

Research Paper

Establishment of a core collection of *Cynodon* based on morphological data

Establecimiento de una colección núcleo de Cynodon basada en datos morfológicos

C.Q. HUANG, T. LONG, the late C.J. BAI, W.Q. WANG, J. TANG AND G.D. LIU

Tropical Crops Genetic Resources Institute, Chinese Academy of Tropical Agricultural Sciences/Key Laboratory of Crop Gene Resources and Germplasm Enhancement in Southern China, Ministry of Agriculture and Rural Affairs, Haikou, Hainan Province, China. catas.cn

Abstract

In a field plot study conducted in Danzhou, Hainan province, China, a total of 537 wild *Cynodon* accessions from 22 countries and classified into 11 groups according to taxonomy and origin, were characterized in terms of 11 phenotypic traits in order to construct a core collection. For this, the optimal strategy was developed by screening within the following method levels: (i) 7 sampling proportions (5, 10, 15, 20, 25, 30 and 35%); (ii) 3 sampling methods (preferential sampling, deviation sampling and random sampling); (iii) 5 clustering methods [single linkage, completed linkage, median linkage, unweighted pair-group average (UPGMA) and Ward's method]; (iv) 3 genetic distances (Euclidean distance, Mahalanobis distance and principal component distance); and (v) 3 sampling proportions within groups (simple, logarithmic and square root proportions). Mean difference percentage, variance difference percentage, coincidence rate of range and variation coefficient changing rate were the criteria adopted for evaluating how well the core collection represented the original collection. The correlation between the original and core collections was determined for comparison. The core collections were validated with the sample distribution diagram of the main components. Results showed that the optimal sampling method for constructing a *Cynodon* core collection was preferential sampling, the optimal sampling proportion being 20%. The optimal sampling proportion within groups was the square root proportion, the optimal genetic distance was Mahalanobis distance and the optimal clustering method was UPGMA. The proposed core collection of *Cynodon* is composed of 108 accessions; it was constructed following the optimal sampling strategy identified and retained the original collection's phenotypic diversity, phenotypic trait correlations and phenotypic group structure. Thus, this collection could be considered a representative sample of the entire resource.

Keywords: Germplasm conservation, optimal statistical strategy, phenotypic selection.

Resumen

Con el objetivo de construir una colección núcleo, en un estudio en parcelas de campo localizadas en Danzhou, provincia de Hainan, China, se caracterizaron un total de 537 accesiones de *Cynodon* silvestres procedentes de 22 países, clasificadas en 11 grupos según taxonomía y origen, teniendo en cuenta 11 caracteres fenotípicos. Para el efecto, se determinó la estrategia óptima evaluando y seleccionando los métodos más apropiados en los siguientes niveles metodológicos: (i) 7 proporciones de muestreo (5, 10, 15, 20, 25, 30 y 35%); (ii) 3 métodos de muestreo (preferencial, de desviación y aleatorio); (iii) 5 métodos de aglomeración (enlace simple, enlace completo, método de la mediana, UPGMA (unweighted pair-group method using arithmetic averages) y método de Ward); (iv) 3 distancias genéticas (euclidiana, Mahalanobis y componente principal); y (v)

Correspondence: C.Q. Huang & G.D. Liu, Tropical Crops Genetic Resources Institute, Chinese Academy of Tropical Agricultural Sciences/Key Laboratory of Crop Gene Resources and Germplasm Enhancement in Southern China, Ministry of Agriculture and Rural Affairs, Haikou 571101, China.
E-mail: huangchunjiong44@163.com; liuguodao2008@163.com

3 proporciones de muestreo dentro de grupos (proporciones simples, logarítmicas y de raíz cuadrada). El porcentaje de diferencia de la media, el porcentaje de diferencia de la varianza, la tasa de coincidencia del rango y la tasa de cambio del coeficiente de variación fueron los criterios adoptados para evaluar la representatividad de la colección núcleo respecto a la colección original. Para la comparación se determinó la correlación entre las colecciones original y núcleo. Las colecciones núcleo fueron validadas con el diagrama de distribución muestral de los componentes principales. Los resultados mostraron que para construir una colección núcleo de *Cynodon*, el óptimo método de muestreo es el preferencial, siendo el 20% la óptima proporción de muestreo. La óptima proporción de muestreo dentro de los grupos fue la proporción de raíz cuadrada, la óptima distancia genética fue la distancia de Mahalanobis y el óptimo método de aglomerado fue UPGMA. La colección núcleo de *Cynodon* propuesta consiste de 108 accesiones; se construyó siguiendo la estrategia de muestreo óptima identificada y retuvo la diversidad fenotípica de la colección original, las correlaciones de caracteres fenotípicos y la estructura fenotípica de los grupos. Por lo tanto, esta colección puede considerarse una muestra representativa de todo el recurso evaluado.

Palabras clave: Conservación de germoplasma, estrategia estadística óptima, selección fenotípica.

Introduction

Cynodon Rich., a C4-type herbaceous genus of the family Poaceae, subfamily Chlorioideae, tribe Cynodonteae, is one of the main 3 warm-season turf grasses as well as a fine pasture (Harlan and de Wet 1969) that can be divided into 9 species and 10 varieties (Taliaferro 1995). Most *Cynodon* species originated from the African region; they are mainly distributed in warm and humid tropical or subtropical regions, with some in temperate regions (Rochecouste 1962; Harlan et al. 1970; Wofford and Baltensperger 1985). Two *Cynodon* taxa, *Cynodon dactylon* (L.) Pers. var. *dactylon* and *C. transvaalensis* Burt Davy, are important to the turf grass industry (Kenworthy et al. 2007). *C. dactylon* var. *dactylon* is a cosmopolitan species with a latitudinal distribution range of 53° N–45° S. In terms of elevation, this species can grow from sea level to the Himalayas at an elevation of 4,000 m. As a result of its widespread distribution and prevalence, it is frequently referred to as ‘common’ Bermuda grass. *C. transvaalensis*, often referred to as African Bermuda grass, is indigenous to only the Transvaal region of South Africa (Harlan et al. 1970). Two species, namely *C. dactylon* and *C. radiatus* Roth, plus an as yet undetermined *Cynodon* taxon, exist in China. Domestic and foreign scholars have made progress in the collection and genetic diversity evaluation of *Cynodon* germplasm resources (Wu et al. 2004, 2006; Wang et al. 2009; Tiwari et al. 2016; Zheng et al. 2017). These studies provided a good basis for the construction of a core collection of *Cynodon*.

Currently, core collections are mainly concentrated on crops or horticultural plants (Ortiz et al. 1998; Li et al. 2002; Martínez et al. 2017). Few studies have focused on pasture species with constructed core collections being mainly *Lolium perenne*, perennial *Medicago*, annual *Medicago*, *Cajanus cajan* and *Pennisetum glaucum* (Basigalup et al. 1995; Charmet and Balfourier 1995; Diwan et al. 1995; Reddy et al. 2005; Bhattacharjee et al. 2007). However, few investigations have involved a *Cynodon* core collection.

Anderson (2005) constructed a *C. dactylon* core collection of 169 genotypes with the phenotypic data for 598 *C. dactylon* samples. Jewell et al. (2011) performed EST-SSR amplification on the DNA of 690 Australian samples of *C. dactylon*. The SSR data obtained were combined with information on the geographical origin of the samples to construct the Australian *C. dactylon* core collection with the core material representing 13% of the original collection selected by the hierarchical cluster sampling method. Zheng et al. (2014) used phenotypic data for 831 *C. dactylon* samples to construct a preliminary core collection, including 208 genotypes.

In recent years, the Tropical Crops Genetic Resources Institute of the Chinese Academy of Tropical Agricultural Sciences has assembled a large array of *Cynodon* germplasm. The collected accessions were studied in terms of morphology, molecular biology, stress resistance, turf value, feed value and so on (Huang et al. 2013; 2014) and provided a broad range of material for systematic research and genetic breeding of *Cynodon*. Nevertheless, many investigations experience difficulties with preservation, evaluation and identification. There is an urgent need to examine existing germplasm resources to hasten the identification of particular germplasm for breeding studies. Construction of a *Cynodon* core collection is crucial for germplasm resource innovation, effective protection and utilization, species improvement and breeding of new varieties.

Materials and Methods

Plant material

The plant material used was 537 wild *Cynodon* accessions collected from 22 different countries and regions from 2006 to 2014 by the Prataculture Research Office in the Tropical Crops Genetic Resources Institute of the Chinese Academy of Tropical Agricultural Sciences (CATAS). These accessions included 476 *C. dactylon*, 59 *C. radiatus*, 1 *C. transvaalensis*

and 1 as yet unidentified taxon (*Cynodon* sp., with 2-flowered spikelets). The sources are shown in Table 1.

Experiment setup

Plants were established in December 2013 in 1-m² plots in the experimental fields of the Tropical Crops Genetic Resources Institute of the Chinese Academy of Tropical Agricultural Sciences (CATAS), in Danzhou, Hainan, China. A randomized block design was used; each material was replicated 3 times.

Morphological traits

The morphological traits were measured, after standardization cuts in March 2014 and March 2015, on 2 occasions, i.e. June 2014 and June 2015. The following 11 traits were assessed: 7 quantitative traits (erect branch leaf length and width; stolon leaf length and width; stolon stem diameter and length; and turf height), which were measured 15 times in each plot to obtain mean values; and 4 qualitative traits (leaf hair; leaf posture; leaf color; and stolon stem color), which were measured 3 times (Huang et al. 2012). The measuring method was as follows: erect branch leaf length and width: the fourth mature leaf from the top to the base of the erect branch was randomly selected to measure the length and width of the leaf at the widest point with a Vernier caliper. Stolon leaf length and width: the fourth mature leaf from the top to the base of

the stolon was randomly selected to measure the length and width of the leaf at the widest point with a Vernier caliper. Stolon stem diameter and length: the fourth stem of the stolon from the top to the base was selected to measure the diameter and length with a Vernier caliper. Turf height: the natural height of the accession was measured using a rule. Leaf hair, leaf posture, leaf color and stem color were measured by ocular estimate.

Screening of sampling method, sampling proportion, clustering method and genetic distance

The sampling methods compared were: preferential sampling; deviation sampling; and random sampling (Hu et al. 2000; Li et al. 2004). Sampling proportions from the entire collection were at 7 levels (5, 10, 15, 20, 25, 30 and 35%). Five clustering methods (single linkage method, completed linkage method, median linkage method, unweighted pair-group average method and Ward's method) (Ward 1963; Sibson 1973) were adopted. The genetic distances were: Euclidean distance; Mahalanobis distance; and principal component distance. Consequently, 315 *Cynodon* core subsets were constructed by combining 2 of 4 factors, i.e. sampling method, entire sampling proportion, clustering method and genetic distance. By evaluating the core subsets, the relative effectiveness of optimal sampling method, entire sampling proportion, clustering method and genetic distance for selecting a core collection from the entire collection was assessed.

Table 1. Sources and grouping of *Cynodon* accessions used in the present study.

Group number	Species	Number. of accessions	Source/origin
1	<i>C. dactylon</i>	174	South China (136 in Hainan, 17 in Guangdong, 21 in Guangxi)
2	<i>C. dactylon</i>	139	East China (10 in Anhui, 27 in Fujian, 10 in Shanghai, 29 in Jiangsu, 11 in Jiangxi, 33 in Shandong, 19 in Zhejiang)
3	<i>C. dactylon</i>	36	Central China (19 in Henan, 10 in Hubei, 7 in Hunan)
4	<i>C. dactylon</i>	4	North China (1 in Hebei, 1 in Beijing, 2 in Tianjin)
5	<i>C. dactylon</i>	48	Southwest China (1 in Chongqing, 16 in Guizhou, 6 in Sichuan, 25 in Yunnan)
6	<i>C. dactylon</i>	28	Northwest China (1 in Gansu, 20 in Xinjiang, 7 in Shanxi)
7	<i>C. dactylon</i>	47	Abroad (3 in Vietnam, 3 in Thailand, 3 in Cambodia, 1 in Malaysia, 2 in East Timor, 1 in Myanmar, 2 in Indonesia, 5 in Sri Lanka, 1 in India, 1 in Grenada, 2 in Costa Rica, 2 in Venezuela, 3 in Colombia, 6 in Congo, 2 in Zambia, 1 in South Africa, 1 in Australia, 5 in Papua New Guinea, 1 in Liberia, 1 in Burundi, 1 in Côte d'Ivoire)
8	<i>C. radiatus</i>	55	China (48 in Hainan, 1 in Guangdong, 1 in Shanghai, 2 in Jiangsu, 1 in Shandong, 1 in Tianjin, 1 in Yunnan)
9	<i>C. radiatus</i>	4	Abroad (2 in Vietnam, 1 in Grenada, 1 in Papua New Guinea)
10	<i>Cynodon</i> sp. (unid.)	1	Southwest China (1 in Yunnan)
11	<i>C. transvaalensis</i>	1	South America (1 in Colombia)

Screening of sampling proportion within groups

The original germplasm samples, 537 *Cynodon* accessions, were first divided into 4 groups with the hierarchical grouping method, initially on the basis of botanical names: *C. dactylon*, *C. radiatus*, *Cynodon* sp. (unid.) and *C. transvaalensis*. These 4 major groups were further divided into 11 subgroups depending on geographical origin. The above screened sampling method, entire sampling proportion, clustering method and genetic distance with the original germplasm grouping were used to screen the 3 sampling proportions within groups (simple proportion, logarithmic proportion and square root proportion) for the construction of the core collection. The number of core materials of each subgroup was calculated. The specified quantity of core materials of the 11 groups was sampled with the screened sampling method, entire sampling proportion, clustering method and genetic distance to be evaluated. Variance analysis of the evaluation parameters was conducted to determine the optimal sampling proportion within groups.

Construction of *Cynodon* core collection

Phenotypic data for *Cynodon* were processed by the screened sampling method, entire sampling proportion, sampling proportion within group, clustering method and genetic distance and core materials were then extracted to construct a phenotypic core collection of *Cynodon*.

Evaluation of *Cynodon* core collection

To determine whether the constructed core collection retained the genetic diversity of the original collection and thus was representative of the original collection, 4 evaluation parameters (mean difference percentage, variance difference percentage, coincidence rate of range and variation coefficient changing rate; [Hu et al. 2000](#)) were selected. The core collection was considered representative/effective only when the mean difference percentage was <20%, and the coincidence rate of range was ≥80%. Moreover, a small mean difference percentage resulted in a large variance difference percentage. The coincidence rate of range and the variation coefficient changing rate showed strong genetic diversity of the original collection. The 4 evaluation parameters were calculated using the following formulae ([Hu et al. 2000](#)):

$$MD = (S_t/n) \times 100\%$$

$$VD = (S_F/n) \times 100\%$$

$$CR = \frac{1}{n} \sum_{i=1}^n \frac{R_{C(i)}}{R_{I(i)}} \times 100\%$$

$$VR = \frac{1}{n} \sum_{i=1}^n \frac{CV_{C(i)}}{CV_{I(i)}} \times 100\%$$

where:

MD = the mean difference percentage, VD = variance difference percentage, CR = coincidence rate of range, VR = variation coefficient changing rate;

i represents the *i*th trait;

n is the total number of traits;

S_t is the number of traits with significant difference when the *t*-test between core collection and original collection is performed;

S_F is the number of traits with significant differences when the F-test between core collection and original collection is performed;

R_{C(i)} is the range of the *i*th trait of the core collection group;

R_{I(i)} is the range of the *i*th trait of the original population;

CV_{C(i)} is the coefficient of variation of the *i*th trait of the core collection; and

CV_{I(i)} is the coefficient of variation of the *i*th trait of the original population.

Confirmation of *Cynodon* core collection

To evaluate the representativeness of the constructed core collection, principal component analysis of the original and core collections was conducted. Taking the first and second principal components as the horizontal and vertical axes, respectively, a 2D distribution diagram of their principal components was drawn to compare their population structures. The correlations of the 11 phenotypic traits of both the original and core collections were analyzed to decide whether the correlations of the traits in the original collection were retained in the core collection ([Ortiz et al. 1998](#)).

Data processing and analysis

Microsoft Office Excel 2007 was used to process the data. SPASS 19.0 was used for correlation analysis and principal component analysis.

Results

Screening of sampling method, entire sampling proportion, clustering method and genetic distance

Without the grouping of the original collection, 3 sampling methods (preferential sampling, deviation sampling and random sampling), 7 entire sampling proportions (5, 10, 15, 20, 25, 30 and 35%), 5 clustering methods (single linkage, completed linkage, median linkage, UPGMA and Ward's method) and 3 genetic distances (Euclidean distance, Mahalanobis distance and principal component distance) were combined to construct 315 core subsets. These 315 core subsets were evaluated using mean difference percentage, variance difference percentage, coincidence rate of range and variation coefficient changing rate. The 4 evaluation parameters were subjected to analysis of variance, and the results are shown in Table 2.

Screening of sampling method

The constructed core subsets by the 3 sampling methods were comprehensively evaluated using the mean difference percentage, variance difference percentage, coincidence rate of range and variation coefficient changing rate (Table 2). The mean coincidence rates of range of the constructed core subsets by preferential

sampling and deviation sampling were 100%. The 3 other parameters were significantly better than that of deviation sampling and random sampling. Thus, we concluded that the optimal sampling method was preferential sampling.

Screening of genetic distance

Table 2 shows that the average mean difference percentage range of the core subsets constructed by Euclidean distance, Mahalanobis distance and principal component distance was in the range of 2.86–4.16%. The average variance percentage range was 65.3–65.6%, the average coincidence rates of range values were all 100% and the average variation coefficient changing rate range was 136–158%. These 3 factors were not significantly different in the 4 evaluation parameters, so any of the 3 genetic distances could be used. Mahalanobis genetic distance was considered the optimal genetic distance.

Screening of clustering method

Table 2 also demonstrates that the minimum average mean difference percentage of the core subsets constructed with the 5 clustering methods was 0.87%, and the maximum was 4.33%. The value range of the average difference percentage was 62.3–68.8%, the average coincidence rates of range were all 100%, and the average

Table 2. Evaluation parameters of the core subsets constructed with different construction strategies.

Item	Construction strategy	Average mean difference percentage (%)	Average variance difference percentage (%)	Average coincidence rate of range (%)	Average variation coefficient changing rate (%)
Sampling method	Preferential sampling	0.26A ¹	69.4A	100A	162Aa
	Deviation sampling	7.19B	61.6B	100A	139Ab
	Random sampling	1.30A	44.2C	79.3B	102Bc
Genetic distance	Euclidean distance	2.86	65.3	100	136
	Mahalanobis distance	4.16	65.6	100	158
	Principal component distance	4.16	65.6	100	158
Clustering method	Single linkage	0.87	63.9	100	136
	Completed linkage	4.98	67.5	100	135
	Median linkage	5.41	64.9	100	160
	UPGMA	3.03	68.8	100	137
	Ward's method	4.33	62.3	100	135
Sampling proportion (%)	5	0.00A	76.0A	98.4	273Aa
	10	0.59A	70.7AB	99.4	150ABab
	15	1.47AB	70.7AB	99.4	137BCbc
	20	3.81AB	69.2AB	99.4	129CDcd
	25	4.99AB	61.6BC	99.5	122DEde
	30	7.33B	55.1C	99.5	118Eef
	35	7.33B	50.4C	99.7	115Eef

¹Within parameters and columns, upper-case letters indicate significant differences at $P < 0.01$ and lower-case letters show significant differences at $P < 0.05$.

range of variation coefficient changing rate was 135–160%. The 4 evaluation parameters of the 5 clustering methods showed no significant differences, so any of the clustering methods could be selected to construct the core collection. In this research we used the commonly used unweighted pair-group average (UPGMA) method as the optimal clustering method.

Preliminary screening of sampling proportion

The evaluation parameters of each of the core subsets exhibited large differences when different entire sampling proportions, namely, 5, 10, 15, 20, 25, 30 and 35%, were used to construct the *C. dactylon* core subsets (Table 2). Comprehensive analysis of the 4 evaluation parameters revealed that 5 and 10% were optimal for all 4 evaluation parameters, with no significant differences between them. Additionally, 15 and 20% showed significant differences from 5% only for the average variation coefficient changing rate. The entire parameter values of 25, 30 and 35% were inferior to the other 4 proportions. Theoretically, 5 and 10% of the original population should be selected as the entire sampling proportions but are rarely used in similar research. Since 15 and 20% showed little significant difference from 5 and 10% we selected them as alternatives for further screening in subsequent tests.

Screening of sampling proportion within groups and entire sampling proportions

Sampling proportions within groups were set as simple proportion, square root proportion and logarithmic proportion, while the entire sampling proportions were 10 and 20%. We combined entire sampling proportions and sampling proportions within groups in 6 ways to screen and calculate the number of core collections to be extracted from

each group. Table 3 shows that the numbers of *Cynodon* accessions contained in the 9th, 10th and 11th groups were 4, 1 and 1. To ensure that the special material from the original collection was maintained by the core collection, only square root proportion in the 3 kinds of sampling proportions within groups could extract core materials from all 3 groups. Thus, square root proportion was selected as the sampling proportion within groups.

Moreover, 10 and 20% constructions and the evaluation parameters of the core subsets were compared. Table 4 demonstrates that the 2 constructions showed only small differences and differences were inconsistent, so 20% was used as the optimal sampling proportion.

Cynodon core collection

The phenotypic data for 537 *Cynodon* samples in the collection were analyzed using the above optimal core collection construction strategy for screening. The core materials were extracted to construct a core collection of 108 samples, which consisted of 89 *C. dactylon*, 17 *C. radiatus*, 1 *Cynodon* sp. (unid.) and 1 *C. transvaalensis* accession (Table 5).

Evaluation of core collection

Correlation analysis of phenotypic traits. Correlation analysis on the 11 phenotypic traits of the original and core collections was conducted (Table 6) and showed that 28 pairs of phenotypic traits in the original *Cynodon* collection were highly significantly ($P < 0.01$) correlated and 4 pairs of phenotypic traits were significantly ($P < 0.05$) correlated. In the core collection, 24 pairs of phenotypic traits were highly significantly ($P < 0.01$) correlated and 4 pairs of traits were significantly ($P < 0.05$)

Table 3. Numbers of core materials extracted in each group of *Cynodon* with different sampling proportions.

Group number	Number. of accessions	Simple proportion sampling		Logarithmic proportion sampling		Square root proportion sampling	
		10%	20%	10%	20%	10%	20%
1	174	17	35	9	18	11	22
2	139	14	28	8	17	10	20
3	36	4	7	6	12	5	10
4	4	0	1	2	5	2	4
5	48	5	10	7	13	6	12
6	28	3	6	6	11	4	9
7	47	5	9	7	13	6	12
8	55	6	11	7	14	6	13
9	4	0	1	2	5	2	4
10	1	0	0	0	0	1	1
11	1	0	0	0	0	1	1

Table 4. Evaluation parameters of the constructed core subsets with the combination of two sampling proportions.

Sampling proportion (%)	Mean difference percentage (%)	Variance difference percentage (%)	Coincidence rate of range (%)	Variation coefficient changing rate (%)
10	9.09	81.8	92.7	142
20	9.09	90.9	99.7	137

Table 5. Core collection for *Cynodon* accessions.

Species	Collection site	Total number of accessions	Species	Collection site	Total number of accessions
<i>C. dactylon</i>	Hainan, China	17	<i>C. dactylon</i>	Xinjiang, China	5
<i>C. dactylon</i>	Guangdong, China	2	<i>C. dactylon</i>	East Timor	1
<i>C. dactylon</i>	Guangxi, China	4	<i>C. dactylon</i>	Cambodia	1
<i>C. dactylon</i>	Anhui, China	2	<i>C. dactylon</i>	Sri Lanka	2
<i>C. dactylon</i>	Fujian, China	3	<i>C. dactylon</i>	Thailand	1
<i>C. dactylon</i>	Jiangsu, China	6	<i>C. dactylon</i>	India	1
<i>C. dactylon</i>	Jiangxi, China	2	<i>C. dactylon</i>	Colombia	1
<i>C. dactylon</i>	Shandong, China	6	<i>C. dactylon</i>	Costa Rica	1
<i>C. dactylon</i>	Shanghai, China	2	<i>C. dactylon</i>	Congo	1
<i>C. dactylon</i>	Henan, China	6	<i>C. dactylon</i>	Côte d'Ivoire	1
<i>C. dactylon</i>	Hubei, China	3	<i>C. dactylon</i>	South Africa	1
<i>C. dactylon</i>	Hunan, China	2	<i>C. dactylon</i>	Zambia	1
<i>C. dactylon</i>	Tianjing, China	1	<i>C. radiatus</i>	Hainan, China	12
<i>C. dactylon</i>	Guizhou, China	6	<i>C. radiatus</i>	Shandong, China	1
<i>C. dactylon</i>	Sichuan, China	1	<i>C. radiatus</i>	Vietnam	2
<i>C. dactylon</i>	Yunnan, China	4	<i>C. radiatus</i>	Grenada	1
<i>C. dactylon</i>	Chongqing, China	1	<i>C. radiatus</i>	Papua New Guinea	1
<i>C. dactylon</i>	Gansu, China	1	<i>Cynodon</i> sp. (unid.)	Yunnan, China	1
<i>C. dactylon</i>	Shanxi, China	3	<i>C. transvaalensis</i>	Colombia	1
			Total		108

correlated. Approximately 85% of the phenotypic trait correlations in the original collection was maintained in the core collection.

Principal component analysis

To further determine how representative the core collection was of the original collection, principal component analysis of the phenotypic data of the 2 groups was conducted (Table 7). Four principal components were selected with a standard eigenvalue >1. The eigenvalue, proportion and accumulating contribution rate of the principal components in both

core and original collections were similar. The eigenvalues of principal components 1–3 were slightly higher than those of the original collection. The 2D scatter diagram (Figure 1), based on the core collection for the first and second principal components, showed that the geometrical shapes and features for the core collection's principal components were similar to those of the original collection. The collection materials of both groups were concentrated mainly in the left part of the scatter diagram. They were distributed sporadically in the lower part of the diagram, thereby indicating that the genetic structures of the original collections were retained well by the core collection.

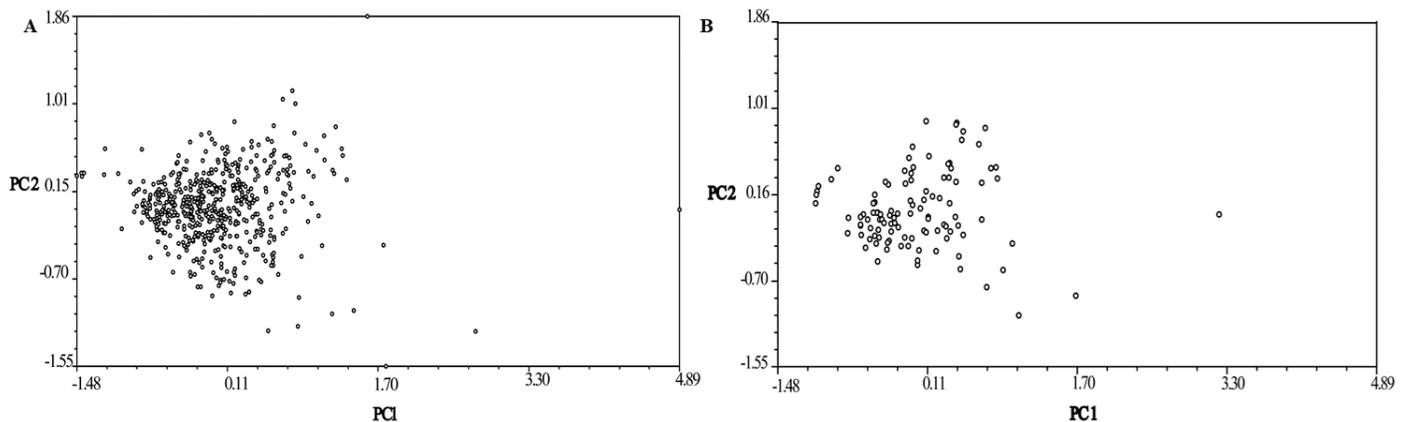
Table 6. Correlation analysis¹ of phenotypic traits in original and core collections.

Item	Erect stem leaf length	Stolon stem leaf length	Erect stem leaf width	Stolon stem leaf width	Stolon stem diameter	Stolon stem internode length	Turf height	Leaf color	Leaf hair	Leaf posture	Stolon stem color
Erect stem leaf length		0.79**	0.55**	0.53**	0.66**	0.38**	0.80**	-0.16	0.41**	-0.13	-0.06
Stolon stem leaf length	0.60**		0.44**	0.46**	0.49**	0.37**	0.70**	-0.27**	0.37**	-0.12	-0.15
Erect stem leaf width	0.47**	0.27**		0.84**	0.46**	0.50**	0.41**	0.19	0.18	0.03	-0.09
Stolon stem leaf width	0.41**	0.37**	0.80*		0.44**	0.51**	0.41**	0.04	0.16	0.03	-0.09
Stolon stem diameter	0.50**	0.27**	0.33**	0.33**		0.17	0.49**	0.08	0.21*	-0.23*	0.01
Stolon stem internode length	0.26**	0.10*	0.48**	0.41**	0.04		0.28**	-0.03	0	0.01	-0.05
Turf height	0.69**	0.52**	0.29**	0.25**	0.36**	0.13**		-0.14	0.52**	-0.12	-0.07
Leaf color	-0.07	-0.09*	0.08	0.08	0.16**	-0.09*	-0.04		-0.07	0	-0.24*
Leaf hair	0.26**	0.25**	0.11*	0.08	0.18*	0	0.31**	-0.02		-0.11	-0.22*
Leaf posture	-0.07	-0.03	0.11**	0.08	-0.08	0.03	-0.02	-0.02	-0.08		-0.02
Stolon stem color	0.03	-0.11**	-0.05	-0.05	0.13**	-0.01	-0.03	0.02	-0.15**	-0.01	

¹Data in the upper section indicate the correlation coefficients among traits of the core collection and data in the lower section indicate the correlation coefficients among traits of the original collection.

Table 7. Eigenvalues and accumulating contribution rates of components for the original and core collections in principal component analysis.

Principal component	Original collection			Core collection		
	Eigenvalue	Contribution rate (%)	Accumulating contribution rate (%)	Eigenvalue	Contribution rate (%)	Accumulating contribution rate (%)
PC1	3.4	30.9	30.9	3.89	35.34	35.34
PC2	1.47	13.37	44.27	1.49	13.52	48.86
PC3	1.16	10.51	54.79	1.35	12.27	61.13
PC4	1.1	10.03	64.82	1.04	9.43	70.56

**Figure 1.** Distribution diagram of the original collection (A) and core collection (B) principal components based on phenotypic data. Horizontal axis indicates the first principal component quantity for each collection, and the vertical axis shows the second principal component quantity.

Discussion

Construction strategy of core collection

Core collection construction is a continuous sampling process employing a certain sampling strategy. A good sampling method can remove the genetic redundancy of the furthest group and maximize the variation in the core collection, while conserving the genetic diversity of the original group (Brown 1989). The main methods for constructing a core collection are: completely random sampling; grouping first and then completely random sampling; and clustering first and then sampling based on genetic structure (Spagnoletti Zeuli and Qualset 1993). Given the uneven genetic diversity of resource distribution, core collections with different diversity structures can result with random and cluster samplings (Spagnoletti Zeuli and Qualset 1993). Most studies show that the sampling effect of cluster sampling is better than that of random sampling (Spagnoletti Zeuli and Qualset 1987; Diwan et al. 1995). Cluster sampling can remove a certain proportion of samples from genetic material close to the genetic distance to reduce genetic redundancy and maintain the genetic structure of the original collection; thus, the sampling results of cluster sampling are better than those of random sampling. In the present research, 3 sampling methods (multiple clustering preferential sampling method, multiple clustering variability sampling method and non-clustered completely random sampling method) were compared. The multiple clustering preferential sampling method proved most suitable for *Cynodon* core collection construction.

Sampling proportion is also a key step in constructing a core collection and is divided into the entire sampling proportion and sampling proportion within groups. High sampling proportion may result in samples with high redundancy, or important core collection materials may be lost (Brown 1989). In the construction of different plant core collections the selected proportion can range from 5 to 30% of the total collection of the species, but commonly about 10% is used (Lindroth et al. 2002). In this study, we used 20% as the sampling proportion for construction of the *Cynodon* core collection and compared 3 sampling proportions within groups (simple proportion, logarithmic proportion and square root proportion). Square root proportion proved to be the optimal proportion, which was in accordance with the screening results of Zheng et al. (2014), who used phenotypic traits to construct a *C. dactylon* preliminary core collection.

Evaluation of core collection

Correlation among phenotypic traits is an intrinsic characteristic of a species, which is the external manifestation of the correlation among genetic materials within the species. Sampling should not change the correlation among the intrinsic traits of the species; therefore, a well-selected core collection should maintain the correlations among the phenotypic traits of the original collection (Ortiz et al. 1998; Xu et al. 2006). Given that not all germplasm of the original population is held in the phenotypic core collection, and the degrees of freedom of the 2 differ, the correlation coefficients cannot be compared. To determine whether the significantly correlated traits in the original population were also significantly correlated in the phenotypic core collection we performed correlation analysis based on the 11 phenotypic traits of the original and core collections. The findings revealed that the core collection maintained more than 85% of the trait correlations of the original collection.

Principal component analysis is a data analysis technique that transforms multivariates and multi-indexes into few variables and indexes; this technique is widely used in biodiversity research, including the construction of core collections (Zheng et al. 2014). The 2D scatter diagram based on the scores of the germplasm material in the first and second principal components closely reflected the distribution of the germplasm material in the original population, thereby intuitively reflecting the genetic structure of the original population. The distances among the germplasm materials in the scatter diagram reflect their genetic similarity, with close distance indicating high similarity, and long distance reflecting a high degree of dissimilarity. The high contribution rates of the first and second principal components resulted in acceptably accurate reflection of the genetic distribution in the original population. On the basis of a comparison of the 2D distribution diagrams of the principal components of both original and core collections, the core collection was shown to closely reflect the genetic make-up of the original collection.

Conclusions

Developing the optimal strategy regarding sampling method, entire sampling proportion and sampling proportion within groups, genetic distance and clustering method, a core collection of *Cynodon*, including 108 accessions, based on 11 phenotypic traits from 537 wild accessions, was constructed. This core collection is

important for germplasm resources innovation, efficient conservation and utilization and should be useful for species improvement and breeding of varieties.

Acknowledgments

The authors gratefully acknowledge the efforts of the editors and the anonymous reviewers for their valuable comments to improve this manuscript. This work was supported by high level talent project of basic and applied basic research plan of Hainan province (Natural Science Field) (No. 2019RC313) and central public-interest scientific institution basal research fund for Chinese Academy of Tropical Agricultural Sciences (No. 1630032014028).

References

(Note of the editors: All hyperlinks were verified 14 July 2020.)

- Anderson WF. 2005. Development of a forage bermudagrass (*Cynodon* spp.) core collection. *Grassland Science* 51:305–308. doi: [10.1111/j.1744-697X.2005.00040.x](https://doi.org/10.1111/j.1744-697X.2005.00040.x)
- Basigalup DH; Barnes DK; Stucker RE. 1995. Development of a core collection for perennial *Medicago* plant introductions. *Crop Science* 35:1163–1168. doi: [10.2135/cropsci1995.0011183X003500040042x](https://doi.org/10.2135/cropsci1995.0011183X003500040042x)
- Bhattacharjee R; Khairwal IS; Bramel PJ; Reddy KN. 2007. Establishment of a pearl millet [*Pennisetum glaucum* (L.) Br.] core collection based on geographical distribution and quantitative traits. *Euphytica* 155:35–45. doi: [10.1007/s10681-006-9298-x](https://doi.org/10.1007/s10681-006-9298-x)
- Brown AHD. 1989. Core collections: A practical approach to genetic resources management. *Genome* 31:818–824. doi: [10.1139/g89-144](https://doi.org/10.1139/g89-144)
- Charmet G; Balfourier F. 1995. The use of geostatistics for sampling a core collection of perennial ryegrass populations. *Genetic Resources and Crop Evolution* 42:303–309. doi: [10.1007/BF02432134](https://doi.org/10.1007/BF02432134)
- Diwan N; McIntosh MS; Bauchan GR. 1995. Methods of developing a core collection of annual *Medicago* species. *Theoretical and Applied Genetics* 90:755–761. doi: [10.1007/BF00222008](https://doi.org/10.1007/BF00222008)
- Harlan JR; de Wet JMJ. 1969. Sources of variation in *Cynodon dactylon* (L.) Pers. *Crop Science* 9:744–748. doi: [10.2135/cropsci1969.0011183X000900060031x](https://doi.org/10.2135/cropsci1969.0011183X000900060031x)
- Harlan JR; de Wet JMJ; Rawal KM; Felder MR; Richardson WL. 1970. Cytogenetic studies in *Cynodon* L.C. Rich. (Gramineae). *Crop Science* 10:288–291. doi: [10.2135/cropsci1970.0011183X001000030023x](https://doi.org/10.2135/cropsci1970.0011183X001000030023x)
- Hu J; Zhu J; Xu HM. 2000. Methods of constructing core collections by stepwise clustering with three sampling strategies based on the genotypic values of crops. *Theoretical and Applied Genetics* 101:264–268. doi: [10.1007/s001220051478](https://doi.org/10.1007/s001220051478)
- Huang CQ; Liu GD; Bai CJ; Wang WQ; Tang J; Yu DG. 2012. A study on the morphological diversity of 475 accessions of *Cynodon dactylon*. *Acta Prataculturae Sinica* 21:33–42. (In Chinese.) [bit.ly/30g2kZ4](https://doi.org/10.1016/j.scienta.2013.07.039)
- Huang CQ; Liu GD; Bai CJ; Wang WQ. 2013. Genetic relationships of *Cynodon arcuatus* from different regions of China revealed by ISSR and SRAP markers. *Scientia Horticulturae* 162:172–180. doi: [10.1016/j.scienta.2013.07.039](https://doi.org/10.1016/j.scienta.2013.07.039)
- Huang CQ; Liu GD; Bai CJ; Wang WQ. 2014. Genetic analysis of 430 Chinese *Cynodon dactylon* accessions using sequence-related amplified polymorphism markers. *International Journal of Molecular Sciences* 15:19134–19146. doi: [10.3390/ijms151019134](https://doi.org/10.3390/ijms151019134)
- Jewell MC; Zhou Y; Loch DS; Godwin ID; Lambrides CJ. 2011. Maximizing genetic, morphological, and geographic diversity in a core collection of Australian bermudagrass. *Crop Science* 52:879–889. doi: [10.2135/cropsci2011.09.0497](https://doi.org/10.2135/cropsci2011.09.0497)
- Kenworthy KE; Martin DL; Taliaferro CM. 2007. Growth habit determination of genotypes of African bermudagrass. *Hortscience* 42:1513–1516. doi: [10.21273/hortsci.42.7.1513](https://doi.org/10.21273/hortsci.42.7.1513)
- Li CT; Shi CH; Wu JG; Xu HM; Zhang HY; Ren YL. 2004. Methods of developing core collections based on the predicted genotypic value of rice (*Oryza sativa* L.). *Theoretical and Applied Genetics* 108:1172–1176. doi: [10.1007/s00122-003-1536-1](https://doi.org/10.1007/s00122-003-1536-1)
- Li ZC; Zhang HL; Zeng YW; Yang ZY; Shen SQ; Sun CQ; Wang XK. 2002. Studies on sampling schemes for the establishment of core collection of rice landraces in Yunnan, China. *Genetic Resources and Crop Evolution* 49:67–74. doi: [10.1023/A:1013855216410](https://doi.org/10.1023/A:1013855216410)
- Lindroth RL; Osier TL; Barnhill HRH; Wood SA. 2002. Effects of genotype and nutrient availability on phytochemistry of trembling aspen (*Populus tremuloides* Michx.) during leaf senescence. *Biochemical Systematics and Ecology* 30:297–307. doi: [10.1016/S0305-1978\(01\)00088-6](https://doi.org/10.1016/S0305-1978(01)00088-6)
- Martínez IB; de la Cruz VM; Nelson MR; Bertin P. 2017. Establishment of a core collection of traditional Cuban *Theobroma cacao* plants for conservation and utilization purposes. *Plant Molecular Biology Reporter* 35:47–60. doi: [10.1007/s11105-016-0999-6](https://doi.org/10.1007/s11105-016-0999-6)
- Ortiz R; Ruia-Tapia EN; Mujica-Sanchez A. 1998. Sampling strategy for a core collection of Peruvian quinoa germplasm. *Theoretical and Applied Genetics* 96:475–483. doi: [10.1007/s001220050764](https://doi.org/10.1007/s001220050764)
- Reddy LJ; Upadhyaya HD; Gowda CLL; Singh S. 2005. Development of core collection in pigeonpea [*Cajanus cajan* (L.) Millspaugh] using geographic and qualitative morphological descriptors. *Genetic Resources and Crop Evolution* 52:1049–1056. doi: [10.1007/s10722-004-6152-7](https://doi.org/10.1007/s10722-004-6152-7)
- Rocheouste E. 1962. Studies on the biotypes of *Cynodon dactylon* (L.) Pers. I. Botanical investigations. *Weed Research* 2:1–23. doi: [10.1111/j.1365-3180.1962.tb00177.x](https://doi.org/10.1111/j.1365-3180.1962.tb00177.x)
- Sibson R. 1973. SLINK: An optimally efficient algorithm for the single-link cluster method. *The Computer Journal* 16:30–34. doi: [10.1093/comjnl/16.1.30](https://doi.org/10.1093/comjnl/16.1.30)

- Spagnoletti Zeuli PL; Qualset CO. 1987. Geographical diversity for quantitative spike characters in a world collection of durum wheat. *Crop Science* 27:235–241. doi: [10.2135/cropsci1987.0011183X002700020022x](https://doi.org/10.2135/cropsci1987.0011183X002700020022x)
- Spagnoletti Zeuli PL; Qualset CO. 1993. Evaluation of five strategies for obtaining a core subset from a large genetic resource collection of durum wheat. *Theoretical and Applied Genetics* 87:295–304. doi: [10.1007/BF01184915](https://doi.org/10.1007/BF01184915)
- Taliaferro CM. 1995. Diversity and vulnerability of bermuda turfgrass species. *Crop Science* 35:327–332. doi: [10.2135/cropsci1995.0011183X003500020006x](https://doi.org/10.2135/cropsci1995.0011183X003500020006x)
- Tiwari AK; Kumar G; Kadam GB; Saha TN. 2016. Optimization of ISSR-PCR system and assessing genetic diversity amongst turf grass (*Cynodon dactylon*) mutants. *Indian Journal of Agricultural Sciences* 87:251–256. krishi.icar.gov.in/jspui/handle/123456789/8774
- Wang ZY; Yuan XJ; Zheng YQ; Liu JX. 2009. Molecular identification and genetic analysis for 24 turf-type *Cynodon* cultivars by sequence-related amplified polymorphism markers. *Scientia Horticulturae* 122:461–467. doi: [10.1016/j.scienta.2009.05.031](https://doi.org/10.1016/j.scienta.2009.05.031)
- Ward JH Jr. 1963. Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association* 58:236–244. doi: [10.1080/01621459.1963.10500845](https://doi.org/10.1080/01621459.1963.10500845)
- Wofford DS; Baltensperger AA. 1985. Heritability estimates for turfgrass characteristics in bermudagrass. *Crop Science* 25:133–136. doi: [10.2135/cropsci1985.0011183X002500010033x](https://doi.org/10.2135/cropsci1985.0011183X002500010033x)
- Wu YQ; Taliaferro CM; Bai GH; Anderson MP. 2004. AFLP analysis of *Cynodon dactylon* (L.) Pers. var. *dactylon* genetic variation. *Genome* 47:689–696. doi: [10.1139/g04-032](https://doi.org/10.1139/g04-032)
- Wu YQ; Taliaferro CM; Bai GH; Martin DL; Anderson JA; Anderson MP; Edwards RM. 2006. Genetic analyses of Chinese *Cynodon* accessions by flow cytometry and AFLP markers. *Crop Science* 46:917–926. doi: [10.2135/cropsci2005.08.0256](https://doi.org/10.2135/cropsci2005.08.0256)
- Xu HM; Mei YJ; Hu J; Zhu J. 2006. Sampling a core collection of Island cotton (*Gossypium barbadense* L.) based on the genotypic values of fiber traits. *Genetic Resources and Crop Evolution* 53:515–521. doi: [10.1007/s10722-004-2032-4](https://doi.org/10.1007/s10722-004-2032-4)
- Zheng YQ; Guo Y; Fang SJ; Xu YN; Chen JB; Liu JX. 2014. Constructing pre-core collection of *Cynodon dactylon* based on phenotypic data. *Acta Prataculturae Sinica* 23:49–60. (In Chinese.) bit.ly/2DKJgup
- Zheng Y; Xu S; Liu J; Zhao Y; Liu J. 2017. Genetic diversity and population structure of Chinese natural bermudagrass [*Cynodon dactylon* (L.) Pers.] germplasm based on SRAP markers. *PLoS ONE* 12:e0177508. doi: [10.1371/journal.pone.0177508](https://doi.org/10.1371/journal.pone.0177508)

(Received for publication 13 August 2018; accepted 8 June 2020; published 30 September 2020)

© 2020



Tropical Grasslands-Forrajés Tropicales is an open-access journal published by *International Center for Tropical Agriculture (CIAT)*, in association with *Chinese Academy of Tropical Agricultural Sciences (CATAS)*. This work is licensed under the Creative Commons Attribution 4.0 International ([CC BY 4.0](https://creativecommons.org/licenses/by/4.0/)) license.