

Geographical distribution, diversity and gap analysis of East African sorghum collection conserved at the ICRISAT genebank

Hari D Upadhyaya^{*,1,2,3}, Narsimha Reddy K¹, Vetriventhan M¹, Murali Krishna G⁴, Irshad Ahmed M⁴, Manyasa E⁵, Thimma Reddy M¹, Shailesh Kumar Singh¹

¹Genebank, International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru 502 324, Telangana, India

²Department of Agronomy, Kansas State University, Manhattan, KS 66506, USA

³The UWA Institute of Agriculture, the University of Western Australia, Crawley, WA 6009, Australia

⁴Remote sensing and GIS Unit, International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru 502 324, Telangana, India

⁵International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Nairobi, Kenya

*Corresponding author: h.upadhyaya@cgiar.org

Abstract

The aim of the investigation was to assess the geographical distribution, diversity and gaps in sorghum collection from East African countries conserved at the ICRISAT genebank. The collection represents a total of 12,750 accessions including 11,672 landraces, 877 breeding materials, six improved cultivars, and 195 wild accessions. Passport data and FloraMap, a GIS software were used to assess geographical distribution and identify gaps. Range, mean, variance and phenotypic diversity index were estimated using GENSTAT 13.1. to assess the diversity in the collection. Cultivated sorghums classified into races and intermediate races based on spikelet and panicle morphology differed significantly, and races showed more variation than intermediate races for days to 50% flowering in postrainy season, plant height in rainy season, basal tillers per plant, panicle length, seed width and 100 seed weight. A total of 153 districts located in 50 provinces of 10 East African countries were found as the geographical gaps. Probably due to large variation for maturity, timing of collecting mission, and accessibility to the area under sorghum cultivation, both North and South Sudan were found as the major gaps with seven and 50 districts, respectively. The wild sorghum collection from East African countries belongs to *S. bicolor*, *S. halepense*, *S. lanceolatum*, *S. macrochaeta*, *S. purpureosericeum* and *S. versicolor*. Remaining species of genus *sorghum* were considered as taxonomic gaps. The gaps identified in the present study need to be explored on a priority basis to collect and conserve most diverse sorghum germplasm.

Submitted 7 Sep 2016; Revised 17 Feb 2017; 21 Feb 2017.

Key Words: Genetic resources; Geographical gap; Germplasm; Landrace; Variability.

Introduction

Sorghum [*Sorghum bicolor* (L.) Moench] is the fifth most important cereal crop grown in both temperate and tropical regions in Africa, Asia, Europe and the Americas. It is a versatile crop capable of growing well under contrasting climatic conditions and high elevations up to 40° latitude in East Africa (Espinoza and Kelly, 2002; Srinivasa Rao et al., 2014). Sorghum is a staple for over 500 million of the world's poorest people, used for food, feed and fodder purposes (Mace et al., 2013). Its alternative uses include production of beer, alcohol, syrup, bakery items, industrial starch, etc. Sorghum is grown on an estimated area of 44.79 m ha with a production of 70.83 Mt (FAO, 2014, accessed on 11 June 2016). Area wise, Sudan (8.4 m ha), India (5.8 m ha), Nigeria (5.4 m ha), Niger (3.6 m ha), USA (2.6 m ha) and Mexico (2.0 m ha) are the major sorghum growing countries in the world. In East Africa, it is grown over an estimated area of 12.3 m ha in Sudan (8.4 m ha), Ethiopia (1.83 m ha), Tanzania (0.80 m ha), Uganda (0.37 m ha), Somalia (0.24 m ha), Eritrea (0.25 m ha), Kenya (0.21 m ha), Rwanda (0.11 m

ha), Malawi (0.09 m ha) and Burundi (0.03 m ha) (FAO, 2014).

Crop losses due to extremes in the environment have risen steadily over the past several decades. Climate models predicted an increased incidence of floods, droughts, and extreme temperatures (Dwivedi et al., 2013; Mickelbart et al., 2015). Genetic erosion was estimated at about 72% and the importance of diversity loss is becoming more important (Hammer et al., 1996). Genetic erosion occurs due to several factors, including replacement of landraces by improved cultivars, concomitant natural catastrophes (droughts, floods, fire hazards, etc.), human settlements, overgrazing, climate change, destruction of plant habitats for irrigation projects, etc. (Upadhyaya and Gowda, 2009). Genetic loci that ensure productivity in challenging environments exist within the germplasm of crops. Therefore, collection and conservation of crop genetic resources are most important, so as to ensure the availability of a wide genetic base to present and future crop improvement programs. Landraces have been recogni-

sed as a source of traits of local adaptation, stress tolerance, yield stability and seed nutrition (Dwivedi et al., 2016). To achieve near completion of species diversity assembling, there is an urgent need for critical assessment of existing collections for status, genetic diversity and gaps, and for launching germplasm collection missions in un- and under-explored areas. Geographic Information Systems (GIS), such as FloraMap, DIVAGIS, ARCGIS, Maxent and remote sensing have enabled better understanding of species distributions, crop cultivation and the representativeness of germplasm collections. The potential of new science tools has opened up data repositories at no cost (NASA, 2014; USGS, 2014). Using these GIS software and spatial data, geographical and taxonomic gaps were identified in different crops to enrich the existing collections for diversity (Jones et al., 1997; Maxted et al., 2008; Upadhyaya et al., 2013, 2014a, 2015).

The genebank at the International Crops Research Institute for the Semi-Arid Tropics, (ICRISAT) India, conserves 39,923 accessions of sorghum germplasm from 93 countries, including 12,750 accessions from 10 East African countries. East Africa, being the primary center of diversity for sorghum, identification of gaps with respect to geographical and species diversity provides important sources for diversity of cultivated, wild and weedy relatives that will enrich the genebank collection. Therefore, this study aimed to assess the geographical distribution, diversity and gaps in the sorghum germplasm collection from East African countries conserved at the ICRISAT genebank, to enrich the sorghum gene pool.

Results

Current status of collection

The world collection of sorghum germplasm from East African countries (12,750 accessions) conserved at the ICRISAT genebank, India, is from a wide range of latitudes ranging from 17.03°S (Malawi) to 19.33°N (Sudan). The collection includes accessions from Burundi (140), Eritrea (193), Ethiopia (4,392), Kenya (998), Malawi (425), Rwanda (291), Somalia (447), Sudan (3,171), Tanzania (821) and Uganda (1,872) (Table 1). Initially, ICRISAT assembled the sorghum germplasm by (i) introducing already collected germplasm from various organizations located in different countries; and (ii) launching systematic germplasm collection missions in East African countries.

Germplasm introduced

A total of 10,764 accessions originated in 10 East African countries were introduced from 31 organizations located in 14 countries (Table 1). A maximum of 3,121 accessions originating from seven countries were introduced from the Rockefeller Foundation, New Delhi, India. Another major donor organization was the Ethiopian sorghum improvement project, Nazareth (now Melkasa), which donated 2,338 accessions originating in Eritrea (6) and Ethiopia (2,332). Other donors were Gezira Agricultural Research Station, Wad Medani, Sudan (927); Royal Botanical Gardens, Kew, United Kingdom (792); Agricultural Research Corporation, Wad Medani, Sudan (634); Food and Agriculture Organization (FAO), Rome, Italy (422); and Mayaguez Institute of Tropical Agriculture, Mayaguez, USA (377). All other organizations donated less than 300 accessions to ICRISAT. All accessions from Eritrea were introductions into the collection.

Germplasm collected

During 1974 - 1996, ICRISAT and its partners have launched 216 collection missions in 62 countries for the germplasm of its mandate crops and collected more than 33,000 samples. Nineteen missions launched in East African countries resulted in a total of 1,986 sorghum samples (Table 1). Five collection missions launched in Tanzania resulted in 415 accessions, followed by two collection missions each in Ethiopia (including Eritrea) (131), Malawi (294), Somalia (391), Sudan (160), and Uganda (448) and one mission each in Burundi (72), Kenya (2), and Rwanda (73). ICRISAT collaborated for germplasm collection with Institut Des Sciences Agronomiques Du Burundi, Bujumbura in Burundi (Prasada Rao and Mengesha, 1982a); Ethiopian Sorghum Improvement Project, Nazareth in Ethiopia (Prasada Rao and Mengesha, 1981); University of Nairobi in Kenya; International Bureau of Plant Genetic Resources (IBPGR) Italy and Ministry of Agriculture and Natural Resources in Malawi (Appa Rao, 1979); Institut Des Agronomiques Du Rwanda (ISAR), Butare in Rwanda (Prasada Rao and Mengesha, 1982b), Food and Agriculture Organization (FAO), Agriculture Research Station, Bonka and Central Agricultural Research Station, Afgoy in Somalia (Ramanatha Rao, 1979); IBPGR, Italy; Ministry of Agriculture, International Sorghum and Millet Collaborative Research Support Program (INTSORMIL) and Norwegian Church Aid Sudan Program in Sudan (Prasada Rao and Mengesha, 1979); IBPGR, Italy, University of Dar es Salaam, International Institute of Tropical Agriculture (IITA)/United States Agency for International Development (USAID), Ministry of Agriculture, Zanzibar and Tanzanian Agricultural Research Institute (TARI), Ilonga, Tanzania in Tanzania (Prasada Rao and Mengesha 1979; Remanandan and Mengesha, 1981) and Makerere University, Kampala, Uganda in Uganda (Singh et al., 1991; Gopal Reddy et al., 1993).

Sources of germplasm

Majority of the sorghum accessions from East African countries assembled at ICRISAT genebank were received from various institutes (9,970 accessions) located in various countries, while remaining were from farmers' field (2,490 accessions), wild habitats (193 accessions), farm stores (70 accessions), local markets (16 accessions) and threshing floors (11 accessions).

Intensity of germplasm collection

The sorghum collection from East African countries represents a total of 971 geographical sites of germplasm collection in Burundi (38), Eritrea (5), Ethiopia (171), Kenya (156), Malawi (50), Rwanda (21), Somalia (47), Sudan (134), Tanzania (112), and Uganda (237). The average number of samples per geographical site (18 x 18 sq km) in the entire collection from East African countries was nine. The average number of sorghum samples collected per geographical site was 37 in Eritrea, 20 in Ethiopia, 11 in Sudan, eight in Somalia, six in Malawi, four each in Kenya, Uganda and Tanzania, and three in Burundi.

Biological status of the collection

The sorghum collection from East African countries includes 11,672 landraces, 877 breeding materials, 195 wild accessions, and six improved cultivars (Table 1). The cultivated sorghum collection includes 12,555

Table 1. Status of sorghum germplasm assembled from East African countries at ICRISAT genebank, India.

Country	Collections		No. of accessions introduced	Total accs.	Biological status of collection				Landraces with latitude data
	Year of collection	Collection missions launched			No. of accessions collected	Improved cultivars	Breeding materials	Landraces	
Burundi	1982	1	72	68		1	135	4	112
Eritrea				193		5	184	4	183
Ethiopia	1981	1	129	4261	3	75	4278	36	3416
	1984	1	2						
Kenya	1976	1	2	996	1	14	961	22	629
Malawi	1979	1	288	131		5	401	19	291
	1983	1	6						
Rwanda	1982	1	73	218		1	289	1	76
Somalia	1979	1	120	56		1	442	4	383
	1987	1	271						
Sudan	1979	1	130	3011	1	384	2729	57	1417
	1983	1	30						
Tanzania	1978	1	51	406	1	62	741	17	403
	1979	1	53						
	1981	1	44						
	1985	1	5						
	1987	1	262						
Uganda	1991	1	126	1424		329	1512	31	1004
	1993	2	322						
Total		19	1986	10764		877	11672	195	7914

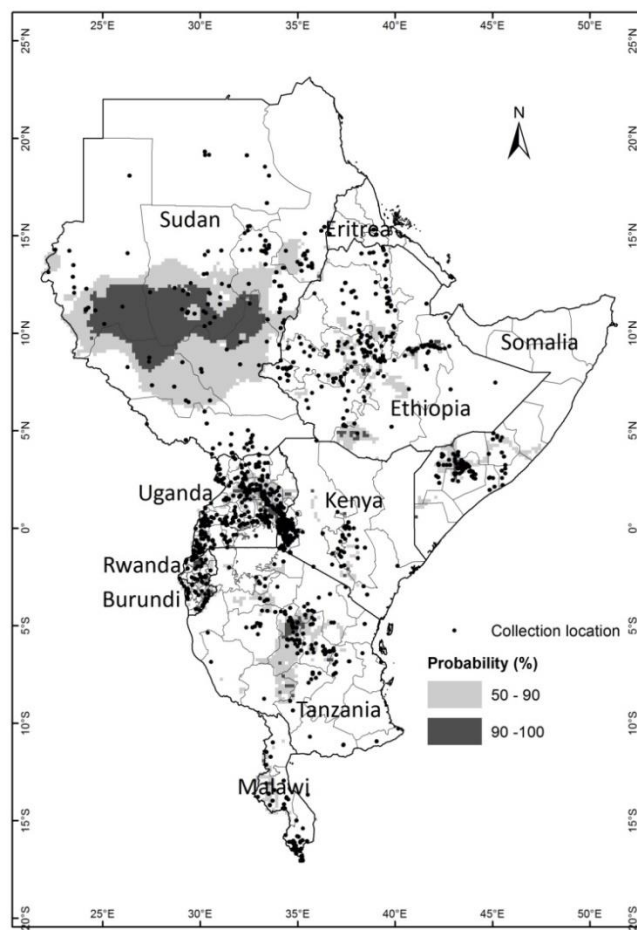


Fig. 1. Geographical distribution and the gaps (districts shaded) identified in the sorghum germplasm collection from East African countries conserved at ICRISAT genebank, India.

Table 2. Geographical distribution of races and wild sorghum germplasm from East Africa at ICRISAT genebank, India.

Race/Species	1*	2	3	4	5	6	7	8	9	10	Total
Basic races											
<i>Bicolor</i>		9	144	29	8	2	1	569	110	93	965
<i>Guinea</i>	3	2	14	8	259			39	373	37	735
<i>Caudatum</i>	107	30	387	749	19	220	12	995	105	1286	3910
<i>Kafir</i>			6	1				23	3	20	53
<i>Durra</i>	4	92	1870	5	5	44	424	229	51	16	2740
Intermediate races											
<i>Guinea-bicolor</i>	2		10	1	18			10	10	4	55
<i>Caudatum-bicolor</i>	3	7	142	28	15	1		131	20	52	399
<i>Kafir-bicolor</i>			2					11	1	3	17
<i>Durra-bicolor</i>	1	21	1078	6	1	1	1	48	5	14	1176
<i>Guinea-caudatum</i>	12	2	207	108	60	2	2	528	97	245	1263
<i>Guinea-kafir</i>										2	2
<i>Guinea-durra</i>		2	8	2	12	1		26	6	2	59
<i>Kafir-caudatum</i>			3	1	1			40	4	5	54
<i>Durra-caudatum</i>	3	22	460	38	8	19	3	312	17	46	928
<i>Kafir-durra</i>	1		10					2		11	24
No information		2	15					151	2	5	175
Cultivated total	136	189	4356	976	406	290	443	3114	804	1841	12555
Wild											
<i>S. bicolor</i> ssp. <i>arundinaceum</i>											
<i>race aethiopicum</i>								5			5
<i>race arundinaceum</i>				3	4			2	3	1	13
<i>race virgatum</i>							3	3			6
<i>race verticilliflorum</i>	3		12	4	2			7	1	17	46
<i>S. bicolor</i>	1		25	14	7	1		35	5	11	99
ssp. <i>drummondii</i>											
<i>S. purpureosericeum</i> ssp. <i>deccanense</i>							1		1		2
<i>S. purpureosericeum</i> ssp. <i>dimidiatum</i>								2			2
<i>S. halepense</i>			1								1
<i>S. lanceolatum</i>								3			3
<i>S. macrochaeta</i>										1	1
<i>S. versicolor</i>									1		1
No information			2	1	6				6	1	16
Wild total	4	4	36	22	19	1	4	57	17	31	195
Grand total	140	193	4392	998	425	291	447	3171	821	1872	12750

*1 = Burundi, 2 = Eritrea, 3 = Ethiopia, 4 = Kenya, 5 = Malawi, 6 = Rwanda, 7 = Somalia, 8 = Sudan, 9 = Tanzania, 10 = Uganda

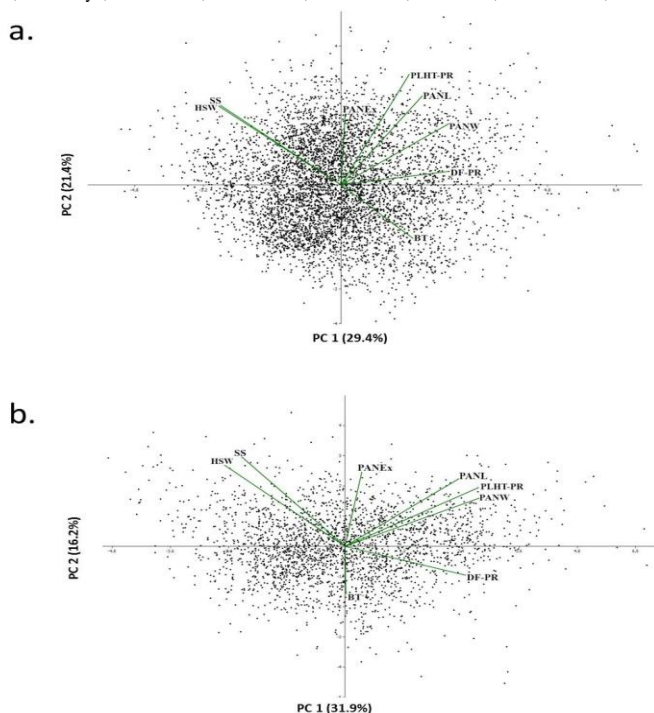


Fig 2. Scatter plots for races (a) and intermediate races (b) showing first two Principal Components (PCs) and percentage of variation. (DF-PR=Days to 50% flowering-postrainy, PLHT=Plant height-postrainy, BT=Basal tillers per plant, PANEx=Panicle exertion, PANL=Panicle length, PANW=Panicle width, SS=Seed width and HSW=100 seed weight).

Table 3. Geographical gaps identified in the world collection of sorghum from East Africa at ICRISAT genebank, India.

Country	Province/State (district)
Burundi	Cankuzo (Cendajuru, Gisagara, Kankuzo, Kigamba); Cibitoke (Mabay); Gitega (Itaba); Kirundo (Gitobe); Makamba (Kayogoro, Mabanda, Makamba); Muyinga (Buhinvuza, Butihinda, Gashoho, Mwakiro); Ngozi (Gashikanwa, Marangara, Nivamurenza); Rutana (Bukemba, Gtanga, Mpinga-Kayove, Musongati); Ruyigi (Bultezi, Butaganzwa2, Bweru)
Eritrea	-
Ethiopia	Amhara (West Gojam); Oromiya (Arsi, Borena); Southern (Bench)
Kenya	Eastern (Kitui, Machakos, Makueni); Rift Valley (Kajiado, Samburu)
Malawi	Central (Dowa, Kasungu, Lilongwe); Northern (Mizimba)
Sudan	Central (Blue Nile, White Nile); Darfur (Northern Darfur, Southern Darfur); Eastern (Kassala); Kordofon (North Korodofon, South Korodofon), Jonglei (Akobo, Ayod, Duk, Fangak, Khorflus, Nyirol, Twic East, Urur); Lakes (Cueibet, Rumbek Center, Rumbek East, Rumbek North, Wulu, Yirol East, Yirol West); Northern Bahr el Ghazal (Aweil Center, Aweil East, Aweil North, Aweil South, Aweil West); Unity (Abiemnhom, Guit, Koch, Leer, Mayendit, Mayom, Panyijar, Panyijiar, Pariang, Rubkona); Upper Nile (Baliet, Fashoda, Longochuk, Luakpiny/Nasir, Maban, Maiwut, Malakal, Manyo, Melut, Panyikang, Renk, Ulang); Warrap (Abyei, Gogrial East, Gogrial West, Tonj East, Tonj North, Twic); Western Bahr el Ghazal (Jur River, Wau)
Rwanda	Byumba (Buyoga, Giti, Gituza, Kinyami, Muhura, Rutare, Tumba); Gisenyi (Kibirira, Ramba); Gitarama (Bulinga, Kayenzi, Muambira, Mugina, Mukingi, Mushubati, Nyabikenke, Nyakabanda, Nyamabuye, Runda, Rutobwe, Taba); Kibungu (Kabarondo, Kayonza, Muhazi); Kigali (Gikomero, Gikoro, Mbogo, Mugambazi, Musasa, Rubungo, Rushashi, Rutongo, Shyorongi, Tare); Ruhengeri (Ndusu)
Somalia	Bakool (Rab Dhuure, Tiyeeglow); Galguduud (El Dere); J. Dhexe (Saakow); J. Hoose (Afmadow)
Uganda	Iganga (Bukooli, Busiki); Kamuli (Bulamogi); Lira (Kioga, Moroto); Masindi (Bujenje, Buruli, Kibanda); Moroto (Kadam (Chekwii), Pian); Pallisa (Butebo); Soroti (Kapelebyong); Tororo (Bunyole, Kisoko (West Budama))
Tanzania	Arusha (Hanang); Mara (Bunda); Mbeya (Mbeya); Shinyanga (Kahama, Kisoko); Singida (Southern Manyoni)

Table 4. Frequency distribution (%) for different qualitative traits of sorghum germplasm from East Africa assembled and characterized at ICRISAT genebank, India.

Trait	Frequency distribution (%)
Plant pigmentation	Pigmented (97.5), Tan (2.4), No data (0.1)
Nodal tillering	Absent (5.4), Present (94.4), No data (0.2)
Midrib color	Brown (0.03), Dull green (14.9), White (79.9), Yellow (5.1) No data (0.07)
Panicle-compactness	Compact elliptic (8.1), Compact oval (4.3), Loose drooping branches (4.5), Loose stiff branches (10.8), Semi-compact elliptic (33.3), Semi-compact oval (2.3), Semi-loose drooping branches (1.1), Semi-loose stiff branches (32.9), Very loose drooping branches (2.0), Very loose stiff branches (0.6), No data (0.10),
Glume color	Black (29.0), Brown (2.7), light brown (5.1), Light red (3.3), Partly straw and brown (7.7), Partly straw and purple (7.5), Purple (6.4), Red (7.5), Reddish brown (5.5), Straw (24.9), White (0.01), Yellow (0.01), No data (0.1)
Glume covering	Grain fully covered (1.5), Grain uncovered (1.1), Half grain covered (36.3), One-fourth grain covered (50.5), Three-fourth grain covered (10.5), No data (0.02)
Seed color	Brown (9.8), Chalky white (3.6), Grey (5.4), Light brown (5.0), Light red (10.4), Purple (1.6), Red (3.8), Reddish brown (28.7), Straw (15.4) Straw and red mixed (0.03), White (9.2), White and red mixed (0.01), Yellow (6.8), No data (0.26)
Seed lustre	Lustrous (49.3), Non-Lustrous (50.3), No data (0.4)
Seed subcoat	Absent (68.9), Present (31.0), No data (0.1)
Endosperm texture	Completely corneous (1.1), Completely starchy (27.8), Mostly corneous (7.4), Mostly starchy (36.7), Partly corneous (26.7), No data (0.3)
Threshability	Difficult to thresh (4.5), Freely threshable (74.2), Partly threshable (20.3), No data (1.0)

Table 5. Range of variation for agronomic traits of different sorghum races from East African countries at ICRISAT genebank, India.

Race/country	1*	2	3	4	5	6	7	8	9	10
Basic races	42-170	41-154	75-615	70-415	1-10	0-52	5-80	1-33	0.8-6	0.4-7.0
<i>Bicolor</i>	46-145	49-105	95-560	110-375	1-5	0-52	10-80	4-33	0.8-3.5	0.4-4.4
<i>Guinea</i>	44-159	52-132	100-570	110-380	1-6	0-51	7-64	4-30	1.5-4.5	1.2-4.9
<i>Caudatum</i>	42-165	41-154	75-615	70-390	1-10	0-49	9-55	1-33	0.8-5	0.6-6.5
<i>Kafir</i>	53-138	55-112	115-340	95-250	1-3	10-25	17-39	4-14	2.5-4	1.6-4.1
<i>Durra</i>	47-170	48-116	125-580	105-415	1-5	0-50	5-51	3-28	1-6	0.5-7.5
Intermediate races	44-190	42-139	70-590	60-580	1-5	0-71	7-52	2-40	0.8-5	0.6-7.0
<i>Guinea-bicolor</i>	55-150	54-110	190-520	140-320	1-5	3-35	19-50	6-24	2-4	1.4-4.2
<i>Caudatum-bicolor</i>	47-151	52-124	95-540	100-380	1-4	0-49	10-45	3-30	0.8-4.5	0.8-6.1
<i>Kafir-bicolor</i>	48-128	50-97	175-560	140-320	1-2	0-20	21-29	5.5-12	2.5-3.5	1.8-3.2
<i>Durra-bicolor</i>	44-168	50-130	90-580	80-580	1-5	0-71	8-51	3-40	1.5-4	0.6-7.0
<i>Guinea-caudatum</i>	44-158	42-139	115-590	85-420	1-5	0-48	12-52	2-29	0.8-5	0.9-5.7
<i>Guinea-durra</i>	47-152	52-101	145-490	90-325	1-4	3-39	14-32	4-15	2-3.5	2.0-5.6
<i>Kafir-caudatum</i>	51-149	46-98	145-420	110-275	1-4	0-40	14-40	4.5-19	2.5-4.5	1.7-4.9
<i>Durra-caudatum</i>	44-190	45-130	70-580	60-360	1-5	0-49	7-49	4-35	1.8-5	1.0-6.7
<i>Kafir-durra</i>	57-140	54-84	105-500	105-350	1-2	3.5-35	18-27	7.5-14	2.5-3.5	1.5-3.6
Countries										
Burundi	58-150	59-126	145-450	125-350	1-5	0-35	10-30	5-16	2-3	1.0-3.1
Eritrea	44-147	48-88	95-465	110-330	1-4	0-39	6-41	3-23	1.8-4.5	1.7-6.6
Ethiopia	42-170	50-116	70-590	60-580	1-5	0-71	7-55	3-40	0.8-4.5	0.4-7.0
Kenya	48-153	53-149	160-540	100-360	1-10	0-38	9-45	2-35	1-4.5	0.9-4.1
Malawi	56-138	52-117	140-530	100-295	2-5	0-36	12-54	4-30	1.8-4.5	1.2-4.9
Rwanda	96-140	63-107	180-500	85-310	1-3	0-33	11-28	5-27	2.5-3	1.3-3.3
Somalia	59-95	58-106	215-420	120-310	1-4	0-37	5-28	4-16	2-3.5	2.4-4.6
Sudan	45-190	41-108	100-615	85-380	1-6	0-48	7-44	1-25	1.8-6	1.3-7.5
Uganda	46-160	49-154	170-550	90-400	1-5	0-52	8.5-64	3-33	1-3.6	0.6-4.4
Tanzania	66-162	65-135	220-570	115-395	1-5	0-37	7-80	4-28	0.8-5	0.7-4.7
Entire collection	42-190	41-154	70-615	60-580	1-10	0-71	5-80	1-40	0.8-6	0.4-7.5

*1 = Days to 50% flowering-rainy, 2 = Days to 50% flowering-postrainy, 3 = Plant height (cm)-rainy, 4 = Plant height (cm)-postrainy, 5 = Basal tillers per plant, 6 = Panicle exertion (cm), 7 = Panicle length (cm), 8 = Panicle width (cm), 9 = Seed width (mm), 10 = Seed weight (g).

Table 6. Mean values for agronomic traits of basic and intermediate races of sorghum races from East African countries at ICRISAT genebank, India.

Race/Country	1*	2	3	4	5	6	7	8	9	10
<i>Bicolor</i>	104.9ab	74.6bc	388.8abc	248.2ab	2.3b	21.7a	29.4a	13.9a	2.5e	2.0e
<i>Guinea</i>	103.6ab	89.9a	400.0ab	247.6ab	2.8a	18.0ab	27.7ab	11.9ab	2.6de	2.2de
<i>Caudatum</i>	104.7ab	75.6bc	340.6cde	207.5c	2.1bc	17.0ab	20.1de	8.9c	3.0bc	2.6bcd
<i>Kafir</i>	64.7e	69.8bc	167.1g	132.4e	1.6cd	18.3ab	24.8abcd	7.8c	3.2b	2.9abc
<i>Durra</i>	90.8bc	76.6bc	400.5ab	246.5ab	1.8bcd	20.2ab	18.7e	8.3c	3.1bc	3.3a
<i>Guinea-bicolor</i>	92.9bc	72.4bc	351.9bcd	212.5bc	2.7a	19.7ab	26.2abc	11.9ab	2.8bcd	2.5cde
<i>Caudatum-bicolor</i>	89.2bc	71.5bc	327.5de	211.0bc	1.9bcd	19.3ab	23.9bcde	9.6bc	2.9bc	2.8abcd
<i>Kafir-bicolor</i>	69.8de	66.9cd	294.4ef	213.8bc	1.4d	12.7b	23.4bcde	8.3c	2.9bc	2.5cde
<i>Durra-bicolor</i>	111.7a	79.7b	414.2a	264.7a	1.7bcd	18.8ab	27.8ab	13.7a	2.8cd	2.4cde
<i>Guinea-caudatum</i>	94.5bc	75.1bc	324.4de	203.9c	2.2b	16.0ab	22.6bcde	9.7bc	2.9bc	2.9abc
<i>Guinea-durra</i>	76.2cde	70.8bc	312.2def	197.1c	2.1bc	18.9ab	22.5bcde	8.6c	2.9bc	3.3a
<i>Kafir-caudatum</i>	81.9cd	61.0d	273.6f	166.0d	1.8bcd	14.3ab	22.9bcde	8.2c	3.5a	3.1ab
<i>Durra-caudatum</i>	94.1bc	73.7bc	363.6abcd	224.1bc	1.8bcd	19.3ab	21.9cde	9.1bc	3.1bc	3.2ab
<i>Kafir-durra</i>	82.9cd	73.6bc	347.1bcde	227.9bc	1.7bcd	20.1ab	24.1bcde	9.6bc	3.0bc	2.7bcd
Basic races	98.7a	77.1a	370.9a	227.7a	2.1a	18.5b	20.5b	9.1b	3.0a	2.9a
Intermediate races	99.3a	75.8b	365.0b	230.3a	1.9b	18.1a	24.4a	10.9a	2.9b	2.8b
F value	100.2	10.5	75.0	22.8	0.9	4.7	70.8	158.2	63.2	33.0
Probability	***	**	***	***	ns	*	***	***	***	***
Countries										
Burundi	124.9ab	92.9b	342.1c	234.3c	1.9e	13.0f	18.3f	7.9c	2.8d	1.9f
Eritrea	87.8f	65.0f	365.7b	176.8e	1.9e	19.3a	18.7ef	7.6c	3.0b	3.7a
Ethiopia	99.0e	77.7d	408.0a	256.6b	1.7f	20.5a	23.7b	10.5a	3.0b	3.0c
Kenya	109.2d	79.9d	367.7b	232.5c	2.1d	15.8cd	19.6ef	8.8b	3.0b	2.5d
Malawi	83.2g	83.3c	361.9b	207.0d	3.2a	20.0a	25.5a	10.9a	2.6e	2.1ef
Rwanda	122.0b	83.9c	339.1c	232.0c	1.4g	10.9g	20.1de	9.5b	2.9c	2.2e
Somalia	73.4h	74.9e	344.1c	199.4d	2.1d	15.0de	8.1g	6.2d	2.8c	3.1b
Sudan	83.1g	65.0f	291.9e	182.4e	1.9e	17.2bc	21.3cd	8.0c	3.3a	3.6a

Uganda	114.5c	78.1d	346.3c	207.5d	2.6b	18.5ab	21.8c	10.6a	2.6e	2.0f
Tanzania	127.8a	97.9a	415.1a	267.5a	2.4c	13.6ef	23.4b	11.0a	2.6e	2.2e
Entire collection	98.9	76.7	369	228.5	2.0	18.4	21.8	9.6	3.0	2.8

*1 = Days to 50% flowering-rainy, 2 = Days to 50% flowering-postrainy, 3 = Plant height (cm)-rainy, 4 = Plant height (cm)-postrainy, 5 = Basal tillers per plant, 6 = Panicle exertion (cm), 7 = Panicle length(cm), 8 = Panicle width (cm), 9 = Seed width (mm), 10 = Seed weight (g)
Means were tested following Newman-Keuls test and variances were tested using Levene's test *P<0.05; **P<0.01; ***P<0.001
Number of *guinea-kafir* accessions are few, therefore not included in analysis

Table 7. Shannon-Weaver diversity index (H') for different traits of sorghum races from East African countries at ICRISAT genebank, India.

	*1	2	3	4	5	6	7	8	9	10	Mean	SD
Races												
<i>Bicolor</i>	0.590	0.617	0.599	0.605	0.415	0.610	0.585	0.567	0.551	0.624	0.576	0.061
<i>Guinea</i>	0.562	0.586	0.622	0.616	0.551	0.600	0.616	0.578	0.516	0.590	0.584	0.033
<i>Caudatum</i>	0.606	0.598	0.631	0.625	0.525	0.602	0.611	0.549	0.478	0.624	0.585	0.051
<i>Kafir</i>	0.165	0.410	0.298	0.298	0.403	0.567	0.541	0.517	0.424	0.563	0.419	0.134
<i>Durra</i>	0.560	0.613	0.612	0.610	0.387	0.620	0.564	0.554	0.616	0.623	0.576	0.072
<i>Guinea-bicolor</i>	0.435	0.465	0.565	0.560	0.506	0.552	0.413	0.499	0.508	0.484	0.499	0.051
<i>Caudatum-bicolor</i>	0.561	0.554	0.600	0.624	0.490	0.605	0.621	0.561	0.656	0.619	0.589	0.048
<i>Kafir-bicolor</i>	0.320	0.391	0.320	0.423	0.287	0.391	0.391	0.452	0.423	0.452	0.385	0.058
<i>Durra-bicolor</i>	0.539	0.627	0.602	0.630	0.391	0.590	0.634	0.608	0.661	0.608	0.589	0.077
<i>Guinea-caudatum</i>	0.575	0.583	0.624	0.626	0.528	0.584	0.608	0.573	0.422	0.637	0.576	0.063
<i>Guinea-durra</i>	0.460	0.543	0.538	0.606	0.511	0.613	0.574	0.523	0.571	0.535	0.547	0.046
<i>Kafir-caudatum</i>	0.488	0.448	0.464	0.522	0.307	0.569	0.524	0.397	0.564	0.602	0.489	0.089
<i>Durra-caudatum</i>	0.576	0.602	0.623	0.621	0.438	0.614	0.623	0.533	0.440	0.625	0.570	0.074
<i>Kafir-durra</i>	0.346	0.415	0.501	0.501	0.260	0.469	0.415	0.415	0.346	0.555	0.422	0.088
Mean	0.484	0.532	0.543	0.562	0.429	0.570	0.551	0.523	0.513	0.581	0.529	
SD	0.128	0.087	0.111	0.098	0.095	0.065	0.085	0.063	0.096	0.057	0.073	
Countries												
Burundi	0.536	0.539	0.587	0.600	0.446	0.557	0.605	0.504	0.508	0.581	0.546	0.050
Eritrea	0.547	0.607	0.573	0.591	0.395	0.626	0.587	0.508	0.458	0.613	0.551	0.075
Ethiopia	0.597	0.633	0.617	0.624	0.354	0.616	0.635	0.552	0.466	0.625	0.572	0.092
Kenya	0.603	0.601	0.614	0.600	0.380	0.598	0.609	0.512	0.509	0.633	0.566	0.078
Malawi	0.506	0.565	0.613	0.619	0.470	0.612	0.609	0.590	0.559	0.572	0.571	0.050
Rwanda	0.594	0.597	0.577	0.618	0.316	0.541	0.579	0.505	0.430	0.601	0.536	0.095
Somalia	0.585	0.591	0.630	0.624	0.425	0.602	0.446	0.593	0.542	0.625	0.566	0.074
Sudan	0.549	0.568	0.616	0.621	0.510	0.606	0.627	0.562	0.542	0.624	0.582	0.042
Uganda	0.567	0.569	0.628	0.621	0.463	0.598	0.595	0.602	0.543	0.618	0.580	0.049
Tanzania	0.546	0.602	0.621	0.630	0.436	0.576	0.609	0.554	0.521	0.606	0.570	0.059
Mean	0.563	0.587	0.607	0.615	0.419	0.593	0.590	0.548	0.508	0.610	0.564	
SD	0.032	0.027	0.021	0.013	0.058	0.027	0.053	0.039	0.043	0.020	0.015	

* 1 = Days to 50% flowering-rainy, 2 = Days to 50% flowering-postrainy, 3 = Plant height (cm)-rainy, 4 = Plant height (cm)-postrainy, 5 = Basal tillers per plant, 6 = Panicle exertion (cm), 7 = Panicle length (cm), 8 = Panicle width (cm), 9 = Seed width (mm), 10 = Seed weight (g)

accessions belonging to five basic and 10 intermediate races (Table 2). Harlan and de Wet (1972) classified the cultivated sorghums into five basic races (*bicolor*, *guinea*, *caudatum*, *kafir* and *durra*) and 10 intermediate races (*guinea-bicolor*, *caudatum-bicolor*, *kafir-bicolor*, *durra-bicolor*, *guinea-caudatum*, *guinea-kafir*, *guinea-durra*, *kafir-caudatum*, *durra-caudatum* and *kafir-durra*) based on the spikelet and panicle morphology. The races are, for the most part, easily identifiable by spikelet morphology alone. Intermediate races involving *guinea*, for example, have glumes that open partially and seeds that twist noticeably, but not as much as in pure *guinea*. The collection from Uganda (15), Sudan (14), Ethiopia (14), Tanzania (13), Kenya (12), and Malawi (11) represented more than 10 races. Among the races, *caudatum* (31%) and *durra* (21.8%) were predominant in the total collection from East Africa. Race *guinea* in Malawi (64%) and Tanzania (46.4%); race *caudatum* in Burundi (78.7%), Kenya (76.7%), Rwanda (75.9%), Sudan (32%) and Uganda (69.9%) and race *durra* in Eritrea (48.7%), Ethiopia (42.9%) and Somalia (95.7%), were found in high frequencies. Races

guinea, *kafir*, *guinea-bicolor*, *caudatum-bicolor*, *kafir-bicolor*, *kafir-caudatum*, *kafir-durra* and *guinea-kafir* represented <20 per cent of the total collection that was characterized for racial classification (38,343 accessions).

Only two accessions were of *guinea-kafir* in the collection. The wild relatives collection from East Africa consists of *S. bicolor*, *S. halepense*, *S. lanceolatum*, *S. macrochaeta*, *S. purpureosericeum* and *S. versicolor* (Table 2).

Gaps in the collection

Geographical gaps

Probably due to large variation for maturity, timing of collecting mission, and accessibility to the area under sorghum cultivation, both North and South Sudan were found as the major gaps with seven and 50 districts, respectively. Rwanda is another important country showing 35 districts as gaps. However, the size of the gaps depends largely on size of the districts. Among the provinces, 12 districts each in

Upper Nile province in South Sudan and Gitarama province in Rwanda were found as major gaps. Area of sorghum, including all patterns of its cultivation during the year 2014 in East African countries and the number of geographical sites of past collections, indicated one collection site per 849 ha sorghum area in Burundi, 50,094 ha in Eritrea, 10,729 ha in Ethiopia, 1,369 ha in Kenya, 1,860 ha in Malawi, 5,238 ha in Rwanda, 5,301 ha in Somalia, 62,519 ha in Sudan, 3,376 ha in Tanzania and 3,330 ha in Uganda (FAOSTAT, 2014). In the entire collection from East African countries, an area of 13,435 ha represented one geographic site. Results reveal large geographical gaps in the sorghum germplasm collections from East African countries, particularly in Eritrea, Ethiopia and Sudan. Probability maps developed using FloraMap software indicated a total of 153 districts located in 50 provinces of 10 East African countries as the geographical gaps (Fig. 1, Table 3).

Taxonomic gaps

The genus sorghum comprising about 22 species provides a rich source of diversity for sorghum improvement (Harlon and De Wet, 1972; de Wet, 1978). The wild relatives collection from East Africa at the ICRISAT genebank consists of only six predominant species, namely, *S. bicolor*, *S. halepense*, *S. lanceolatum*, *S. macrochaeta*, *S. purpureosericeum* and *S. versicolor* (Table 2). *S. bicolor* subspecies *drummondii* (99 accessions) and subspecies *arundinaceum* (70 accessions) were predominant in the collection. Accessions of all other species were found in negligible number (<5 accessions). Among the countries, Sudan was found as an important source with 57 accessions of wild relatives' followed by Ethiopia (40 accessions) and Uganda (31 accessions). *S. bicolor* subsp. *drummondii* was found in eight countries, *S. bicolor* subspecies *arundinaceum* race *verticilliflorum* in seven countries, and race *arundinaceum* in five countries. All other species were found in less than three countries. Though Ethiopia being the origin place and considered as the rich source for sorghum diversity, it is represented by only two species – *S. bicolor* ssp. *arundinaceum* race *verticilliflorum* and *S. halepense* (Table 2). Therefore, the wild relatives' collection of sorghum from East Africa indicates major taxonomic gaps.

Diversity in the collection

Qualitative traits

Qualitative traits are very useful in distinguishing accessions. A high proportion (97.5%) of the accessions in the collection had pigmentation (Table 4). Tan plant color, which is said to be associated with resistance to leaf diseases and grain weathering (Frederiksen and Duncan, 1982; Duncan et al., 1991), was found in 2.4 per cent of the collection. Sudan was found as an important source for tan plant color accessions (5.6%). All accessions from Burundi, Rwanda and Somalia produced pigmented plants. All accessions from Somalia and 90% accessions from all other countries produced basal tillers. Four midrib colors (brown, dull green, white and yellow) were found in the collection of which white midrib color was predominant (79.9% accessions), followed by dull green (14.9%) and yellow midrib (5.1%). One accession each from Ethiopia (IS 11861) and Malawi (IS 21549) produced brown midrib leaves, the trait associated with low lignin content and increased digestibility for livestock (Oliver et al., 2005; Gressell, 2008). More than 60% accessions from all countries produced leaves with white midribs. Ten panicle

types (panicle compactness and shape) with predominance of semi-compact elliptic type (33.3%) followed by semi-loose stiff branches (32.9%) were observed in the collection. High frequency of compact elliptic (30%) and compact oval (64%) panicles were observed in the collection from Somalia. Seventy-nine per cent of accessions from Malawi produced panicles with semi-loose stiff branches. Twelve glume colors were observed in the collection with predominance of black glume accessions (29%). One accession each from Kenya (IS 21336) and Malawi (IS 21631) produced white glumes. Fifty-one per cent of total accessions produced grains covered one-fourth with glumes. High frequency of accessions from Tanzania (17.1%) produced uncovered grains. Twelve seed colors were found in the collection from East Africa and the collection from Ethiopia represented all seed colors. The collection from Malawi (55%) and Tanzania (26.6%) were found to be important sources for white seed sorghum, which is preferred by farmers. The proportion of lustrous (49.3%) and non-lustrous (50.3%) accessions was almost the same in the collection. Maximum accessions (95%) from Kenya produced non-lustrous seeds. Absence of seed sub-coat was predominant in the collection (69%) with maximum frequency of accessions from Malawi (66.1%). Only 1.1% accessions produced completely corneous seed and 27.8% produced completely starchy seeds. All accessions from Rwanda were freely threshable and produced completely starchy seeds. More than 50% of accessions from all other East African countries were freely threshable.

Quantitative traits

Range

Basic races varied widely than intermediate races for days to 50% flowering in postrainy season, plant height in rainy season, basal tillers per plant, panicle length, seed width and 100 seed weight (Table 5). Race *bicolor* varied widely for panicle exertion and panicle length; *caudatum* for days to 50% flowering in postrainy season, plant height in rainy season and basal tillers per plant; *durra* for days to 50% flowering in rainy season seed width and seed weight; *durra-bicolor* for plant height in postrainy season and panicle exertion and width and *durra-caudatum* for days to 50% flowering in rainy season. Variation was relatively less in accessions belonging to race *kafir*, *kafir-bicolor*, *kafir-caudatum* and *kafir-durra* than in other races. Among all races, *bicolor*, *caudatum*, *durra*, *durra-bicolor* and *durra-caudatum* were found to be important sources for one or more traits under study. *Caudatum* for early flowering and short height in both seasons and more basal tillers per plant; *durra* for larger seeds; *durra-bicolor* for tall height and high panicle exertion, were recognized as promising sources for high variability.

The collection from Ethiopia varied widely for plant height in both rainy and postrainy seasons, panicle exertion, panicle width and 100 seed weight. Days to 50% flowering in rainy season and seed width in collections from Sudan; days to 50% flowering in postrainy season in the Ugandan collection; and panicle length and seed width in Tanzanian collection, varied widely. The landraces from Ethiopia for early flowering and short plant height in both seasons, high panicle exertion, panicle width and 100 seed weight; Kenya for high basal tiller number; Sudan for early flowering in postrainy season, tall height in rainy season and larger seeds and Tanzania for long panicles, were found as promising source countries for variability.

Means

The Newman-Keuls test of significance for mean values indicated significant differences between basic and intermediate races for days to 50% flowering in post-rainy season, plant height in rainy season, basal tillers per plant, panicle exertion, length and width, seed width and 100 seed weight (Table 6). Germplasm of basic races flowered late in the post-rainy season, grew tall in the rainy season and produced more basal tillers, small panicles and larger seeds than those of intermediate races. The Newman-Keuls test indicated significant differences among races (Table 6). Accessions of *bicolor* flowered late, produced long and stout panicles with small seeds. *Kafir* flowered significantly earlier than other races in rainy season and grew short in both seasons. *Durra* and *guinea-durra* produced significantly larger seeds than all other races. *Durra-bicolor* grew tall in both seasons and produced stout panicles. *Kafir-caudatum* flowered earlier than all other races and produced larger seeds. Accessions from Eritrea flowered early and grew short in the post-rainy season and differed significantly from those of other countries. The collection from Ethiopia, which is considered as the primary center for diversity produced highly exerted stout panicles than those from other countries. Sorghum germplasm from Malawi produced more basal tillers, highly exerted long and stout panicles, and differed significantly from that of other countries. Accessions from Somalia flowered early in the rainy season and those from Sudan in the post-rainy season. Accessions from Tanzania flowered late and grew tall in both seasons and differed from those of other countries.

Variations

Levene's test of significance for variances revealed highly significant variances for all traits under study, except basal tillers per plant, revealing high heterogeneity in accessions under study (Table 6).

Phenotypic diversity

Principal component analysis (PCA) was carried out using standardized data of 10 quantitative traits. The first 3 PCs captured 64.1%, 65.7% and 62.1% of total variation in the entire collection, basic and intermediate races, respectively data not shown. PC1 was the important one and explained 29.7% of variation in the entire collection, 29.4% in basic races (Fig. 2a), and 31.9% of variation in intermediate races (Fig. 2b) and differentiated accessions by high positive factor loadings by days to 50% flowering, plant height, panicle length and panicle width. Seed width and seed weight showed high negative factor loadings in all groups. PC2 explained 20.1% of variation in the entire collection, 21.4% in basic races, and 16.2% in intermediate races and differentiated accessions by high positive loadings by all traits except days to 50% flowering. PC3 explained 14.4% of variation in entire collection, 14.9% in basic races, and 14.0% in intermediate races and differentiated accessions mainly due to high positive loadings by basal tillers per plant, panicle exertion and panicle length. Days to 50% flowering and plant height showed high negative loadings.

The Shannon-Weaver diversity index (H') was calculated for all traits of races and countries to compare phenotypic diversity for 10 quantitative traits (Table 7). The diversity index values (H') were variable among traits. Mean diversity index (H') ranged from 0.429 ± 0.095 for basal tillers per plant to 0.581 ± 0.057 for 100 seed weight. Mean diversity

over all traits varied from 0.385 ± 0.058 in *kafir* to 0.589 ± 0.048 in *caudatum-bicolor*. Landraces of *durra-bicolor* were highly diverse for flowering and plant height in post-rainy season, panicle length and width, seed width and weight. The collection – from Eritrea for panicle exertion; Ethiopia for flowering in post-rainy season and panicle length; Kenya for days to 50% flowering in rainy season and seed weight; Malawi for seed width; Somalia for plant height in rainy season; Sudan for basal tillers per plant; Uganda for panicle width; and Tanzania for plant height in post-rainy season – was highly diverse.

Discussion

The adequacy of the collections to provide useful genes for current and future crop improvement programs will depend on the collecting efforts in obtaining good representation of the existing genetic diversity within crop species and on the quality of preservation procedures and facilities. Though the global germplasm collections are large, there are still sizeable gaps in the *ex situ* collections of many major crops including sorghum (Jones et al., 1997; Maxted et al., 2008; Upadhyaya et al., 2009, 2010, 2012, 2014a, 2015). Generally, the goal of genebanks is to preserve maximum genetic diversity and that is feasible through identification and exploration of gaps in the collections. But the lack of geo-referenced data for collections and its inaccuracy are the major constraints for spatial analysis and identification of gaps. Therefore, updating passport data for location information and geo-reference data and their validation are essential for the identification of gaps in existing collections using spatial analysis.

Sorghum originated in the north-eastern quadrant of Africa, where the greatest variability in wild and cultivated species is found. It was probably domesticated in Ethiopia between 5,000 and 7,000 years ago. From the center of origin, it was distributed along trade and shipping routes throughout Africa, and through the Middle East to India at least 3,000 years ago (de Wet and Harlan, 1971; Dogget, 1988). Fuller (2002) reported Ethiopia as the potential region for sorghum diversity. Germplasm collection reports at ICRISAT revealed the distribution of sorghum in Bubanza, Bururi, Buzumbura and Kirundo provinces in Burundi (Prasada Rao and Mengesha, 1982a); Kigali, Gitarama and Butare in Rwanda (Prasada Rao and Mengesha, 1982b); Gambella region in Ethiopia (Prasada Rao and Mengesha, 1981); Gezira, Kasala and Blue Nile provinces in Eastern Sudan (Prasada Rao and Mengesha, 1980) and Singida, Mbeya provinces in Tanzania (Prasada Rao and Mengesha, 1979). Appa Rao (1979) reported the cultivation of sorghum in maize fields and backyards in Malawi for chewing purpose. Ramanatha Rao (1979) reported that the occurrence of *durra* sorghum is common in Somalia.

The size of collections from East African countries at the ICRISAT genebank is large. More accessions per geographical site was observed in Eritrea (37), Ethiopia (20) and Sudan (11), indicating high intensity of the collection probably because of cultivation of highly diverse sorghum in the primary center of diversity, during the collection season. However, the number of geographical collection sites are very less when compared to the area under sorghum cultivation in entire East Africa (1:13,435 ha), Sudan (1:62,519 ha), Eritrea (1:50,094 ha), and Ethiopia (1:10,729 ha), indicating major geographical gaps.

Being the primary center of diversity for sorghum, large variation was observed in the collection from East African

countries. Highly significant variances for all traits under study, except basal tillers per plant, indicated highly significant differences among the accessions under study. Upadhyaya et al. (2014b) reported large variation for most of the morphoagronomic traits in the world collection of sorghum at ICRISAT. In the present study, basic races varied widely compared to intermediate races for most of the traits, and variation was relatively less in landraces belonging to race *kafir*, *kafir-bicolor*, *kafir-caudatum* and *kafir-durra* than those of other races. To add the trait-specific germplasm to the existing collection from East Africa, gaps in Ethiopia for plant height, panicle exertion, panicle width and 100 seed weight; Sudan for days to 50% flowering and seed size; Uganda for days to 50% flowering and Tanzania for panicle length and seed size, need to be explored. Races, *guinea*, *kafir*, *guinea-bicolor*, *caudatum-bicolor*, *kafir-bicolor*, *kafir-caudatum*, *kafir-durra* and *guinea-kafir*, which represented <20% in the total collection need to be collected from East African countries for good representation of races and countries and for increased diversity in the collection. Kamala et al. (2014) reported the distribution of *guinea* in Eastern and Western Africa and *kafir* in South-Eastern Africa.

The sorghum collection from East African countries serves as an important source for beneficial genotypes. Forty sorghum germplasm accessions that are conserved at the ICRISAT genebank were released directly as 42 cultivars in 18 countries, and thirteen of them are from five East African countries. IS 8193 that originated in Uganda has been released as IS 8193 in Rwanda and as KARI Mtama 2 in Kenya. IS 18758, a popular *guinea-caudatum* landrace that originated in Ethiopia has been released as E 35-1 in Burkina Faso and as Gambella 1107 in Burundi. IS 18758 is an excellent source for desirable plant type, high grain yield, good grain quality and resistance to leaf diseases (Reddy et al., 2004; Upadhyaya et al., 2014b). At ICRISAT, several promising sources for different biotic and abiotic stresses and other useful traits were identified in the collection from East African countries. Examples include: IS 18758 from Ethiopia for anthracnose, leaf blight and rust; IS 18551 from Ethiopia for stable resistance to shoot fly; IS 3121 from Kenya for grain mold resistance; IS 2333 from Sudan for grain mold and downy mildew resistance; IS 3547 from Sudan for grain mold, downy mildew, anthracnose and rust resistance; IS 9955 from Sudan for high seed protein content (21%); IS 19604 from Sudan for salinity tolerance; IS 9830 and 9951 from Sudan for *Striga* resistance and IS 2266, IS 3572 and IS 9890 from Sudan for high soluble sugars (>16%) in stalk juice; IS 9108 from Kenya for post-flowering drought tolerance (Reddy et al., 2004; Reddy et al., 2008; Upadhyaya et al., 2017). Prasada Rao and Mengesha (1980) reported Dabar and Feterita landraces from Sudan as the bird tolerant. Prasada Rao and Mengesha (1982b) reported that sorghum cultivated at low temperatures prevailing at higher altitudes in Rwanda could be a good source for cold tolerance. As evidenced in the present study, sorghum germplasm from Sudan for tan plant color, from Somalia for high basal tillering, from Malawi and Tanzania for white seeded sorghum, and from Rwanda for easily threshable sorghums, could be good sources. Therefore, there is a need to explore the gaps in these countries and evaluate the entire sorghum collection from East Africa systematically to identify more beneficial accessions.

A pointed collection mission launched in Gambella region in Ethiopia revealed the cultivation of *Zera-zera* sorghums on banks of the Bari river by the Agnawks community (Prasada Rao and Mengesha, 1981). White Nile riverside was also a

potential area for *Zera-zera* sorghum cultivation. Due to the tremendous importance of *Zera-zeras*, IBPGR, Italy recommended their intensive collection from the Sudano-Ethiopian border (Prasada Rao and Mengesha, 1981). *Zera-zera* sorghums from Ethiopia were found promising for high yield, grain quality, grain mold and charcoal rot resistance. Therefore, these *Zera-zeras* were extensively used in sorghum breeding programs. IS 18484, a converted line of IS 3541 *Zera-zera*, was released as variety CSV 4, by the All India Coordinated Sorghum Improvement Program (AICSSIP), in India (Prasada Rao and Mengesha, 1981).

Sorghum is a highly variable genus with about 22 species (Kamala et al., 2014). Crop wild relatives (CWR) possess higher levels of stress resistance and useful adaptive and nutritional traits. For example, Kamala et al. (2014) reported *S. halepense* as a source of resistance for downy mildew. Mote (1984) reported the *S. purpureosericeum* as a source for higher levels of resistance to the shoot fly. The wild relatives' collection from East African countries at ICRISAT genebank is poorly represented with only 195 accessions belonging to six species (Table 2). Except *S. bicolor* subspecies *drummondii* (99 accessions) and subspecies *arundinaceum* (70 accessions) all other species were found in negligible numbers (<5 accessions). Globally, genetic resources have primarily been conserved with utmost attention given to collecting and maintaining landraces and the collection of wild relatives was not a priority (Frankel and Hawkes, 1975; Smith et al., 2003; Maxted and Kell, 2009). None of the collection missions launched by ICRISAT was exclusively for wild sorghums. Therefore there is a need to launch systematic collection missions in East African countries, exclusively for wild relatives of sorghum to fill taxonomical gaps in the collection, before we lose them forever due to several reasons such as natural calamities, urbanization, etc. (Upadhyaya and Gowda, 2009). Past germplasm collection reports at ICRISAT revealed the occurrence of *S. purpureosericeum* near Imbo plain in Burundi; *S. bicolor* subspecies *drummondii* in Rwanda; *S. halepense* and *S. arundinaceum* in Uganda; *S. purpureosericeum* and *S. verticelliflorum* in shire valley, Phalombe plains, Mzimba and Rumphu in Malawi; and *S. versicolor*, *S. purpureosericeum* and *S. arundinaceum* in Morogoro and Iringa provinces in Tanzania (Appa Rao, 1979; Prasada Rao and Mengesha, 1979, 1980, 1982a; Gopal Reddy et al., 1993). Remanandan and Mengesha (1981) reported the occurrence of wild sorghum on field bunds and cultivated sorghum fields near Mbeya and also suggested exploring the Zanzibar islands in Tanzania on a priority basis. Prasada Rao and Mengesha (1979) reported the occurrence of wild sorghums in high rainfall regions of Southern Tanzania, mainly in Mtwara region.

The collection under study is from a wide latitudinal distribution of 17.03°S to 19.33°N. Therefore, exploration of gaps identified in this study could be very useful for collecting highly diverse germplasm adaptable to diverse climates (Damon, 1962). The gaps may be prioritized 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 the knowledge of sorghum cultivation in the districts. Past collection reports and catalogues may be considered while preparing route maps for exploring the identified gaps. It is also important to collect all collection-related information, including geo-reference data, while collecting germplasm samples in order to facilitate future mapping efforts. The collection under study can be

accessed following the Standard Material Transfer Agreement (SMTA) of International Treaty on Plant Genetic Resources for Food and Agriculture (ITPGRFA) and passport and characterization data can be accessed at www.ICRISAT.org.

Materials and Methods

Gap analysis

The ICRISAT genebank conserves 39,923 sorghum germplasm accessions from 93 countries. Of this 12,750 were assembled/collected from 10 East African countries namely Burundi, Eritrea, Ethiopia, Kenya, Malawi, Rwanda, Somalia, Sudan, Tanzania and Uganda. To identify geographical and taxonomic gaps, we used 7,914 landraces having georeferenced data and representing 10 East African countries. FloraMap, a window-based GIS software developed at the International Center for Tropical Agriculture (CIAT) in Cali, Columbia, was used to predict the occurrence of sorghum in individual countries of East Africa (Jones and Gladkov, 1999). The FloraMap system is based on calculating the probability that a climate record belongs to a multivariate normal distribution described by the climates at the collection points of a calibration set of organisms. With its user-friendly software linked to agroclimatic and other databases, biodiversity specialists can create maps showing the most likely distribution of any particular species in nature. The basic input in the FloraMap software is the coordinates (Latitude and Longitude) of the collection site with a unique identifier. While working on the dataset, equal weights were allocated to the three climatic variables (rainfall, temperature and diurnal temperature), and an exponential transformation with a power of 0.3 was applied to the monthly rainfall data. Maps showing the districts with high probability (more than 50%) for sorghum occurrence with few and without, collection sites were considered as geographical gaps in the collection.

Diversity assessment

To assess the diversity in the collection from East African countries, 7,863 accessions having characterization data retrieved from ICRISAT's sorghum characterization database were used. Accessions were characterized in rainy (June–November) and postrainy (October to March) seasons during 1977–2013, in vertisols, in batches of 500–1000 at ICRISAT, Patancheru, India (17° 25'N latitude, 78° 00'E longitude and 545 m.a.s.l.). Crop was raised in an augmented block design with three control accessions (IS 2205, IS 18758, IS 33844) in each block, repeated after every block of 20 accessions. Each accession was sown in one 4-m long row with spacing of 75 cm between rows, and the crop was thinned after two weeks leaving plant-to-plant distance of about 10 cm. Fertilizers were applied at the rate of 80 kg ha⁻¹ N and 40 kg ha⁻¹ P₂O₅ in both the seasons. Appropriate crop production/protection measures were taken to raise a healthy crop.

Observations on days to 50% flowering and plant height (cm) were recorded during both rainy and postrainy seasons, while remaining traits such as number of basal tillers per plant, panicle exertion, panicle length and width, seed width and 100 seed weight were recorded only during postrainy seasons. Observations were also recorded on 11 qualitative traits (plant pigmentation, nodal tillers, midrib color, panicle compactness and shape, glume color, glume covering, seed

color, seed lustre, seed sub-coat, endosperm texture and threshability) following descriptors for sorghum (IBPGR and ICRISAT, 1993). Days to 50% flowering were recorded on a plot basis while the remaining traits (except seed traits) were recorded on five randomly selected representative plants and averaged. Average width of 10 seeds at the widest point was recorded for seed width, while 100-seed weight was weighed from bulk sample of each accession.

Data analysis

The range and means were calculated for all traits, for basic and intermediate races and for each country of origin. The means for different traits were compared using the Newman-Keuls procedure (Newman, 1939; Keuls, 1952) and variances were tested using Levene's test (Levene, 1960). Principal Component Analysis (PCA) was performed for 10 quantitative traits and Shannon-Weaver diversity index (H') (Shannon and Weaver, 1949) was used to measure and compare phenotypic diversity for each trait. The diversity index (H') was estimated for both quantitative and qualitative traits, for each race, and for each country of origin using GENSTAT 13.1. (VSN International, 2010). Low value of H' indicates lack of genetic diversity in the collection.

Conclusion

Sorghum is an important cereal crop grown in both temperate and tropical regions in Africa, Asia, Europe and the Americas for food, feed, fodder and bioenergy purposes. The genebank at ICRISAT, India, conserves 39,923 accessions of sorghum germplasm from 93 countries, including 12,750 from 10 East African countries. The collection from East Africa is large, but with several gaps, showed wide variation for important traits. Races showed more variation than intermediate races for days to 50% flowering in postrainy season, plant height in rainy season, basal tillers per plant, panicle length, seed width and 100 seed weight. Race *kafir* and intermediate races, such as *kafir-bicolor*, *kafir-caudatum* and *kafir-durra*, were less diverse as compared to other races for the traits studied. Races and intermediate races differed significantly for almost all the quantitative traits studied. A total of 153 districts located in 50 provinces of 10 East African countries were found to be the geographical gaps. Both North and South Sudan were found as major gaps with seven and 50 districts, respectively. Rwanda is another important country showing 33 districts as gaps. Wild species such *S. halepense*, *S. lanceolatum*, *S. macrochaeta*, *S. purpureosericeum* and *S. versicolor* are present in the East African sorghum collection conserved in the genebank, while other species of sorghum were considered as taxonomic gaps. The gaps identified in the present study need to be explored on a priority basis to collect highly diverse sorghum germplasm.

Acknowledgements

The authors sincerely acknowledge the contribution of all former and present staff of the genebank at ICRISAT, Patancheru, India, in collection, assembly, conservation and characterization of sorghum germplasm from East African countries.

References

Appa Rao S (1979) Germplasm collecting mission to Malawi. Genetic Resources Progress Report 2.

- International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru, India.
- Damon EG (1962) The cultivated sorghums of Ethiopia. Ethiopia College of Agriculture Mech Arts Expt Station Bull 6.
- De Wet JMJ (1978) Systematics and evolution of sorghum sect. Sorghum (Gramineae) Am J Bot. 65(4):477-484.
- De Wet JMJ and Harlan JR (1971) The origin and domestication of *sorghum bicolor*. Econ Bot. 25:128-135.
- Doggett H (1988) Sorghum, 2nd ed. Longman, Harlow, Essex, UK.
- Duncan RR, Bramel-Cox PJ, Miller FR (1991) Contribution of introduced sorghum germplasm to hybrid development in the USA. In: Shands HL, Weisner LE (eds.), Use of plant introduction in cultivar development, Part-1. CSSA Special Publication No. 17. Crop Science Society of America, Madison, WI, USA, pp. 69-102.
- Dwivedi SL, Ceccarelli S, Blair MW, Upadhyaya HD, Ashok KA, Ortiz R (2016) Landrace germplasm for improving yield and abiotic stress adaptation. Trends Plant Sci. 21:31-42.
- Dwivedi S, Sahrawat K, Upadhyaya HD, Ortiz R (2013) Food, nutrition and agrobiodiversity under global climate change. In: Sparks DL (eds), Advances in Agronomy. Academic Press, Elsevier Inc., USA, 1-128.
- Espinoza L, Kelley J (2002) Grain sorghum production handbook. In: Espinoza L, Kelley J (eds), Grain Cooperative Extension Service. University of Arkansas, Arkansas, USA, pp. 1-74.
- FAO (Food and Agriculture Organization of the United Nations) (2014) <http://www.fao.org>
- Frankel OH, Hawkes JG (1975) Crop genetic resources for today and tomorrow. International Biological Programme 2. Cambridge University Press, Cambridge, UK.
- Frederiksen RA, Duncan RR (1982) Sorghum disease in North America. In: de Milliano WAJ, Frederiksen RA, Bengston GD (eds), Sorghum and millets diseases: A second world review. ICRISAT, Patancheru, AP, India, pp. 85-88.
- Fuller D (2002) Fifty years of archaeobotanical studies in India. Laying a solid foundation. In: Settar S, Korisetar R (ed). Indian archaeology in retrospect. Vol III. Archaeology and interactive disciplines. Delhi, India, pp. 247-363.
- Gopal Reddy V, Remanandan P, Sabiiti EN, Mengesha MH (1993) Sorghum, finger millet and pigeonpea germplasm collection in Uganda. Genetic Resources Progress Report 77. International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru, India.
- Gressell J (2008) Transgenics are imperative for biofuel crops. Plant Sci 174(3):246-263.
- Harlan JR, de Wet JMJ (1972) A simplified classification of cultivated sorghum. Crop Sci. 12:172.
- Hammer K, Knupffer H, Xhuveli L, Perrino P (1996) Estimating genetic erosion in landraces-two case studies. Genet Resour Crop Ev. 43:329-336.
- IBPGR, ICRISAT (1993) Descriptors for sorghum [*Sorghum bicolor* (L.) Moench]. International Board for Plant Genetic Resources, Rome, Italy. International Crops Research Institute for the Semi-Arid Tropics, Patancheru, India.
- Jones PB, Beebe SE, Tohme J, Galwey NW (1997) The use of geographical information systems in biodiversity exploration and conservation. Biodivers Conserv. 6:947-958.
- Jones PG, Gladkov A (1999) In: Jones Annie L and CD-ROM series (eds), FloraMap: A computer tool for predicting the distribution of plants and other organisms in the wild. Version 1. Cali, Colombia: Centro Internacional de Agricultura Tropical (CIAT).
- Kamala V, Mourya M, Dwivedi SL, Upadhyaya HD (2014) Wild sorghums - Their potential use in crop improvement. In: Yi-Hong Wang, Upadhyaya HD, Chittaranjan Kole (eds.), Genetics, genomics and breeding of sorghum. CRC Press, Taylor & Francis Group, Boca Raton, London, New York, pp 56-89.
- Keuls M (1952) The use of the "Studentized range" in connection with an analysis of variance. Euphytica 1:112-122.
- Levene H (1960) Robust tests for equality of variances. In: Olkin I. (eds.), Contributions to probability and statistics: Essays in honour of Harold Hotelling. Stanford University Press, Stanford, pp. 278-292.
- Mace ES, Tai S, Gilding EK, Li Y, Prentis PJ, Bian L, Campbell BC, Hu W, Innes DJ, Han X, Cruickshank A, Dai C, Frère C, Zhang H, Hunt CH, Wang X, Shatte T, Wang M, Su Z, Li J, Lin X, Godwin ID, Jordan DR, Wang J (2013) Whole-genome sequencing reveals untapped genetic potential in Africa's indigenous cereal crop sorghum. Nat Commun. 4: 2320.
- Maxted N, Dulloo E, Ford-Lloyd BV, Iriodo JM, Jarvis A (2008) Gap analysis: a tool for complementary genetic conservation assessment. Divers Distrib. 14(6):1018-1030.
- Maxted N, Kell SP (2009) Establishment of a global network for the in situ conservation of crop wild relatives: Status and needs. FAO Commission on Genetic Resources for Food and Agriculture, Rome, Italy, p 266.
- Mickelbart MV, Hasegawa PM, Bailey-Serres J (2015) Genetic mechanisms of abiotic stress tolerance that translate to crop yield stability. Nat Rev Genet 16(4):237-251.
- Mote UN (1984) Sorghum species resistant to shoot fly. Indian J Entomol. 46:241-243.
- NASA (National Aeronautics and Space Administration) (2014) <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.

- Oliver AL, Pedersen JF, Grant RJ, Klopfenstein TJ (2005) Comparative effects of the sorghum *BMR-6* and *BMT-12* genes: I Forage sorghum yield and quality. *Crop Sci.* 45:2234-2239.
- Prasada Rao KE, Mengesha MH (1979) Sorghum and millets germplasm collection in Tanzania. Genetic Resources Progress Report 9. International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru, India.
- Prasada Rao KE, Mengesha MH (1980) Sorghum and millets germplasm collection in Eastern Sudan. Genetic Resources Progress Report 16. International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru, India.
- Prasada Rao KE, Mengesha MH (1981) A pointed collection of Zera-Zera sorghum in the Gambella area of Ethiopia. Genetic Resources Progress Report 33. International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru, India.
- Prasada Rao KE, Mengesha MH (1982a) Sorghum germplasm collection in Burundi. Genetic Resources Progress Report 51. International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru, India.
- Prasada Rao KE, Mengesha MH (1982b) Sorghum germplasm collection in Rwanda. Genetic Resources Progress Report 46. International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru, India.
- Ramanatha Rao V (1979) Germplasm collection mission in Somalia. Genetic Resources Progress Report 12. International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru, India.
- Reddy BVS, Kumar AA, Reddy PS, Elangovan M (2008) Sorghum germplasm: diversity and utilization. In: Bantilan MCS, Deb UK, Gowda CLL, Reddy BVS, Obilana AB, Evenson RE (eds), Sorghum genetic enhancement: Research process, dissemination and impacts. International Crops Research Institute for the Semi-Arid Tropics, Patancheru, AP, India, pp. 153-169.
- Reddy BVS, Rao P, Deb UK, Stenhouse JW, Ramaiah B, Ortiz R (2004) Global sorghum genetic enhancement processes at ICRISAT. In: Bantilan MCS, Deb UK, Gowda CLL, Reddy BVS, Obilana AB, Evenson RE (eds), Sorghum genetic enhancement: Research process, dissemination and impacts. International Crops Research Institute for the Semi-Arid Tropics, Patancheru, AP, India, pp. 65-102.
- Remanandan P, Mengesha MH (1981) Pigeonpea germplasm collection mission in Tanzania. Genetic Resources Progress Report 38. International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru, India.
- Shannon CE, Weaver W (1949) The mathematical theory of communication. Univ. Illinois Press, Urbana.
- Singh AK, Rusoke DG, Latigo MWO, Mukuru SZ, Appa Rao S (1991) Legumes germplasm collection in Uganda (groundnut, pigeonpea and cowpea). Genetic Resources Progress Report 66. International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru, India.
- Smith RD, Dickie JB, Linington SH, Pritchard HW, Probert RJ (2003) Seed conservation: Turning science into practice. Royal Botanic Gardens, Kew, UK.
- Srinivasa Rao P, Reddy BVS, Nagaraj N, Upadhyaya HD (2014) Sorghum production for diversified uses. In: Yi-Hong Wang, Upadhyaya HD, Chittaranjan K (eds), Genetics, genomics and breeding of sorghum. CRC Press, Taylor & Francis Group, Boca Raton, London, New York.
- Upadhyaya HD, Dwivedi SL, Vetriventhan M, Krishnamurthy L, Singh SK (2017) Post-flowering drought tolerance using managed stress trials, adjusted to flowering, and mini core collection in sorghum. *Crop Sci.* 12 (February):1-12. DOI: 10.2135/cropsci2016.04.0280
- Upadhyaya HD, Gowda CLL (2009) Managing and enhancing the use of germplasm - strategies and methodologies. Technical Manual No.10. ICRISAT, Patancheru, Andhra Pradesh, India, p. 236.
- Upadhyaya HD, Reddy KN, Irshad Ahmed M, Gowda CLL (2010) Identification of gaps in pearl millet germplasm from Asia conserved at the ICRISAT genebank. *Plant Genet Resour.* 8(3):267-276.
- Upadhyaya HD, Reddy KN, Irshad Ahmed M, Gowda CLL (2012) Identification of gaps in pearl millet germplasm from East and Southern Africa conserved at the ICRISAT genebank. *Plant Genet Resour.* 10(3):202-213.
- Upadhyaya HD, Reddy KN, Irshad Ahmed M, Gowda CLL, Haussmann Bettina (2009) Identification of geographical gaps in the pearl millet germplasm conserved at ICRISAT genebank from West and Central Africa. *Plant Genet Resour.* 8(1):45-51.
- Upadhyaya HD, Reddy KN, Irshad Ahmed M, Gowda CLL, Thimma Reddy M, Senthil Ramachandran (2015) Identification of gaps in pigeonpea germplasm from East and Southern Africa conserved at the ICRISAT genebank. *Indian J of Plant Genet Resour.* 28(2):180-188.
- Upadhyaya HD, Reddy KN, Pundir RPS, Sube Singh, Gowda CLL, Irshad Ahmed M (2013) Diversity and geographical gaps in *Cajanus scarabaeoides* (L.) Thou. germplasm conserved at the ICRISAT genebank. *Plant Genet Resour.* 11(1):3-14. doi:10.1017/S1479262111000736.
- Upadhyaya HD, Reddy KN, Sube Singh, Gowda CLL, Irshad Ahmed M, Vinod Kumar (2014a) Diversity and gaps in *Pennisetum glaucum* subsp. *monodii* (Maire) Br. germplasm conserved at the ICRISAT genebank. *Plant Genet Resour.* 12(2):226-235.
- Upadhyaya HD, Sharma S, Dwivedi SL, Singh SK (2014b) Sorghum genetic resources: Conservation and diversity assessment for enhanced utilization in sorghum improvement. In: Yi-Hong Wang, Upadhyaya HD, Chittaranjan K (eds), Genetics,

genomics and breeding of sorghum. CRC Press, Taylor & Francis Group, Boca Raton, London, New York, pp 28-55.

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

VSN International (2010) GenStat software for windows. Release 13.1. VSN International Ltd., Hemel Hempstead, UK.