© NIAB 2016
ISSN 1479-2621

Plant Genetic Resources; 1–12
doi:10.1017/S147926211600023X

Status, genetic diversity and gaps in sorghum germplasm from South Asia conserved at ICRISAT genebank

H. D. Upadhyaya[1,2,3]*, K. N. Reddy[1], M. Vetriventhan[1], Murali Krishna Gumma[4], M. Irshad Ahmed[4], M. Thimma Reddy[1] and Shailesh Kumar Singh[1]

[1]*Genebank, International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru 502 324, Telangana, India,* [2]*Department of Agronomy, Kansas State University, Manhattan, KS 66506, USA,* [3]*The UWA Institute of Agriculture, the University of Western Australia, Crawley, WA 6009, Australia and* [4]*Remote Sensing and GIS Unit, International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru 502-324, Telangana, India*

Received 22 January 2016; Accepted 31 May 2016

Abstract

The genebank at ICRISAT, India that serves as a world repository for sorghum germplasm conserves 39,234 accessions from 93 countries, including 6249 from seven South Asian countries: Afghanistan (6), Bangladesh (9), India (6101), the Maldives (10), Nepal (8), Pakistan (90) and Sri Lanka (25). A total of 5340 georeferenced accessions were used to identify gaps, and 5322 accessions that were characterized at ICRISAT were used to assess the diversity in the collection. Accessions of basic races varied widely than those of intermediate races for flowering in the postrainy season, plant height in both rainy and postrainy seasons, panicle exsertion, panicle length and width, seed size and 100 seed weight. Landraces from India were late flowering, tall and produced stout panicles and larger seeds. Landraces from Pakistan flowered early in both seasons and produced stout panicles and those from Sri Lanka were late flowering and tall in both seasons, produced more basal tillers and stout panicles. A total of 110 districts in 20 provinces of India, 13 districts in three provinces of Pakistan, three districts in Bangladesh and five districts in four provinces of Sri Lanka were identified as geographical gaps. *Sorghum bicolor* subsp. *verticilliflorum, S. halepense* and *S. propinquum* were identified as taxonomic gaps in the collection. Therefore, it is suggested to explore the districts identified as gaps to enrich the variability in the world collection of sorghum at ICRISAT.

Keywords: accession, collection, genetic resources, geographical gap, landrace

Sorghum [*Sorghum bicolor* (L.) Moench] is the fifth most important cereal crop in terms of area (44.80 m-ha) and production (70.83 million tonnes) (FAO, 2014, accessed during January 2016). It is, grown in about 100 countries by resource-limited farmers in the Semi-Arid Tropics (SAT). Globally, India, Nigeria, Mexico, USA, Argentina and Ethiopia are the major sorghum producing countries. In South Asia, it is grown over a total of 6.53 m-ha, in India (5.82 m-ha), Pakistan (0.17 m-ha), Sri Lanka (96 ha) and Bangladesh (70 ha) (FAO, 2014). It is mainly grown for food and fodder in Africa, Asia and Central America, while it is an important animal feed in the Americas and Australia. Alternative uses include production of beer, alcohol, syrup, bakery items, industrial starch, etc.

Diverse populations and species rich ecosystems have greater potential to adapt to climate change. Therefore, access to a wide genetic base is the key for a sustainable crop improvement programme, especially under a climate change scenario. To ensure the availability of a wide genetic base for crop improvement programmes both in the present and in the future, there is a need to assemble and

---

*Corresponding author. E-mail: h.upadhyaya@cgiar.org

conserve as many crop genetic resources as possible before we lose them forever due to replacement of landraces by improved cultivars besides several other factors (Upadhyaya and Gowda, 2009a). The existing global collections in international genebanks are large and represent the wide diversity in different crops, including sorghum. However, a critical assessment of these collections for diversity, identifying gaps and launching germplasm collection missions in unexplored and underexplored areas is important to further enrich the variability and achieve near completion of assembling species diversity to support future crop improvement programs. Unfortunately, international efforts to collect plant genetic resources in general have declined in recent years (FAO, 2009). Geographic Information Systems (GIS) have enabled better understanding of species distribution, crop cultivation and the representativeness of germplasm collections. Using GIS software, remote sensing images, passport information and characterization data of existing germplasm collections, several geographical, trait-diversity and taxonomic gaps were identified in different crops (Jones *et al.*, 1997; Maxted *et al.*, 2008; Upadhyaya *et al.*, 2009b, 2010, 2012, 2013, 2014, 2015). Spatial information derived from satellite imagery is critical to find geographical gaps in areas where the crop is cultivated. It is a known fact that crop geographies express local climate based traits for crop improvement.

The genebank at ICRISAT, India that serves as a world repository for sorghum germplasm conserves 39,234 accessions from 93 countries including 6249 from South Asian countries. South Asia represents the southern region of the Asian continent comprising Afghanistan, Bangladesh, Bhutan, India, the Maldives, Nepal, Pakistan and Sri Lanka. A critical assessment for gaps in the existing sorghum collection from South Asian countries is yet to be done. Therefore, the aim of the present paper is to assess the status, diversity and identify geographical and taxonomic gaps in the sorghum collection from South Asian countries.

Materials and methods

Passport information and characterization data of collections from South Asian countries, including Afghanistan (6), Bangladesh (9), India (6101), the Maldives (10), Nepal (8), Pakistan (90) and Sri Lanka (25), were used in this study. Bhutan has no representation in the sorghum collection at ICRISAT.

Identification of gaps

A total of 5340 landraces assembled from Afghanistan (1), Bangladesh (5), India (5259), Pakistan (52) and Sri Lanka (23) for which there is georeferenced data, were considered to identify geographical gaps in the collections from these countries. Passport data, particularly, information on the precise location of the collecting site and the corresponding geographic coordinates was updated by referring back to all related records, collection reports and catalogues. Using Microsoft Encarta^R, an electronic atlas (MS Encarta^R Interactive World Atlas, 2000), the geographic coordinates were retrieved for accessions with location information. The accuracy of the coordinates was verified by plotting accessions on the political map of each country. The Moderate Resolution Imaging Spectroradiometer (MODIS) is a satellite-based remote sensing platform, which acquires imagery on a daily basis composited to 8 and 16 d at a spatial resolution of 250 and 500 m. This study used imagery at 16 d intervals over the 2014–2015 cropping season. A vegetation index (NDVI – Normalized Difference Vegetation Index) estimated based on the red and infra-red bands acquired by the satellite were used to identify the crop domains. The NDVI is a direct indicator of vegetation vigour during the cropping season. The 16 d interval stack of imagery provided a profile of vegetation cover based on its growth condition during the cropping season indicating a specific vegetation type (dominant crop) (Gumma *et al.*, 2014). A rigorous validation was also carried out at the field level to maintain the accuracy of the information extracted. Major crops in the South Asian region including sorghum were mapped using the method described above (Gumma *et al.*, 2011). Sorghum crop extent was extracted from this layer and used for the present study. The sorghum collection that had georeferenced data (5340 accessions) was overlaid on the present sorghum area (2014–2015) and the gaps in the collection were identified. Districts (gaps) were identified using crop map and germplasm collection sites. Districts where less than five germplasm samples had been collected about 30 years ago were identified as a potential gap.

Characterization of germplasm

A total of 5322 accessions were characterized in different years during 1977–2013, in batches of 500–1000 in vertisols in the rainy (June–November) and postrainy (October–March) seasons at ICRISAT, Patancheru, India (17°25′N latitude, 78°00′E longitude and 545 m a.s.l.). Each accession was sown in one 4 m long row with a spacing of 75 cm between rows. The crop was thinned after 2 weeks leaving approximately 10 cm distances between plants so as to accommodate about 40 plants in a row. Fertilizers were applied at the rate of 80 kg/ha N and 40 kg/ha P_2O_5 in both the seasons. Every year, landraces were grown in an augmented design using one of the three checks (IS 2205, IS 18758 and IS 33844) repeated for every block of 20 test accessions. Life-saving irrigations were provided in the rainy seasons while the crop was irrigated at regular intervals to provide sufficient moisture to the growing crop in the postrainy seasons. The crop was protected from weeds, pests and diseases for a good crop.

Observations on eight quantitative traits (days to 50% flowering, plant height, number of basal tillers per plant, panicle exsertion, panicle length and width, seed size and 100 seed weight) and ten qualitative traits (plant pigmentation, midrib colour, panicle compactness and shape, glume colour, glume covering, seed colour, seed lustre, seed sub-coat, endosperm texture and threshability) were recorded following descriptors for sorghum (IBPGR and ICRISAT, 1993). Observations on days to 50% flowering and plant height were recorded in both the rainy and postrainy seasons, whereas observations on all other traits were recorded only during the postrainy season. Number of days from sowing to when 50% plants start flowering in an accession was recorded as days to 50% flowering. Mean height of five representative plants from the base to the tip of the panicle was recorded as plant height. Length of the peduncle from the ligule of the flag leaf to the base of the inflorescence was recorded as panicle exsertion. The mean length of five panicles selected randomly from base to tip and mean width of panicles in natural position at the widest portion was recorded in centimetres. Width of the seed at the broadest point was recorded in millimeters. Weight of 100 seeds drawn from the plot yield was recorded in grams.

Data analysis

The mean, range and variances were calculated for eight quantitative characters for all races and countries. The means for different traits were compared using the Newman–Keuls procedure (Newman, 1939; Keuls, 1952) and homogeneity of variances was tested by Levene's test (Levene, 1960). Principal component analysis (PCA) of eight quantitative traits was performed using GENSTAT 13.1. (VSN International, 2010). The Shannon–Weaver diversity index (*H*′) (Shannon and Weaver, 1949) was used to measure and compare the phenotypic diversity for each trait, race and country. Phenotypic proportions were estimated for ten qualitative traits (Snedecor and Cochran, 1980).

Results

Status of the collection

The sorghum germplasm collection from South Asian countries conserved at the genebank at ICRISAT, India, includes 6101 accessions from India, 90 from Pakistan, 25 from Sri Lanka, ten from the Maldives, nine from Bangladesh, eight from Nepal and six from Afghanistan (Table 1). The collection is from a wide range of latitudes ranging from 6 98°N (Sri Lanka) to 36 71°N (Afghanistan). The collection includes 5737 landraces, 452 breeding materials, 40 improved cultivars and 20 wild accessions belonging to *S. bicolor* subsp. *drummondii* (9), *S. halepense* (5) and *S. purpureosericeum* (3) and three wild accessions that have no species information. Sorghum germplasm at ICRISAT was assembled by introducing germplasm that had already been collected from various organizations located in different countries and by subsequently launching systematic germplasm collection missions in South Asian countries.

Germplasm introduced

A total of 3575 accessions were introduced from 38 organizations located in eight countries (Table 1). Organizations located in India were the major donors of sorghum germplasm to ICRISAT (3393 accessions). The Rockefeller Foundation, New Delhi, India, was an important donor providing 2723 accessions originating in Afghanistan (5), India (2697), Nepal (7) and Pakistan (14). Other donors in India included Andhra Pradesh Agricultural University, Hyderabad (136); Dr Punjabrao Deshmukh Krishi Vidyapeeth, Akola (83); G B Pant University of Agriculture and Technology, Pantnagar (6); National Botanical Research Institute, Lucknow (12); Gujarat Agricultural University, Surat (4); Haryana Agricultural University, Hisar (1); Indian Agricultural Research Institute (IARI), New Delhi (33); Kerala Agricultural University, Palghat (2); Mahatma Phule Krishi Vidyapeeth, Mohol (5); Marathwada Agricultural University, Parbhani (83); National Bureau of Plant Genetic Resources (NBPGR), New Delhi (90); National Research Center for Sorghum, Rajendranagar (186); Tamil Nadu Agricultural University, Coimbatore (21) and University of Agricultural Sciences, Dharwad (8).

Outside India, Bangladesh Agricultural Research Institute (BARI), Sher-E-Banglanagar, Bangladesh (9 accessions); National Agricultural Research Centre (NARC), Islamabad, Pakistan (9); International Rice Research Institute (IRRI), Los Banos, Philippines (2); International Plant Genetic Resources Institute (IPGRI)/APO, Singapore (10); University of Peradeniya, Peradeniya, Sri Lanka (2); Royal Botanical Gardens, Kew, UK (10); Colorado State University, Fort Collins (3), Mayaguez Institute of Tropical Agriculture, Mayaguez (107); USDA-ARS-Tropical Agriculture Research Station, Mayaguez (1) and United States Development Agency (USDA) and ARS Plant Genetic Resources, Griffin (29), USA, had donated sorghum germplasm to ICRISAT. All accessions originating in Afghanistan, Bangladesh, the Maldives and Nepal were introductions in the collection at the ICRISAT genebank (Table 1).

Table 1. Status of sorghum germplasm assembled from South Asia at ICRISAT genebank, India

Country	Collections			No. of accessions introduced	Total accessions
	Year of collection	Collection missions launched	No. of accessions collected		
Afghanistan				6	6
Bangladesh				9	9
India	1975	1	143		143
	1976	2	5		5
	1977	3	158		158
	1978	1	260		260
	1979	2	87		87
	1980	5	183		183
	1981	6	76		76
	1982	1	6		6
	1983	2	6		6
	1986	1	183		183
	1987	3	175		175
	1988	4	210		210
	1989	5	379		379
	1990	2	122		122
	1992	1	6		6
	1993	1	176		176
	1994	1	160		160
	1996	1	270		270
India total		42	2605	3496	6101
Maldives				10	10
Nepal				8	8
Pakistan	1989	1	46	44	90
Sri Lanka	1980	1	23	2	25
Total		44	2674	3575	6249

Germplasm collected

During 1975–1996, ICRISAT and its partners had launched 44 collection missions in South Asian countries for its mandate crops and collected 2674 samples of sorghum (Table 1). Forty-two collection missions were launched in India alone and one mission each in Pakistan and Sri Lanka and 2605, 46 and 23 accessions respectively were collected. In India, ICRISAT collaborated with NBPGR, New Delhi for 18 collection missions, with Tamil Nadu Agricultural University, Coimbatore for four missions (Gopal Reddy *et al.*, 1993; Gopal Reddy and Verma, 1996), one mission each with C.S. Azad University of Agriculture and Technology, Kanpur and Indian Grassland and Fodder Research Institute, Jhansi; Haryana Agricultural University, Hisar (Pundir, 1980); Andhra University, Visakhapatnam; All India Coordinated Research Project on Oil Seeds (AICORPO), Akola (Ramanatha Rao and Rajagopal Reddy, 1981); Andhra Pradesh Agricultural University, Hyderabad (Remanandan 1983); Rajasthan Agricultural University, Durgapura (Appa Rao, 1978); University of Agricultural Sciences, Hebbal and Orissa University of Agriculture and Technology, Bhubaneswar. ICRISAT had collaborated with the National Agricultural Research Centre (NARC), Islamabad in Pakistan during 1989 and in Sri Lanka with the Central Agricultural Research Institute during 1980 (van der Maesen, 1980).

Sources of germplasm

Sorghum germplasm samples were assembled from different sources. In the entire collection of 6249 accessions, the maximum of 3570 samples were received from various institutes, 2624 samples were collected from farmer's fields, 20 samples from threshing floors, 15 from farm stores and 20 were taken from wild habitats.

Intensity of the collection

The cultivated sorghum collection which has georeferenced data (5340) represents 763 geographical sites of germplasm collection in India, 31 in Pakistan, 16 in Sri Lanka, four in Bangladesh and one in Afghanistan. The average number of samples per geographical site was one in Afghanistan, Bangladesh and Sri Lanka and two in Pakistan and seven in India. FAO statistics for area of sorghum cultivation during the year 2014 and the representation in the world collection at ICRISAT revealed one collection site per 18 ha of sorghum cultivation in Bangladesh, per 7628 ha in India, per 5516 ha in Pakistan and per 6 ha in Sri Lanka (FAO, 2014). These results indicate a low intensity of collection in all the countries under study suggesting the identification of gaps and exploration to fill the gaps.

Taxonomy of the collection

The sorghum germplasm collection from India represents *S. bicolor* subsp. *bicolor* and subsp. *drummondii* (Steud.) de Wet; *S. halepense* (L.) Pers. and *S. purpureosericeum* (A. Rich.) Aschers. & Schweinf. Deccanense Garber (Table 2). All basic (*bicolor, guinea, caudatum, kafir* and *durra*) and intermediate races (*guinea–bicolor, caudatum–bicolor, kafir–bicolor, durra–bicolor, guinea–caudatum, guinea–kafir, guinea–durra, kafir–caudatum, durra–caudatum* and *kafir–durra*) were found in the collection. The collection from India represented all basic and intermediate races with predominance of *durra* (3548), *guinea* (774), *durra–bicolor* (558), *durra–caudatum* (359), *bicolor* (333), *caudatum* (130), *guinea–caudatum* (122) and *caudatum–bicolor* (120). All other races and intermediate races represented less than 60 accessions. The collection from Pakistan represented *S. bicolor* subsp. *drummondii* (2 accessions).

Diversity in the collection

Qualitative traits

Plant pigmentation is very useful in differentiating accessions and a high proportion of the collection (98.3%) had pigmentation (Table S1). Only 1.7% of the accessions had tan plant colour, which is said to be associated with resistance to leaf diseases and grain weathering (Frederiksen and Duncan, 1982; Duncan *et al*., 1991). Out of three midrib colours observed in the collection, the white midrib colour was predominant (69.8%). Ten panicle types (panicle compactness and shape) with predominance of the semi-compact elliptic type were observed in the collection. Other types observed in considerable proportion include compact elliptic (29.5%) and semi-loose stiff branches (14.3%). Grains of more than 50% of the accessions were half covered with glumes. Only 5.8% of the accessions produced uncovered grains. Eleven seed colours were observed in the collection of which straw seed colour was found in large proportion (45.4%) followed by white colour (31.5%). More than 90% of the accessions produced lustrous seeds with a sub-coat. Only 4.5% of the accessions produced completely corneous seeds and 28% produced mostly starchy seeds. Only 5.4% accessions were found as difficult to thresh.

Quantitative traits

Range. A high range of variation was observed for different quantitative traits of basic and intermediate races as groups, individual races and for the collections from different countries (Table S2). Accessions of basic races varied widely than those of intermediate races for plant height in both seasons, days to 50% flowering, panicle exsertion, length and width and seed size and 100 seed weight in the postrainy season. Promising sources for various quantitative traits were race specific. Among the basic races, *bicolor* (45 d) and *caudatum* (33 d) were found as promising sources for early flowering, *durra* for short height (95 cm) and *guinea* (655 cm) for tall height in the rainy season, *durra* for long panicles (90 cm) and large seeds (7.3 g 100 seeds^{-1}) and *guinea* for stout panicles (51 cm). Landraces belonging to intermediate races varied more widely for days to flowering in the rainy season and number of basal tillers per plant, than basic races (Table S2). Among the intermediate races, the promising sources were *guinea–durra* for flowering in the rainy season; *guinea–caudatum* for flowering in the postrainy season, *durra–bicolor* for plant height in both seasons, basal tillers per plant, panicle exsertion and length and seed size; *guinea–bicolor* for panicle width and *durra–caudatum* for 100 seed weight. Landraces from India varied widely for all the traits under study.

Means. The Newman–Keuls test was performed to compare mean values of all the traits for basic and intermediate races as groups, individual races and countries (Table S3). The basic and intermediate races were significantly different for all the traits, except plant height in the postrainy season. Among the races, *guinea* and *guinea–caudatum* flowered later than other races in both seasons. *Bicolor* in postrainy, *guinea* in the rainy season and *guinea–durra* in both seasons grew taller than other races. Basal tillering per plant and panicle exsertion was highest in *bicolor* and differed significantly from all other races. *Guinea* produced

Table 2. Geographical distribution of sorghum races and wild relatives germplasm from South Asia at ICRISAT genebank, India

Species/sub-species/race	Afghanistan	Bangladesh	India	Maldives	Nepal	Pakistan	Sri Lanka	Total
S. bicolor (L.) Moench								
Bicolor	1		333			13	2	349
Guinea		5	774		5	1	3	788
Caudatum	1	1	130	7		2	10	151
Kafir	1		12					13
Durra			3548	2	2	44	3	3599
Guinea–bicolor			22			1		23
Caudatum–bicolor	2	1	120		1	5		129
Kafir–bicolor			9					9
Durra–bicolor			558			7		565
Guinea–caudatum		1	122	1		6	2	132
Guinea–kafir			1					1
Guinea–durra			56					56
Kafir–caudatum			7			1		8
Durra–caudatum	1	1	359			8	5	374
Kafir–durra			32					32
S. bicolor subsp. *drummondi* (*Steud.*) *de Wet*			7			2		9
S. halepense (L.) Pers.			5					5
S. purpureosericeum (A. Rich.) Aschers. & Schweinf. Deccanense Garber			3					3
No information			3					3
Total	6	9	6101	10	8	90	25	6249

significantly longer and stouter panicles than all other races. The races *durra* and *guinea–durra* produced larger seeds. Landraces from Bangladesh flowered late in the rainy season, grew tall in the postrainy season and produced highly exserted long and thick panicles. Landraces from India were late, tall and produced stout panicles and larger seeds. Landraces from Pakistan flowered early in both seasons and produced stout panicles and those from Sri Lanka were late and tall in both seasons, produced more basal tillers and stout panicles.

Variances. Levene's test of significance for variances revealed highly significant variances for all the traits under study, except days to flowering in the rainy season, which was significant at the 5% level of probability suggesting significant differences among the landraces under study.

Phenotypic diversity. PCA was carried out using standardized data of eight quantitative traits; the first three PCs captured 65.64, 66.49 and 72.89% of the total variation in the entire collection, basic and intermediate races, respectively. The Shannon–Weaver diversity index (H') was calculated for the entire South Asian collection and for each race and country of origin to compare phenotypic diversity for ten quantitative traits (Table 3). A low H' indicates extremely unbalanced frequency classes for an individual trait and lack of genetic diversity in the collection. The diversity index values (H') were variable among traits. In the entire collection, the diversity index (H') ranged from 0.414 ± 0.026 for basal tillers per plant to 0.569 ± 0.029 for plant height in the postrainy season (Table 3). The mean diversity across the traits was maximum in *durra–bicolor* ($H' = 0.591 \pm 0.018$) and lowest in *kafir* ($H' = 0.276 \pm 0.001$). The mean diversity over countries varied from 0.396 ± 0.032 for basal tillers per plant to 0.541 ± 0.044 for days to flowering in the postrainy season. The *guinea* race was highly diverse for days to flowering and plant height in the rainy season, panicle exsertion, length and width. Among the countries, India was highly diverse for all traits except panicle width and basal tillers per plant.

Geographical gaps

In South Asian countries, sorghum is grown either as sole crop or mixed with chickpea, green gram, black gram, soybean, groundnut, safflower, linseed, cucumber,

Table 3. Shannon–Weaver diversity index (H') for agronomic traits of sorghum germplasm from South Asia at ICRISAT genebank, India

Race/country	1[a]	2	3	4	5	6	7	8	9	10	Mean	SE±
Races												
Bicolor	0.578	0.589	0.536	0.617	0.512	0.618	0.588	0.593	0.378	0.583	0.559	0.023
Guinea	0.609	0.457	0.631	0.626	0.498	0.618	0.629	0.616	0.595	0.603	0.588	0.019
Caudatum	0.455	0.555	0.568	0.563	0.527	0.592	0.612	0.561	0.635	0.552	0.562	0.015
Kafir	0.276	0.276	0.276	0.276	0.276	0.276	0.276	0.276	0.276	0.276	0.276	0.000
Durra	0.593	0.611	0.621	0.629	0.382	0.534	0.610	0.588	0.630	0.633	0.583	0.024
Guinea–bicolor	0.581	0.506	0.556	0.550	0.308	0.563	0.529	0.572	0.568	0.484	0.522	0.026
Caudatum–bicolor	0.540	0.592	0.565	0.629	0.478	0.608	0.608	0.584	0.573	0.595	0.577	0.014
Durra–bicolor	0.587	0.620	0.606	0.609	0.432	0.592	0.621	0.610	0.626	0.607	0.591	0.018
Guinea–caudatum	0.517	0.550	0.575	0.590	0.483	0.576	0.624	0.567	0.570	0.612	0.566	0.013
Guinea–durra	0.550	0.519	0.555	0.618	0.290	0.574	0.573	0.405	0.588	0.600	0.527	0.032
Durra–caudatum	0.598	0.590	0.599	0.622	0.414	0.579	0.571	0.555	0.541	0.601	0.567	0.019
Kafir–durra	0.521	0.458	0.538	0.497	0.363	0.579	0.573	0.407	0.482	0.538	0.496	0.022
Mean	0.534	0.527	0.552	0.569	0.414	0.559	0.568	0.528	0.539	0.557	0.535	0.014
SE±	0.027	0.028	0.027	0.029	0.026	0.027	0.028	0.031	0.032	0.028		
Country												
Bangladesh	0.458	0.413	0.458	0.292	0.458	0.217	0.458	0.413	0.292	0.413	0.387	0.028
India	0.607	0.613	0.630	0.632	0.440	0.565	0.608	0.513	0.646	0.629	0.588	0.021
Pakistan	0.603	0.563	0.482	0.590	0.357	0.398	0.407	0.527	0.517	0.563	0.501	0.027
Sri Lanka	0.430	0.576	0.435	0.479	0.328	0.521	0.300	0.300	0.359	0.392	0.412	0.030
Mean	0.525	0.541	0.501	0.498	0.396	0.425	0.443	0.438	0.454	0.499	0.472	0.015
SE±	0.047	0.044	0.044	0.076	0.032	0.078	0.064	0.053	0.080	0.058		

[a] 1 = Days to 50% flowering-rainy, 2 = Days to 50% flowering-postrainy, 3 = Plant height (cm)-rainy, 4 = Plant height (cm)-postrainy, 5 = Number of basal tillers per plant, 6 = Panicle exsertion (cm), 7 = Panicle length (cm), 8 = Panicle width (cm), 9 = Seed size (mm) and 10 = 100 Seed weight (g).

watermelon, etc., in the rainy and postrainy seasons. A total of 131 districts located in 27 provinces of four South Asian countries were found as the geographical gaps (Fig. 1, Table 4). Country wise, 110 districts located in 20 provinces of India, 13 districts located in three provinces of Pakistan, three districts in Bangladesh and five districts located in four provinces of Sri Lanka were the gaps. Uttar Pradesh in India with relatively low representation (238 accessions) in the collection had 27 districts as gaps in sorghum germplasm. Other provinces showing more than five districts as gaps include Bihar, Madhya Pradesh, Rajasthan, Odisha, Tamil Nadu and West Bengal.

Taxonomic gaps

The wild relatives' collection from South Asian countries at the ICRISAT genebank includes *S. bicolor* subsp. *drummondii* (Steud.) (9 accessions) of primary genepool, *S. halepense* (L.) Pers. (5 accessions) of secondary genepool and *S. pururiosericeum* (A. Rich.) Aschers. & Schweinf. Deccanense Garber (3 accessions) of tertiary genepool (Acheampong *et al.*, 1984) (Table 2). Kamala *et al.* (2014) reported the distribution of *S. bicolor* subsp. *drummondii, S. bicolor* subsp. *verticilliflorum* and *S. halepense* in India and *S. propinquum* in India and Sri Lanka. Therefore, *S. bicolor* subsp. *verticilliflorum* and *S. halapense, S. propinquum* and other wild relatives may be considered as important taxonomic gaps in the collection.

Discussion

The extent of coverage of the total diversity of different crops in *ex situ* collections is difficult if not impossible to estimate with any real precision, but it varies considerably according to the crop and several other factors. There are still sizeable gaps in the *ex situ* collections of many major crops including sorghum. The collection of germplasm is not the only way to fill gaps. In recent years, we have made extensive efforts at ICRISAT, by comparing databases of ICRISAT mandate crops with collections at other

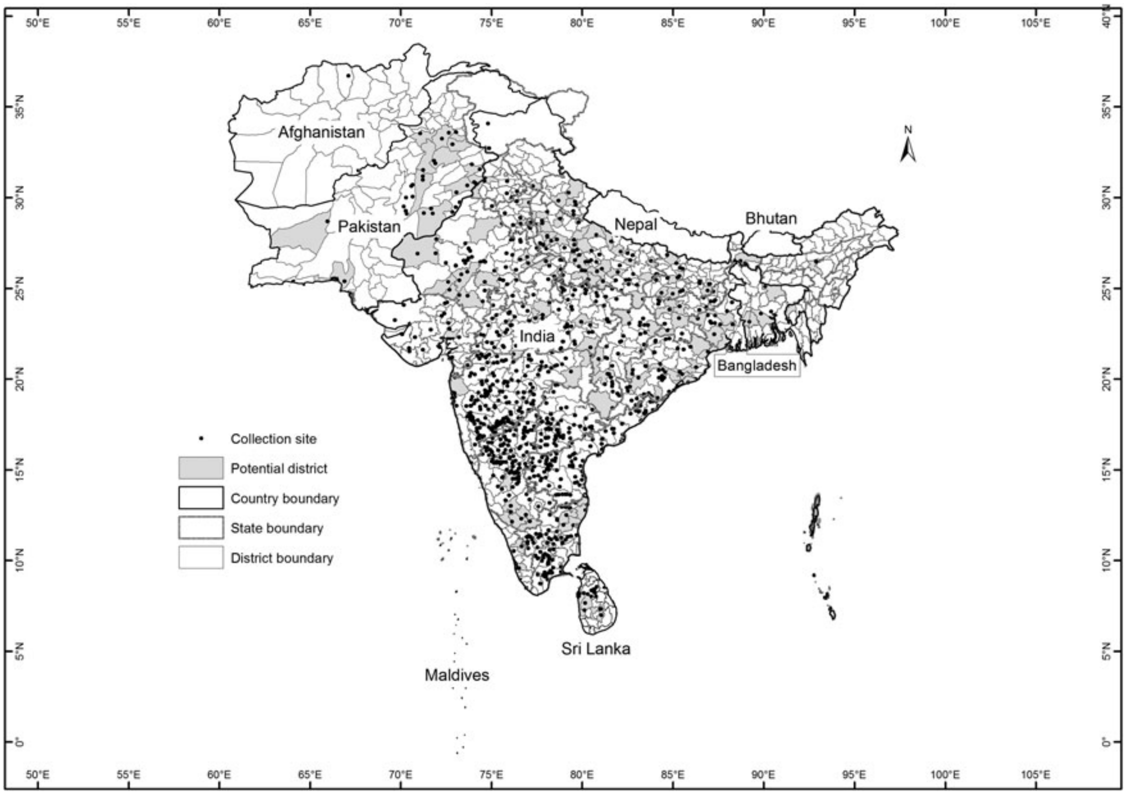

Fig. 1. Geographical distribution and the gaps (districts shaded) identified in the sorghum collection from South Asia at ICRISAT genebank, India.

important genebanks globally. Based on the comparison we added more than 4000 unique accessions of mandate crops from 12 countries, including 2052 samples of sorghum from nine countries. Afterwards, we have identified gaps in the collections of mandate crops. Exploring the gaps identified in this study will enrich the diversity of sorghum collection at ICRISAT genebank.

The adequacy of the collections to provide useful genes for current and future programmes will be dependent on how successful the collecting efforts have been in obtaining good representation of the existing genetic diversity within crop species and close relatives and on the quality of preservation procedures and facilities. Many species are threatened by the vagaries of the climate as well as other changes. Therefore, the goal of collection missions should be to preserve the highest level of useful genetic diversity that is feasible through identification of gaps in the collections and exploring the areas identified as gaps.

The success of gap analysis depends on the quality of the input data. In many genebanks, most of the older germplasm collections do not have complete passport information, particularly, the georeferenced data (latitude and longitude) of the collection sites. This poses a problem while assessing the geographical completeness of collections (Upadhyaya *et al*., 2010). The inaccuracy of georeferenced data is an additional constraint. Updating passport data for location information and georeferenced data and their validation is essential for the identification of gaps using spatial analysis. As is seen in the present study, the usefulness of satellite images with temporal resolution to map the distribution of seasonal crops is well established.

Extraction of spatial information on the vegetative cover, specifically Land Use Land Cover (LULC) using remote sensing has paved the way for the availability of such information accurately and economically. It has the advantage of providing information over large areas which are inaccessible. Studying crop diversity and identifying *in situ* conservation areas requires land use information, specifically crop domain maps, which can be prepared using satellite imagery at regular intervals, as well as long-term changes. Any collection mission requires readily available information on crop cultivation for planning germplasm collection missions. The production domains of crops will help in prioritizing locations for phased collection. A satellite imagery which provides the accurate location as well as the area can help in optimizing trips for germplasm collection. Realizing the potential of these new science tools, premier scientific organizations like the National Aeronautics and Space Administration (NASA) – https://earthdata.

Table 4. Geographical gaps identified in sorghum germplasm from South Asian countries at ICRISAT genebank, India

Country	State/Province	District
India	Assam	Nagaon
	Bihar	Aurangabad, Banka, Buxar, Gaya, Patna, Purba Champaran, Vaishali
	Chhattisgarh	Dantewara, Jashpur, Kanker, Korea, Raj Nandgaon
	Gujarat	Anand, Dangs
	Haryana	Jhajjar, Karnal
	Himachal Pradesh	Sirmaur
	Jharkhand	Chatra, Godda, Hazaribag, Jamtara, Latehar, Ranchi, Simdega
	Karnataka	Chamrajnagar, Davangere, Hassan, Kodagu, Mandya, Mysore
	Kerala	Kottayam, Palakkad
	Madhya Pradesh	Balaghat, Bhopal, Burhanpur, Datia, Harda, Jabalpur, Morena, Vidisha
	Maharashtra	Chandrapur, Gondiya, Thane
	Odisha	Baleshwar, Bauda, Cuttack, Gajapati, Jajapur, Kalahandi, Khordha, Mayurbhanj, Nayagarh, Nuaparha, Puri
	Pondicherry	Karaikal
	Punjab	Firozpur
	Rajasthan	Bhilwara, Churu, Jaisalmer, Jalor, Karauli, Pali, Sirohi, Udaipur
	Tamilnadu	Ariyalur, Kanchipuram, Karur, Perambalur, Tiruvallur, Tiruvannamalai, Villupuram
	Tripura	West Tripura
	Uttar Pradesh	Akbarpur, Aligarh, Ambedkarnagar, Auraiya, Baghpat, Bahraich, Banda, Bulandshahr, Chitrakoot, Deoria, Etah, Farrukhabad, Fatehpur, Gautambudh Nagar, Ghazipur, Gonda, Hamirpur, Hardoi, Hathras, Kheri, Meerut, Moradabad, Pratapgarh, Rae Bareli, Sant Ravi Das Nagar(Bhadohi), Shahjahanpur, Sonbhadra
	Uttaranchal	Almora, Chamoli, Garhwal, Udham Singh Nagar
	West Bengal	Barddhaman, Hugli, Jalpaiguri, Pashchim Medinipur, Purba Medinipur, Puruliya
Pakistan	Baluchistan	Kharan, Lasbela
	N.W.F.P.	Kohat
	Punjab	Attok, Bahawalpur, Jhelum, Kasur, Mainwali, Muzaffargarh, Rawalpindi, Sahiwal, Sargodha, Sheikhupura
Bangladesh	NA	Faridpur, Jessore, Mymensingh
Sri Lanka	Central	Kandy
	North Western	Kurunegala, Puttalam
	Uva	Badulla
	Western	Gampaha

nasa.gov/about/daacs/daac-lpdaac and the United States Geological Survey (USGS) have opened up their data repositories at no cost (http://earthexplorer.usgs.gov/, https://lpdaac.usgs.gov/) (NASA, 2014; USGS, 2014). This type of spatial information can be used for multiple purposes such as identification of crop domains, identification of gaps in germplasm collections, zones of adaptation, introduction of improved cultivars and intensification of domains.

Representation of sorghum germplasm from South Asia for wide latitudes (from 6.98° N in Sri Lanka to 36.71°N in Afghanistan), indicates adaptation of landraces to diverse climates in this region. Srinivasa Rao *et al.* (2014) reported that sorghum can be grown up to 40° latitude on either side of the equator. Espinoza and Kelley (2002) reported sorghum as being a versatile crop capable of growing well under contrasting climatic conditions. Therefore, there is a suggestion to increase the variability for adaptive traits by filling the gaps in the sorghum collection from South Asian countries.

The sorghum collection from South Asian countries at the ICRISAT genebank indicates low representation from all countries except India. However, as per the FAO statistics for sorghum cultivation and the number of collection sites, there is one collection site for 7628 ha in India and for 5516 ha in Pakistan suggesting exploration of the gaps identified in the present study (FAO, 2014). As the remote sensing images indicate the area under sorghum cultivation in South Asia during 2014–2015, the identified gaps can be considered as potential areas for sorghum germplasm. Low representation of *bicolor* and *caudatum* races, which were identified as sources for early flowering also emphasize the need to collect these races during future collection missions. Among the intermediate races, except for *durra–bicolor* and *durra–caudatum*, all other races need to be collected for good representation in the world collection. Landraces from Pakistan flowered early and those from Sri Lanka flowered late in both seasons indicating reduced sensitivity to climate related traits. Landraces of the *guinea* race were highly diverse for flowering, plant height and panicle number in the collections from all countries except India. These need to be collected to enrich the world collection at the ICRISAT genebank for diversity.

Unlike the cultivated species, which tend to have lower levels of diversity, the Crop Wild Relatives (CWR) forms an important source for higher levels of stress resistance and useful adaptive and nutritional traits. Kamala *et al.* (2014) reported *S. halepense* as a good source for a higher level of downy mildew resistance. Mote (1984) reported *S. purpureosericeum* of Indian origin as a source for higher levels of resistance to the shoot fly. A total of 458 wild accessions belonging to 13 species accounting for 1.2% of the total collection (39,234 accessions) are being conserved at the ICRISAT genebank. The existing wild relatives' collection from South Asian countries at the ICRISAT genebank indicated the presence of very few accessions of *S. bicolor* subsp. *drummondii* (Steud.) (9 accessions), *S. halepense* (L.) Pers. (5 accessions) and *S. pururiosericeum* (A. Rich.) Aschers. & Schweinf. Deccanense Garber (3 accessions). The collection indicated conservation of only a fraction of the total genetic variability that exists in wild relatives of sorghum and much collecting of diversity is still required. Kamala *et al.* (2014) reported the occurrence of *S. bicolor* subsp. *drummondii* as a weed wherever sorghum is cultivated in Asia; *S. halepense*, a rhizomatous perennial in Eurasia east of India; *S. propinquum* in Sri Lanka and eastern India to Myanmar and further east to the islands of South East Asia and *S. purpureosericeum* in Asia. Lazarides *et al.* (1991) reported the distribution of parasorghum species in Asia. Globally, genetic resources have primarily been conserved using *ex situ* methods, with most attention having been given to collecting and maintaining landraces; the collection of wild relatives has not been a priority (Frankel and Hawkes, 1975; Smith *et al.*, 2003; Maxted and Kell, 2009). ICRISAT had launched only a few collection missions exclusively for wild relatives. Therefore, there is a need to launch collection missions in the region under study, exclusively for wild relatives of sorghum and other mandate crops, to fill taxonomical gaps in the collection at the ICRISAT genebank, before we lose them forever due to fast urbanization, irrigation projects, natural calamities, over grazing, etc. (Upadhyaya and Gowda, 2009a).

India being the secondary centre of diversity for sorghum, the geographical gaps identified in South Asia may be considered as the potential areas for diverse sorghum (Fuller, 2002). Figure 1 indicates the reduction in traditional growing areas of sorghum and new areas of cultivation, where there is less collection. The gaps identified in the present study can be prioritized for exploration. Generally, prioritization is done by the collecting team at the time of actual launch of the collection mission depending upon the threat to diversity, availability of resources and accessibility to the target region in consultation with local government officials, NARS scientists, extension officers and non-governmental organizations, who will have a knowledge of sorghum cultivation in the districts. All reports and other publications on past collections should be considered while preparing route maps for exploring the identified gaps. While collecting germplasm, it is important to collect complete passport information including georeferenced data facilitating future mapping efforts. The passport and characterization data for the collection under study can be accessed at www.genesys-pgr.org. Seeds of all accessions under study are available at ICRISAT genebank, Patancheru, India, following the Standard Material Transfer Agreement (SMTA) of International Treaty on Plant Genetic Resources for Food and Agriculture (ITPGRFA).

Supplementary material

The supplementary material for this article can be found at http://dx.doi.org/10.1017/S147926211600023X.

Acknowledgements

Authors sincerely acknowledge the contribution of all former and present staff of genebank at ICRISAT, Patancheru, India in collection, assembly, conservation, characterization and documentation of sorghum germplasm.

References

Acheampong EN, Murthi A and William JT (1984) *A World Survey of Sorghum and Millets Germplasm*. Tome, Italy: International Plant Genetic Resources (IPGRI).

Appa Rao S (1978) Collection of pearl millet in Rajasthan, India. Genetic Resources Progress Report 4. Patancheru, India: International Crops Research Institute for the Semi-Arid Tropics (ICRISAT).

Duncan RR, Bramel-Cox PJ and Miller FR (1991) Contribution of introduced sorghum germplasm to hybrid development in the USA. In: Shands HL and Weisner LE (eds) *Use of Plant Introduction in Cultivar Development Part-1*. CSSA Special

Publication no. 17. Madison, WI, USA: Crop Science Society of America, pp. 69–102.

Espinoza L and Kelley J (2002) Grain sorghum production handbook. In: Espinoza L and Kelley J (eds.) *Grain Cooperative Extension Service*. Arkansas, USA: University of Arkansas, pp. 1–74.

FAO (Food and Agriculture Organization of the United Nations) (2009) Second Report on the State of the world's Plant Genetic Resources for Food and Agriculture (draft). Rome: FAO. Available at: ftp://ftp.fao.org/docrep/fao/meeting/017/ak528e.pdf

FAO (Food and Agriculture Organization of the United Nations) (2014) Available at: http://www.fao.org

Frankel OH and Hawkes JG (1975) *Crop Genetic Resources for Today and Tomorrow*. International biological programme 2. Cambridge, UK: Cambridge University Press.

Frederiksen RA and Duncan RR (1982) Sorghum disease in North America. In: de Milliano WAJ, Frederiksen RA and Bengston GD (ed.). *Sorghum and Millets Diseases: A Second World Review*. Patancheru, AP, India: ICRISAT, pp. 85–88.

Fuller DQ (2002) Fifty years of archaeobotanical studies in India. Laying a solid foundation. In: Settar S and Korisettar R (eds). *Indian Archaeology in Retrospect, vol. III. Archaeology and Interactive Disciplines*. New Delhi, India: Indian Council of Historical Research, pp. 247–364.

Gopal Reddy V, Mathur PN, Prasada Rao KE and Mengesha MH (1993) Rabi sorghum collection in Northern Karnataka and adjoining areas of Andhra Pradesh. Genetic Resources Progress Report 74. Patancheru, India: International Crops Research Institute for the Semi-Arid Tropics (ICRISAT).

Gopal Reddy V and Verma VD (1996) Rabi sorghum collection in Maharashtra and adjoining areas of Karnataka. Genetic Resources Progress Report 85. Patancheru, India: International Crops Research Institute for the Semi-Arid Tropics (ICRISAT).

Gumma MK, Pyla K, Thenkabail P, Reddi V, Naresh G, Mohammed I and Rafi I (2014) Crop dominance mapping with IRS-P6 and MODIS 250-m time series data. *Agriculture* 4: 113–131.

Gumma MK, Thenkabail PS and Nelson A (2011) Mapping irrigated areas using MODIS 250 meter time-series data: a study on krishna river basin (India). *Water* 3: 113–131.

IBPGR and ICRISAT (1993) *Descriptors for Sorghum [Sorghum bicolor (L.) Moench]*. Rome, Italy: International Board for Plant Genetic Resources, Patancheru, India: International Crops Research Institute for the Semi-Arid Tropics.

Jones PB, Beebe SE, Tohme J and Galwey NW (1997) The use of geographical information systems in biodiversity exploration and conservation. *Biodiversity Conservation* 6: 947–958.

Kamala V, Mourya M, Dwivedi SL and Upadhyaya HD (2014) Wild sorghums-Their potential use in crop improvement. In: Wang Y-H, Upadhyaya HD and Kole C (eds) *Genetics, Genomics and Breeding of Sorghum*. Boca Raton, London, New York: CRC Press, Taylor & Francis group, pp. 56–89.

Keuls M (1952) The use of the 'Studentized range' in connection with an analysis of variance. *Euphytica* 1: 112–122.

Lazarides M, Hacker JB and Andrew MH (1991) Taxonomy, cytology and ecology of indigenous Australian sorghum (*Sorghum bicolor* (L.) Moench Andropogonae, Poaceae). *Australian Systemic Botany* 4: 591–635.

Levene H (1960) Robust tests for equality of variances. In: Olkin I (ed.). *Contributions to Probability and Statistics: Essays in Honour of Harold Hotelling*. Stanford: Stanford University Press, pp. 278–292.

Maxted N, Dulloo E, Ford-Lloyd BV, Iriodo JM and Jarvis A (2008) Gap analysis: a tool for complementary genetic conservation assessment. *Diversity and Distributions* 14: 1018–1030.

Maxted N and Kell SP (2009) *Establishment of a Global Network for the in Situ Conservation of Crop Wild Relatives: Status and Needs*. Rome, Italy: FAO Commission on Genetic Resources for Food and Agriculture, p. 266.

Mote UN (1984) Sorghum species resistant to shoot fly. *Indian Journal of Entomology* 46: 241–243.

MS Encarta^{R} Interactive World Atlas (2000) *1995–1999 Microsoft Corporation*. One Microsoft way, Redmond, WA 98052-6399.

NASA (National Aeronautics and Space Administration) (2014) Available at: https://earthdata.nasa.gov/about/daacs/daac-lpdaac.

Newman D (1939) The distribution of range in samples from a normal population expressed in terms of an independent estimate of standard deviation. *Biometrika* 31: 20–30.

Pundir RPS (1980) Pulse germplasm collection in Punjab state of India. 4–17 April 1980. Genetic Resources Progress Report 22. Patancheru, India: International Crops Research Institute for the Semi-Arid Tropics (ICRISAT).

Ramanatha Rao V and Rajagopal Reddy C (1981) Collection of groundnut and pearl millet germplasm in Maharashtra, India. Oct-Nov 1981. Genetic Resources Progress Report 39. Patancheru, India: International Crops Research Institute for the Semi-Arid Tropics (ICRISAT).

Remanandan P (1983) Pigeonpea germplasm collection in Orissa, India. Jan 1983. Genetic Resources Progress Report 49. Patancheru, India: International Crops Research Institute for the Semi-Arid Tropics (ICRISAT).

Shannon CE and Weaver W (1949) *The Mathematical Theory of Communication*. Urbana: Univ. Illinois Press.

Smith RD, Dickie JB, Linington SH, Pritchard HW and Probert RJ (2003) *Seed Conservation: Turning Science into Practice*. Kew, UK: Royal botanic gardens.

Snedecor GW and Cochran WG (1980) *Statistical Methods*. 7th edn. Ames: Iowa State University Press.

Srinivasa Rao P, Reddy BVS, Nagaraj N and Upadhyaya HD (2014) Sorghum production for diversified uses. In: Wang Y-H, Upadhyaya HD and Kole C (eds) *Genetics, Genomics and Breeding of Sorghum*. Boca Raton, London, New York: CRC Press, Taylor & Francis Group, pp. 1–28.

Upadhyaya HD and Gowda CLL (2009a) Managing and enhancing the use of germplasm – strategies and methodologies. Technical Manual no.10. Patancheru, Andhra Pradesh, India: ICRISAT. p. 236.

Upadhyaya HD, Reddy KN, Irshad Ahmed M and Gowda CLL (2010) Identification of gaps in pearl millet germplasm from Asia conserved at the ICRISAT genebank. *Plant Genetic Resources: Characterization and Utilization* 8: 267–276.

Upadhyaya HD, Reddy KN, Irshad Ahmed M and Gowda CLL (2012) Identification of gaps in pearl millet germplasm from East and Southern Africa conserved at the ICRISAT genebank. *Plant Genetic Resources: Characterization and Utilization* 10: 202–213.

Upadhyaya HD, Reddy KN, Irshad Ahmed M, Gowda CLL and Bettina H (2009b) Identification of geographical gaps in the pearl millet germplasm conserved at ICRISAT genebank from West and Central Africa. *Plant Genetic Resources: Characterization and Utilization* 8: 45–51.

Upadhyaya HD, Reddy KN, Irshad Ahmed M, Gowda CLL, Thimma Reddy M and Ramachandran S (2015) Identification of gaps in pigeonpea germplasm from East and Southern

Africa conserved at the ICRISAT genebank. *Indian Journal of Plant Genetic Resources* 28: 180–188.

Upadhyaya HD, Reddy KN, Pundir RPS, Singh S, Gowda CLL and Irshad Ahmed M (2013) Diversity and geographical gaps in *Cajanus scarabaeoides* (L.) Thou. germplasm conserved at the ICRISAT genebank. *Plant Genetic Resources: Characterization and Utilization* 11: 3–14.

Upadhyaya HD, Reddy KN, Singh S, Gowda CLL, Irshad Ahmed M and Kumar V (2014) Diversity and gaps in *Pennisetum glaucum* subsp. *monodii* (Maire) Br. germplasm conserved at the ICRISAT genebank. *Plant Genetic Resources: Characterization and Utilization* 12: 226–235.

USGS (United States Geological Survey) (2014) Available at: http://earthexplorer.usgs.gov/, https://lpdaac.usgs.gov/

van der Maesen (1980) Germplasm collection in Sri Lanka. Genetic Resources Progress Report 15. Patancheru, India: International Crops Research Institute for the Semi-Arid Tropics (ICRISAT).

VSN International (2010) *GenStat Software for Windows*. Release 13.1. Hemel Hempstead, UK: VSN International Ltd.