Methods of Constructing Core Set Using Agro-morphological Traits in Foxtail Millet [*Setaria italica* (L.) Beauv]

Jayarame Gowda*, D Rekha, S Bharathi and M Krishnappa
Project Coordination Cell (Small Millets), University of Agricultural Sciences, GKVK, Bangalore-56006, Karnataka, India

The two different methods were employed to construct core sets of foxtail millet germplasm using data on geographical distribution and agro-morphological traits and the variation generated for different traits in core sets developed by two methods was compared with the entire collection. First method was grouping the accessions under distinct morphological traits followed by clustering and Principal Component Analysis which resulted in C_1 having 156 accessions; second was by using a software called Power Core which yielded C_2 consisting of 78 accessions. In both the core sets formed, maximum accessions were from Asian continent. Newman-Keuls test for means inferred that both the core sets were true representatives of entire collection. Similarly, Levene's test for homogeneity of variances revealed non-significant differences between entire set, C_1 and C_2. The range for the traits studied in both C_1 and C_2 were almost similar to the range of the traits in the entire collection. Likewise, the chi-square test for frequency distribution analysis for different morphological traits indicated that the variation available in the entire collection was preserved in both the core sets. The Shannon-Weaver diversity indices of 11 quantitative and 12 qualitative traits of entire collection, C_1 and C_2 indicated the presence of diversity of entire collection in the core sets.

Key Words: Core set, Diversity index, PowerCore, Foxtail millet

Introduction

Foxtail millet [*Setaria italica* (L.) Beauv] is one of the oldest crops cultivated for food grain and fodder. It is an indispensable crop of rainfed areas in semi-arid regions of India. It can withstand severe moisture stress and can adjust to wide range of soil conditions. China is considered to be the native home of foxtail millet. The nearest relative of *Setaria italica* is *S. viridis* and both have same genomic constitution. It is presumed that the cultivated species was most probably derived from *S. viridis* in china and later spread to Africa, Europe and Asia in pre-historic times.

The foxtail millet collection exceeding 1,400 accessions assembled and maintained at Project Co-ordinating Unit (Small Millets), All India Coordinated Small Millets Improvement Project (AICSMIP), Bangalore represents good diversity from various regions within and outside the country. The size of the large germplasm collection is an obstacle to their evaluation and exploitation. Their utilization in breeding programme could be increased if more information is available on the amount and kind of variation present in these collections. However, in most cases, the resources needed to characterize these accessions phenotypically and genotypically are meagre and unavailable.

The management and use of large germplasm collection could be enhanced if a limited number of genetically diverse accessions within the collection are selected as core collection (Frankel, 1984) and given priority to evaluation and hybridization (Brown, 1989). Core collection which is a subset of entire collection should include the maximum genetic variation contained in the entire collection with a minimum repetitiveness. There are different methods proposed for constructing core set and also for evaluating representatives of core set. But how many of these methods meet the goals of core set depends on the species, the composition of the collection and the type of characters of interest.

In the present study, two different methods of core formation *i.e.*, one by grouping accessions under distinct morphological traits followed by clustering and Principal Component Analysis and another by using PowerCore software were followed to develop core sets using entire collection of foxtail millet. The amount and the kind of variation generated for quantitative and qualitative traits in core sets developed by two methods were compared with the entire collection.

Materials and Methods

a) Evaluation of Entire Germplasm

A total of 1,478 accessions of foxtail millet maintained at National Active Germplasm Site (NAGS), Project Co-ordinating Unit (Small Millets), Bangalore were evaluated at the main research station, Gandhi Krishi Vigyana Kendra, Bangalore, over years from 2002 to 2006. The

*Author for Correspondence: E-mail: jg_gene@rediffmail.com

accessions were grown in a single row of 3 m length with a spacing of 22.5 cm between rows and 10 cm between plants within a row. Data on 25 descriptors were recorded following the procedures given in descriptors of foxtail millet (Anon., 1985).

Observations were recorded on five competitive plants in each accession and averaged for eleven quantitative traits, namely, days to 50 per cent flowering, plant height (cm), number of basal tillers, flag leaf length (cm), flag leaf width (cm), peduncle length (cm), ear length (cm), panicle exertion, days to maturity, grain yield per plant (g) and 1000-grain weight (g) and for twelve qualitative traits *i.e.* plant pigmentation at flowering, blade pubescence, sheath pubescence, senescence, inflorescence shape, inflorescence compactness, grain color, grain shape and apical sterility in panicle.

b) Formation and Evaluation of Core Set

Method 1: Data on 11 quantitative and 12 qualitative traits were used to form Core Set 1 based on scores of Principal Component Analysis and is designated as C_1. In this method, initial grouping of accessions was done using distinct morphological traits, *viz.,* plant pigmentation, grain colour, grain shape and inflorescence shape. The groups thus formed were further subjected to clustering analysis using SYSTAT 9 package. Clusters having large number of accessions were further subjected to Principal Component Analysis (PCA) using package SPLUS 2000. Using PCA scores, 156 accessions (10%) were selected for inclusion in C_1.

Method 2: A software called PowerCore (v.1.0) developed by Genetic Resources Division, Rural Development Administration, Republic of Korea, was used to form Core Set 2 of 78 accessions using both quantitative and qualitative traits of entire collection and designated as C_2. This new method is used for the establishment of core and allele mining sets by the Advanced M (Maximization) Strategy implemented through a Modified Heuristic Algorithm. It minimizes the loss of useful alleles and effectively selects accessions with highest diversity reducing the repeated alleles.

c) Comparison of Entire and Core Sets

The accessions in the entire collection, C_1 and C_2 were classified according to their place of origin and their percentages were calculated and compared. The quantitative data were subjected to statistical analysis to estimate mean, range, coefficient of variability and variance and the same were compared to determine whether the core sets formed represents the entire collection for the variability present.

The mean data of entire collection and two sets of core were compared using Newman-Keuls procedure (Newman, 1939; Keuls 1952). Levene's test was used to test the homogeneity of variances of the entire collection and two sets of core (Levene, 1960).

The ratio of different sub descriptors of qualitative trait of entire collection, C_1 and C_2 were compared using chi-square test (χ^2). The Shannon- Weaver diversity index (H') (Shannon and Weaver, 1949) of the entire collection and core sets which gives the measures of diversity was estimated for all the traits studied.

Results and Discussion

The accessions of entire collection of foxtail millet showed considerable variability for all the traits studied. The PCA scores of entire collection yielded 156 accessions which were included in C_1. This C_1 constituted around 10.55 per cent of the entire collection. Similarly, using power core statistical package, C_2 was formed which comprised of 78 accessions and constituted around 5.27 per cent of the entire collection. The accessions of entire collection, C_1 and C_2 were grouped based on geographical distribution and their percentages are presented in Table 1.

The distribution of different accessions of foxtail millet were first classified under different continents and then into countries and states. Among the continents, maximum share to entire collection were from Asia (1,392) followed by USA (39), Africa (9) and Europe (3). In Asian continent, major share was from India (1,363) followed by China (25), Bangladesh (3) and Pakistan (1). Among the different states of India, Uttar Pradesh (553) contributed more followed by Andhra Pradesh (164) and Karnataka (145). In the entire collection, 197 accessions were from unknown regions of India and 35 accessions were from unknown countries.

With respect to C_1, Asia (144) had the maximum number of accessions followed by USA (6) and Europe (1). Within Asian continent, India (140) again contributed more number of accessions. Among states of India, Uttar Pradesh contributed 64 accessions followed by Andhra Pradesh (14), Karnataka (10) and 29 accessions from unknown region of India.

Only Asian continent contributed to the C_2. Source of remaining one accession was unknown. Among the countries of Asia, India (70) ranked first whereas among

Table 1. Number and percentage of accessions contributed to entire collection (E) core 1 (C_1) and core 2 (C_2) from different continents and states within India in foxtail millet

Continent		Countries/States	E		C_1		C_2	
I. Asia	1.	India						
	a	Andhra Pradesh	164	11.10	14	8.97	5	6.41
	b	Bengal	25	1.69	-	-	-	-
	c	Bihar	57	3.86	5	3.21	5	6.41
	d	Chattisgarh	10	0.68	1	0.64	-	-
	e	Gujarat	11	0.74	-	-	-	-
	f	Himachal Pradesh	8	0.54	-	-	-	-
	g	Jammu and Kashmir	12	0.81	2	1.28	1	1.28
	h	Karnataka	145	9.81	10	6.41	6	7.69
	i	Kerala	7	0.47	1	0.64	-	-
	j	Madhya Pradesh	34	2.30	4	2.56	-	-
	k	Maharashtra	28	1.89	3	1.92	3	3.85
	l	NEFA	13	0.88	1	0.64	1	1.28
	m	Orissa	11	0.75	1	0.64	-	-
	n	Punjab	10	0.68	-	-	-	-
	o	Rajasthan	2	0.14	-	-	-	-
	p	Tamil Nadu	76	5.14	5	3.21	4	5.13
	q	Uttar Pradesh	553	37.42	64	41.03	31	39.74
	r	Unknown (India)	197	13.33	29	18.59	14	17.95
		Total (India)	1363	92.23	140	89.74	70	89.74
	2	China	25	1.69	3	1.92	6	7.69
	3	Pakistan	1	0.07	-	-	-	-
	4	Bangladesh	3	0.20	1	0.64	1	1.28
		Total (Asia)	1392	94.19	144	92.30	77	98.71
II. Europe								
	1	USSR	2	0.14	-	-	-	-
	2	Turkey	1	0.07	-	-	-	-
		Total (Europe)	3	0.21	-	-	-	-
III. N. America								
	1	USA	39	2.64	6	3.85	-	-
		Total (America)	39	2.64	6	3.85	-	-
IV. Africa								
	1	Ethiopia	4	0.27	-	-	-	-
	2	Kenya	5	0.34	1	0.64	-	-
		Total (Africa)	9	0.61	1	0.64	-	-
V. Unknown source			35	2.37	5	3.21	1	1.28
		Grand total	1478		156		78	

states, Uttar Pradesh had 31 accessions followed by Karnataka (6), Andhra Pradesh (5) and source of one accession was not known. Thus, the number of accessions present in C_1 and C_2 represented adequately with the number of accessions present in the entire collection. However, entries of C_1 corresponded more with the entries of entire collection than the C_2.

Mean, range, variance and co-efficient of variability of eleven quantitative traits of entire collection, C_1 and C_2 are presented in Table 2. The difference between the means of entire collection and C_1 were non significant for all the eleven traits considered for forming core set. The difference was also non significant for all the traits except plant height in C_2. The homogeneity test for variances between entire collection and C_1 was non significant for all the traits. It was also non significant for all the traits between entire collection and C_2. The range for the traits studied in both C_1 and C_2 were almost similar to the range of the traits in the entire collection. From these variability characteristics, it is evident that two cores formed from two different methods, *viz.*, PCA scores and power core was adequate representative of the entire collection.

In all the three sets (entire, C_1 and C_2), scores were recorded for different sub descriptors of 12 qualitative traits and are presented in Table 3. In order to test whether the ratio of sub descriptors of qualitative traits of core sets formed were in accordance with the ratio of entire collections, chi-square (χ^2) test was done. The χ^2 values for both C_1 and C_2 were non significant indicating that both the methods were successful in forming cores that are true representative of entire collection.

Shannon-Weaver diversity index (H') which indicates the presence of genetic diversity for an individual trait

Table 2. Comparison of mean, range, variance and co-efficient of variability for the quantitative traits in the entire collection and core sets of foxtail millet

S.No.	Characters	Range			Mean			Difference	
		E	C_1	C_2	E	C_1	C_2	C_1	C_2
1.	Days to 50% flowering	33-69	33-69	33-69	49.242	49.391	48.372	NS	NS
2.	Plant height (cm)	52.20-184.00	95.20-184.00	52.20-180.00	142.430	142.858	132.180	NS	*
3.	Number of basal tillers	1.00-12.20	1.00-12.20	1.00-12.00	3.915	3.737	3.908	NS	NS
4.	Flag leaf length (cm)	16.00-47.50	16.00-47.50	17.00-47.50	27.683	28.182	29.114	NS	NS
5.	Flag leaf width (cm)	0.78-4.40	0.85-3.50	0.85-4.40	1.512	1.534	1.597	NS	NS
6.	Peduncle length (cm)	13.40-56.50	13.80-56.50	15.25-56.50	28.055	28.126	29.858	NS	NS
7.	Ear length (cm)	3.60-24.50	3.80-24.50	3.60-23.80	14.766	15.068	15.025	NS	NS
8.	Panicle exertion	1.50-29.00	1.50-29.00	2.70-29.00	13.315	13.135	14.060	NS	NS
9.	Days to maturity	72-110	72-110	72-110	90.800	91.083	89.244	NS	NS
10.	Grain yield per plant (g)	2.10-23.80	2.80-20.00	2.10-23.80	9.751	10.238	10.014	NS	NS
11.	1000-grain weight (g)	1.90-4.00	2.00-4.00	2.00-3.90	3.069	3.034	2.999	NS	NS

Table 2. Contd.

S.No.	Characters	Variance			Difference		CV (%)		
		E	C_1	C_2	C_1	C_2	E	C_1	C_2
1.	Days to 50% flowering	13.926	24.046	35.276	NS	NS	7.578	9.928	12.279
2.	Plant height (cm)	195.518	197.378	701.881	NS	NS	10.379	9.834	20.043
3.	Number of basal tillers	2.009	3.008	5.225	NS	NS	36.204	46.410	58.491
4.	Flag leaf length (cm)	30.401	44.970	51.114	NS	NS	19.917	23.795	24.557
5.	Flag leaf width (cm)	0.078	0.103	0.355	NS	NS	18.471	20.922	35.110
6.	Peduncle length (cm)	19.711	33.055	62.864	NS	NS	15.825	20.441	26.555
7.	Ear length (cm)	5.973	9.915	10.970	NS	NS	16.551	20.897	22.044
8.	Panicle exertion	11.016	17.704	24.649	NS	NS	24.927	32.034	35.311
9.	Days to maturity	45.274	50.296	68.524	NS	NS	7.410	7.786	9.276
10.	Grain yield per plant (g)	15.212	16.884	18.786	NS	NS	39.999	40.135	43.282
11.	1000-grain weight (g)	0.133	0.150	0.192	NS	NS	11.883	12.765	14.611

NS-Non Significant at P=0.05

was estimated for all the 23 traits in the three sets (entire, C_1 and C_2) (Table 4). The comparison between average H' values of 11 quantitative and 12 qualitative traits of entire collection and C_1; entire collection and C_2 indicated the presence of diversity of entire collection in both C_1 and C_2. Thus, the diversity of entire collection was well captured in both the cores, representing the entire collection. But when the Shannon-weaver Index of both the cores was compared, the diversity pattern showed that for each of the qualitative variables except for three, the diversity index was higher for C_2 (PowerCore) compared to C_1 obtained from PCA scoring method indicating the preciseness of power core. This might be accounted for the decrease in the redundancy of the alleles in the accessions of the core set formed using Powercore which resulted in decreased sample size (5%) compared to PCA scoring method (10%).

Comparing the two core forming techniques, PCA score method aims at improving the percentage of sampled diversity without modifying the relative intensity of the selection (10%) as proposed by Brown (1989). In contrast, this technique does not consider the qualitative traits which symbolize principally the phylogenetic diversity. On the other hand, PowerCore has many advantages over the PCA technique. Firstly, PowerCore takes into account the uniqueness in the value of an accession for each character including qualitative traits during the filling of the diversity cells and retains all classes in the core collection. Secondly, increase in number of classes leads to more accessions being selected to fill in the range thus giving more weight or capturing more diversity in the particular character.

A sampling method to obtain a core set should maximize the diversity in the core set thus by reducing the repetitiveness of the identical genotypes. The frequencies of the rare classes should increase with reduction in most represented classes. The core sets thus extracted can be used to widen the genetic base of their breeding population by identifying new yield promoting and other desirable alleles. Core sampling also helps in

Table 3. Descriptors, descriptor states, score code and phenotypic proportions in entire collection (E) and core sets (C_1 and C_2) of foxtail millet

S.No.	Descriptor	Descriptor state	Score code	E	C_1	C_2	χ^2	
							C_1	C_2
1	Plant pigmentation	Non-pigmented	0	1292	139	64	NS	NS
	at flowering	Pigmented	1	186	17	14		
2	Blade pubescence	Essentially glabrous	1	1469	153	77	NS	NS
		Medium pubescent	5	7	2	1		
		Strongly pubescent	9	180	17	10		
3	Sheath pubescence	Essentially glabrous	1	154	17	8	NS	NS
		Medium pubescent	5	1144	122	60		
		Strongly pubescent	9	180	17	10		
4	Senescence	Actively growing	1	160	27	11	NS	NS
		Dead	9	1318	129	67		
5	Inflorescence lobes	Absent	0	250	28	16	NS	NS
		Short	3	979	90	43		
		Long	7	244	36	18		
		Large and thick	9	5	2	1		
6	Inflorescence bristles	Absent	0	27	4	4	NS	NS
		Very short	1	146	18	10		
		Short	3	289	28	17		
		Medium	5	496	51	22		
		Long	7	508	52	23		
		Carrying a spikelet	9	12	3	2		
7	Lobe compactness	Loose	3	27	5	4	NS	NS
		Medium	5	173	16	12		
		Compact	7	1000	104	47		
		Spongy	9	276	31	15		
8	Inflorescence shape	Cylindrical	1	1378	143	70	NS	NS
		Pyramidal	2	99	13	8		
		Obovate	3	1	0	0		
9	Inflorescence compactness	Loose	3	208	24	11	NS	NS
		Medium	5	508	51	25		
		Compact	7	341	31	20		
		Spongy	9	421	50	22		
10	Fruit colour	Red	1	52	6	7	NS	NS
		Black	2	22	2	2		
		White	3	1018	113	51		
		Yellow	4	386	35	18		
11	Grain shape	Oval	1	1310	136	64	NS	NS
		Elliptical	2	168	20	14		
12	Apical sterility in panicle	Absent	0	708	78	46	NS	NS
		Present	1	770	78	32		

NS- Non Significant at P=0.05

defining a specific characteristics in which portion of the entire collection is likely to be found. The two methods used to construct foxtail core sets were found to preserve the variation of the entire collection as shown by their non significant differences thereby representing the total diversity of the entire collection. These core sets can be extensively examined for all economically important traits. The data generated will provide the information on genetic variability in foxtail millet. This core set developed should be revised periodically as additional accessions and information becomes available.

Acknowledgement

We gratefully acknowledge Dr. PN Mathur, Scientist, Bioversity International, New Delhi for providing

Table 4. Shannon-Weaver diversity index (H') for 11 quantitative and 14 qualitative characters in the entire collection (E) and core sets (C_1 and C_2) of foxtail millet

S.No.	Character	Shannon–Weaver Index		
	Quantitative character	E	C_1	C_2
1	Days to 50% flowering	2.092	2.321	2.453
2	Plant height (cm)	2.132	2.100	2.660
3	Number of basal tillers	2.293	2.371	2.496
4	Flag leaf length (cm)	2.615	2.721	2.802
5	Flag leaf width (cm)	1.712	1.782	2.254
6	Peduncle length (cm)	2.071	2.198	2.497
7	Ear length (cm)	2.254	2.416	2.495
8	Panicle exertion	2.263	2.416	2.606
9	Days to maturity	2.195	2.139	2.430
10	Grain yield per plant (g)	2.635	2.627	2.652
11	1000 grain weight (g)	2.614	2.590	2.717
	Mean ± SE	2.261 ± 0.08	2.335 ± 0.08	2.551 ± 0.04
	Qualitative character			
1	Plant pigmentation at flowering	0.378	0.344	0.470
2	Blade pubescence	0.040	0.355	0.270
3	Sheath pubescence	0.690	0.675	0.699
4	Senescence	0.343	0.461	0.407
5	Inflorescence lobes	0.890	1.020	1.068
6	Inflorescence bristles	1.393	1.577	1.608
7	Lobe compactness	0.906	0.935	1.063
8	Inflorescence shape	0.251	0.287	0.555
9	Inflorescence compactness	1.339	1.339	1.347
10	Fruit colour	0.885	0.890	1.264
11	Grain shape	0.465	0.474	0.623
12	Apical sterility in panicle	0.692	0.693	0.677
	Mean ± SE	0.689 ± 0.12	0.754 ± 0.11	0.837 ± 0.12
	Overall mean ± SE	1.48 ± 0.78	1.55 ± 0.79	1.70 ± 0.85

PowerCore (v. 1.0) software and his guidance in forming core.

References

Anonymous (1985) *Descriptors for Setaria italica and Setaria pumila.* International Board for Plant Genetic Resources, Rome, Italy, **18** p.

Frankel OH (1984) Genetic perspective of germplasm conservation. In: W Arber, K Llimensee, WJ Peacock and P Starlinger (eds) *Genetic Manipulations: Impact of Man and Society.* Cambridge University Press, Cambridge, England, pp 161-170.

Brown AHD (1989) Core collections: a practical approach to genetic resources management. *Genome* **321:** 818-824.

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

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

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

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