ISSN 1479-2621

Plant Genetic Resources **4**(1); 2–12
DOI: 10.1079/PGR200599

Methodology to establish a composite collection: case study in lentil

Bonnie J. Furman*

International Center for Agricultural Research in the Dry Areas (ICARDA), P.O. Box 5466, Aleppo, Syrian Arab Republic

Received 2 August 2005; Accepted 1 November 2005

Abstract

The International Center for Agricultural Research in the Dry Areas (ICARDA) is participating in a large-scale programme, Subprogram 1 of the Consultative Group on International Agricultural Research (CGIAR) Generation Challenge Program, that aims to explore the genetic diversity of the global germplasm collections held by the CGIAR research centres. This project will identify a 'composite collection' of germplasm for individual crops, representing the range of diversity of each crop species and its wild relatives, and characterize each composite set using anonymous molecular markers, mainly simple sequence repeats (SSRs). The overall goal of this project is to study diversity across given genera and identify genes for resistance to biotic and abiotic stresses that can be used in crop improvement programmes. ICARDA was responsible for creating the composite collection for lentil. ICARDA has the global mandate for lentil and houses the largest global collection of this crop with 10,509 accessions. From this collection, a global composite collection of 1000 lentil accessions was established with the aim to represent genetic diversity and the agro-climatological range of lentil. Accessions for the composite collection were compiled from landraces, wild relatives, and elite germplasm and cultivars. The methodology presented here combined classical hierarchical cluster analyses using agronomic traits and two-step cluster analyses using agro-climatological data linked to the geographical coordinates of the accessions' collection sites. Genotyping for 30 SSR loci will be carried out for all 1000 accessions. Plants grown for DNA analysis will be harvested and progeny will be evaluated under field conditions at ICARDA.

Keywords: cluster analysis; composite collection; core collection; genetic diversity; lentil

Introduction

Lentil (*Lens culinaris* Medik.) is an important cool-season crop and dietary mainstay in North Africa, West Asia and the Indian Subcontinent (Erskine, 1996; Robertson and Erskine, 1997). Although traditionally cultivated from the Atlantic coastal regions of Spain and Morocco to South Asia, cultivation has more recently extended into North and South America (Sharma *et al.*, 1995). It is an important source of dietary protein (25%) in both human and animal diets, second only to soybeans as a source of usable protein (Consultative Group on International Agricultural Research (CGIAR), 2005). Lentil ranks seventh among grain legumes and is grown on over 3.5 million hectares in over 48 countries with a total production of over 3 million metric tons (Food and Agriculture Organization of the United Nations (FAO), 2004). The major lentil-producing regions are South Asia (58% of the area) and the West Asia–North Africa region (37% of the hectarage of developing countries).

Lentils originated in the Fertile Crescent of the Near East and date back to the very beginning of agriculture (Ladizinsky, 1979; Webb and Hawtin, 1981; Zohary and Hopf, 1988; Harlan, 1995); archeological records place domesticated lentil in Syria and Turkey as early as

* Corresponding author. E-mail: b.furman@cgiar.org

ca 8500 BC (Cubero, 1981). The genus *Lens* comprises seven taxa within four species (Ferguson *et al.*, 2000) including the cultivated type, *Lens culinaris* Medikus subsp. *culinaris*. Cultivated lentil includes two varietal types: small-seeded *microsperma* and large-seeded *macrosperma* (Barulina, 1930). *L. culinaris* subsp. *culinaris* is grouped with its three putative wild progenitors, *L. culinaris* Medikus subsp. *orientalis* (Boiss.) Ponert, *L. culinaris* Medikus subsp. *odemensis* (Ladizinsky) Ferguson, Maxted, van Slageren & Robertson, and *L. culinaris* subsp. *tomentosus* (Ladizinsky) Ferguson, Maxted, van Slageren & Robertson. The remaining wild species include *L. nigricans* (M.Bieb.) Godron, *L. ervoides* (Brign.) Grande and *L. lamottei* Czefr. All members of *Lens* are self-pollinating diploids ($2n = 2x = 14$; Sharma *et al.*, 1995). The haploid genome size of the cultivated genome is 4063 Mbp (Arumuganathan and Earle, 1991).

Lentil is a cool-season annual grain legume. The most suitable conditions for lentil production include areas with limited rainfall and drier growing season (Johnson and Jimmerson, 2003). Average lentil yields tend to be low due to vulnerability of local landraces to abiotic and biotic stresses (Robertson and Erskine, 1997; A. Sarker, personal communication). Constraints to lentil production include high temperature and low moisture in the spring and cold temperatures in the winter. The major biotic stresses include rust caused by *Uromyces vicia-fabae* (Pers.) Schroet., vascular wilt caused by *Fusarium oxysporum* f. sp. *lentis* Schlecht. and Aschochyta blight caused by *Aschochyta fabae* Speg. f. sp. *lentis* (Robertson and Erskine, 1997). Agronomic constraints include lodging and pod dehiscence. Lentil improvement has stressed the use of landraces which are adapted to specific growing environments and may provide useful genes for breeding stress-tolerant varieties adapted to the varied target areas and farming systems (Solh and Erskine, 1981; Valkoun *et al.*, 1995). Wild *Lens* species are also a potential source for improvement of lentil for stress tolerance (Ferguson and Robertson, 1999).

The International Center for Agricultural Research in the Dry Areas (ICARDA) has a global mandate for research on lentil improvement. As such, ICARDA houses the world collection of *Lens*, totalling 10,509 accessions. The ICARDA lentil genetic resources collection includes 8789 accessions of landraces and cultivars from 70 different countries representing 12 geographic regions, 1146 ICARDA breeding lines and 574 accessions of six wild *Lens* taxa representing 23 countries (Fig. 1). The majority of the collection (53%) consists of accessions from West Asia and North Africa, the centre of origin and primary diversity (Zohary and Hopf, 1988; Ferguson and Erskine, 2001), while South Asia represents an additional 26%. Accessions in the collection were obtained from ICARDA collection missions (46%), 59 donor institutions (44%) and ICARDA's breeding programme (11%).

In 2004, ICARDA began participation in a large-scale programme that aims to explore the genetic diversity of the global germplasm collections held by the CGIAR research centres. Subprogram 1 of the CGIAR Generation

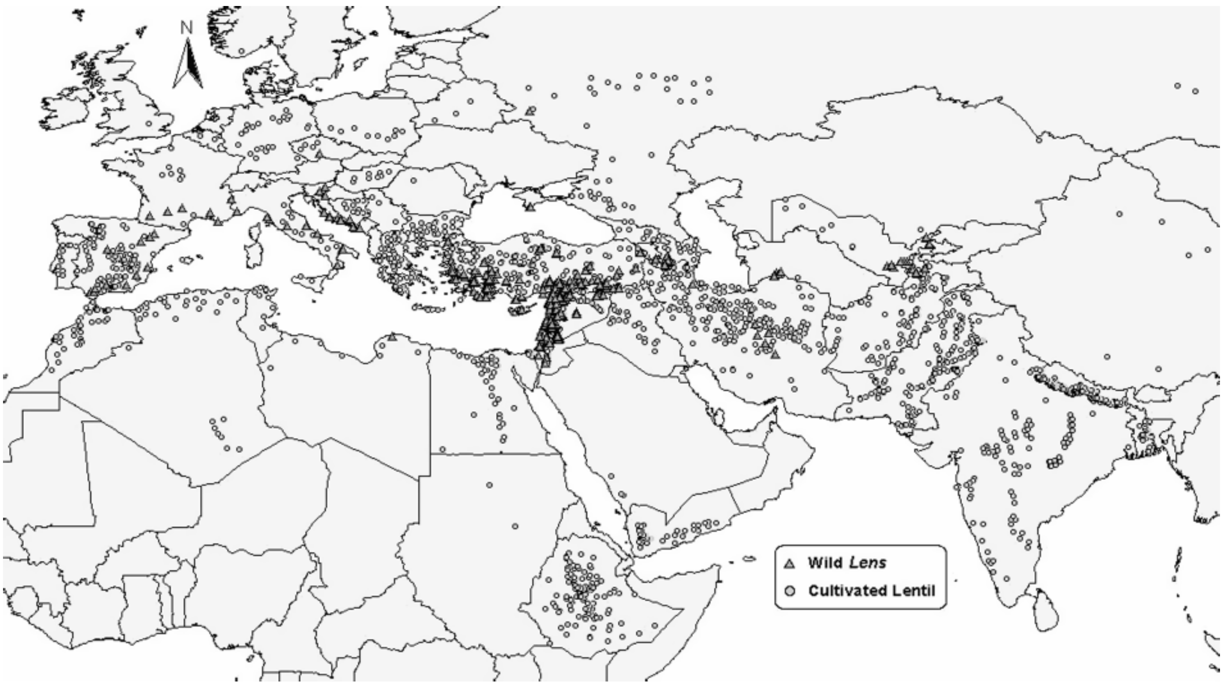

Fig. 1. Distribution of ICARDA lentil collection.

https://doi.org/10.1079/PGR200599 Published online by Cambridge University Press

Challenge Program (http://www.generationcp.org) will identify a global composite collection of germplasm for individual crops and characterize each set using anonymous molecular markers, mainly simple sequence repeat (SSR) markers. This will allow researchers to study diversity across a given genus and identify genes for resistance to biotic and abiotic stresses that can be used in crop improvement programmes. The 'composite collection' developed will be similar in size and structure to the 'core collection' conceptualized by Frankel and Brown (1984) and Brown (1989a, b). A core collection is defined as a limited set of accessions derived from an existing larger collection chosen to ensure representative variation of the larger collection, and was suggested as a means to utilize collections too large to evaluate effectively. It was proposed that such a core collection should contain approximately 10% of the accessions found in the larger collection, which would allow for approximately 80% of the total alleles represented (Brown, 1989a). A number of different strategies for selecting representatives of a core collection have been put forward and may be adapted for the development of Generation Challenge Program global composite collections (Spagnoletti Zeuli and Qualset, 1993; Crossa *et al*., 1995; Schoen and Brown, 1995; van Hintum, 1995, 1999; Yonezawa *et al*., 1995; Greene and Morris, 2001; Tai and Miller, 2001; Upadhyaya *et al*., 2001).

ICARDA was responsible for creating a global composite germplasm collection of 1000 accessions of lentil which will then be characterized utilizing microsatellite (SSR) markers and evaluated under field conditions for growth and morphological characters. The composite collection developed here differs from the classical definition of a core collection in that it also includes additional accessions chosen for containing specific beneficial traits (i.e. resistances to biotic and abiotic stresses). It also differs in that it is a representative sample to be utilized for further analysis and not maintained as a singular working collection while designating the remaining accessions into an unutilized 'reserve collection'. The objective of this paper is to present the methodology and results in the development of ICARDA's global composite collection of lentil as part of the Generation Challenge Program SP1.

Materials and methods

A candidate set of lentil accessions was derived from the global collection at ICARDA. The candidate accessions were chosen by identifying material with sufficient seed for distribution, evaluation data, and which are FAO designated and thus property of the global community, allowing for unrestricted distribution. The candidate set comprised a total of 7345 cultivated lentil accessions from 65 different countries representing 12 geographic regions, and included cultivated landraces, cultivars, breeding materials, 45 accessions of unknown type and 238 wild *Lens* accessions from 12 different countries (Table 1). The wild *Lens* accessions considered were limited to the three putative progenitor species, *L. culinaris* subsp. *orientalis*, *L. culinaris* subsp. *odemensis* and *L. culinaris* subsp. *tomentosus*.

Characterization and evaluation data were available for all cultivated candidate accessions and were separated into four separate data sets representing different evaluation years. Within each of these sets, data were further separated by geographic location/region, giving a total of 50 data sets for analysis. These 50 data sets were then subjected to hierarchical cluster analysis based on 12 phenological and agronomic characteristics (Table 2) following Upadhyaya *et al*. (2001) using SPSS version 12.0.1. Approximately 10% of the accessions in each cluster within a set were then selected randomly and included in the ICARDA composite collection (Fig. 2).

In addition, to ensure that cultivated lentil's full agro-climatological range was represented, accessions were subjected to a two-step cluster analysis using agro-climatological data linked to the geographical coordinates of the accessions' collection sites (Table 3). Two hundred clusters were produced, and one accession was selected randomly from each cluster, unless a cluster was already represented from the previous methodology (Table 4). Researchers at ICARDA also included 64 accessions of breeding material and landraces important to lentil improvement for their resistances to biotic and abiotic stresses.

Representatives from the three putative wild progenitors were also included. A total of 238 accessions were separated by species and by geographical origin within species. A similar hierarchical cluster analysis was carried out for phenological and morphological characters. Approximately 5% of the accessions within a cluster were randomly chosen. An additional 16 accessions were chosen by researchers for their resistances to biotic and abiotic stresses.

For each evaluation year, means, variances and ranges of each trait were calculated for both the candidate and chosen accessions of cultivated lentil using GenStat version 7.2. Diversity analysis to calculate genotypic coefficients of variation (GCV) for each of the 12 phenological and agronomic traits was carried out for each year's data of both the candidate and chosen accessions as described by Empig *et al*. (1970) according to the formula: GCV $= (\sqrt{\delta^2 g})/x \times 100$, where $\delta^2 g$ is the genotypic variance and x is the sample mean.

Genotypic variance was calculated by the formula PV − EV, where PV is the phenotypic variance of all

https://doi.org/10.1079/PGR200599 Published online by Cambridge University Press

Table 1. Composition of candidate set and composite collection

Region	Candidate set		Composite collection	
	No. of countries	No. of accessions	No. of countries	No. of accessions
Africa	1	2	1	1
Arabian Penninsula	2	62	2	13
Asia	4	2172	4	224
Central Asia and the Caucasus	5	24	5	12
E. Europe	14	148	12	37
N. Africa	6	616	6	107
N. America	5	71	4	21
Oceania	1	1	0	0
Russia	1	75	1	14
S. America	6	363	5	34
W. Asia	10	2726	10	372
W. Europe	10	336	9	66
Unknown	–	45	–	7
ICARDA breeding material	–	704	–	64
Wild *Lens*	12	238	9	28
Total	65	7583	59	1000

accessions excluding check varieties and EV is the environmental variance calculated as the variance of the check variety. To determine the representativeness of the chosen accessions from each data set, 95% and 99% confidence intervals were calculated for means and variances of each trait for each year using GenStat Version 7.2. The chosen accessions are considered as representative of the candidate accessions for a data set if the mean and variance of a trait of the candidate accessions fell within the respective confidence intervals of that trait of the chosen accessions. Shannon Weaver Indices for morphological and qualitative traits (Table 2) were also calculated for the candidate set and composite collection using POPGENE version 1.32 (Yeh *et al.*, 1997) to provide a non-biased estimate of the amount of variation maintained in the composite collection.

Table 2. Quantitative and qualitative traits of lentil accessions at ICARDA

Trait	Abbreviation	Description
Phenological and agronomic traits		
Days to 50% flowering	dflr	Number of days from planting to the stage when 50% of the plants have begun to flower
Days to maturity	dmat	Number of days from planting to the stage when 90% of the plants have matured
Plant height	ptht	Mean of five plants at pod-filling stage (in cm)
Height of first pod	hlp	Mean height of first pod from ground level at harvest (in cm)
Number of pods per plant	ppp	Mean number of pods from five plants at harvest
Number of seeds per pod	spd	Mean number of 10 pods each from five plants at harvest
Biological yield	byld	Biological yield in kg/ha
Seed yield	syld	Seed yield in kg/ha
Straw yield	styld	Straw yield in kg/ha
Harvest index	hi	Harvest index as a percentage
Number of seeds per plant	spp	Mean number of seeds from five plants at harvest
100-seed weight	hsw	Weight of 100 seeds in g
Morphological and qualitative traits		
Leaflet size	lfs	Size of fully extended leaflet on lower flowering node (score)
Leaf pubescence	lfp	Amount of leaf hairiness (score)
Tendril length	tl	Length of tendril during pod-filling stage (score)
Flowers per peduncle	fpi	Number of flowers per peduncle (score)
Flower ground colour	fgc	Ground colour of standard petal (code)
Lodging	lod	Amount of lodging (score)
Pod dehiscence	pdh	Amount of shattering 1 week after maturity (score)
Pod shedding	pss	Amount of pod loss 1 week after maturity (score)

https://doi.org/10.1079/PGR200599 Published online by Cambridge University Press

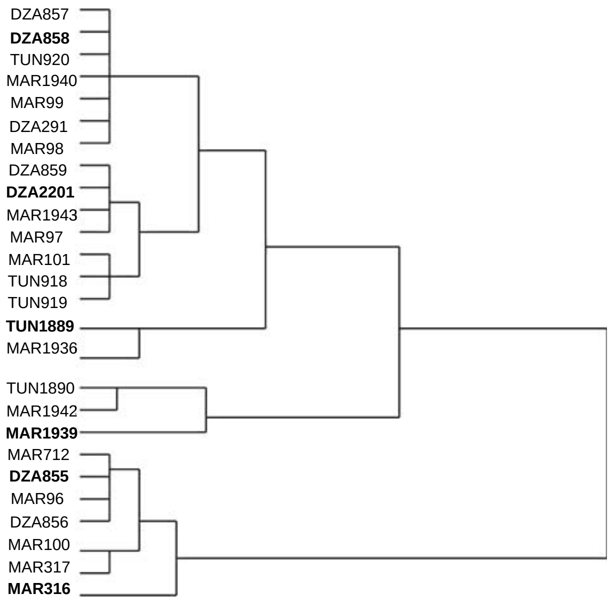

Fig. 2. Example cluster for North African accessions from 1980 evaluation year. Bold accessions were randomly chosen for inclusion into the composite collection.

Results and discussion

A global composite collection of 1000 lentil accessions was created at ICARDA consisting of landraces, wild relatives, and elite germplasm and cultivars. The methodology presented here combined classical hierarchical cluster analyses using agronomic traits and two-step cluster analyses using agro-climatological data linked to the geographical coordinates of the accessions' collection sites, ensuring that the resulting composite collection contains representative genetic diversity and that the agro-climatological range of the species is represented.

The composite collection represented the candidate set of accessions both in distribution and type. The composite collection consists of 972 cultivated lentil accessions representing 59 different countries from 11 different regions (Table 1). In the candidate set, 2726 (37.1%) were from West Asia, 2172 (29.6%) from Asia, 616 (8.4%) from North Africa, 484 (6.6%) from Europe and 434 (5.9%) from the Americas. In the composite collection, 372 (38.3%) were from West Asia, 224 (23%) from Asia, 107 (8.4%) from North Africa, 103 (10.6%) from Europe and 55 (5.7%) from the Americas (Fig. 3). The candidate set contained 6161 (81.1%) cultivated landrace accessions, 894 (11.8%) breeding lines, 93 (1.2%) cultivars, 207 (2.7%) accessions of unknown origin and 238 (3.1%) wild *Lens* accessions (Fig. 3). The composite collection consists of 829 (82.9%) cultivated landrace accessions, 78 (7.8%) breeding lines, 25 (2.5%) cultivars and 41 (4.1%) accessions of unknown status (Fig. 4). A total of 28 (2.8%) wild *Lens* accessions representing nine countries were also included.

Evaluation of the lentil candidate accessions was carried out over 4 years during 1980, 1987, 1992 and 1997. Each year's data were thus used separately for hierarchical cluster analyses and calculation of means, variances and GCV. Comparative data were not calculated for those accessions chosen from agro-climatological data or for those accessions chosen for their resistances to biotic and abiotic stress, as evaluation data for these groups were carried out over a number of years, making comparison difficult. As a result, it is not possible to assess the entire candidate and composite collections for validation purposes. For individual evaluation years, however, genetic variances, means, ranges and GCV were similar or higher for all traits (Table 5). GCV was only slightly lower for plant height, height of lowest pod and 100-seed weight for the chosen accessions from the 1980 composite accessions, and for height of lowest pod for the 1987 composite accessions. Variation ranges maintained in the composite accessions were high for most traits (Table 5). The percentage of the range maintained was higher than 80% for the majority

Table 3. Agro-climatological data linked to the geographical coordinates of accessions' collection sites

Data type	Abbreviation	Description
Longitude	lon_dd	Longitude of collection site (decimal degrees)
Latitude	lat_dd	Latitude of collection site (decimal degrees)
Altitude	alt_dem	Altitude of collection site (m)
Monthly precipitation	precyr	Amount precipitation for each of 12 months (mm)
Monthly max. temperature	tmaxyr	Maximum temperature for each of 12 months (°C)
Monthly min. temperature	tminyr	Minimum temperature for each of 12 months (°C)
Monthly potential evapotranspiration		Potential evapotranspiration for each of 12 months
Aridity index		Aridity index for each of 12 months
Aridity class	ariclass	Aridity class (code)
Agro-climatic zone	acz	Agro-climatic zone linking to UNESCO-ACZ classification
Soil type		FAO code of soil

https://doi.org/10.1079/PGR200599 Published online by Cambridge University Press

Table 4. Example results from two-step clustering using agro-climatological data (see Table 3 for details of data types)

Cluster	Accession	Origin	Province	lon_dd	lat_dd	alt_dem	precyr	tmaxyr	tminyr	acz	ariclass
1	1665	AFG	Bamian	67.83333	34.8	2486	108.9	14.4	−0.4	SH-K-M	A
	1876	AFG	Bamian	67.774	34.8	2660	266.1	12.5	−2.4	SH-K-M	SA
	1892	AFG	Ghazni	68.0	34.1	2891	336.7	13.4	−1.9	SH-K-M	SA
	1894	AFG	Ghowr	66.28333	34.6	2985	327.8	11.5	−4.9	SH-K-M	SA
	2455	AFG	Ghazni	68.0	34.1	2891	336.7	13.4	−1.9	SH-K-M	SA
2	1754	AFG	Vardak	68.71667	34.5	2300	423.5	15.8	1.5	SA-K-W	SA
	1773	AFG	Ghazni	68.26667	33.6	2583	310.5	15.4	0.1	SA-K-W	SA
	1775	AFG	Vardak	68.48	33.7	2406	289.0	16.5	1.4	SA-K-W	SA
	1778	AFG	Ghazni	68.26667	33.6	2583	310.5	15.4	0.1	SA-K-W	SA
3	**71310**	PAK	NWF	72.58333	36.3	5040	824.0	4.3	−8.4	H-K-M	H
4	87	ARM	Ashtarak	44.35	40.3	1173	349.4	16.0	4.6	SH-K-M	SA
	128	TUR	Van	43.76667	39.0	1729	367.3	14.5	2.3	SH-K-M	SA
	223	IRN	East Azerbaijan	46.31667	38.7	1406	354.4	17.8	5.2	SH-K-M	SA
	599	ARM	Ashtarak	44.35	40.3	1173	349.4	16.0	4.6	SH-K-M	SA
	619	ARM	Ashtarak	44.35	40.3	1173	349.4	16.0	4.6	SH-K-M	SA
	3396	ARM	Kafan	46.4	39.2	800	521.3	17.7	6.5	SH-K-M	SH
	3398	ARM	Ashtarak	44.35	40.3	1173	349.4	16.0	4.6	SH-K-M	SA
	5572	IRN	East Azerbaijan	46.31667	38.7	1406	354.4	17.8	5.2	SH-K-M	SA
	70182	ARM	Ashtarak	44.35	40.3	1173	349.4	16.0	4.6	SH-K-M	SA
	107419	IRN	East Azerbaijan	47.13333	38.9	1378	371.7	16.2	3.5	SH-K-M	SA

Bold accessions were randomly chosen for inclusion into the composite collection.

https://doi.org/10.1079/PGR200599 Published online by Cambridge University Press

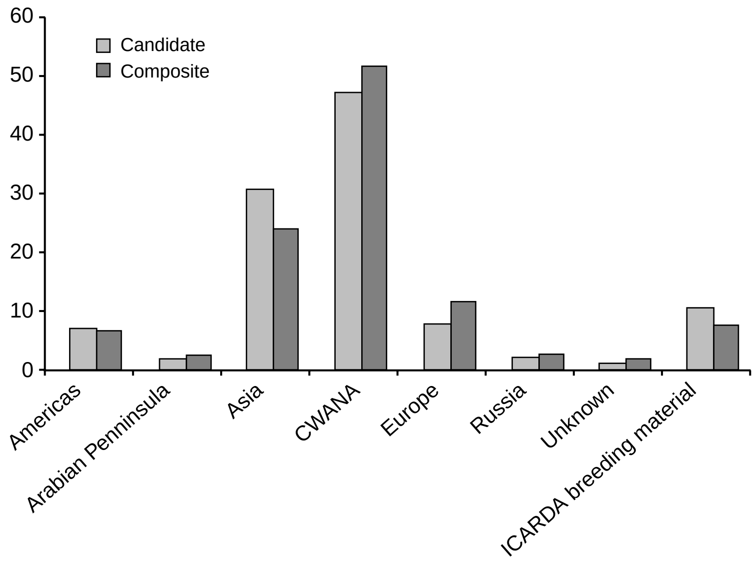

Fig. 3. Area represented by percentage of accessions represented in the ICARDA candidate and composite collections. CWANA, Central and Western Asia and North Africa.

of the traits. These results suggest that a sufficient amount of the variation found in the candidate set was maintained in the composite collection for phenological and agronomic characters.

Representativeness of the chosen accessions was estimated for phenological and agronomic characters for each evaluation year (Table 6). The chosen accessions are considered representative of the candidate accessions if the means and variances of the candidate accessions fall within the confidence intervals of the means and variances of the chosen accessions. The chosen accessions were representative for the majority of traits for all evaluation years. Low representation was noted for variance of four yield traits in 1980, two yield traits in both 1992 and 1997, and two phenological traits in 1992. Low representation was only noted for means of days to maturity in 1987 and harvest index in 1992. Shannon Weaver Indices were also calculated for morphological characters as an unbiased estimate of representativeness of the composite collection as the selection of accessions was not based on morphological diversity. Morphological characters are generally not affected by environment and were thus calculated for all years combined. Shannon Weaver Indices for the candidate set and composite collection were similar or higher for all traits (Table 7).

The hierarchical cluster analysis ensured that an acceptable level of variation found in the candidate set is maintained in the composite collection for those accessions chosen by this method. In addition, accessions

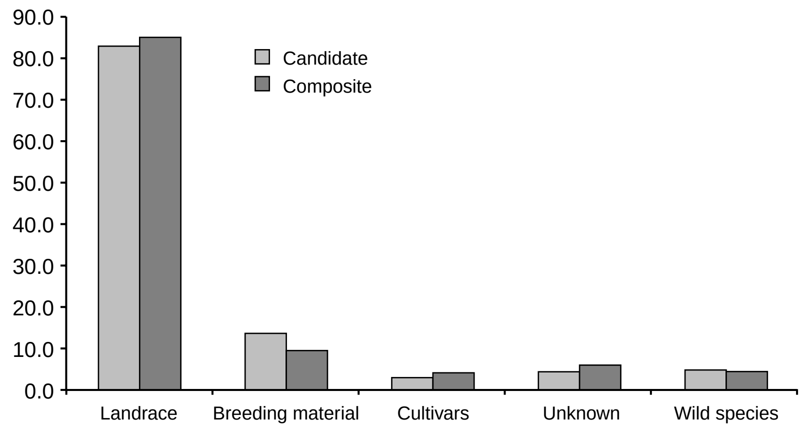

Fig. 4. Type represented by percentage of accessions represented in the ICARDA candidate and composite collections.

https://doi.org/10.1079/PGR200599 Published online by Cambridge University Press

Table 5. Mean, variance, range and GCV for candidate and composite accessions by evaluation year (see Table 2 for details of traits)

Trait	dflr	dmat	ptht	hlp	ppp	spd	byld	syld	styld	hi	spp	hsw
Candidate 1980												
Observations	4028	1943	1943	1941		3910	3990	3892	3876	3876		4002
Mean	132.82	164.26	25.83	13.63		1.53	3677.39	1142.60	2543.62	32.02		3.21
Variance	138.05	104.80	20.63	16.86		0.06	2423200.50	262113.71	1517772.74	87.08		1.41
Range	62	52	35	27		1	10233	3208	8078	76.1		7.52
GCV	0.09	0.05	0.15	0.23		0.14	0.29	0.21	0.36	0.22		0.36
Composite 1980												
Observations	469	217	21/	218		469	469	469	469	469		469
Mean	133.88	164.14	26.45	13.93		1.52	3867.23	1180.40	2689.58	31.83		3.23
Variance	152.75	92.24	20.88	15.46		0.06	3159161.66	345778.38	2049231.39	122.61		1.28
Range	62	41	24	18		1	10163	2945	8028	76.1		6.32
GCV	0.09	0.05	0.14	0.21		0.14	0.36	0.32	0.43	0.29		0.34
Candidate 1987												
Observations	996	996	996	996		992	953	996	953	953		994
Mean	116.81	162.30	33.34	17.43		1.34	3151.60	1090.83	2060.78	34.69		3.86
Variance	72.56	71.46	52.29	32.82		0.07	1305507.13	316908.91	651739.16	162.59		1.58
Range	44	34	47	31		1.4	6787	3647	5409	76.9		6.26
GCV	0.07	0.05	0.20	0.31		0.17	0.32	0.48	0.35	0.35		0.32
Composite 1987												
Observations	163	163	163	163		163	163	163	163	163		163
Mean	118.15	163.19	34.00	18.23		1.36	3239.10	1067.43	2173.18	33.15		3.71
Variance	72.42	70.56	50.56	30.79		0.71	1549374.83	318061.99	851715.56	167.80		1.59
Range	35	34	42	24		1	6650	2520	5303	67		6
GCV	0.13	0.12	0.21	0.30		0.19	0.36	0.49	0.40	0.38		0.35
Candidate 1992												
Observations	1674	1674	1674	1673	1661	1661	1640	1660	1661	1640	1661	1669
Mean	122.70	162.31	22.71	10.98	25.53	1.00	2330.79	991.18	1340.02	42.36	24.67	4.29
Variance	30.27	47.37	18.65	13.11	158.79	0.08	884396.81	190547.93	317932.79	56.24	163.05	3.01
Range	41	42	32	24	118.7	1.7	6968	3537	3773	72.9	116	8.54
GCV	0.04	0.04	0.17	0.27	0.29	0.20	0.33	0.34	0.34	0.14	0.32	0.40
Composite 1992												
Observations	167	167	164	164	164	164	164	164	164	164	164	164
Mean	124.04	164.52	23.80	11.60	25.37	0.98	2369.08	939.68	1429.40	39.92	23.65	4.26
Variance	46.16	72.57	28.64	18.55	219.16	0.09	1102709.90	199401.43	482059.30	92.29	192.44	3.19
Range	34	42	28	21	117.4	1.4	5171	2170	3690	62.2	103	7.61
GCV	0.05	0.05	0.21	0.32	0.42	0.23	0.38	0.38	0.43	0.21	0.41	0.42
Candidate 1997												
Observations	372	372	369	369	371	371	365	365	365	363	371	364
Mean	119.34	155.19	26.27	13.65	25.08	1.25	2965.49	717.60	2259.20	24.23	31.15	2.97
Variance	27.41	56.71	35.77	22.51	253.90	0.06	1490500.31	173877.49	907416.09	104.59	418.90	1.32
Range	26	33	32	24	86.3	1.4	5754	1772	4603	89.1	129	5.54
GCV	0.04	0.05	0.15	0.23	0.12	0.15	0.40	0.57	0.42	0.36	0.41	0.34

Table 5. *Continued*

Trait	dflr	dmat	ptht	hlp	ppp	spd	byld	syld	styld	hi	spp	hsw
Composite 1997												
Observations	40	40	40	40	40	40	40	40	40	40	40	40
Mean	120.48	157.48	25.58	13.03	23.08	1.31	2683.55	621.93	2061.63	24.45	31.53	2.63
Variance	40.31	106.05	57.79	36.13	323.74	0.05	1726285.95	192829.66	1264569.21	253.69	725.03	1.52
Range	23	30	28	21	75.6	0.8	5472	1665	4424	85.9	124	4.65
GCV	0.05	0.06	0.24	0.37	0.36	0.13	0.48	0.70	0.54	0.61	0.69	0.42

from a total of 200 agro-climatological clusters were represented in the composite collection (127 accessions already included from the hierarchical cluster procedure and an additional 73 accessions, representing 73 clusters, not previously included). Although the amount of variation maintained cannot be calculated for these accessions, landrace materials are adapted to micro-environments (Solh and Erskine, 1981; Valkoun *et al.*, 1995) and would thus suggest that additional variation would result from their inclusion into the composite collection.

Researchers at ICARDA suggested inclusion of 64 accessions of landraces, released cultivars and breeding materials for their resistances to a number of stresses affecting lentil production. These materials include resistance to high temperatures (Summerfield *et al.*, 1985), drought (Hamdi *et al.*, 1992; Silim *et al.*, 1993a, b; Hamdi and Erskine, 1996), boron deficiency (Yau and Erskine, 2000), lodging (Erskine and Goodrich, 1991; Ibrahim *et al.*, 1993), rust (Robertson *et al.*, 1996), fusarium wilt (Robertson *et al.*, 1996; Bayaa *et al.*, 1997), downy mildew (Abou-Zeid *et al.*, 1995) and Ascochyta blight (Robertson *et al.*, 1996). A number of these accessions possessed multiple resistances (Robertson *et al.*, 1996). An additional 18 wild *Lens* accessions included in the composite collection contained resistance to drought (Hamdi and Erskine, 1996), cold (Hamdi *et al.*, 1996; Robertson *et al.*, 1996), wilt (Bayaa *et al.*, 1995) and Ascochyta blight (Bayaa *et al.*, 1994; Robertson *et al.*, 1996). The inclusion of these materials should further ensure that important alleles will be available for Generation Challenge Program researchers in their efforts to mine valuable genes for crop improvement.

The global composite collection of lentil developed at ICARDA should provide ample diversity to realize the ultimate goal of the Generation Challenge Program. This collection has been planted for DNA extraction and analysis of microsatellite diversity will be completed for 30 SSR markers (Hamwieh *et al.*, 2005). Integral to this procedure will be future phenotypic characterization. Plants grown for DNA analysis will be harvested and progeny will be evaluated under field conditions at ICARDA. The information gathered will allow for the selection of a 'reference sample' for advanced characterization. Ultimately, these procedures will hopefully lead to the identification of important traits and the genes that control them. Methodologies for allele mining might include the use of tilling and ecotilling to identify allelic variants of candidate genes associated with beneficial traits. Scientists will utilize these materials for functional and comparative genomics, molecular mapping, gene cloning and applied plant breeding, providing the best means to incorporate these traits into plants and into farmers' fields.

https://doi.org/10.1079/PGR200599 Published online by Cambridge University Press

Table 6. Representativeness of composite collection based on confidence intervals (see Table 2 for details of traits)

	1980		1987		1992		1997	
Trait	Mean	Variance	Mean	Variance	Mean	Variance	Mean	Variance
dflr	++	++	++	++	+	–	++	++
dmat	++	++	–	++	+	–	++	–
ptht	+	++	++	++	+	+	++	+
hlp	++	++	+	++	++	–	++	+
ppp	nd	nd	nd	nd	++	–	++	++
spd	++	++	++	++	++	++	++	++
byld	+	–	++	++	++	+	++	++
syld	++	–	++	++	++	++	++	++
styld	+	–	++	++	++	–	++	++
hi	++	–	++	++	–	–	++	–
spp	nd	nd	nd	nd	++	++	++	–
hsw	++	++	++	++	++	++	++	++

+ , Representative at 0.99; ++, representative at 0.95; nd, no data.

Table 7. Shannon Weaver Indices (I′) for morphological and qualitative traits of candidate set and composite collection (see Table 2 for details of traits)

Trait	lfs	lfp	tl	fpi	fgc	lod	pdh	pss	Ī
Candidate collection									
n	3042	3042	3042	996	1368	3042	2599	2599	2488
I′	1.03	0.82	0.41	0.81	1.23	1.19	1.27	1.31	0.90
Composite collection									
n	368	368	368	164	204	368	331	331	314
I′	1.06	0.85	0.38	0.84	1.25	1.20	1.33	1.32	1.03

Ī, mean index over all traits.

Acknowledgements

The author gratefully acknowledges the assistance of Dr Jan Valkoun, Head GRU, ICARDA, for his support of this project, Dr Murari Singh, Senior Biometrician, CBSU, ICARDA, for his statistical advice and programming in GenStat, Drs A. Sarker, Senior Lentil Breeder and W. Erskine, ADG Research, ICARDA, for suggesting important ICARDA breeding materials for inclusion into the composite collection, and Drs M. Baum and W. Erskine for reviewing this manuscript. This work was supported by a commissioned grant received from the GCP.

References

Abou-Zeid N, Erskine W and Bayaa B (1995) Preliminary screening of lentil for resistance to downy mildew. *Arab Journal of Plant Protection* 13: 17–19.

Arumuganathan K and Earle ED (1991) Nuclear DNA content of some important plant species. *Plant Molecular Biology* 9: 208–218.

Barulina H (1930) Lentils of the USSR and other countries. *Bulletin of Applied Botany, Genetics and Plant Breeding* 40(Suppl): 265–304.

Bayaa B, Erskine W and Hamdi A (1994) Response of wild lentil to *Ascochyta fabae* fsp *lentis* from Syria. *Genetic Resources and Crop Evolution* 41: 61–65.

Bayaa B, Erskine W and Hamdi A (1995) Evaluation of a wild lentil collection for resistance to vascular wilt. *Genetic Resources and Crop Evolution* 42: 231–235.

Bayaa B, Erskine W and Singh M (1997) Screening lentil for resistance to fusarium wilt: methodology and sources of resistance. *Euphytica* 98: 69–74.

Brown ADH (1989a) Core collections: a practical approach to genetic resources management. *Genome* 31: 818–824.

Brown ADH (1989b) The case for core collections. In: Brown ADH, Frankel OH and Marshal RD (eds) *The Use of Plant Genetic Resources*. Cambridge: Cambridge University Press, pp. 136–155.

Consultative Group on International Agricultural Research (CGIAR) (2005) Research & impact, areas of research: lentils, http://www.cgiar.org/impact/research/lentils.html/ (accessed 14 July 2005).

Crossa J, DeLacy IH and Taba S (1995) The use of multivariate methods in developing a core collection. In: Hodgkin T, Brown AHD, Van Hintum ThJL and Morales EAV (eds) *Core Collections of Plant Genetic Resources*. Chichester: John Wiley & Sons, pp. 77–92.

Cubero JI (1981) Origin, taxonomy and domestication. In: Webb C and Hawtin G (eds) *Lentils*. Slough: Commonwealth Agricultural Bureaux, pp. 15–38.

https://doi.org/10.1079/PGR200599 Published online by Cambridge University Press

Empig L, Lantican T and Escuro PB (1970) Heritability estimates of quantitative characters in *Phaseolus aureus* Robx. *Crop Science* 10: 240–242.

Erskine W (1996) Lessons for breeder from land races of lentil. *Euphytica* 93: 107–112.

Erskine W and Goodrich WJ (1991) Variability in lentil growth habit. *Crop Science* 31: 1040–1044.

Ferguson ME and Robertson LD (1999) Morphological and phenological variation in the wild relatives of lentil. *Genetic Resources and Crop Evolution* 46: 3–12.

Ferguson M and Erskine W (2001) Lentils (*Lens* L). In: Maxted N and Bennett SJ (eds) *Plant Genetic Resources of Legumes in the Mediterranean*. Dordrecht: Kluwer Academic Publishers, pp. 125–131.

Ferguson ME, Maxted M, van Slageren M and Robertson LD (2000) A re-assessment of the taxonomy of *Lens* Mill. (Leguminosae, Papilionoideae, Vicieae). *Botanical Journal of the Linnean Society* 133: 41–59.

Food and Agriculture Organization of the United Nations (FAO) (2004) Food and Agriculture Organization statistical databases (FAOSTAT), Rome, http://faostat.fao.org/ (accessed 14 July 2005).

Frankel OH and Brown AHD (1984) Current plant genetic resources—a critical appraisal. In: Chopra VL, Joshi BC, Sharma RP and Bansal HC (eds) *Genetics: New Frontiers*, Vol. IV. New Delhi: Oxford & IBH Publishing Co., pp. 1–13.

Greene SL and Morris JB (2001) The case for multiple-use plant germplasm collections and a strategy for implementation. *Crop Science* 41: 886–892.

Hamdi A and Erskine W (1996) Reaction of wild species of the genus *Lens* to drought. *Euphytica* 91: 173–179.

Hamdi A, Erskine W and Gates P (1992) Adaptation of lentil seed yield to varying moisture supply. *Crop Science* 32: 987–990.

Hamdi A, Küsmenoğlu I and Erskine W (1996) Sources of winter hardiness in wild lentil. *Genetic Resources and Crop Evolution* 43: 63–67.

Hamwieh A, Udupa SM, Choumane W, Sarker A, Dreyer F, Jung C and Baum M (2005) A genetic linkage map of *Lens* sp. based on microsatellite and AFLP markers and the localization of *Fusarium* vascular wilt resistance. *Theoretical and Applied Genetics* 110: 669–677.

Harlan JR (1995) Agricultural origins and crop domestication in the Mediterranean region. *Diversity* 2(1/2): 14–16.

Ibrahim M, Erskine W, Hanti G and Fares A (1993) Lodging in lentil as affected by plant population soil moisture and genotype. *Experimental Agriculture* 29: 201–206.

Johnson JB and Jimmerson J (2003) Lentils. Briefing no. 61. Bozeman, MT: Agricultural Marketing Policy Center, Montana State University.

Ladizinsky G (1979) The origin of lentil and its wild genepool. *Euphytica* 28: 179–187.

Robertson LD and Erskine W (1997) Lentil. In: Fuccillo D, Sears L and Stapleton P (eds) *Biodiversity in Trust*. Cambridge: Cambridge University Press, pp. 128–138.

Robertson LD, Singh KB, Erskine W and Abd El Moneim AM (1996) Useful genetic diversity in gemplasm collections of food and forage legumes from West Asia and North Africa. *Genetic Resources and Crop Evolution* 43: 447–460.

Schoen DJ and Brown AHD (1995) Maximising genetic diversity in core collections of wild relatives of crop species. In: Hodgkin T, Brown AHD, Van Hintum ThJL and Morales EAV (eds) *Core Collections of Plant Genetic Resources*. Chichester: John Wiley & Sons, pp. 55–76.

Sharma SK, Dawson IK and Waugh R (1995) Relationships among cultivated and wild lentils revealed by RAPD analysis. *Theoretical and Applied Genetics* 91: 647–654.

Silim SN, Saxena MC and Erskine W (1993a) Adaptation of lentil to the Mediterranean environment I: factors affecting yield under drought conditions. *Experimental Agriculture* 29: 9–19.

Silim SN, Saxena MC and Erskine W (1993b) Adaptation of lentil to the Mediterranean environment II: response to moisture supply. *Experimental Agriculture* 29: 21–28.

Solh M and Erskine W (1981) Genetic resources. In: Webb C and Hawtin G (eds) *Lentils*. Slough: Commonwealth Agricultural Bureaux, pp. 53–68.

Spagnoletti Zeuli PL and Qualset CO (1993) Evaluation of five strategies for obtaining a core subset from a large genetic resource collection of durum wheat. *Theoretical and Applied Genetics* 87: 295–304.

Summerfield RI, Roberts IH, Erskine W and Ellis RH (1985) Effects of temperature and photoperiod on flowering in lentils (*Lens culinaris* Medic). *Annals of Botany* 56: 659–671.

Tai PYP and Miller JD (2001) A core collection for *Saccharum spontaneum* L from the world collection of sugarcane. *Crop Science* 41: 879–885.

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.

Valkoun J, Robertson LD and Konopka J (1995) Genetic resources at the heart of ICARDA mission throughout the Mediterranean region. *Diversity* 2(1/2): 23–24.

van Hintum ThJL (1995) Hierarchical approaches to the analysis of genetic diversity in crop plants. In: Hodgkin T, Brown AHD, Van Hintum ThJL and Morales EAV (eds) *Core Collections of Plant Genetic Resources*. Chichester: John Wiley & Sons, pp. 23–34.

van Hintum ThJL (1999) The general methodology for creating a core collection. In: Johnson RC and Hodgkin T (eds) *Core Collections for Today and Tomorrow*. Rome: International Plant Genetic Resources Institute, pp. 10–17.

Webb C and Hawtin G (1981) Introduction. In: Webb C and Hawtin G (eds) *Lentils*. Slough: Commonwealth Agricultural Bureaux, pp. 1–5.

Yau S-K and Erskine W (2000) Diversity of boron-toxicity tolerance in lentil growth and yield. *Genetic Resources and Crop Evolution* 47: 55–61.

Yeh FC, Yang RC, Boyle TBJ, Ye ZH and Mao JX (1997) *POPGENE, the User-friendly Shareware for Population Genetic Analysis*. Edmonton: Molecular Biology and Biotechnology Centre, University of Alberta.

Yonezawa K, Nomura T and Morishima H (1995) Sampling strategies for use in stratified germplasm collections. In: Hodgkin T, Brown AHD, Van Hintum ThJL and Morales EAV (eds) *Core Collections of Plant Genetic Resources*. Chichester: John Wiley & Sons, pp. 35–54.

Zohary D and Hopf M (1988) *Domestication of Plants in the Old World*. Oxford: Clarendon Press.