한작지(Korean J. Crop Sci.), 66(4): 443~451(2021)
DOI : https://doi.org/10.7740/kjcs.2021.66.4.443

ISSN 0252-9777(Print)
ISSN 2287-8432(Online)

Original Research Article

Development of EST-SSRs and Assessment of Genetic Diversity in Germplasm of the Finger Millet, *Eleusine coracana* (L.) Gaertn.

Xiaohan Wang[1], Myung Chul Lee[1], Yu-Mi Choi[1], Seong-Hoon Kim[1], Seahee Han[2], Kebede Taye Desta[1,3], Hye-myeong Yoon[1], Yoonjung Lee[1], Miae Oh[1], Jung Yoon Yi[1], Myoung-Jae Shin[1,†], and Kyung-Min Kim[4,†]

ABSTRACT Finger millet (*Eleusine coracana*) is widely cultivated in tropical regions worldwide owing to its high nutritional value. Finger millet is more tolerant against biotic and abiotic stresses such as pests, drought, and salt than other millet crops; therefore, it was proposed as a candidate crop to adapt to climate change in Korea. In 2019, we used expressed sequence tag simple sequence repeat (EST-SSR) markers to evaluate the genetic diversity and structure of 102 finger millet accessions from two geographical regions (Africa and South Asia) to identify appropriate accessions and enhance crop diversity in Korea. In total, 40 primers produced 116 alleles, ranging in size from 135 to 457 bp, with a mean polymorphism information content (PIC) of 0.18225. Polymorphism was detected among the 40 primers, and 13 primers were found to have PIC values > 0.3. Principal coordinate and phylogenetic analyses, based on the combined data of both markers, grouped the finger millet accessions according to their respective collection areas.Therefore, the 102 accessions were classified into two groups, one from Asia and the other from Africa. We have conducted an in-depth study on the finger millet landrace pedigree. By sorting out and using the molecular characteristics of each pedigree, it will be useful for the management and accession identification of the plant resource. The novel SSR markers developed in this study will aid in future genetic analyses of *E. coracana*.

Keywords : phylogenetic analysis, population structure, selfing plants

Millet is a small grain crop that is mainly cultivated in semi-arid tropical regions of Africa and South Asia. The millet grains are used for human food and the stalks are normally used as animal fodder in these cultivation regions (Baryeh, 2002). Traditionally grown millet species include pearl millet, finger millet [*Eleusine coracana* (L.) Gaertn], foxtail millet [*Setaria italica* (L.) P. Beauvois], Japanese barnyard millet [*Echinochloa esculneta* (A. Braun) H. Scholz], Indian Barnyard millet [*Echinochloa frumetacea* (Roxb.) Link], kodo millet [*Paspalum scrobiculatum* L.], little millet [*Panicum sumatrense* Roth ex Roem. & Schult.] proso millet [*Panicum miliaceum* L.] and Tef [*Eragrostis tef* (Zucc.) Trotter] (Goron & Raizada, 2015). Finger millet grain is more nutritious than other cereals, providing large amounts of protein, minerals, and vitamins. Finger millet has a high protein content and has been reported to possess a fairly high biological value due to particularly rich in methionine and other essential amino acids. Finger millet also contains fat (1.29 percent) mainly composed of two polyunsaturated fatty acids as linoleic acid and α-linolenic acid (Fernandez *et al.*, 2003; Pandey & Kumar, 2005). Furthermore, finger millet has a wide adaptability to diverse environment such as drought area. The proso millet and foxtail millet are cultivated as a major millet crop in Korea due to its good tolerance to abiotic stresses. For this reason, finger millet is used as a candidate crop for coping with severe climate

[1)]Junior Scientist, National Agrobiodiversity Center, National Institute of Agricultural Sciences, Rural Development Administration, Jeonju 54874, Korea
[2)]Associate Researcher, Honam National Institute of Biological Resources, Mokpo 58762, Korea
[3)]Professor, Department of Applied Chemistry, Adama Science and Technology University, Adama 1888, Ethiopia
[4)]Professor, School of Applied Biosciences, College of Agriculture and Life Science, Kyungpook National University, Daegu 41566, Korea

[†]Corresponding author: ; Myoung-Jae Shin; (Phone) +82-63-238-4891; (E-mail) smj1204@korea.kr
Kyung-Min Kim; (Phone) +82-53-950-5711; (E-mail) kkm@knu.ac.kr
<Received 9 October, 2021; Revised 1 November, 2021; Accepted 17 November, 2021>

change in Korea.

The selection of target trait using different morphological and biochemical traits or genetic diversity based on molecular markers is essential in finger millet breeding approach. Molecular markers are often used to analyze the genetic diversity and genetic differentiation of plant populations (Govindaraj *et al*., 2015). Molecular markers are one of the tools for precise, efficient and cost effective germplasm characterization and management through the detection of DNA polymorphisms in most areas of plant genetics. The advantages of DNA markers are well documented (Powell *et al.*, 1996) and these include: high information content, reproducibility and locus specificity. Such markers that have been used in finger millet include Inter Simple Sequence Repeats (ISSRs) (Gupta *et al.*, 2010), Amplified Fragment Length Polymorphisms (AFLPs) and SSRs (Dida *et al.*, 2007; Arya *et al.*, 2009), EST-SSRs and SNPs (Kumar *et al.*, 2016; Gimode *et al.*, 2016). Among the DNA markers, expressed sequence tags (ESTs)-SSR are very useful in species diversity research due to the primer target sequences residing in the expressed genes which are relatively well conserve.

Due to the complexity of the genome and the economically underdeveloped areas as the main growing areas, molecular works of finger millet have been little attention from the research community and are less information of genomic and EST sequences than other millets such as sorgum, foxtail millet and proso millet. There has been tremendous increase in availability of ESTs and it has been confirmed that EST-SSR based markers can be easily developed from existing database. (Gupta *et al.*, 2010). EST-derived SSRs are probably tightly linked with functional genes that may control certain agronomic traits. They are thus a potential tool for marker-assisted selection (MAS) breeding and genetic diversity of genetic resources.

The objective of this study was to assess the level of genetic diversity and determine the population structure on finger millet accessions that were introduced from Asian countries and Kenya using EST-derived SSRs markers. And to characterize the possible relationships of genetic classification with the germplasm origin or collection sites of these accessions.

MATERIALS AND METHODS

Plant materials and DNA extraction

In this study, we obtained seeds for a total of 102 *E. coracana* landraces, including 42 genotypes from Asian countries provided by the National Agrobiodiversity Center of the National Institute of Agricultural Sciences (NAS), Rural Development Administration (RDA), Korea, and 60 genotypes from the Kenya Agricultural and Livestock Research Organization (KALRO), Kenya. The details of each genotype are provided in Table 1 (Table 1). Six young leaves were collected from each individual at four-leaf stage after 15 days of sowing and genomic DNA was extracted by Genomic Plus DNA Prep kit according to the manufacturer's instruction (Inclone Co, Korea). The DNA quality was first checked on 1% agarose gel electrophoresis and the quantification and concentrations of DNA were estimated using EpochTM microplate spectrophotometers (BioTek Instruments Inc., USA). The extracted DNA of 102 finger millet genotypes was adjusted to a concentration of 50 ng/μL and stored at -20°C until further use.

Database mining and primer design

Database mining and primer design used data deposited in the NCBI (as on March 17, 2019) to derive finger millet sequences. A total of 1,956 EST sequences were retrieved from NCBI (http://www.ncbi.nlm.gov/nuccore/?term=finger+millet+EST) to search simple sequence repeats. EST-SSRs assembling using SeqMan DNA Star Lasergene version 7.1 (DNASTAR Inc, Madison, WI). The microsatellite primer pairs identification using SSR locator v1 software (Da Maia *et al*., 2008) and simple sequence repeat identification tool (SSRIT) (http://www.gramene.org/gramene/searches/ssrtool). The amplified product size was controlled within the range of 100 to 300 bp. For designing primers, optimum primer length was 20 mer (ranged 18-25 mer), optimum annealing temperature was 60°C (range was 55-62°C) and the rest of parameters was default value. In order to test whether the 40 EST-SSRs used in the study are concentrated in a limited position, linkage disequilibrium analysis (LD) was performed on 40 markers using TASSEL v5.2.64.

Table 1. List of Finger millet used in the study with their origin and accession number.

Sample No.	Accession No.	Origin[a]	Sample No.	Accession No.	Origin
1	IT109527	NEP	52	GBK 000514	UGA
2	IT136281	NEP	53	GBK 000527	UGA
3	IT136283	NEP	54	GBK 000543	ZAM
4	IT136284	NEP	55	GBK 000545	ZAM
5	IT136286	NEP	56	GBK 000561	ZAM
6	IT136288	NEP	57	GBK 000563	ZIM
7	IT200202	NEP	58	GBK 000599	ZIM
8	IT200204	NEP	59	GBK 000605	ZIM
9	IT200206	NEP	60	GBK 000608	ZIM
10	IT208543	NEP	61	GBK 000609	KEN
11	IT215793	NEP	62	GBK 000620	KEN
12	IT219072	NEP	63	GBK 000624	KEN
13	IT235490	NEP	64	GBK 000638	KEN
14	IT235528	NEP	65	GBK 000648	KEN
15	IT288847	NEP	66	GBK 029632	KEN
16	IT288848	NEP	67	GBK 029683	KEN
17	IT302265	NEP	68	GBK 029691	KEN
18	IT235713	NEP	69	GBK 029696	KEN
19	IT251839	NEP	70	GBK 029721	KEN
20	IT153442	IND	71	GBK 029744	KEN
21	IT153444	IND	72	GBK 029753	KEN
22	IT153447	IND	73	GBK 029761	KEN
23	IT153448	IND	74	GBK 029779	KEN
24	IT153449	IND	75	GBK 029782	KEN
25	IT153451	IND	76	GBK 029783	KEN
26	IT153453	IND	77	GBK 029790	KEN
27	IT153456	IND	78	GBK 029808	KEN
28	IT153459	IND	79	GBK 029848	KEN
29	IT270217	IND	80	GBK 029864	KEN
30	IT303961	IND	81	GBK 036782	KEN
31	IT303964	IND	82	GBK 036827	KEN
32	IT303966	IND	83	GBK 037025	KEN
33	IT303967	IND	84	GBK 037055	KEN
34	IT303968	IND	85	GBK 037143	KEN
35	IT303969	IND	86	GBK 039132	KEN
36	IT303970	IND	87	GBK 039146	KEN
37	IT303971	IND	88	GBK 039203	KEN
38	IT303972	IND	89	GBK 039240	KEN
39	IT235685	IND	90	GBK 039296	KEN
40	IT235687	IND	91	GBK 040471	KEN
41	IT262503	CHN	92	GBK 043019	KEN
42	IT235689	PAK	93	GBK 043055	KEN
43	GBK 000368	UGA	94	GBK 043057	KEN
44	GBK 000420	UGA	95	GBK 043067	KEN
45	GBK 000453	UGA	96	GBK 043120	KEN
46	GBK 000468	UGA	97	GBK 043122	KEN
47	GBK 000470	UGA	98	GBK 043128	KEN
48	GBK 000471	UGA	99	GBK 043158	KEN
49	GBK 000483	UGA	100	GBK 043159	KEN
50	GBK 000492	UGA	101	GBK 043165	KEN
51	GBK 000504	UGA	102	GBK 043910	KEN

[a]NEP, Nepal; IND, India; CHN, China; PAK, Pakistan; UGA, Uganda; ZAM, Zambia; ZIM, Zimbabwe; KEN, Kenya.

Table 2. Description of forty EST-SSR markers developed in this study and genetic parameters among 102 finger millet genotypes.

Primer name	Forward Primer (5'-3')	Reverse Primer (3'-5')	GenBank No.	Repeat motif	Tm (°C)	N_A[a]	PIC[b]	H_E[c]
FM-EST 20	TGAGAGTTTGTGATGATAGTC	GGTGAAGAAAAATAAAACAA	EB187410.1	(A)28-(AG)3-(AT)3-(TTC)8	45	3	0.22	0.24
FM-EST 25	CGTCGTAGAGAACTACCTC	TCGACAGAGAATAAGATCAC	FD661924.1	(TTC)8	52	2	0.09	0.1
FM-EST 27	CTCTGTTTCTTGATTTTGTC	ATCTCTACTCCGACCTAAAC	FD661922.1	(GCG)7	50	3	0.37	0.49
FM-EST 33	ACCAAGAACCCTAGTCTCT	GACATCTCCTTCAACCTC	FD661919.1	(TCTCC)4	55	3	0.30	0.37
FM-EST 41	GGACTGGCTAGGGTTC	TCGGTAGATCTTAGAAGACA	FD661913.1	(AGCCCA)3	59	4	0.30	0.34
FM-EST 46	ATCTTAGCTTTCCTGTCTTT	GACTCCCTGTTCGAGAG	FD661908.1	(AGA)6	53	2	0.06	0.06
FM-EST 54	AGTTTGGTGTGTGTTATAGG	GTGAAACATGCATGAACTA	FD661904.1	(GT)9	48	2	0.37	0.38
FM-EST 57	TTACTAGGCTACTACGATGG	GCAGGTAGCTCTTGTACTT	FD661902.1	(GGC)6	53	3	0.06	0.06
FM-EST 62	TGTCGGACTACGGAAC	TATCCTTGATCTTGTCTTTG	FD661898.1	(GT)4-(TA)3	52	2	0.01	0.01
FM-EST 69	CTGGTTCGTCGACTACTAC	TATGCACTTACATGGTTACA	FD661888.1	(GA)3	49	2	0.03	0.03
FM-EST 81	ACTGATTCGTCTTATGGAAT	GATCATAGGTGAAACAGAAA	FD661876.1	(A)21	49	2	0.37	0.48
FM-EST 97	GTAATTAGCTCAGTGCCTTA	CGATTAAAGGTACGTAAAGA	FD661866.1	(A)60	57	3	0.40	0.04
FM-EST 109	TTATTTAAAATCCTGGTTTG	CTTTTAAATTTCGGGAAC	FD661857.1	(GAGGGG)2-(C)19	60	3	0.33	0.39
FM-EST 125	TATTACCATTGGAATCTTTC	CAGTACCACATCCTTTGAG	FD661849.1	(TGTTGC)2-(CTGTTA)2	55	2	0.01	0.01
FM-EST 126	AAACCCTCGACGACTC	GTACATTTCCAGGTCACAT	CX265222.1	(CCGCAG)3-(TCGGCG)2	52	2	0.11	0.11
FM-EST 130	GGTACTAGTTCTAGATCGC	AGAAACTGTAATGTGGATTG	CX265173.1	(T)24-(ACTGCT)2	50	3	0.09	0.09
FM-EST 132	TTCCGCTCCGACAAG	GGTTGAAGACCAGGTACA	CX265159.1	(CAC)5-(CCA)4-(GCT)4	55	3	0.31	0.37
FM-EST 137	CTCTCCTTCACCATCG	GACTTATTCGGACTGGAC	CX265101.1	(TAGGGT)2-(GGCGAC)2	54	3	0.39	0.47
FM-EST 138	CCACTACTACTTGCTGTGAC	AGAGTTGGACGAGGAAT	CX265096.1	(CGGCTG)2-(CCTCCG)2	56	3	0.13	0.14
FM-EST 145	GGCGTGTTTATAATTTGTAT	TTTTATATAACGAGGAATCG	CX265023.1	(GTGA)3-(CATATG)2	53	4	0.05	0.05
FM-EST 146	AACTTTTCCTCCCTCCT	GACCCTCATGAACAGTATT	CX265019.1	(CCT)4-(TCCG)3	61	2	0.06	0.06
FM-EST 150	CAGATAGTCCAACAACAACT	ACCGAGTTCATTATGTTAGA	CX264942.1	(CAA)13	55	3	0.27	0.31
FM-EST 155	GTACTGAAGAATAGCAGCAG	CCTTGTCGGTCAGGTA	CX264904.1	(GGT)4-(TGGCGG)2	53	4	0.21	0.23
FM-EST 157	GAGAGGAGAAGGAGGG	TAAAACCCTAAAAGACCTTC	CX264880.1	(TCGCGG)2-(GCG)4	52	2	0.06	0.06
FM-EST 159	GTGGGTCCATTTCTTTAC	ACTCCAACTCCTCTTCTTC	CX264861.1	(CGG)8	56	4	0.08	0.08
FM-EST 160	ATCGTCATCAAGCTAAGATA	GCTGACCCTGAACGC	CX264838.1	(ACGCCG)2-(CGGCGC)2	53	4	0.36	0.39
FM-EST 170	GATCGTGTTGATGTTCG	CTGAACAGCAGGAACAG	CX264732.1	(GCTGGC)2-(CCTGGC)2	56	4	0.18	0.19
FM-EST 172	CTAATCAATTAAGCTCATCG	GGACTTGAGGCAGTTG	CX264730.1	(ACGCCG)2-(CGGCGC)2	64	4	0.13	0.14
FM-EST 199	AGTGGTGCTATGGTGG	GTCGTACTTGAACACTAGGA	EB187498.1	(CTC)4-(CAACAG)2	51	3	0.04	0.04
FM-EST 200	CATCTAGATTGATTCGAGC	CTATTTCAGGGAGTATGGA	EB187488.1	(T)45	53	3	0.42	0.53
FM-EST 208	AAGCATCTAGATTGATTCG	ACAGCTATGACCATGATTAC	EB086270.1	(T)30	57	3	0.31	0.36
FM-EST 209	CATCTAGATTGATTCGAGC	CCTTTTGGGAAAGGTT	EB086246.1	(T)40-(A)13	52	5	0.24	0.27
FM-EST 211	TCGCGATCGCATCTA	ACAGCTATGACCATGATTAC	EB086202.1	(T)24	52	2	0.01	0.01
FM-EST 212	TTCCTGTGTGAAATTGTTAT	CGACAAAAACACAAAAAG	EB086191.1	(TTGGGG)2-(A)13	56	2	0.03	0.03
FM-EST 213	GGTAGAGGCGTAGAGAGTA	TCTCACTCATCTCTCACACT	EB086166.1	(GTGA)3-(GTGAGT)2	55	4	0.34	0.42
FM-EST 214	CATCTAGATTGATTCGAGC	TTTTCCCGGTGTTTC	FD661810.1	(T)74-(CCAAAA)2	50	2	0.01	0.01
FM-EST 215	CCAAACCCCCTTAAAT	GTTCACACACACTCTTTCTT	FD661809.1	(CCTTCC)2-(CCCGGG)2	51	4	0.21	0.23
FM-EST 217	CTAGATTGAAGCAGTTGGT	CCGTTTCTAAATTAATCTGT	FD661944.1	(A)35	61	2	0.05	0.05
FM-EST 228	AAATTTCAATATCCAACATC	CATACCTAGGCTGATGATCT	FD661795.1	(TCC)4-(CTCCGA)2	54	3	0.11	0.11
FM-EST 230	ACAAAAATCATCAGTCCTAA	ATTGTATGTCACATCTGGTT	FD661794.1	(GAGTTG)2-(TCAGAT)2	50	2	0.17	0.17
Mean						2.9	0.18	0.20

[a]NA: number of alleles per locus, [b]PIC: polymorphic information content, [c]HE: expected heterozygosity.

Finger millet genotyping using EST-SSR markers

The SSR-makers were amplified in a 20 μL total volume containing 50 ng of genomic DNA, 2 μL of each EST-SSR primer (10 pmol), 4 μL of 5x reaction Buffer (Inclone Co, Korea), 1 U of Taq DNA polymerase (Inclone Co, Korea), 1.6 μL of dNTP (2.5 mM), and 11 μL nuclease-free water. DNA amplifications were performed in PTC-100 thermal controller (MJ Research Watertown, MA, USA). The PCR profile was: initial denaturation of 3 min at 94°C, followed by 35 cycles at 95°C for 40 s, 30 s at the annealing temperature (Table 2) and 45s at 72°C, and a final extension of 10 min at 72°C. PCR products of clear, stable and specific bands

with expected length (100-250bp) on 1.5% of the agarose gel were considered as successful PCR amplifications. PCR products were electrophoresis by Fragment Analyzer^{TM} 96-capillary Automated CE System using DNF-900 double-stranded DNA Reagent Kit (Advanced analytical, USA) according to the manufacturer's instruction.

Data Analysis

The blast function of the NCBI database was used to annotate EST-SSRs using the default parameters. Different parameters such as the number of observed alleles (NA), expected heterozygosity (HE), and polymorphism information content (PIC) were analyzed using the Cervus v3.0 software (Araneda *et al.*, 2004; Kalinowski *et al.*, 2007). The neighbor-joining method was used to construct the phylogenetic tree using the DARwin6.0 software (Perrier, 2006). Principal component analysis (PCA) was performed using the facto-extra package (https://github.com/kassambara/facoextra) in R v4.1. Population structure analysis was performed using the STRUCTURE v2.3.1 software (Pritchard *et al.*, 2000). The admixture model used a burn-in of 50,000 and 100,000 iterations for 1-10 K populations, with three independent runs each. This set of parameters exhibited a convergence point, at which summary statistics attained equilibrium. The ad hoc statistic ΔK was used to determine the optimal number of subpopulations (Evanno *et al.*, 2005). The neighbor-joining method (Saitou & Nei, 1987) was used to clustering.

RESULTS

EST-SSRs selection and genetic diversity analysis

A total of 1,927 ESTs were used to evaluate the presence of SSR motifs. The SeqMan DNA Star Lasergene v7.1 software (DNASTAR Inc, Madison, WI, USA) was used to obtain consensus sequences from overlapping EST clusters to eliminate redundant sequences; 1,412 singletons were generated. To search for SSR motifs containing 2-6 nucleotides across these unigenes, we applied SSRIT (http://www.gramene.org/gramene/searches/ssrtool) and the SSR Locator V.1 software (da Maia *et al*., 2008) which generated a total of 268 SSR motifs with repeated sequences longer than 16 bp and an expected size ranging from 100 to 300 bp. In a preliminary experiment, we randomly selected 12 germplasm resources to amplify the 268 primer pairs, and 40 primers were used to detect polymorphisms.

The results showed a low level of linkage disequilibrium (LD). In all SSR marker loci, a total of 562 locus pairs were detected. The significance threshold is set to $p < 0.001$. There are 5 pairs that are significant and have square correlation coefficients of $0.1 < (r\ 2) < 0.5$, including FM-EST 130 and FM-EST 33, FM-EST 152 and FM-EST 46, FM-EST 215 and FM-EST 46, FM-EST 212 and FM-EST 152, FM-EST 230 and FM-EST 152.

The genetic variability of 102 *E. coracana* L. genotypes, representing diverse germplasm collected from Asia and Africa, was assessed using 40 EST-SSR markers. The primer sequence information, product sizes, and repeat motifs of these 40 EST-SSR markers are described in Table 2. A total of 116 alleles were detected for the 40 EST-SSR loci among 102 finger millet genotypes. The number of alleles (NA) per EST-SSR marker locus ranged from 2 to 5. These alleles were expressed at all loci, with an average of 2.9 alleles per locus. The average HE value for the germplasm examined in this study was 0.198 (0.01-0.53) (Table 2); typically, higher HE values indicate higher genetic variability. The average PIC value was 0.18 (0.01-0.39) (Table 2).

Structure analysis and phylogenetic analysis

We performed population structural analysis using the Structure Harvester program, with the number of subpopulations (K) ranging from 2 to 10. The results showed that ΔK reached a maximum (383.809466) at K = 2, indicating the most suitable K value (Fig. 1A), followed by 5 (ΔK = 23.226126). The 102 accessions collected in various regions were thus classified into two subpopulations (Fig. 1B). Similarly, the PCA analysis also clearly showed that the 102 accessions were clustered in two regions (Fig. 1C). Individuals in groups I and II were distributed in Asia (Nepal, India, China, and Pakistan) and Africa (Uganda, Zambia, Zimbabwe, and Kenya), respectively (Fig. 1D). The germplasm was clustered into two groups (Fig. 2). Group I included 42 accessions collected from the five Asian countries, and group II included 60 accessions collected from the five African countries (Fig. 2, Table 1). The F statistic between the two groups is 0.194, which shows that the degree of genetic differentiation among populations

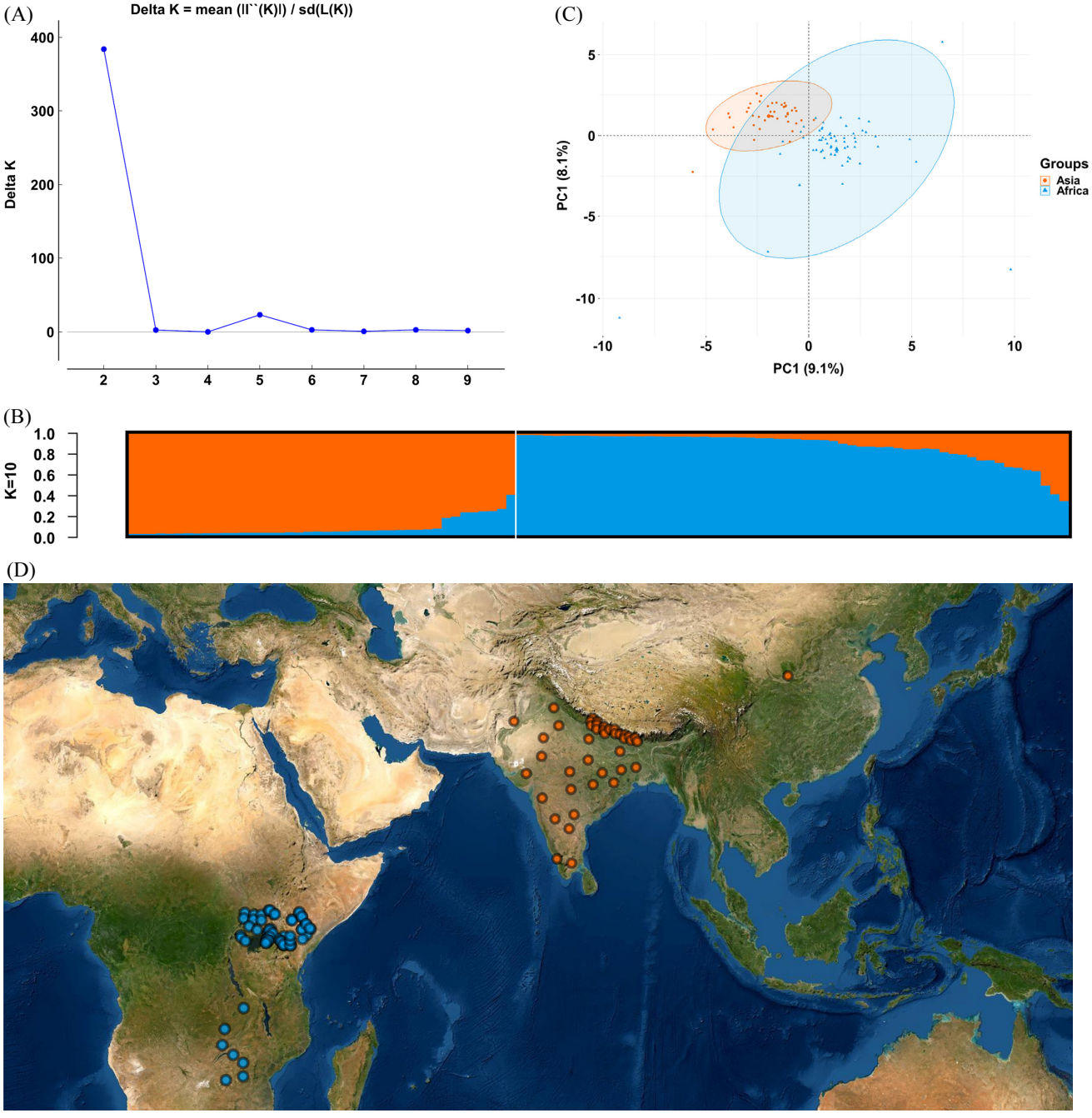

Fig. 1. Structure analysis of 102 *E. coracana* individuals based on microsatellite data. (A) Estimation of population using mean of estimated lnP(D) and lnP(D)-derived delta K (log probability of data) with cluster number (K) ranging from one to ten. (B) Two estimated clusters of the 102 *E. coracana* individuals represented by different colors inferred by STRUCTURE analysis. (C) Principal component analysis (PCA) based on SSR data. (D) Geographical distribution for K = 2.

is high. In addition, with a global Nei's (1973) genetic diversity of 0.214, the genetic diversity of group II is 0.203, which is higher than group I (0.175).

DISCUSSION

Due to their specificity and highly conserved nature, EST-SSR markers are an important genotyping tool used to study genetic relationships and cross transferability among

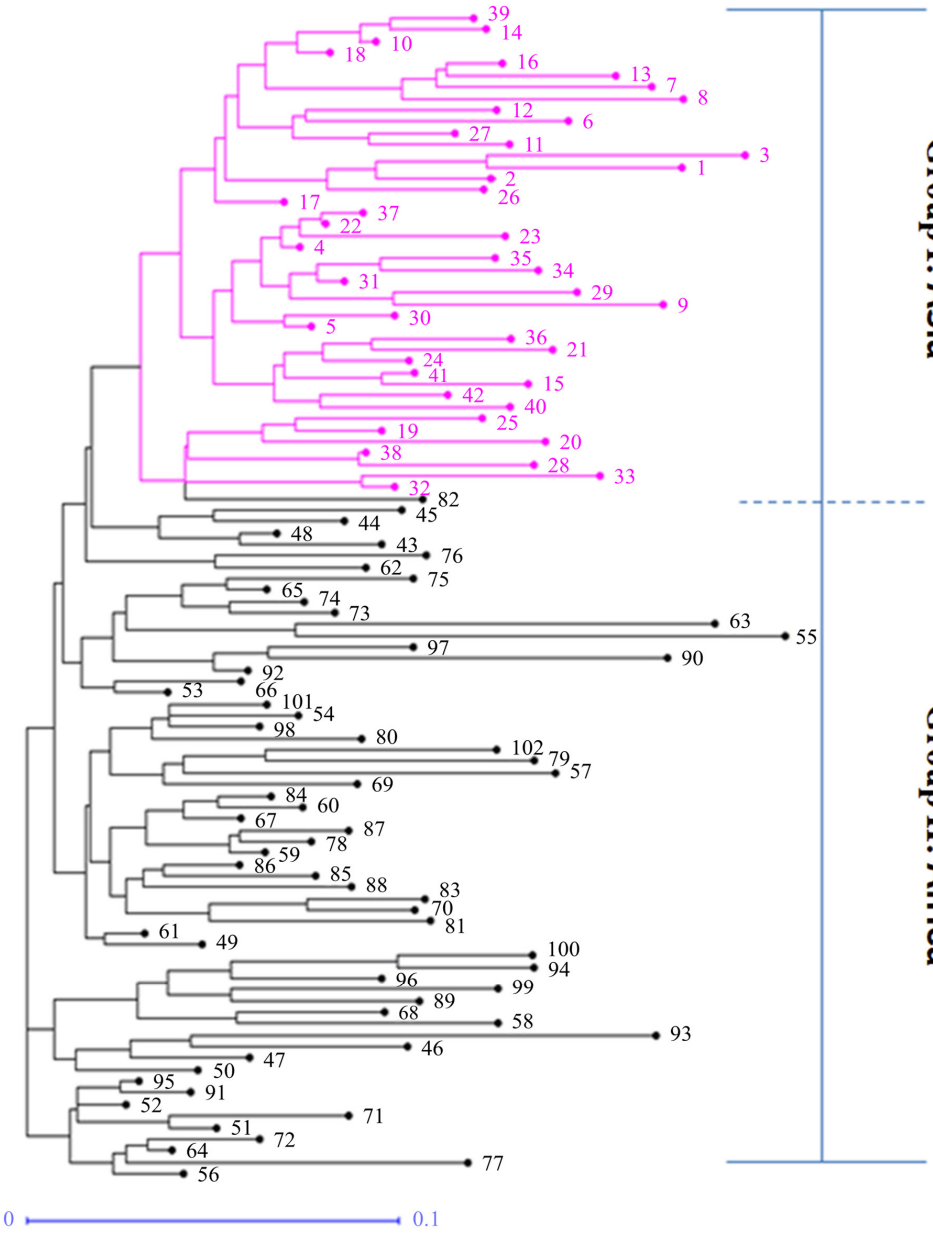

Fig. 2. Phylogenetic analysis showing the genetic relationships among the 102 finger millet accessions using 40 EST-SSR markers. Group I: collection of Asia, Group II: Collection of Africa.

crop species (Cordeiro *et al.*, 2001). The finger millet whole-genome sequencing project was started late because the main finger millet production areas are mainly in underdeveloped countries, and because the genome is a complex allotetraploid (AABB). The first finger millet genome draft was released in 2017 (Hittalmani *et al.*, 2017). No online genome viewer tool has been announced so far. Therefore, EST-SSR marker is still one of the most commonly used tools for exploring finger millet genetic information. For this reason, in this study, we developed EST-SSR markers for finger millet using publicly available genomic resources for switchgrass, which shares the same family and genus as finger millet. Switchgrass has a high percentage of molecular marker cross-species transferability (Pandey & Kumar, 2005) which supports the application of cross-species EST-SSRs to develop markers for minor crops.

Among 1,927 designed EST-SSRs, tri-nucleotide repeat motifs were found in higher proportion (41%) followed by

di-nucleotides (33%). Among the 351 finger millet EST sequences published earlier in dbEST NCBI, 463 SSRs, di-nucleotides (11.4%), tri-nucleotides (75.4%), repeats, respectively (Arya *et al*., 2009) were found. In another finger millet record, di-nucleotides (54.6%), tri-nucleotides (40.7%), the proportion of di-nucleotides is higher than our result (Naga *et al.*, 2012). The high frequency of tri-nucleotide expansions and deletions is due to the fact that these changes will not disturb the open reading frames.

The genetic structure reveals strong geographic differences. According to delta K, determine the optimal number of clusters K = 2 (Fig. 1), which corresponds to the two sampled geographic regions. This is consistent with previous research. Dida *et al.* (Dida *et al*., 2008) used 45 SSR markers to classify finger millet cultivars into African *coracana* subpopulation and Asian *coracana* subpopulation. Kumar *et al*. (Kumar *et al*., 2016) used GBS data to analyze the population structure of 113 finger millet cultivars, and divided all germplasm into 3 subpopulations. Most of the subpopulation 1 is the East African subpopulation, and the subpopulation 2 is the Asian subpopulation. Subpopulation 3 is the hybrid Indaf varieties of African and Asian germplasm in the Finger Millet Improvement Project. The genetic background of finger millet cultivars is relatively single. Gimode *et al*. (Gimode *et al*., 2016) performed Kompetitive Allele-Specific PCR (KASP) assays on 92 SNPs and clearly distinguished wild species and cultivated species. The wild species are divided into 4 subpopulations, but the cultivated species cannot be distinguished. The phylogenetic tree shows that the genetic distance between wild species and cultivated species is much larger than the intraspecific genetic distance of cultivated species. The finger millet cultivar (*E. coracana* L. subsp. *coracana*) was domesticated from the wild East African *E. coracana* subsp. *africana* subspecies about 5000 years ago (De Wet *et al.*, 1984). About 3000 years ago, finger millet was introduced to India and then spread further in Asia. Thus, East Africa is considered to be the center of domestication, and India is the secondary center of origin. Our results are consistent with this history of finger millet domestication and migration. The efficiency of the EST-SSR markers was assessed according to genetic relatedness among the accessions based on their grouping pattern in the neighbor-joining tree (Fig. 2). It revealed two distinct groups among the finger millet accessions. The F-statistic (Fst) show that the degree of genetic differentiation among 2 groups is high. The genetic diversity of the African group is higher than that of the Asian group, which is in line with the argument that East Africa is the center of origin. Because in general, the closer to the center of origin, the higher the diversity.

In summary, we developed 40 EST-SSR markers for finger millet utilizing cross genomic resources. And using these markers to analyze the genetic diversity and population structure of finger millet. Our findings indicated low genetic differentiation among finger millet germplasm collections. We believe that this is because the reproductive mode of finger millet (selfing or high inbreeding) promotes the formation of homozygotes and makes favorable alleles exclude rare alleles. This leads to a decrease in genetic diversity and an increase in the genetic differentiation of the two groups. The decrease in genetic diversity may make it difficult for species to withstand the drastic climate changes in recent years. Breeders' work should focus on increasing the genetic diversity of finger millet to ensure food security in arid regions of Africa and Asia. With the development of next-generation sequencing (NGS), finger millet has also completed whole-genome sequencing (Hittalmani *et al*., 2017). We will also perform genotyping by sequencing (GBS) of finger millet landraces and wild species in the future, and make a linear comparison of the genomes of cultivated and wild species. The research results will help select suitable genotypes for breeding programs.

ACKOWLEDGEMENTS

This study was supported by grant allocated to M.-J.S. from National Academy of Agricultural Sciences (Project No. PJ014242), the Rural Development Administration Korea, Republic of Korea. This study was supported by the 2021 Postdoctoral Fellowship Program (X.W.) of the National Institute of Agricultural Sciences, RDA, Republic of Korea.

REFERENCES

Araneda, C., C. Correa, N. Lam, E. Uribe, M. Camiruaga, and P. Iturra. 2004. Asignación de paternidadmaternidad y parentesco

en un plantel de avestruces (Struthio camelus), utilizando seis loci microsatélites polimórficos. Av. Produc. Anim 29 : 3-14.

Arya, L., M. Verma, V. Gupta, and J. Karihaloo. 2009. Development of EST SSRs in finger millet (Eleusine coracana ssp coracana) and their transferability to pearl millet (Pennisetum glaucum). Journal of Plant Biochemistry and Biotechnology 18 : 97-100.

Babu, B. K., N. Senthil, S. M. Gomez, K. R. Biji, N. S. Rajendraprasad, S. S. Kumar, and R. C. Babu. 2007. Assessment of genetic diversity among finger millet (Eleusine coracana (L.) Gaertn.) accessions using molecular markers. Genetic Resources and Crop Evolution 54 : 399-404.

Baryeh, E. A. 2002. Physical properties of millet. Journal of Food Engineering 51 : 39-46.

Bradbury, P. J., Z. Zhang, D. E. Kroon, T. M. Casstevens, Y. Ramdoss, and E. S. Buckler. 2007. TASSEL: software for association mapping of complex traits in diverse samples. Bioinformatics 23 : 2633-2635.

Cordeiro, G. M., R. Casu, C. L. McIntyre, J. M. Manners, and R. J. Henry. 2001. Microsatellite markers from sugarcane (Saccharum spp.) ESTs cross transferable to erianthus and sorghum. Plant science 160 : 1115-1123.

Da Maia, L. C., D. A. Palmieri, V. Q. De Souza, M. M. Kopp, F. I. F. de Carvalho, and A. Costa de Oliveira. 2008. SSR locator: tool for simple sequence repeat discovery integrated with primer design and PCR simulation. International journal of plant genomics 2008.

De Wet, J., K. P. Rao, D. Brink, and M. Mengesha. 1984. Systematics and evolution of Eleusine coracana (Gramineae). American journal of Botany 71 : 550-557.

Dida, M. M., S. Ramakrishnan, J. L. Bennetzen, M. D. Gale, and K. M. Devos. 2007. The genetic map of finger millet, Eleusine coracana. Theoretical and Applied Genetics 114 : 321-332.

Dida, M. M., N. Wanyera, M. L. H. Dunn, J. L. Bennetzen, and K. M. Devos. 2008. Population structure and diversity in finger millet (Eleusine coracana) germplasm. Tropical Plant Biology 1 : 131-141.

Evanno, G., S. Regnaut, and J. Goudet. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Molecular ecology 14 : 2611-2620.

Fernandez, D. R., D. J. Vanderjagt, M. Millson, Y.-S. Huang, L.-T. Chuang, A. Pastuszyn, and R. H. Glew. 2003. Fatty acid, amino acid and trace mineral composition of Eleusine coracana (Pwana) seeds from northern Nigeria. Plant Foods for Human Nutrition 58 : 1-10.

Gimode, D., D. A. Odeny, E. P. de Villiers, S. Wanyonyi, M. M. Dida, E. E. Mneney, and S. M. de Villiers. 2016. Identification of SNP and SSR markers in finger millet using next generation sequencing technologies. PloS one 11 : e0159437.

Goron, T. L. and M. N. Raizada. 2015. Genetic diversity and genomic resources available for the small millet crops to accelerate a New Green Revolution. Frontiers in plant science 6 : 157.

Govindaraj, M., M. Vetriventhan, and M. Srinivasan. 2015. Importance of genetic diversity assessment in crop plants and its recent advances: an overview of its analytical perspectives. Genetics research international 2015.

Gupta, R., K. Verma, D. Joshi, D. Yadav, and M. Singh. 2010. Assessment of genetic relatedness among three varieties of finger millet with variable seed coat color using RAPD and ISSR markers. Genet Eng Biotechnol J 2 : 1-9.

Hittalmani, S., H. B. Mahesh, M. D. Shirke, H. Biradar, G. Uday, Y. R. Aruna, and A. Mohanrao. 2017. Genome and transcriptome sequence of finger millet (Eleusine coracana (L.) Gaertn.) provides insights into drought tolerance and nutraceutical properties. BMC genomics 18 : 1-16.Kalinowski, S. T., Taper, M. L., and T. C. Marshall. 2007. Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. Molecular ecology 16 : 1099-1106.

Kumar, A., D. Sharma, A. Tiwari, J. Jaiswal, N. Singh, and S. Sood. 2016. Genotyping-by-sequencing analysis for determining population structure of finger millet germplasm of diverse origins. The plant genome 9 : 1-15.

Naga, B. L. R. I., L. N. Mangamoori, and S. Subramanyam. 2012. Identification and characterization of EST-SSRs in finger millet (Eleusine coracana (L.) Gaertn.). Journal of Crop Science and Biotechnology 15 : 9-16.

Pandey, P. and G. Kumar. 2005. Finger Millet: A flair for human nutrition in Uttaranchal. Indian Farmers Digest 38 : 28-30.

Perrier, X. 2006. DARwin software. http://darwin. cirad. fr/darwin:

Powell, W., G. C. Machray, and J. Provan. 1996. Polymorphism revealed by simple sequence repeats. Trends in plant science 1 : 215-222.

Pritchard, J. K., M. Stephens, and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. Genetics 155 : 945-959.

Saitou, N. and M. Nei. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Molecular biology and evolution 4 : 406-425.