domestication of *V. aconitifolia* is also apparently Indian, whereas that

of *V. angularis* and *V. umbellata* is Far Eastern. The green gram is already a popular food throughout Asia and other parts of the world. The present level of its consumption can be expected to increase. The black gram, although very popular in India, is less likely to generate sufficient demand to stimulate production significantly outside its traditional areas. The azuki bean has generated interest as a pulse outside traditional areas of production and consumption, and consumer demand for it could increase in the near future. Perhaps the most interesting future exists for rice bean, which has a high food value and tolerance to biotic and abiotic stresses. It possibly has the highest yielding capacity of any of the Asian *Vigna* and could become a useful crop, if a sizeable consumer demand were built up. Moth bean has a future in India as a pulse crop. *V. trilobata* is probably most useful as a forage crop in semi-arid conditions. The fullest possible range of landraces and cultivars needs to be collected and conserved together with the conspecific wild-related species. The wild germplasm resources have a great potential for widening the genetic base of the *Vigna* gene pool by interspecific hybridization. The available genetic resources with valuable characters will therefore be required to make extended cultivation economically attractive.

1.10 Grass Pea

Grass pea presents a fascinating paradox; it is both a lifesaver and a destroyer. It is easily cultivated and can withstand extreme environments from drought to flooding. However, when eaten as a large part of the diet over a long enough period (which is often the case during famine), it can permanently paralyse adults from the knees down and cause brain damage in children, a disorder named lathyrism. Grass pea has a long history in agriculture. It was first domesticated some 7000–8000 years ago in the eastern Mediterranean region and has a history of cultivation in southern parts of Europe, North Africa and across Asia. Today it is mostly grown in India, Pakistan, Bangladesh and Ethiopia. More recently, grass pea has become popular as a forage crop in Kazakhstan, Uzbekistan, South Africa and Australia.

Recently ICARDA at Aleppo, Syria, together with Ethiopian breeders, has undertaken a project to develop cultivars with low neurotoxin levels. The role of diversity in breeding programmes was instantly clear: the toxins found in African and Asian grass pea plants are seven times more toxic than Middle Eastern types. The Centre for Legumes in Mediterranean Agriculture (CLIMA) in Australia has also recently produced a low-toxin grass pea variety. The use of grass pea diversity in breeding has shown how the genetic resources of a crop can be used to improve its nutritional value for human health. The ICARDA scientists used the diversity found in the world's largest collection of grass pea and its relatives, stewarded at the ICARDA seed bank in Syria, with more than 3000 accessions. Large *Lathyrus* collections are also conserved in France, NBPGR in India, Bangladesh and Chile. Despite this research, much additional work is needed in order to produce locally adapted, low-toxin varieties and to distribute these to the farming community. Furthermore, there is a need to expand the molecular research work in species identification and their proper utilization in grass pea breeding.

1.11 Horsegram

Horsegram (*Macrotyloma uniflorum*) is one of the lesser known grain legume species. The whole seeds of horsegram are generally utilized as cattle feed. However, it is consumed as a whole seed, as sprouts, or as whole meal in India. It is quite a popular legume, especially in southern Indian states such as Karnataka, Tamil Nadu, Andhra Pradesh, northwestern Himalayan states and Uttarakhand. The chemical composition is comparable with more commonly cultivated legumes. Like other legumes, horsegram is deficient in methionine and tryptophan, though it is an excellent source of iron and molybdenum. Horsegram is also known to have many therapeutic effects – not scientifically proven – though it has been recommended in ayurvedic medicine to treat renal stones, piles, oedema, etc. A total of 1721 accessions of horsegram are being conserved in different gene banks of the world. Of these collections, about 95% are conserved at NBPGR, New Delhi, India, and its regional research station, Thrissur, Kerala, is designated as an active site for the conservation and evaluation of horsegram germplasm. No worthwhile genomic resource information on horsegram is available.

References

- Agbicodo, E. M., Fatokun, C. A., Bandyopadhyay, R., Wydra, K., Diop, N. N., Muchero, W., et al. (2010). Identification of markers associated with bacterial blight resistance loci in cowpea [*Vigna unguiculata* (L.) Walp.]. *Euphytica*, 175, 215–226.
- Maxted, N., Shelagh, K., Ford-Lloyd, B., Dulloo, E., & Toledo, A. (2012). Toward the systematic conservation of global crop wild relative diversity. *Crop Science*, 52(2), 774–785.
- Rispail, N., Kalo, P., Kiss, G. B., Ellis, T. H. N., Gallardo, K., Thompson, R. D., et al. (2010). Model of legumes to contribute to faba bean breeding. *Field Crops Research*, 115, 253–269.
- Upadhyaya, H. D., & Ortiz, R. (2001). A mini core collection for capturing diversity and promoting utilization of chickpea genetic resources in crop improvement. *Theoretical and Applied Genetics*, 102, 1292–1298.