Bangladesh J. Bot. **53**(2): 319-325, 2024 (June) DOI: https://doi.org/10.3329/bjb.v53i2.74314

GENETIC DIVERSITY AND CONSTRUCTION OF CORE COLLECTION OF *STERCULIA* GERMPLASM

PEISI LUO, LIJUN HUANG, WENYAN LI, YOU WEI, JING ZHAO, FUCHANG ZHUO, JINGMEI TANG, FANGNAN KONG AND JING ZHOU*

Guangxi South Subtropical Agricultural Science Research Institute, Guangxi Academy of Agricultural sciences, Chongzuo 532200, Guangxi, China

Keywords: Sterculia, Germplasm, ISSR, Genetic diversity, Core collection

Abstract

Amplification of 107 *Sterculia* accessions with 12 ISSR primers resulted in 235 bands, the percentage of polymorphic sites was 100%, with the average number of effective alleles (*Ne*) of 1.2464, the average Nei's gene diversity index (*H*) of 0.1556, the average Shannon's information index (*I*) of 0.2563, and the average polymorphism information content (PIC) of 0.8594, indicating that the collected 107 accessions had high genetic diversity. The genetic diversity was unevenly distributed among *Sterculia* samples collected from different provinces, and the highest value was observed in samples from Guangdong Province, followed by those from Yunnan Province, and the lowest was from Guangxi Province. The core collection construction of the 107 *Sterculia* accessions was carried out by using the SCR and LDSS methods, respectively, and the results showed that the LDSS method was more suitable for *Sterculia* germplasm. Different core collections were constructed based on sampling proportions of 35, 30, 25, 20 and 15%, respectively, the *t*-test and principal coordinates analysis results showed that the genetic diversity of the original germplasm can be fully represented at a sampling proportion of 30%.

Introduction

The genus *Sterculia* belongs to the Malvaceae family, with approximately 300 species distributed in the tropical and subtropical regions of the world. There are 23 *Sterculia* species and 1 variety in China, which are distributed in Yunnan, Guizhou, Sichuan, Guangxi, and Guangdong provinces(Editorial Committee of Chinese Flora 2007). *Sterculia* plants have a wide range of medicinal and edible value（Li *et al.* 2012，Oppong *et al.* 2018）.As the strategies for protection, development, and utilization of tropical plant germplasms have been incorporated into the national strategy(China Rural Science and Technology Editorial Committee 2020), some research institutes in China have focused on the investigation, conservation, and evaluation of *Sterculia* germplasm (Li *et al.* 2021). However, due to a lack of knowledge on the genetic background of *Sterculia* germplasm, these works are largely aimless. In this study, inter-simple sequence repeat (ISSR) molecular markers were used to analyze the genetic diversity of *Sterculia* germplasm, and a core collection was constructed. This study provides reference for the investigation, collection, preservation, evaluation, development, and utilization of *Sterculia* germplasm.

Materials and Methods

The 107 *Sterculia* accessions investigated in this study were collected from Guangxi, Yunnan, and Guangdong provinces (Table 1), and were planted and preserved in the fruit tree germplasm nursery of the Guangxi South Subtropical Agricultural Science Research Institute.

*Author for correspondence: < <443868870@qq.com> >.

Table 1. Origin of 107 *Sterculia* accessions.

No.	Germ-plasm name	Species	Origin	No.	Germ-plasm name	Species	Origin	No.	Germ-plasm name	Species	Origin	No.	Germ-plasm name	Species	Origin
1	NN-1	Sn	1	28	TD-7	Sn	2	55	LZ-13	Sn	2	82	LZ-40	Sl	2
2	NN-2	Sn	1	29	TD-8	Sn	2	56	LZ-14	Sn	2	83	GZ-1	Sn	3
3	NN-3	Sn	1	30	TD-9	Sn	2	57	LZ-15	Sn	2	84	GZ-2	Sn	3
4	NN-4	Sn	1	31	TD-10	Sn	2	58	LZ-16	Sn	2	85	GZ-3	Sn	3
5	HX-1	Sn	1	32	DX-1	Sn	2	59	LZ-17	Sn	2	86	GZ-4	Sn	3
6	HX-2	Sn	1	33	DX-2	Sn	2	60	LZ-18	Sn	2	87	GZ-5	Sl	3
7	HX-3	Sn	1	34	DX-3	Sn	2	61	LZ-19	Sn	2	88	HZ-1	Sl	4
8	HX-4	Sn	1	35	DX-4	Sn	2	62	LZ-20	Sn	2	89	HZ-2	Sl	4
9	HX-5	Sn	1	36	DX-5	Sn	2	63	LZ-21	Sn	2	90	FS-1	Sn	5
10	HX-6	Sn	1	37	DX-6	Sn	2	64	LZ-22	Sn	2	91	FS-2	Sn	5
11	HX-7	Sn	1	38	DX-7	Sn	2	65	LZ-23	Sn	2	92	YF-1	Sl	6
12	HX-8	Sn	1	39	DX-8	Sn	2	66	LZ-24	Sn	2	93	MM-1	Sn	7
13	HX-9	Sl	1	40	DX-9	Sf	2	67	LZ-25	Sn	2	94	MM-2	Sn	7
14	HX-10	Sl	1	41	NM-1	Sn	2	68	LZ-26	Sn	2	95	ZJ-1	Sf	8
15	CZ-1	Sn	2	42	NM-2	Sn	2	69	LZ-27	Sn	2	96	XSBN-1	Sn	9
16	CZ-2	Sn	2	43	LZ-1	Sn	2	70	LZ-28	Sn	2	97	XSBN-2	Sn	9
17	CZ-3	Sn	2	44	LZ-2	Sn	2	71	LZ-29	Sn	2	98	XSBN-3	Sn	9
18	CZ-4	Sn	2	45	LZ-3	Sn	2	72	LZ-30	Sn	2	99	HH-1	Sn	10
19	CZ-5	Sn	2	46	LZ-4	Sn	2	73	LZ-31	Sn	2	100	HH-2	Sn	10
20	PX-1	Sn	2	47	LZ-5	Sn	2	74	LZ-32	Sn	2	101	HH-3	Sl	10
21	PX-2	Sn	2	48	LZ-6	Sn	2	75	LZ-33	Sn	2	102	HH-4	Sn	10
22	TD-1	Sn	2	49	LZ-7	Sn	2	76	LZ-34	Sn	2	103	WS-1	Sn	11
23	TD-2	Sn	2	50	LZ-8	Sn	2	77	LZ-35	Sn	2	104	WS-2	Sn	11
24	TD-3	Sn	2	51	LZ-9	Sn	2	78	LZ-36	Sl	2	105	WS-3	Sn	11
25	TD-4	Sn	2	52	LZ-10	Sn	2	79	LZ-37	Se	2	106	WS-4	Sn	11
26	TD-5	Sn	2	53	LZ-11	Sn	2	80	LZ-38	Se	2	107	WS-5	Sn	11
27	TD-6	Sn	2	54	LZ-12	Sn	2	81	LZ-39	Se	2				

For the species:Sn=*Sterculia nobilis* Smith,Sl=*Sterculialanceolata* Cav.,Sf=*Sterculia foetida* Linn.,Se=*Sterculia nobilis* Smith.For the origin:1=Nanning City, Guangxi Province,2=Chongzuo City, Guangxi Province,3=Guangzhou City, Guangdong Province,4=Huizhou City, Guangdong Province,5=Foshan City, Guangdong Province,6=Yunfu City, Guangdong Province,7=Maoming City, Guangdong Province,8=Zhanjiang City, Guangdong Province,9=Xishuangbanna Prefecture, Yunnan Province,10=Honghe Prefecture, Yunnan Province,11=Wenshan Prefecture, Yunnan Province.

DNA was extracted using a plant genomic DNA extraction kit (Takara Bio Inc.Beijing, China), and the DNA integrity and concentration were detected via 1.2 % agarose gel electrophoresis at 120 V for 30 min. Then, the DNA sample was diluted to 50 ng/μl and stored at -20°C for later use.

Seventy-five ISSR primers were randomly selected from the 100 ISSR primers published by University of British Columbia (UBC, http://www.biotech.ubc.ca/services/naps/primers/Primers.pdf), and were screened using three *Sterculia* samples showing relatively greater phenotypic differences. 12 ISSR primers with good repeatability and high polymorphism (Table 2) were selected and used for amplification of all the *Sterculia* samples. The total volume of the ISSR-PCR system was 25 μl, including 0.13 μl of Taq enzyme (5 U/μl), 0.5 μl of dNTP Mixture (10 mmol/l), 4.5 μl of 10 × PCR Buffer (Mg^{2+} plus), 1 μl of each primer (10 μM), 5 μl of DNA template (10 ng/μl), and 13.87 μl of ddH_2O. The PCR reaction program was as follows: pre-denaturation at 98°C for 3 min; 35 cycles of denaturation at 98°C for 10 s, annealing at 56°C for 10 s, extension at 72°C for 1 min and 30 s; and extension at 72°C for 5 min. The PCR products were detected by electrophoresis on an 8% SDS-PAGE gel, followed by silver staining.

Table 2 Amplification results and polymorphism information of 12 ISSR primers.

Primer	Total No. of bands	No. of polymorphic bands	PPB (%)	*Na*	*Ne*	*H*	*I*	PIC
UBC-808	19	19	100	2.0000	1.1857	0.1220	0.2110	0.8374
UBC-816	14	14	100	2.0000	1.2060	0.1307	0.2213	0.7717
UBC-826	30	30	100	2.0000	1.1363	0.0996	0.1829	0.9110
UBC-846	19	19	100	2.0000	1.1810	0.1094	0.1817	0.7957
UBC-848	19	19	100	2.0000	1.1414	0.1106	0.2022	0.8978
UBC-866	13	13	100	2.0000	1.3735	0.2290	0.3567	0.8521
UBC-870	11	11	100	2.0000	1.3182	0.1924	0.3086	0.7934
UBC-886	15	15	100	2.0000	1.3804	0.2244	0.3464	0.8760
UBC-887	26	26	100	2.0000	1.2328	0.1522	0.2542	0.9057
UBC-888	22	22	100	2.0000	1.2481	0.1432	0.2301	0.8546
UBC-890	27	27	100	2.0000	1.2336	0.1545	0.2594	0.9170
UBC-891	20	20	100	2.0000	1.3200	0.1991	0.3210	0.9006
Mean	19.83	19.83	100%	2.0000	1.2464	0.1556	0.2563	0.8594

The loci showing a DNA band were marked as 1, while those without a DNA band were denoted as 0. The data set was converted into a mathematical matrix used by PopGene32 software to calculate the percentage of polymorphic bands (PPB), observed number of alleles (*Na*), effective number of alleles (*Ne*), Shannon's information index (*I*), Nei's gene diversity (*H*), total genetic diversity (Ht), intra-population genetic diversity (Hs), Gene flow (Nm, and the relative magnitude of genetic differentiation among population (Gst = Ht-Hs/Ht)(Nei and Li 1979). The formula used for calculating the polymorphic index content (PIC) was as follows:$PIC=1-\sum_{j=1}^{n} P_{ij}^{2}$, PIC represents the PIC value of locus *i*, and P_{ij} represents the occurrence frequency of the *j*-th allele of locus *i* (Smith *et al.* 1997).The genetic similarity among accessions was evaluated by calculating the simple matching (SM) similarity coefficient, and the unweighted pair-group method with arithmetic means (UPGMA) algorithm was used for

cluster analysis. Principal coordinates analysis (PCoA) was carried out using NTSYS-pc 2.10 software (Rohlf 2000).

According to the UPGMA clustering results, the sampling proportions were set at 35, 30, 25, 20, and 15%, stepwise clusters with random sampling strategy (SCR) (Hu *et al.* 2000) and least distance stepwise sampling (LDSS) (Wang *et al.* 2007) were comprehensively compared to determine the optimal strategy for core collection construction.

Results and Discussion

Twelve ISSR primers were used to amplify the 107 *Sterculia* accessions, and a total of 235 bands were obtained, with an average of 19.83 bands per primer. The percentage of polymorphic bands of all primers was 100% (Table 2), among which primer UBC -826 produced the largest number of bands (30), while primer UBC-870 produced the least bands (11). For a single primer, the *Ne* ranged from 1.1363 to 1.3804, with an average of 1.2464; the *H* ranged from 0.0996 to 0.2290, with an average of 0.1556; the *I* ranged from 0.1817 to 0.3567, with an average of 0.2563; and the PIC ranged from 0.7717 to 0.9170, with an average of 0.8594. These results indicated that the selected primers harbor a high level of polymorphism among *Sterculia* accessions.

Genetic diversity analysis was carried out on *Sterculia* populations from the three provinces (Table 3). The results showed that the *Ne*, *H*, and *I* of the *Sterculia* population from Guangdong Province were the highest (1.3413, 0.2200, and 0.3584, respectively), followed by those from Yunnan Province (1.3355, 0.2175, and 0.3541, respectively), and the lowest values were found in Guangxi Province (1.2579, 0.1626, and 0.2666, respectively).

Table 3 Genetic diversity among three populations of *Sterculia.*

Population origin	No. of polymorphic bands	PPB (%)	*Na*	*Ne*	*H*	*I*
Guangxi	210	100%	2.0000	1.2579	0.1626	0.2666
Guangdong	150	100%	2.0000	1.3413	0.2200	0.3584
Yunnan	141	95.92 %	1.9592	1.3355	0.2175	0.3541
Mean	167	98.64%	1.9864	1.3116	0.2001	0.3264

The total genetic diversity (Ht) of the three *Sterculia* populations was 0.2799, the genetic diversity within the population (Hs) was 0.2582, and the genetic differentiation index between populations (Gst) was 0.0776. Approximately 7.76% of the genetic variation was between *Sterculia* populations, while 92.24% of the genetic variation was within the population, indicating the genetic differentiation within the population was much greater than that between populations. Moreover, the gene flow (Nm) was 5.9397, suggesting that there was certain gene exchange between different *Sterculia* populations.

Five core collections of *Sterculia* germplasm were constructed by the SCR and LDSS methods, and the optimal sampling method for constructing the core collection was determined by comparing the coefficient of variation (CV) of *Ne*, *H*, and *I*. The results showed that although the CV of *Ne* obtained using the LDSS method was slightly larger than that using the SCR method, the CVs of *H* and *I* using the LDSS method were smaller than those using the SCR method (Table 4), indicating that the LDSS method is more suitable for constructing the core collection of *Sterculia* germplasm compared to the SCR method.

Five candidate core collections were constructed with various sampling proportions (35, 30, 25, 20 and 15%) of the original germplasm using the LDSS method. Compared with the original germplasm, the number of polymorphic loci of each candidate core collection decreased with the decreasing sampling proportions, but the percentage of polymorphic bands did not change (Table 5). The *Ne*, *H*, and *I* of each candidate core collection were higher than those of the original germplasm.

Table 4 Evaluation for different sampling strategies base on genetic diversity.

Sampling proportion (%)	SCR			LDSS		
	Ne	*H*	*I*	*Ne*	*H*	*I*
35	1.2571	0.1673	0.2799	1.2696	0.1758	0.2934
30	1.2661	0.1742	0.2913	1.2806	0.1819	0.3019
25	1.2658	0.1755	0.2947	1.2805	0.1839	0.3065
20	1.2738	0.1830	0.3076	1.2853	0.1874	0.3122
15	1.2866	0.1928	0.3232	1.3067	0.2018	0.3339
AV	1.2699	0.1786	0.2993	1.2845	0.1862	0.3096
STDV	0.0111	0.0097	0.0166	0.0137	0.0097	0.0152
CV	0.0087	0.0544	0.0554	0.0106	0.0521	0.0492

Table 5 Comparison of genetic parameters of candidate core collection.

Sampling proportion (%)	Sample No.	No.r of polymorphic bands	PPB (%)	*Ne*	*H*	*I*
100	107	235	100	1.2464	0.1556	0.2563
35	37	217	100	1.2696(0.4682)	0.1758(0.2418)	0.2934(0.1227)
30	32	211	100	1.2806(0.2982)	0.1819(0.1388)	0.3019(0.0655)
25	27	208	100	1.2805(0.2911)	0.1839(0.1063)	0.3065(0.0405*)
20	21	206	100	1.2853(0.2291)	0.1874(0.0717)	0.3122(0.0236*)
15	16	198	100	1.3067(0.0646)	0.2018(0.0107*)	0.3339(0.0023*)

The data in parentheses are the p-values of the t-test; * indicates significantly different at $p < 0.05$.

At the sampling proportion of 35 and 30%, the genetic diversity indexes of the core collections were not significantly different from those of the original germplasm. When the sampling proportion was reduced to 25 or 20%, the *I* value of these two candidate core collections was significantly different from that of the original germplasm. When the sampling proportion was decreased to 15%, the *H* and *I* values of the candidate core collection were significantly different from those of the original germplasm. In summary, the core collection constructed with a sampling proportion of 30% can represent the genetic diversity of the original germplasm. A total of 32 *Sterculia* accessions were present in the core *Sterculia* collection, including 24 *Sterculia nobilis* Smith samples, five *Sterculia lanceolata* Cav. samples, two *Sterculia foetida* Linn. samples, and one *Sterculia euosma* W. W. Smith sample.The results of PCoA showed that the 32 core germplasm resources were distributed in a scattered pattern among the 107 total samples (Fig 1)，the distribution pattern of the core collection was very similar to that of the initial collection, and more peripheral individuals were selected, providing further evidence of the representativeness of the proposed core.

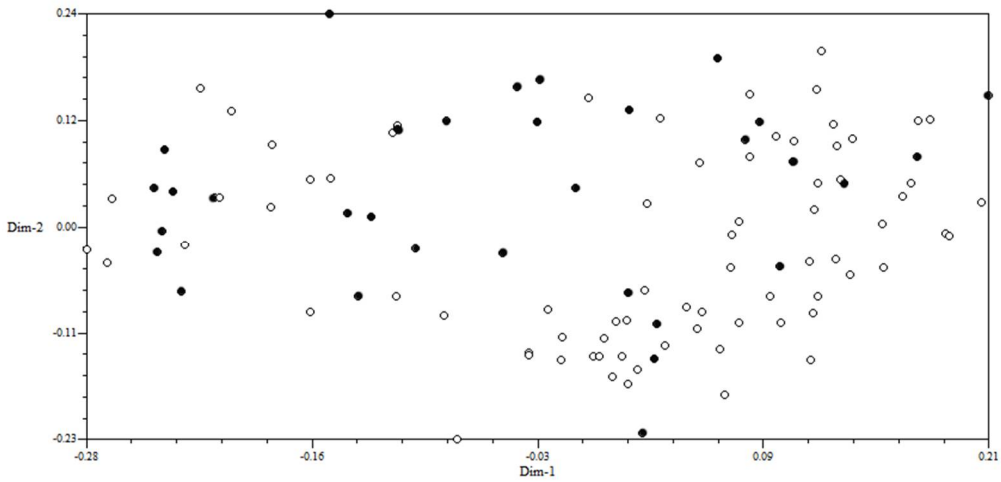

Fig. 1. Principal coordinates analysis (PCoA) for the core and reserved collections at 30% sampling proportion.The 32 accessions of the core collection were represented by black circles and the 75 accessions of the reserved collection were represented by open circles.

Genetic diversity is the basis of species evolution, selection, and recombination,and is the core and fundamental of biodiversity. In this study, 12 out of 75 ISSR primers were screened out and applied to 107 *Sterculia* accessions from three provinces, i.e. , Guangxi, Guangdong, and Yunnan. A total of 235 polymorphic bands were obtained, with the PPB, *Ne*, *H*, and *I* values of 100%, 1.2464, 0.1556 and 0.2563, respectively, which indicated high genetic diversity among the collected *Sterculia* accessions. Comparisons of the genetic diversity among *Sterculia* populations from the three provinces showed that the genetic diversity of the *Sterculia* population from Guangdong Province was the highest, followed by that from Yunnan Province, and the lowest was found in that from Guangxi Province. This indicated that the adaptability of the *Sterculia* population from Guangdong and Yunnan to the environment was higher than that from Guangxi.

Sampling strategy has always been the focus of core collection construction because it determines which accessions are eligible for the core collection. The LDSS method proposed by Wang in 2007 can eliminate the influence of different clustering methods on the core collections(Wang *et al.* 2007), thus it has been widely used in the core collection construction of *Elymus sibiricus* L. (Yan *et al.* 2017), walnut(Wu *et al.* 2020), and *Akebia trifoliata* (Thunb.) Koidz.(Zhang *et al.* 2021). In this study, the SCR and LDSS methods were used to construct the core collection of *Sterculia* germplasm, and the results showed that the LDSS method was more suitable.

The core collection can represent the genetic diversity of the original germplasm to the greatest extent with the least accessions, and the key to core collection construction is to set a reasonable sampling proportion.Li et al. (2003) believed that the sampling proportion should be determined according to the genetic structure and population scale of a species. Species with more accessions may have a smaller sampling proportion, while species with less accessions may have a relatively higher sampling proportion. The population scale of fruit trees is not as large as that of field crops, thus the sampling proportion for core collection construction is usually higher in fruit trees than in other crops(Lang *et al.* 2016). Based on 107 *Sterculia* accessions, the candidate core collections were constructed based on sampling proportions of 35, 30, 25, 20, and 15%. The *t*-test and PCoA results showed that no significant difference in genetic diversity between the original

germplasm and the core collection constructed at a sampling proportion of 30%, indicating that the constructed core collection could fully represent the genetic diversity of the original germplasm.

Acknowledgments

The study was supported by the National Natural Science Foundation of China (31860533), the Special Funds for Guiding Local Scientific and Technological Development by the Central Government (GuikeZY19183010), and the Guangxi Academy of Agricultural Sciences Special Funding Project for Basic Scientific Research Business (Guinongke2021YT163 and Guinongke2022YM23).

References

China Rural Science and Technology Editorial Committee 2020. Developing high-efficiency agriculture with tropical characteristics is an important subject related to national economy and people's livelihood. China Rural Science and Technology. 7: 3.

Editorial Committee of Chinese Flora 2007. Flora of China. Science Press, Beijing. pp. 303-305.

Hu J, Zhu J and Xu HM 2000. Methods of constructing core collections by stepwise clustering with three sampling strategies based on the genotypic values of crops. Theor. Appl. Genet. **101**: 264-268.

Lang BB, Huang CH, Zhu B, Xie M, Zhang WB, Zhong M, Qu XY, Tao JJ and Xu XX 2016. Study on the method of constructing a primary core collection of Jiangxi wild *Actinidia eriantha* based on fruit traits. Journal of Fruit Science. **33**(7): 794-803.

Li WY, Zhou J, Zhang XF, Zhou CX, Huang LJ, Lu MY, Zhao J, Zhuo FC, Luo PS and Wei Y 2021. Comprehensive evaluation of *Sterculia nobilis* Smith germplasm resources based on AHP-FCE mathematical model and DTOPSIS method. J. South. Agric. **52**(6): 1557-1567.

Li ZC, Zhang HL, Cao YS, Qiu ZE, Wei XH, Tang SX, Yu P and Wang XK 2003. Studies on the sampling strategy for primary core collection of Chinese ingenious rice. Acta Agron. Sin. **29**: 20-24.

Ma KP 1993. On the concept of biodiversity. Biodiversity Science. **1**(1): 20-22.

Nei M and Li WH 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. PNAS. **76**: 5269-5273.

Oppong MB, Li Y, Banahene PO, Fang SM and Qiu F 2018. Ethnopharmacology phytochemistry and pharmacology of *Sterculia lychnophora* Hance (Pangdahai) . Chinese Journal of Natural Medicines. **16**(10): 721-731.

Rohlf FJ 2000. NTSYS-pc version 2.1, Numerical Taxonomy and Multivariate Analysis System. Exeter Software, Setauket, New York.

Smith JSC, Chin ECL, Shu H, Smith OS, Wall SJ, Senior ML, Mitchell SE, Kresovich S and Ziegle J 1997. An evaluation of theutility of SSR loci as molecular marker in maize (*Zea mays* L.): comparisons with data from RFLPs and pedigree. Theor. Appl. Genet. **95**: 163-173.

Wang JC, Hu J, Xu HM and Zhang S 2007. A strategy on constructing core collections by least distance stepwise sampling. Theor. Appl. Genet. **115**: 1-8.

Wu T, Chen SY, Xiao LJ, Ning DL, Pan L, He N and Zhu YF 2020. SSR markers based genetic diversity analysis and core collection construction of Walnut germplasm in Yunnan Province. J. Plant Genet. Resour. **21**(03): 767-774.

Yan WH, Ma YB, Zhang JR, Wang K and Meng QW 2017.The analysis of genetic diversity and the construction of core collection for *Elymus sibiricus* L. Prataculture Animal Husbandry.Journal of Grassland and Forage Science.**1**:1-12

Zhang Z, Yang Q, Niu YN, Zhang YX, Dong S, Zhang WW and Wang ZZ 2021. Diversity analysis and establishment of core collection among *Akebia trifoliata* (Thunb.) Koidz. in Qinba mountain area of China using ISSR and SRAP markers. Genet. Resour. Crop Evol. **68**:1085-1102.

(Manuscript received on 01 September, 2022, revised on 23 June, 2024)