ISSN 1479-2621

Plant Genetic Resources **4**(1); 13–19
DOI: 10.1079/PGR2005101

Development of a composite collection for mining germplasm possessing allelic variation for beneficial traits in chickpea

H. D. Upadhyaya[1]*, B. J. Furman[2], S. L. Dwivedi[1], S. M. Udupa[2], C. L. L. Gowda[1], M. Baum[2], J. H. Crouch[1†], H. K. Buhariwalla[1†] and S. Singh[1]

[1] *International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru, 502 324, Andhra Pradesh, India and* [2] *International Center for Agricultural Research in the Dry Areas (ICARDA), P.O. Box 5466, Aleppo, Syrian Arab Republic*

Received 1 September 2005; Accepted 4 November 2005

Abstract

Chickpea is one of the most important grain legume crops in the world. Large collections of genetic resources are maintained in the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT) and International Center for Agricultural Research in the Dry Areas (ICARDA) genebanks. Association mapping using neutral markers has been suggested as a means to identify useful alleles in the vast reservoirs of genetic diversity existing in the germplasm collections that could be associated with the phenotypes among the population individuals. ICRISAT in collaboration with ICARDA developed a global composite collection of 3000 accessions that will be profiled using 50 polymorphic simple sequence repeat (SSR) markers. The data generated through this collaborative effort will be used to define the genetic structure of the global composite collection and to select a reference sample of 300 accessions representing the maximum diversity for the isolation of allelic variants of candidate gene associated with beneficial traits. It is then expected that molecular biologists and plant breeders will have opportunities to use diverse lines in functional and comparative genomics, in mapping and cloning gene(s), and in applied plant breeding to diversify the genetic base of the breeding populations which should lead to the development of broad-based elite breeding lines/cultivars with superior yield and enhanced adaptation to diverse environments.

Keywords: chickpea composite collection; gene mining; Generation Challenge Program; germplasm; molecular markers

The crop

Chickpea (*Cicer arietinum* L.) is the fourth largest grain-legume crop in the world, with a total production of 8.57 Mt from an area of 10.38 Mha (Food and Agriculture Organization, http://faostat.fao.org), predominantly in Asia (89.5%), with minor production from Africa (4.4%), North and Central America (3.4%), Oceania (1.6%) and Europe (1.0%). Although North and Central America and Oceania together only contribute about 5% of world production, their productivity is higher than elsewhere. Asia and Africa show low productivity but contribute about 94% of world production. Major producers are India, Turkey, Pakistan and Iran in Asia; Ethiopia in Africa; and Mexico in the Americas. Chickpea productivity records over the last four decades (1965–1974, 1975–1984, 1985–1994 and 1995–2004) show a consistent increase in

* Corresponding author. E-mail: h.upadhyaya@cgiar.org
† Present address: International Centre for Maize and Wheat Improvement (CIMMYT), Apdo. Postal 6-641, 06 600 Mexico, D.F., Mexico.

India and Mexico, but a decline in Turkey, Pakistan and Iran. Ascochyta blight [*Ascochyta rabiei* (Pass.) Labr.], botrytis grey mould (*Botrytis cinerea* Pers. ex Fr.), fusarium wilt [*Fusarium oxysporum* Schlecht. emend Snyd. & Hans. f. sp. *ciceri* (Padwick) Snyd. & Hans], black root rot [*Fusarium solani* (Mart.). Sacc.], dry root rot [*Rhizoctonia bataticola* (Taubenhaus) E.J. Butler], pod borer (*Helicoverpa armigera* Hübner) and leaf miner (*Liriomyza cicerina* Rondani) are the most important biotic stresses; while drought, salinity and fluctuations in temperature (both extremes) are the abiotic stresses which impose the major constraints to chickpea productivity.

Generation Challenge Program

The revolution in molecular biology, bioinformatics and information technology has provided the scientific community with tremendous opportunities for solving some of the world's most serious agricultural and food security issues, and has led to the formation of the Generation Challenge Program (GCP) 'Unlocking Genetic Diversity in Crops for the Resource-poor' (www.generationcp.org). The GCP aims to utilize molecular tools and comparative biology to explore and exploit genetic diversity housed in existing germplasm collections, with particular focus on drought tolerance. The development of gene-based markers based on information derived from model plants is a key component. An important goal of the GCP is extensive genetic characterization, using molecular markers, of the genetic resources held by the participating institutions. The rationale for dissecting crop genetic diversity is to identify genes which can be used to reduce the impacts of environmental and biotic stresses, thereby enhancing yield and improving nutritional quality of crop products.

Association mapping

The purpose of allele mining is to identify useful alleles present in germplasm collections. The strategy is to establish regions of the genome associated with critical phenotypes by association or linkage disequilibrium mapping. The approach relies on the assumption that an allele responsible for a phenotype, along with the markers which flank the locus, are inherited as a block, and therefore neutral marker-based selection will be predictive of allelic content at critical genes determining favourable phenotype. The prerequisites to perform such association mapping include a dense genetic linkage map, passport information and phenotypic data, an understanding of population structure, and contrasting genotypes for beneficial traits (Kresovich *et al.*, 2002). Association between genotype at DNA markers and phenotype with respect to agronomic characters in a collection of plant genetic resources would allow (i) an assessment of the genetic potential of specific genotypes prior to phenotypic evaluation, (ii) the identification of superior trait alleles in germplasm collection, (iii) high-resolution quantitative trait locus (QTL) mapping, and (iv) some validation of candidate genes responsible for quantitative agronomic traits (Gebhardt *et al.*, 2004). In recent years, several studies conducted in plants have detected DNA markers associated with ecology, geography, disease resistance and quantitative traits (Sun *et al.*, 2001, 2003; Thornsberry *et al.*, 2001; Turpeinen *et al.*, 2001; Ivandic *et al.*, 2002, 2003; Russell *et al.*, 2003; Amirul Islam *et al.*, 2004; Gebhardt *et al.*, 2004; Kraakman *et al.*, 2004; Sabharwal *et al.*, 2004), demonstrating that association mapping can be a viable alternative to standard QTL analyses based on crosses between pairs of lines.

Enhancing germplasm use in crop improvement

Large collections of chickpea germplasm are maintained by two Consultative Group on International Agricultural Research (CGIAR) institutions: the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), India and the International Center for Agricultural Research in the Dry Areas (ICARDA), Syria. The former maintains 17,258 accessions (135 wild, 17,123 cultivated) and the latter 12,647 (304 wild, 12,343 cultivated). Despite such an impressive number of accessions, there has been only limited use of genebank materials for the genetic enhancement of chickpea. For example, during the period 1978 to 2004, the ICRISAT chickpea breeders used just 83 germplasm lines, in contrast to their use of 480 breeding lines for the development of 3430 advanced varieties (ICCV) during the same period. A similar experience applied at ICARDA, where during the same period approximately 250 germplasm lines were used in crosses, compared to approximately 600 breeding lines in generating breeding materials from which 31 varieties were released. India, as the largest producer of chickpea, has a strong improvement programme, which has released 126 cultivars between 1967 and 2003. Pedigree analysis of 86 cultivars developed from crosses has revealed that although 95 progenitors were involved, just 10 of these contributed 35% of the genetic base (Shiv Kumar *et al.*, 2004). The five most frequently used ancestors were Pb7, IP 58, F 8, Rabat and S 26. Furthermore, about 41% of the cultivars developed through crossing have Pb7 as an ancestor. This suggests that many cultivars share a narrow genetic base.

To enhance the utilization of genetic resources in chickpea improvement programmes, Upadhyaya *et al.* (2001)

developed a core subset of 1956 accessions based on geographic distribution and quantitative traits of 16,691 accessions maintained at ICRISAT. A statistical analysis has indicated that both the genetic variation with respect to 13 quantitative traits (days to 50% flowering, plant height and width, days to maturity, number of basal primary branches, number of apical primary branches, number of basal secondary branches, number of apical secondary branches, number of tertiary branches, pods per plant, seeds per pod, 100-seed weight, and seed yield) and phenotypic correlations among traits under the control of co-adapted gene complexes, were preserved in this core collection.

Developing a global composite collection

Through a GCP-commissioned grant, ICRISAT and ICARDA jointly developed a global composite collection of 3000 accessions (Table 1). The composite collection included the 1956 accessions of the ICRISAT core collection, 709 ICARDA cultivated genebank accessions, 39 advanced breeding lines and released cultivars, 35 distinct morphological variants, 20 wild species (*C. echinospermum* and *C. reticulatum*) accessions and 241 accessions carrying specific traits such as tolerance/resistance to biotic and abiotic stress, important agronomic characters (early maturity, multi-seeded pods, double podded, large-seed size, high seed protein, nodulation and responsiveness to high-input conditions) (Table 2). The selection of accessions from ICARDA was based on a classical hierarchical cluster analysis using quantitative traits and two-step cluster analyses using agroclimatological data linked to the geographical coordinates of the accessions' collection site. This global composite collection is composed of 80% landraces, 9% advanced breeding lines, 2% cultivars, 1% wild species and 8% for which precise status is unknown (Table 3). Geographically, 39% of the composite collection originates from South and South-East Asia, 25% from West Asia and 22% from the Mediterranean region. Africa and the Americas each contribute 5% of the collection. We believe that a wide spectrum of genetic diversity has been captured in this global composite collection of 3000 accessions. All accessions are FAO designated and are held in trust in both ICRISAT and/or ICARDA, and are available upon request to chickpea researchers via an appropriate Material Transfer Agreement.

Table 1. Characteristics of germplasm included in a global composite collection of chickpea

Germplasm/traits	No. of accessions
Accessions from ICRISAT	
Core collection	1956
Cultivars/breeding lines	39
Ascochyta blight	13
Botrytis grey mould	8
Stunt	8
Fusarium wilt	50
Collar rot	9
Black root rot	8
Dry root rot	6
Helicoverpa	16
Leaf miner	5
Nematode	8
Low temperature	12
High temperature	4
Drought	10
Salinity	4
Early maturity	25
High protein	10
Multi-seeded pods	7
Seed size	18
High-input responsive	4
Twin pods	8
Nodulation	8
Morphological diversity	35
Accessions from ICARDA	
Based on characterization and evaluation data	599
Based on agroclimatological data	110
Cicer echinospermum	7 (1 from ICRISAT)
Cicer reticulatum	13 (2 from ICRISAT)

Selecting markers for association mapping

From a preliminary screening of 200 simple sequence repeats (SSRs) (Hüttel *et al.*, 1999; Winter *et al.*, 1999; Niroj *et al.*, 2003) on a mini core collection of 211 accessions (Upadhyaya and Ortiz, 2001), 50 polymorphic SSRs were selected (Table 4). Profiling of the 3000 accessions will be carried out for 35 loci at ICRISAT and 15 at ICARDA. An initial analysis at ICRISAT with 35 SSR markers against 288 chickpea accessions (the 211 mini core, along with 20 wild *Cicer* species and 57 kabuli-type accessions) delivered a mean of 25 alleles per locus (range 6–44). The dinucleotide repeat motifs in general detected a lower number of alleles per locus (mean = 11) compared to those based on three nucleotides (mean = 27); this led to differing estimates of gene diversity (0.72 for the dinucleotide repeats and 0.90 for the trinucleotide repeats). A similar analysis at ICARDA based on 15 trinucleotide SSR markers revealed high diversity (mean = 0.91; range 0.77–0.96) at loci showing a high number of alleles per locus (mean = 28.4; range 14–55).

Utilizing diverse accessions for functional genomics, gene tagging and genetic enhancement

The data generated from this collaborative venture will be used to define the genetic structure of the global chickpea

Table 2. Details of the trait-based accessions included in a global composite collection of chickpea

Trait	No. of accessions	Accession's identity
Morphological variants	35	ICC# 1014, 1032, 2828, 4934, 4969, 4991, 4992, 5316, 5319, 5320, 5325, 5780, 5783, 6119, 6146, 8319, 8400, 9816, 10035, 10301, 12031, 12951, 12952, 13812, 14321, 14333, 14335, 14340, 14447, 14872, 16341, 16359, 16949, 4918N, 4957N
Early maturity	25	ICC# 1097, 2023, 2171, 2859, 5810, 6919, 8378, 8618, 8931, 10629, 10822, 10926, 10976, 10981, 10996, 11021, 11039, 11040, 11059, 11160, 12424, 16644, 16947, 17258 and ICCV 96030
Multi-seeded pods	7	ICC# 1315, 1813, 2593, 3832, 3850, 5857, 7225
Double pods	8	ICC# 583, 1237, 5020, 7302, 8284, 9644, 10919, 11510
Seed size	18	Large-seeded types: ICC# 4878, 7344, 7672, 12498, 13787, 14194, 14204, 14205, 14926, 15331, 15554, 15574, 16670, 16674, 16744, 17109 Small-seeded type: ICC 5002, ICC 11255
Nodulation	8	High nodulating: 4948HN, 5003HN Low nodulating: 4948LN, 5003LN Non-nodulating: ICCV 2NN, ICCs 435NN, 4918NN, 4993NN
High-input responsive	4	ICCV# 95605, 95701, 95703, 95705
High protein	10	ICC# 1932, 5384, 5912, 6969, 8397, 10600, 11393, 11599, 12288, 14315
High temperature	4	ICCV# 91902, 95981, 95982, ICC 17256
Low temperature	12	ICC# 3491, 4889, 5623, 6355, 7150, 8508, 8923, 11282, 16349, 16350, 16351 and ICCV 92503
Salinity	4	ICCL# 82115, 83135 and ICCV# 95332, 95334
Drought	10	ICC# 4958, 10386, 10387, 10939, 10942, 10943, 10950, 10951, 10953, 10963
Nematode	8	ICC# 311, 315, 343, 353, 455, 4984, 8932, 10130
Leaf miner	5	ICC# 7980, 7993, 9443, 9483, 11905
Helicoverpa	16	ICC# 12475, 12477, 12479, 12480, 12482, 12486, 12487, 12488, 12491, 12493, 12494, 12495, 12496, 14876 and ICCL 86102 and ICCX 730041
Fusarium wilt	50	ICC# 184, 229, 338, 342, 1246, 1405, 2104, 2595, 4928, 5535, 5901, 11223, 11224, 11312, 11318, 11321, 11322, 11324, 11550, 11554, 12233, 12235, 12237, 12242, 12246, 12248, 12251, 12253, 12258, 12259, 12267, 12268, 12270, 12273, 12289, 12428, 12430, 12431, 12435, 12440, 12450, 12452, 12454, 12472, 14344, 14364, 14366, 14368, 14369, 14371
Dry root rot	6	ICC# 11088, 11315, 12269, 12437, 14440, 14449
Black root rot	8	ICC# 12274, 12275, 14411, 14425, 14426, 14444, 14450, 14451
Collar rot	9	ICC# 344, 542, 618, 684, 1696, 4709, 9934, 14282, 14391
Botrytis grey mould	8	ICC# 1084, 1102, 3540, 4018, 4065, 4075, 6671, 12512
Ascochyta blight	13	ICC# 652, 1929, 3864, 4063, 12955, 12965, 14912, 14915, 14917, 15973, 15975, 15978, 17000
Stunt	8	ICC# 403, 685, 693, 1136, 2546, 3718, 6433, 10495

Table 3. Geographical origin and biological status of the accessions selected to form a global composite collection in chickpea

Region	Country (number of accessions)	Biological status					
		Landraces	Cultivars	Advanced lines	Unknown types	Wild species	Total
Africa	Ethiopia (124), Kenya (1), Malawi (18), Nigeria (2), Sudan (5), Tanzania (2), Uganda (1)	146		6	1		153
North and Central America	Mexico (65), USA (30)	71	2	7	15		95
South America	Chile (45), Colombia (1), Ecuador (1), Peru (3)	26		4	20		50
South and South-East Asia	Bangladesh (32), China (14), ICRISAT-India (151), India (821), Myanmar (25), Nepal (20), Pakistan (98), Sri Lanka (2)	882	46	227	8		1163
West Asia	Afghanistan (147), Armenia (2), Azerbaijan (6), Georgia (2), Iran (553), Iraq (15), Kazakhstan (2), Kyrgyzstan (3), Tajikistan (3), Uzbekistan (7)	733		1	6		740
Russian Federation	Former Soviet Union (45)	39			6		45
Mediterranean	Algeria (20), Cyprus (19), Egypt (18), France (12), Greece (9), ICARDA-Syria (1), Israel (10), Italy (18), Jordan (32), Lebanon (12), Libyan Arab Jamahiriya (2), Morocco (68), Palestine (4), Spain (41), Syrian Arab Republic (184), Tunisia (35), Turkey (162), Yugoslavia (3)	440	2	15	173	20	650
Europe	Bulgaria (17), Former Czechoslovakia (6), Germany (3), Hungary (2), Republic of Moldova (4), Portugal (23), Romania (2), Ukraine (8), UK (1)	47			19		66
Oceania	Australia (3)	2	1				3
Unknown	Unknown (35)	29			6		35
Total	58 countries, ICRISAT and ICARDA, and Unknown	2415	51	260	254	20	3000

Table 4. Fifty SSR primers selected for assessing molecular diversity in a global composite collection in chickpea

Marker[a]	Allele size (bp)	Annealing temperature[b] (°C)
Niroj *et al.* (2003)		
NCPGR4	195	65–60
NCPGR6	253	65–60
NCPGR7	222	65–60
NCPGR12	251	65–60
NCPGR19	306	65–60
Winter *et al.* (1999)		
TA 14	250	65–60
TA 21	347	65–60
TA 22	228	65–60
TA 27	241	65–60
TA 28	300	65–60
TA 46	152	60–55
TA 64	239	65–60
TA 71	225	65–60
TA 72	256	65–60
TA 76 s	206	55–45
TA 113	203	60–55
TA 116	182	65–60
TA 117	248	60–55
TA 118	213	65–60
TA 130	219	65–60
TA 135	192	60–55
TA 142	135	55–45
TA 200	296	65–60
TA 206	373	65–60
TR 2	210	60–55
TR 7	204	65–60
TR 29	220	65–60
TR 31	217	65–60
TR 43	297	65–60
TS 84	230	65–60
TAA 58	276	60–55
TAASH	436	65–60
TA2	175	55
TA80	211	55
Ta203	217	55
TA5	205	55
TA96	192	55
TR1	224	55
TA8	246	55
TA144	241	55
TA11	230	55
TA42	209	55
TA176	233	55
TS45	244	55
TA78	205	55
TA194	132	55
TA3	287	55
Hüttel *et al.* (1999)		
CaSTMS2	234	65–60
CaSTMS15	241	65–60
CaSTMS21	174	65–60

[a] The primers either belong to di- or trinucleotide repeat motifs.
[b] Touchdown temperature during PCR.

composite collection for functional and comparative genomics. The analysis of genetic diversity will help to elucidate population structures that influence the analysis of the associations between markers and phenotypes. Using all available information, a subset of 300 accessions will be selected to capture the maximum diversity for the isolation of allelic variants of candidate genes for traits of economic importance including functional genomics analysis. For example, accessions showing drought (Krishnamurthy *et al.*, 2003) and salinity (Serraj *et al.*, 2004) tolerance have been identified when the chickpea mini core (Upadhyaya and Ortiz, 2001) was evaluated. We hope that involving the diverse accessions in crop improvement programmes will allow the development of genetically broad-based mapping and breeding populations.

Acknowledgements

A commissioned grant received from the GCP to develop the global composite collection of chickpea is gratefully acknowledged.

References

Amirul Islam FM, Beebe S, Munoz M, Tohme J, Redden RJ and Basford KE (2004) Using molecular markers to assess the effect of introgression on quantitative attributes of common bean in the Andean gene pool. *Theoretical and Applied Genetics* 108: 243–252.

Gebhardt C, Ballvora A, Walkemeier B, Oberhagemann P and Schuler K (2004) Assessing genetic potential in germplasm collections of crop plants by marker-trait association: a case study for potatoes with quantitative variation of resistance to late blight and maturity type. *Molecular Breeding* 13: 93–102.

Hüttel B, Winter P, Weising K, Choumane W, Weigand F and Kahl G (1999) Sequence-tagged microsatellite site markers for chickpea (*Cicer arietinum* L.). *Genome* 42: 210–217.

Ivandic V, Hackett CA, Nevo E, Keith R, Thomas WTB and Forster BP (2002) Analysis of simple sequence repeats (SSRs) in wild barley from the Fertile Crescent: associations with ecology, geography, and flowering time. *Plant Molecular Biology* 48: 511–527.

Ivandic V, Thomas WTB, Nevo E, Zhang Z and Forster BP (2003) Association of simple sequence repeats with quantitative trait variation including biotic and abiotic stress tolerance in *Hordeum spontaneum*. *Plant Breeding* 122: 300–304.

Kraakman ATW, Niks RE, van den Berg PMMM, Stam P and van Eeuwjik FA (2004) Linkage disequilibrium mapping of yield and yield stability in modern spring barley cultivars. *Genetics* 168: 435–446.

Kresovich S, Luongo AJ and Schloss SJ (2002) Mining the gold: finding allelic variants for improved crop conservation and use. In: Engels JMM, Rao VR, Brown AHD and Jackson MT (eds) *Managing Plant Genetic Diversity*. Wallingford: CABI, pp. 379–386.

Krishnamurthy L, Kashiwagi J, Upadhyaya HD and Serraj R (2003) Genetic diversity of drought-avoidance root traits in the mini-core germplasm collection of chickpea. *International Chickpea and Pigeonpea Newsletter* 10: 21–24.

Niroj KS, Shokeen B and Bhatia S (2003) Isolation and characterization of sequence-tagged microsatellite sites markers in chickpea (*Cicer arietinum* L.). *Molecular Ecology Notes* 3: 428–430.

Russell JR, Booth A, Fuller JD, Baum M, Ceccarelli S, Grando S and Powell W (2003) Patterns of polymorphism detected in the chloroplast and nuclear genomes of barley landraces sampled from Syria and Jordan. *Theoretical and Applied Genetics* 107: 413–421.

Sabharwal V, Negi MS, Banga SS and Lakshmikumaran M (2004) Mapping of AFLP markers linked to seed coat color loci in *Brassica juncea* (L) Czern. *Theoretical and Applied Genetics* 109: 160–166.

Serraj R, Krishnamurthy L and Upadhyaya HD (2004) Screening chickpea mini core germplasm for tolerance to salinity. *International Chickpea and Pigeonpea Newsletter* 11: 29–32.

Shiv Kumar, Gupta S, Chandra S and Singh BB (2004) How wide is the genetic base of pulse crops? In: Ali M, Singh BB, Shiv Kumar and Dhar V (eds) *Pulses in New Perspective*. Proceedings of the National Symposium on Crop Diversification and Natural Resources Management, 20–22 December 2003. Kanpur: Indian Society of Pulses Research and Development, Indian Institute of Pulses Research, pp. 211–221.

Sun GL, William M, Liu J, Kasha KJ and Pauls KP (2001) Microsatellites and RAPD polymorphisms in Ontario corn hybrids are related to the commercial sources and maturity ratings. *Molecular Breeding* 7: 13–24.

Sun G, Bong M, Nass H, Martin R and Dong Z (2003) RAPD polymorphism in spring wheat cultivars and lines with different level of *Fusarium* resistance. *Theoretical and Applied Genetics* 106: 1059–1067.

Thornsberry JM, Goodman MM, Doebley J, Kresovich S, Nielsen D and Buckler ES IV (2001) *Dwarf8* polymorphisms associate with variation in flowering time. *Nature Genetics* 28: 286–289.

Turpeinen T, Tenhola T, Manninen O, Nevo E and Nissila E (2001) Microsatellite diversity associated with ecological factors in *Hordeum spontaneum* populations in Israel. *Molecular Ecology* 10: 1577–1591.

Upadhyaya HD and Ortiz R (2001) A minicore subset for capturing diversity and promoting utilization of chickpea genetic resources in crop improvement. *Theoretical and Applied Genetics* 102: 1292–1298.

Upadhyaya HD, Bramel PJ and Singh S (2001) Development of a chickpea core subset using geographic distribution and quantitative traits. *Crop Science* 41: 206–210.

Winter P, Pfaff T, Udupa SM, Hüttel B, Sharma PC, Sahi S, Arreguin-Espinoza R, Weigand F, Muehlbauer FJ and Kahl G (1999) Characterization and mapping of sequence-tagged microsatellite sites in the chickpea (*Cicer arietinum* L.) genome. *Molecular and General Genetics* 262: 90–101.