Breeding Science 69: 640–650 (2019)
doi:10.1270/jsbbs.19050

Research Paper

Genetic diversity of leafy amaranth (*Amaranthus tricolor* L.) resources in Vietnam

Duc Chinh Nguyen[1,2], Danh Suu Tran[2], Thi Thu Hoai Tran[2], Ryo Ohsawa[3] and Yosuke Yoshioka*[3]

[1] *Graduate School of Life and Environmental Sciences, University of Tsukuba*, 1-1-1 Tennodai, Tsukuba, Ibaraki 305-8572, Japan
[2] *Plant Resources Center, Vietnamese Academy of Agricultural Sciences*, An Khanh, Hoai Duc, Hanoi 13219, Vietnam
[3] *Faculty of Life and Environmental Sciences, University of Tsukuba*, 1-1-1 Tennodai, Tsukuba, Ibaraki 305-8572, Japan

Leafy amaranths, which are consumed as traditional food in Asia and Africa, are now considered among the most promising vegetables. In Vietnam, leafy amaranths, particularly *Amaranthus tricolor* L., are important summer vegetables due to their excellent nutritional values and high tolerance to biotic and abiotic stresses. However, this species has not been subjected to systematic breeding. Here we describe species identification and evaluation of the genetic diversity of Vietnamese amaranth collection by using *matK* and simple sequence repeats (SSR) markers. Our phylogenetic analysis based on the *matK* marker classified the species of 68% of the accessions, of which 120 belonged to *A. tricolor*. We developed 21 SSR markers, which amplified a total of 153 alleles in 294 *A. tricolor* accessions originating from Vietnam and overseas, with a mean allelic richness of 7.29 per marker, observed heterozygosity of 0.14, expected heterozygosity of 0.38, and polymorphic information content of 0.35. The STRUCTURE and F_{ST} analysis indicated a positive relationship between geographic distance and genetic differentiation among most of the overseas groups and the Vietnamese collection, but not among geographic groups within the Vietnamese collection. Vietnamese amaranths could be divided into two major types, one common in East Asia and the other one unique to Vietnam.

Key Words: chloroplast, indigenous vegetable, landrace, *matK*, SSR marker.

Introduction

The genus *Amaranthus* ($2n = 32$ or 34) contains 60–75 species (Brenner *et al.* 2000, Costea and Demason 2001, Mosyakin and Robertson 1996, Sauer 1967, Waselkov *et al.* 2018), of which at least 17 are edible leafy species and 3 are grain species (grown for seeds) (Grubben and Denton 2004). Amaranth species are widely distributed, but most are found in tropical and warm temperate regions (Sauer 1967, Suresh *et al.* 2014). Approximately 55 species are native to the Americas, and the rest originated from Europe, Asia, Africa, and Australia (Sauer 1967, Waselkov *et al.* 2018). The first recorded use of amaranth grain as a staple food was by the Aztec civilization in the central Mexico (Brenner *et al.* 2000, Sauer 1993), whereas amaranth leaves and stems are used as vegetables in South Asian and African countries (Achigan-Dako *et al.* 2014, Grubben and Denton 2004, Rastogi and Shukla 2013). They can compete with spinach leaves in terms of protein content (Bui *et al.* 1998). Cultivated amaranths are also used as forage and ornamental crops. Some species are weeds (Brenner *et al.* 2000). In the last few decades, amaranth has been considered worldwide as a new millennium crop owing to its excellent nutritional value, good adaptability to severe conditions, and the absence of major diseases (Achigan-Dako *et al.* 2014, Brenner *et al.* 2000, Rastogi and Shukla 2013, Shukla *et al.* 2006).

Vegetable amaranths are a rich and inexpensive source of protein, carotenoids, vitamins, dietary fiber, and a wide range of minerals (Akubugwo *et al.* 2007, Jimenez-Aguilar and Grusak 2017, Shukla *et al.* 2010). Among leafy amaranth species, *Amaranthus tricolor* L. is grown widely owing to the high palatability of its leaves (Srivastava 2017). It was domesticated in South or Southeast Asia and then spread throughout the tropical and temperate regions (Grubben and Van Sloten 1981, Khandaker *et al.* 2008). Like in other Southeast Asian countries, it has been an important summer vegetable in Vietnam for a long time. However, most of the amaranth germplasm grown by farmers comprises landraces with some desired phenotypes, and it has not been the subject of any breeding program. The Plant Resources Center of Vietnam maintains more than 270 amaranth accessions collected from eight ecological areas of Vietnam, which are expected to contain a wide range of genetic variation and desired traits for crop improvement.

Communicated by Katsunori Hatakeyama
Received April 1, 2019. Accepted August 1, 2019.
First Published Online in J-STAGE on October 9, 2019.
*Corresponding author (e-mail: yoshioka.yosuke.fw@u.tsukuba.ac.jp)

Breeding Science
Vol. 69 No. 4 BS

However, the Vietnamese amaranth collection has not been taxonomically classified or evaluated for genetic diversity.

Evaluation of genetic diversity is a key prerequisite for a crop breeding program and is also useful for the management and exploitation of plant genetic resources from gene banks (Shukla *et al.* 2010). However, *Amaranthus* has been widely considered as a "difficult" genus in terms of taxonomy and analysis of genetic variance using morphology alone because of the overall similarity among many species, small diagnostic parts, and the presence of intermediate forms (Achigan-Dako *et al.* 2014, Assad *et al.* 2017, Costea and Demason 2001). To overcome these difficulties, recent analyses of the genetic diversity of amaranths used chloroplast and nuclear DNA sequences (Stetter and Schmid 2017, Viljoen *et al.* 2018, Waselkov *et al.* 2018).

The maturase K gene (*matK*) is one of the most variable protein-coding chloroplast genes in angiosperms (Hilu *et al.* 2003, Yu *et al.* 2011). The very high evolutionary rate of *matK* has made it suitable for phylogenetic analysis at various taxonomic levels, from order to species (Chase *et al.* 2007, Hilu and Liang 1997, Hilu *et al.* 2003, Lahaye *et al.* 2008, Muller *et al.* 2006). Evaluation of eight potential DNA barcodes for biodiversity inventories using >1600 samples collected in southern Africa and >1000 species of Mesoamerican orchids suggested *matK* as a universal barcode for flowering plants (Lahaye *et al.* 2008). In 58 species from 47 families of angiosperm plants, the matK primers showed high amplification (93.1%) and sequencing (92.6%) rates (Yu *et al.* 2011).

Over the last decades, simple sequence repeats (SSRs) have been considered the most powerful molecular markers in genetic analysis due to (1) their versatility, (2) high polymorphism provided by a large number of alleles per locus, (3) their co-dominance suitable for direct assessment of heterozygosity, (4) small amounts of DNA required, and (5) high transferability to related species (Guichoux *et al.* 2011, Kumar *et al.* 2009, Taheri *et al.* 2018, Varshney *et al.* 2005, Vieira *et al.* 2016). However, the development of SSRs is labor-intensive, economically costly, and time-consuming (Taheri *et al.* 2018). Studies on the genetic diversity of amaranths were based only on 11–14 SSR markers (Khaing *et al.* 2013, Kietlinski *et al.* 2014, Lee *et al.* 2008, Mallory *et al.* 2008, Singh *et al.* 2013, Suresh *et al.* 2014, Wang and Park 2013). These markers were developed by Mallory *et al.* (2008) and Lee *et al.* (2008) on the basis of genomic DNA of *A. hypochondriacus* L. The advent of the high throughput sequencing technology along with bioinformatics tools has facilitated fast discovery of thousands of potentially amplifiable microsatellite regions at low cost, even for species with limited genomic information (Guichoux *et al.* 2011, Silva *et al.* 2013, Taheri *et al.* 2018, Zalapa *et al.* 2012).

In this study, we used the chloroplast *matK* marker to identify leafy amaranth species from Vietnam. We then developed SSR markers for *A. tricolor* based on PacBio sequence data obtained from genomic DNA of cultivar 'Biam' to evaluate the genetic diversity of Vietnamese and overseas accessions of *A. tricolor*, an economically important leafy vegetable in Vietnam.

Materials and Methods

Plant materials and DNA isolation

A total of 272 Vietnamese amaranth accessions (hereafter referred to as VA accessions; **Supplemental Table 1**) preserved at the Plant Resources Center in Vietnam, and 182 accessions collected from all over the world (WA accessions; **Supplemental Table 2**) were used. Among VA accessions, 268 belonged to unknown species collected in 8 ecological areas of Vietnam (**Fig. 1**) and 4 accessions were imported from the World Vegetable Center. Among WA accessions, 166 were provided by the USDA National Plant Germplasm System, and 16 are preserved at the University of Tsukuba (9 are commercial cultivars and 7 are provided by the NARO Genebank). The WA collection included 176 accessions of *A. tricolor* and 6 accessions of *A. hypochondriacus*, which was used as an outgroup species in this study. DNA was extracted from the first true leaf of each seedling 14 days after sowing by using a DNeasy Plant Mini Kit (Qiagen, Hilden, Germany) according to the manufacturer's protocol with minor modifications.

Phylogenetic analysis

We used the *matK* marker developed by Ooi *et al.* (1995) for phylogenetic analysis. The *matK* gene was amplified

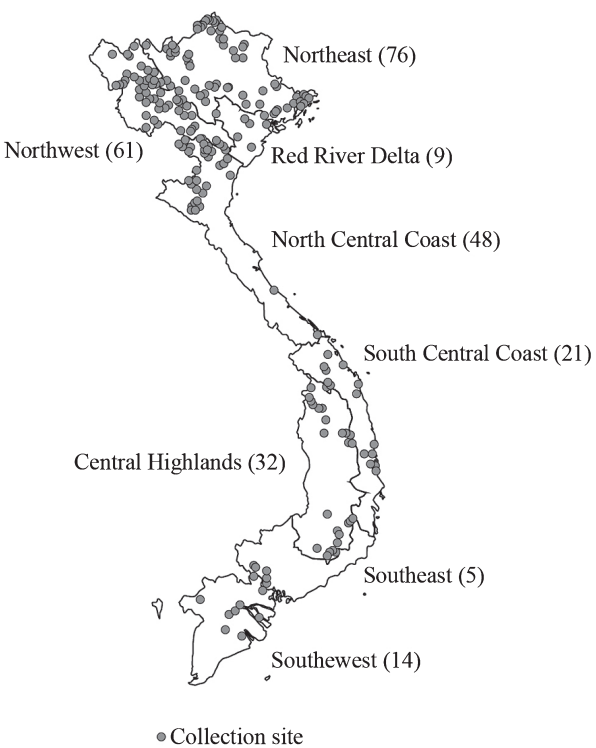

Fig. 1. Map of Vietnam showing the collection sites of amaranth accessions used in this study.

using primers matK-F (5′-CTATATCCACTTATCTTTCAG GAGT-3′) and matK-R (5′-AAAGTTCTAGCACAAGAA AGTCGA-3′) from 272 VA and 27 WA accessions (26 *A. tricolor* and 1 *A. hypochondriacus*). Ampdirect Plus PCR Mastermix (Ampdirect Plus Buffer, Blend Tag Plus DNA Polymerase, and 3 µM forward and reverse primers) was used to amplify each sample in a final volume of 10 µl. The PCR was performed in a thermal cycler under the following conditions: denaturation at 94°C for 2 min, 40 cycles of 94°C for 30 s, 55°C for 30 s, 72°C for 1 min, and a final extension at 72°C for 10 min. The PCR products were cleaned by using ExoSAP-IT Express PCR Product Cleanup, stained with The BigDye Terminator v3.1, and the DNA precipitate was dissolved in 10 µl of HiDi formamide. DNA was sequenced in an Applied Biosystems DNA Sequencing analyzer (model 3130xl). The sequencing data were analyzed with the Sequencing Analysis 5.1.1 and GeneStudio software. Sequences of the 272 VA and 27 WA accessions were aligned by using the ClustalW algorithm in the Molecular Evolutionary Genetics Analysis (MEGA) v7.0 software (Kumar *et al.* 2016) with default settings. The alignment sequences were used to produce a dendrogram with the Maximum Likelihood Algorithm in MEGA7. As references, we also used the *matK* sequences of 45 GRIN (Germplasm Resources Information Network) amaranth accessions (known species) deposited in GenBank (KX079543–KX079587; Viljoen *et al.* 2018).

Development of SSR primers for leafy amaranths

The genome sequence data of *A. tricolor* cultivar ‘Biam’ were obtained by the single molecular real-time (SMRT) cell of PacBio RSII and used to develop SSR markers for leafy amaranth species. SSR isolation and primer design were carried out by using msatcommander software (Faircloth 2008) with the following criteria: di, tri and tetranucleotide repeats, amplified fragments of 100–500 bp, optimal melting temperature of 60.0°C (range, 57.0–62.0°C), optimal GC content of 50%, and low levels of self- or pair-complementarity.

Of 27,690 markers, 817 novel SSR markers were randomly selected, and amplification was assessed using ‘Biam’ DNA. Forward primers were 5′-labeled with the fluorescent dyes 6-FAM, VIC, NED, or PET (Shimizu and Yano 2011). PCR was carried out using a KAPA2G Fast PCR kit in a final volume of 10 µl containing 1 µl of genomic DNA, 2 µl of 5X KAPA2G buffer A, 0.2 µl of 10 mM dNTPs, 0.2 µl of 25 mM $MgCl_2$, 0.02 µl of 5 U/µl KAPA2G FAST DNA polymerase, 6.08 µl of Milli-Q water, and 0.5 µl of primer mix. Reactions were performed in a thermal cycler (Applied Biosystems) as follows: 94°C for 3 min; 30 cycles of 94°C for 20 s, 54°C for 30 s, and 62°C for 30 s; 3 cycles of 94°C for 20 s, 49°C for 10 s, and 72°C for 5 s; and a final extension at 72°C for 10 min. Amplification was assessed by electrophoresis in 8% polyacrylamide gels.

The polymorphism of amplified SSR markers was assessed in eight accessions of *A. tricolor* (WA001, WA002, WA003, WA004, WA011, WA012, WA017, and WA022). Ampdirect Plus PCR Mastermix (Ampdirect Plus Buffer, KAPA2G Fast DNA Polymerase, 0.5 µl of primer mix) was used to amplify each sample in a final volume of 5 µl. Reactions were performed in a thermal cycler (Applied Biosystems) as above. The sizes of the amplified fragments were estimated by using an automated DNA analyzer (model 3130xl) with a GeneScan-600LIZ size standard and GeneMapper v. 4.0 software (all from Thermo Fisher Scientific, Inc., Waltham, MA, USA).

Analysis of genetic diversity and population structure using SSR markers

Using 30 developed SSR markers (**Supplemental Table 3**), we analyzed the genetic diversity of 120 VA accessions of *A. tricolor* (the species was determined using the *matK* marker) and 181 WA accessions (175 *A. tricolor* and 6 *A. hypochondriacus*; of these, 25 *A. tricolor* and 1 *A. hypochondriacus* used in the *matK* analysis). PCR and estimation of the sizes of the amplified fragments were performed as above.

The GenAlEx package v. 6.503 was used to calculate the number of alleles (Na), the number of effective alleles (Ne), observed heterozygosity (Ho), expected heterozygosity (He), information index (I), Nei’s genetic distance (Peakall and Smouse 2006, 2012). Major allele frequency and the polymorphic information content (PIC) were determined and the phylogenetic tree was constructed by using PowerMarker v. 3.25 software (Liu and Muse 2005). Arlequin 3.5.2.2 software (Excoffier and Lischer 2010) was used for analysis of molecular variance (AMOVA) and pairwise population F_{ST} values. Principal coordinates analysis (PCoA) was performed using GenAlEx package v. 6.503. Population structure was inferred by using the STRUCTURE 2.3.4. software package (Pritchard *et al.* 2000). The number of subpopulations (K) was set at 1–9, with 10 replications per K value. The K value was determined by running an admixture and allele frequencies correlated model. Each run started with 1,000,000 burn-ins, followed by 100,000 MCMC (Markov chain Monte Carlo) iterations. The *ad* hoc statistic Delta K was calculated to detect populations by using the online program STRUCTURE HARVESTER (Evanno *et al.* 2005).

Results

Phylogenetic analysis of the amaranth collection by using the matK barcode

Alignment of the *matK* gene region sequences from 272 VA, 27 WA, and 45 GRIN amaranth accessions had a total length of 434 bp per accession and contained 96.8% of constant, 3.2% of variable, and 2.5% of parsimoniously informative sites. Phylogenetic relationships among the 344 amaranth accessions inferred from all nucleotide sites are shown in **Fig. 2**. Leafy amaranths (*A. tricolor*, *A. blitum* L., and *A. viridis* L.), and weed amaranths (*A. spinosus* L.,

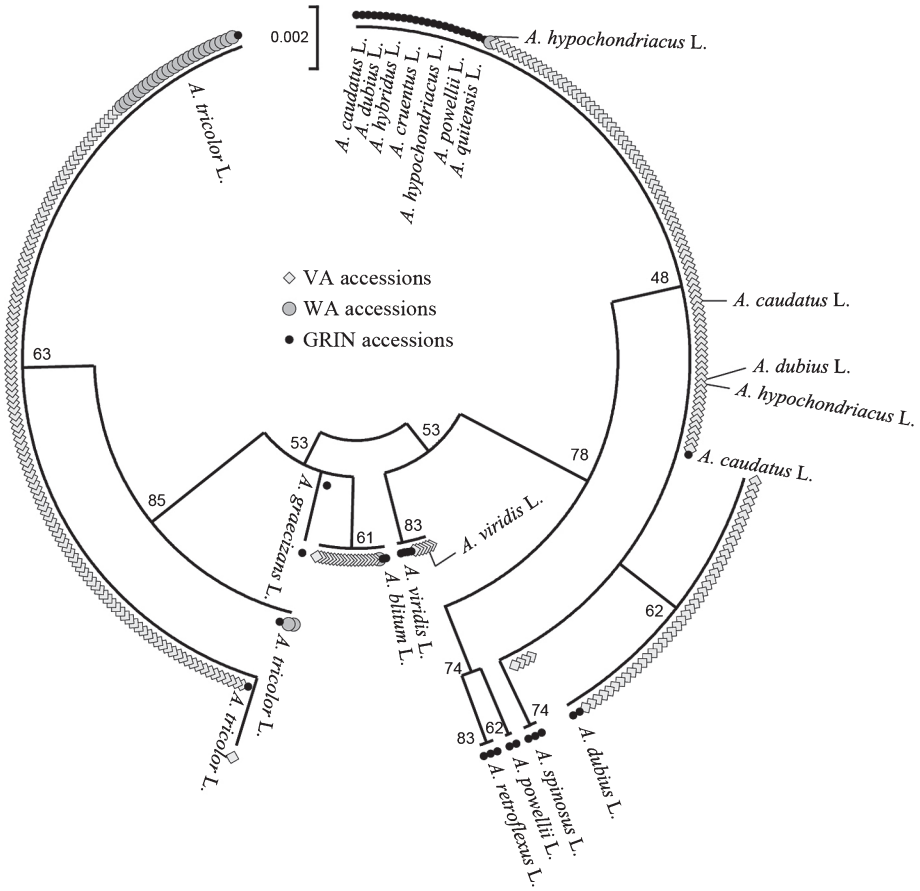

Fig. 2. Phylogenetic dendrogram of 344 amaranth accessions inferred by using the maximum likelihood method based on the Tamura–Nei model in MEGA7 with 1000 bootstrap replicates.

A. retroflexus L., and *A. powellii* S. Watson) clustered with accessions from the same species, except for an *A. powellii* accession (GRIN28), which was classified to into the hybridus complex clade. Three grain amaranths (*A. hypochondriacus*, *A. caudatus* L., and *A. cruentus* L.) and their putative ancestors (*A. hybridus* L. and *A. quitensis* Kunth) clustered together. All GRIN and 25 WA accessions of *A. tricolor* species were placed in the same clade with 120 VA accessions (unknown species) with high bootstrap support (85%). The clade of *A. tricolor* was clearly divided into two subclades, of which one contained three accessions (WA012, WA016, and GRIN41 = WA129). Two GRIN accessions of *A. blitum*, which is another vegetable amaranth species native to Asia, clustered together with 18 VA and one WA (WA008) accessions (bootstrap support: 61%). All GRIN accessions of *A. viridis* were placed in one clade with high bootstrap support (83%) together with one VA accession of *A. viridis* (VA226) and 5 VA unknown-species accessions. We also determined the positions of three weed amaranth species (*A. retroflexus*, bootstrap support of 83%; *A. powellii*, 62%; *A. spinosus*, 74%). However, no VA accession fell into these clades. Two of the four GRIN accessions of *A. dubius* Mart. ex Thell. (a wild relative of the grain amaranths) clustered together with 41 VA accessions (bootstrap support 62%), whereas the other two were placed in the hybridus complex clade, which contained three grain amaranth species and their putative ancestors, with low bootstrap support (48%). The hybridus complex clade also included five GRIN accessions of *A. caudatus*, five GRIN accessions of *A. hybridus*, three GRIN accessions of *A. cruentus*, six GRIN accession of *A. hypochondriacus*, two GRIN accessions of *A. quitensis*, one WA accession of *A. hypochondriacus* (WA10), one VA accession of *A. hypochondriacus* (VA228), one VA accession of *A. dubius* (VA227), one VA accession of *A. caudatus* (VA179), and 84 unknown-species VA accessions.

Analysis of genetic diversity and population structure of A. tricolor using SSR markers

Nine out of 30 selected SSR markers performed poorly in GeneMapper, and one accession (VA276) with a high missing rate (47.6%) was excluded from the following analysis. The 294 accessions of *A. tricolor* were supported as a monophyletic group, with 6 accessions of *A. hypochondriacus* forming an outgroup (**Fig. 3**). The *A. tricolor* group was divided clearly into two subgroups; one of them contained 15 WA accessions, including WA012, WA016, and WA129, in agreement with the result obtained using the *matK* barcode.

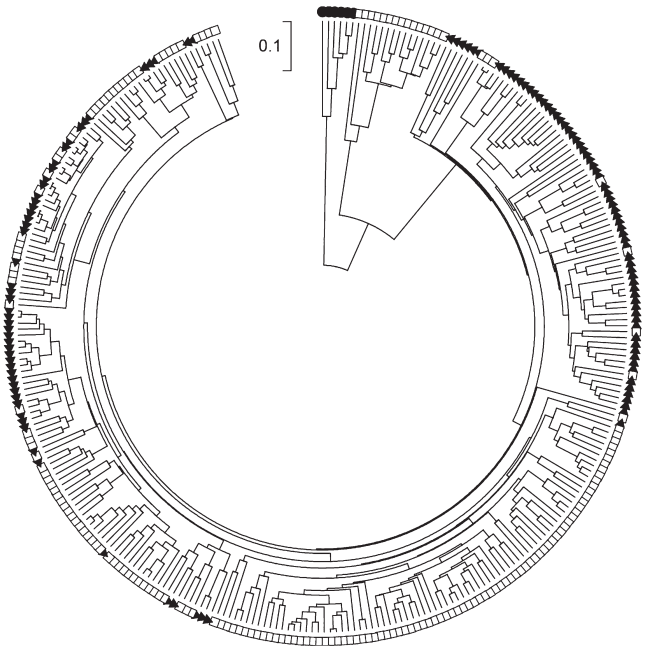

Fig. 3. Phylogenetic tree of 300 amaranth accessions constructed using the neighbor-joining method based on the Nei (1983) distance in PowerMarker. Black triangles, VA *A. tricolor* accessions; open squares, WA *A. tricolor* accessions; black circles, *A. hypochondriacus* accessions (outgroup).

Table 1. Statistics of the 21 SSR markers used for genotyping of 294 accessions of *A. tricolor*

Marker name	Na	Ne	Major allele frequency	Ho	He	PIC
TAM001	7	1.24	0.88	0.04	0.19	0.19
TAM002	2	1.81	0.65	0.14	0.45	0.35
TAM006	20	3.41	0.46	0.56	0.71	0.67
TAM008	4	1.06	0.95	0.02	0.06	0.06
TAM009	4	1.74	0.71	0.00	0.42	0.35
TAM010	9	1.83	0.69	0.36	0.45	0.40
TAM011	3	1.20	0.89	0.01	0.17	0.16
TAM012	12	5.52	0.31	0.16	0.82	0.80
TAM014	8	3.10	0.45	0.11	0.68	0.62
TAM015	6	2.19	0.55	0.25	0.54	0.45
TAM016	2	1.11	0.93	0.00	0.10	0.09
TAM019	13	1.99	0.69	0.23	0.50	0.48
TAM020	6	1.15	0.91	0.02	0.13	0.12
TAM021	2	1.04	0.98	0.00	0.04	0.04
TAM023	6	1.74	0.73	0.25	0.43	0.40
TAM025	13	7.07	0.23	0.48	0.86	0.84
TAM026	4	1.18	0.91	0.02	0.16	0.15
TAM027	6	1.13	0.92	0.02	0.11	0.11
TAM028	6	1.22	0.89	0.14	0.18	0.18
TAM029	6	1.29	0.86	0.00	0.23	0.22
TAM030	14	5.77	0.26	0.20	0.83	0.81
Average	7.29	2.28	0.71	0.14	0.38	0.35

Na, number of alleles; Ne, number of effective alleles; Ho, observed heterozygosity; He, expected heterozygosity; PIC, polymorphism information content.

Table 2. Genetic diversity indices of the WA and Vietnamese collections, and variations among geographic groups

Geographical group	Number of accessions	Na	Ne	I	Ho	He	PIC
Whole collection	294	7.29	2.28	0.84	0.14	0.38	0.35
WA collection	175	6.24	2.18	0.82	0.14	0.38	0.36
Africa	3	2.00	1.79	0.48	0.16	0.28	0.24
America	18	3.38	2.38	0.90	0.09	0.52	0.46
East Asia	54	3.38	1.83	0.57	0.15	0.29	0.27
South Asia	84	4.90	1.86	0.67	0.14	0.31	0.30
Southeast Asia	7	3.33	2.40	0.88	0.21	0.48	0.42
Commercial cultivars	9	2.24	1.80	0.52	0.10	0.30	0.24
Vietnamese collection	119	4.52	2.10	0.69	0.14	0.34	0.30
Northeast	25	3.62	2.16	0.70	0.14	0.35	0.32
Northwest	25	3.00	1.94	0.63	0.19	0.34	0.30
Red River Delta	9	2.71	2.02	0.61	0.16	0.33	0.30
North Central Coast	18	3.05	1.76	0.56	0.11	0.28	0.26
South Central Coast	13	2.24	1.68	0.46	0.12	0.26	0.23
Central Highlands	12	2.90	2.02	0.62	0.13	0.32	0.29
Southeast	5	2.05	1.66	0.48	0.13	0.28	0.25
Southwest	12	2.76	1.87	0.59	0.19	0.32	0.29

Na, number of alleles; Ne, number of effective alleles; I, information index; Ho, observed heterozygosity; He, expected heterozygosity; PIC, polymorphism information content.

The statistics of the 21 SSR markers used for genotyping of 294 accessions of *A. tricolor* are summarized in **Table 1**. These markers amplified a total of 153 alleles, ranging from 2 (TAM002, TAM016, and TAM021) to 20 (TAM006) per marker, with a mean allelic richness (Na) of 7.29, and Ne of 2.28. The average values were: major allele frequency, 0.71; Ho, 0.14; He, 0.38; and PIC, 0.35 (**Table 1**). The number of rare alleles (frequency <5%) was 102 (66.7%).

The genetic diversity indices for the overseas (WA) and Vietnamese collections are shown in **Table 2**. The average number of alleles per marker was 4.52 in the Vietnamese collection, which was considerably less than in the overseas collection (6.24). The average values of almost all other indices (Na, Ne, He, I, and PIC) were also lower in the Vietnamese collection than in the WA one. The only exception was Ho (0.14 in both collections), indicating that most of the samples in the whole collection were highly homozygous. Among the overseas geographical groups, the average number of alleles ranged from 2.0 (Africa) to 4.90 (South Asia). The African collection also had the smallest values of Ne (1.79), I (0.48), He (0.28), and PIC (0.24); except for Ho. The American collection had the largest values of I (0.90), He (0.52), and PIC (0.46). The average values of all indices were higher in the Vietnamese collection than in the commercial cultivar collection; the latter had the smallest or the second smallest values of all the indices, except for He in comparison with overseas groups (**Table 2**). Among Vietnamese geographical groups, the average number of alleles ranged from 2.05 (Southeast) to 3.62 (Northeast). The average values of Na, Ne, I, He, and PIC were the largest in the Northeast. The Southeast and South Central collections had the lowest or the second lowest values of all the indices, except for Ho.

Table 3. Pairwise F_{ST} (below diagonal) and Nei's genetic distance (above diagonal) among overseas geographical groups and the Vietnamese collection

	Africa	America	East Asia	South Asia	Southeast Asia	Commercial cultivars	Vietnam
Africa		0.269	0.043	0.091	0.107	0.089	0.055
America	0.186		0.253	0.222	0.194	0.275	0.260
East Asia	0.015	0.315***		0.085	0.092	0.053	0.042
South Asia	0.089	0.254***	0.144***		0.110	0.129	0.093
Southeast Asia	0.039	0.120	0.152***	0.139**		0.118	0.093
Commercial cultivars	0.031	0.234*	0.069*	0.172***	0.112*		0.086
Vietnam	0.029	0.284***	0.078***	0.163***	0.116***	0.100***	

Significance at the * 5%, ** 1% level, and *** 0.1% level.

The degree of genetic differentiation (F_{ST}) among WA geographical groups and the Vietnamese collection was significant, except for the African group compared with other groups and the American group compared with the Southeast Asian group. Pairwise F_{ST} ranged from 0.015 (Africa vs. East Asia) to 0.315 (America vs. East Asia) (**Table 3**). In comparisons between the Vietnamese and overseas groups, F_{ST} was 0.284 for America, 0.163 for South Asia, 0.116 for Southeast Asia, 0.100 for commercial cultivars, 0.078 for East Asia, and 0.029 for Africa. Nei's genetic distance ranged from 0.042 (Vietnam vs. East Asia) to 0.275 (America vs. commercial cultivars). Similar to pairwise F_{ST}, the Nei's genetic distance between Vietnam and other groups was also the highest for America, and the lowest for Africa and East Asia.

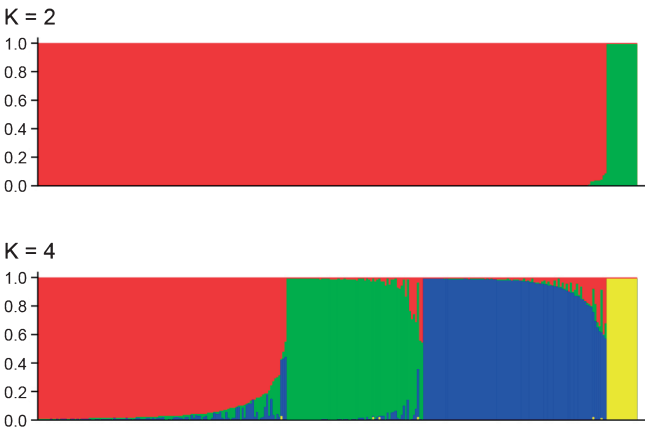

Fig. 4. Population structure of 294 accessions of *A. tricolor* L. determined at different K values by using 21 SSR markers.

Population structure of A. tricolor

In STRUCTURE analysis, the highest likelihood value of ΔK was observed for K = 2, with a secondary peak of ΔK at K = 4 (**Supplemental Fig. 1**), indicating that the accessions could be grouped into two or four subpopulations. At K = 2, all 294 accessions were allocated to two subgroups with membership probabilities (Q) > 0.95 without admixture. Most accessions (279) were assigned to subgroup 1, whereas subgroup 2 contained 15 accessions (5.1% of all accessions), all from overseas (9 from America, 5 from South Asia and one from Southeast Asia; **Fig. 4**, **Table 4**).

At K = 4 with Q > 0.75, 115 accessions (87.8% from East Asia and Vietnam) were assigned to subgroup 1, which included 90.7% of all East Asian accessions; 61 accessions, mostly from Vietnam (96.7%), were allocated to subgroup 2; 83 accessions, mostly from South Asia (92.8%), were assigned to subgroup 3; subgroup 4 was the same as subgroup 2 at K = 2. Eight out of nine commercial cultivars were allocated to subgroup 1 and one to the admixed group (**Fig. 4**, **Table 4**). The result of STRUCTURE analysis was supported by PCoA (**Fig. 5**), which assigned most accessions to four major subgroups; subgroup 1 mostly included accessions from East Asia and Vietnam; subgroup 2 from Vietnam; subgroup 3 from South Asia; and subgroup 4 from America and South Asia. The first and second factors of the PCoA explained approximately 14.63% and 9.86% of the variation in the genetic distance matrix, respectively.

Table 4. Geographic origin of *A. tricolor* accessions assumed by STRUCTURE

Geographical groups	Number of accessions	K = 2		K = 4				
		Subgroup 1[a]	Subgroup 2[a]	Subgroup 1[b]	Subgroup 2[b]	Subgroup 3[b]	Subgroup 4[b]	Admixed
Whole collection	294	279	15	115	61	83	15	20
Africa	3	3		1		1		1
America	18	9	9	4		2	9	3
East Asia	54	54		49	1	1		3
South Asia	84	79	5			77	5	2
Southeast Asia	7	6	1	1	1		1	4
Commercial cultivars	9	9		8				1
Vietnam	119	119		52	59	2		6

[a] Membership probabilities Q > 0.95.
[b] Membership probabilities Q > 0.75.

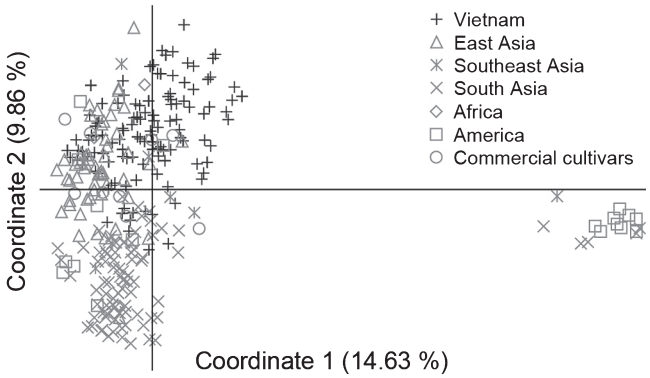

Fig. 5. Principal coordinates analysis (PCoA) of 294 accessions of *A. tricolor* with their origin areas.

Table 5. Hierarchical analysis of molecular variance among the eight Vietnamese geographical groups and the WA collection

Source of variation	d.f.	Percentage of variation
Between the Vietnamese and the WA collections	1	5.78***
Among groups	7	2.69***
Among individuals within geographical groups	285	55.54***
Among loci within individuals	294	35.99***

*** Significant at the 0.1% level for 1000 permutations.

Analysis of genetic diversity and population structure of the Vietnamese collection of A. tricolor using SSR markers

Analysis of molecular variance indicated significant differences at all hierarchical levels (**Table 5**). The percentage of variation among individuals within geographical groups was the largest (55.54%), followed by that among loci within individuals (35.99%). The variance between the Vietnamese and WA collections, and among geographical groups was small (5.78% and 2.69%, respectively).

The F_{ST} between the WA collection and each Vietnamese group was significant and ranged from 0.065 (Central Highlands) to 0.157 (Southeast; **Table 6**). The pairwise F_{ST} among Vietnamese groups ranged from 0.020 (Southeast vs. North Central Coast) to 0.127 (Southeast vs. South Central Coast) and was significant only in some comparisons. Similar to pairwise F_{ST}, Nei's genetic distance tended to be higher between the WA collection and each Vietnamese geographical group than among the Vietnamese groups (**Table 6**).

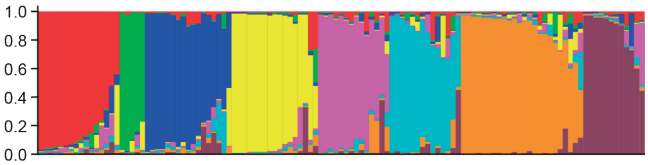

Fig. 6. Population structure of 119 VA accessions of *A. tricolor* determined at K = 8 by using 21 SSR markers.

The STRUCTURE analysis of 119 VA accessions of *A. tricolor* by using 21 SSR markers determined the highest likelihood value of ΔK at K = 8 (**Supplemental Fig. 2**), indicating that the accessions could be grouped into eight subpopulations. With a membership probability threshold of 0.75, 13 accessions were assigned to subgroup 1, 5 to subgroup 2, 12 to subgroup 3, 13 to subgroup 4, 10 to subgroup 5, 9 to subgroup 6, 19 to subgroup 7, and 9 to subgroup 3. The admixed group had the largest number of accessions (29), and each subgroup included accessions from several geographical groups (**Fig. 6**, **Table 7**). In STRUCTURE analysis of all accessions at K = 4, all accessions of subgroups 2, 4 and 7 were assigned to subgroup 1, and all accessions of subgroups 5 and 8, and most accessions of subgroups 1 (84.6%), 3 (75.0%) and 6 (78 %) were assigned to subgroup 2.

Discussion

In the current study, we aligned the chloroplast *matK* marker sequences (434 bp) from 272 VA and 27 WA accessions, and found 96.8% constant and 3.2% variable sites. The length of the aligned sequence was much shorter than in previous studies on amaranth genetic diversity that used the *matK* marker (Viljoen *et al.* 2018, Waselkov *et al.* 2018); however, the percentage of parsimoniously informative sites in our study was higher than that in Viljoen *et al.* (2018): about 2.5% vs. 2.15%. A phylogenetic tree was constructed by combining these sequences with those of 45 GRIN

Table 6. Pairwise F_{ST} (below diagonal) and Nei's genetic distance (above diagonal) among eight Vietnamese geographical groups and the WA collection

	Northeast	Northwest	Red River Delta	North Central Coast	South Central Coast	Central Highlands	Southeast	Southwest	WA
Northeast		0.023	0.044	0.029	0.034	0.035	0.052	0.032	0.057
Northwest	0.025		0.048	0.054	0.053	0.058	0.066	0.025	0.071
Red River Delta	0.049	0.058*		0.072	0.066	0.046	0.085	0.065	0.092
North Central Coast	0.034	0.087***	0.115**		0.055	0.053	0.035	0.073	0.093
South Central Coast	0.047	0.086**	0.096*	0.088*		0.057	0.093	0.056	0.048
Central Highlands	0.029	0.075**	0.038	0.084**	0.082*		0.092	0.046	0.053
Southeast	0.034	0.069	0.097	0.020	0.127	0.104		0.083	0.131
Southwest	0.038	0.021	0.081*	0.122**	0.095*	0.051	0.091		0.060
WA	0.086***	0.108***	0.122**	0.138***	0.070**	0.065*	0.157**	0.088*	

Significance at the * 5%, ** 1% level, and *** 0.1% level.

Table 7. Assignment of Vietnamese *A. tricolor* accessions to subgroups by STRUCTURE analysis

Geographical groups	Number of accessions	Sub[a] 1[b]	Sub 2[b]	Sub 3[b]	Sub 4[b]	Sub 5[b]	Sub 6[b]	Sub 7[b]	Sub 8[b]	Admixed
Vietnamese collection	119	13	5	12	13	10	9	19	9	29
Northeast	25	3		5	2	2	2	4	1	6
Northwest	25	4	1	1	5	3	2	1	1	7
Red River Delta	9	4	1			1		1		2
North Central Coast	18		1	3			1	4	3	6
South Central Coast	13				3		2	7	1	
Central Highlands	12	1		3	1	2	1	1		3
Southeast	5		1						2	2
Southwest	12	1	1		2	2	1	1	1	3

[a] Subgroup.
[b] Membership probabilities Q > 0.75.

amaranth accessions from the study of Viljoen *et al.* (2018). The interrelationships among the GRIN accessions, which include leafy amaranths (*A. tricolor*, *A. blitum*, *A. viridis*, and *A. graecizans*), weed amaranths (*A. spinosus*, *A. retroflexus*, and *A. powellii*), and *A. dubius* in our dendrogram were highly similar to those in the dendrogram in **Fig. 4** in Viljoen *et al.* (2018), which was based on whole chloroplast genome sequences. Most WA accessions of *A. tricolor* were assigned to the same group as the GRIN *A. tricolor* accessions, except that WA008 was in the same clade in our analysis as the *A. blitum* group, which was also determined in another study (our unpublished data). Our phylogenetic analysis determined the species of 68% of the Vietnamese accessions, including 120 accessions of *A. tricolor* with a high bootstrap support (85%), 41 of *A. dubius*, 18 of *A. blitum*, and 6 of *A. viridis*. *Amaranthus tricolor* accounted for the largest percentage of accessions (44%) in the Vietnamese collection, which agrees with a widely held view that *A. tricolor* is one of the most economically important vegetable amaranths in Vietnam. Two of the four GRIN *A. dubius* accessions clustered together with 41 VA accessions (bootstrap support 62%), whereas the other two were placed in the hybridus complex clade. Similar results were obtained by Viljoen *et al.* (2018). Therefore, the 41 unknown VA accessions seem to belong to *A. dubius*. In this study, we identified *A. dubius* and three (*A. blitum*, *A. tricolor* and *A. viridis*) out of seven amaranth species recorded in Vietnam by Pham (1999), which included *A. caudatus*, *A. hybridus*, *A. lividus* (Synonym of *A. blitum*), *A. retroflexus*, *A. spinosus*, *A. tricolor*, and *A. viridis*.

We developed 21 SSR markers from the PacBio sequence data of cultivar 'Biam' and used them to genetically analyze 294 Vietnamese and overseas accessions of *A. tricolor*. Each of these markers amplified a different single locus with relatively high polymorphism and could provide useful information for studies on genetic diversity of leafy amaranths. The 21 markers amplified a total of 153 alleles in 294 accessions with an average of 7.29 alleles per locus, Ho of 0.14, He of 0.38, and PIC of 0.35. The average allele richness in our study was larger than in Mallory *et al.* (2008) and Wang and Park (2013) (4.0 and 4.79, respectively), but smaller than in Khaing *et al.* (2013) and Suresh *et al.* (2014) (12.9 and 11.1, respectively). These differences may be caused not only by using different SSR markers, but also by differences in the sample size, which was smaller in the former two studies and much larger in the latter two studies than in our study. Average genetic diversity indices in our study were smaller than in most previous studies such as Suresh *et al.* (2014) (Ho, 0.29; He, 0.70; PIC, 0.66), Khaing *et al.* (2013) (PIC, 0.71), Lee *et al.* (2008) (Ho, 0.28; He, 0.74), and Mallory *et al.* (2008) (He, 0.71). The cause of this difference may be the much higher number of species used in all of the above studies than in our study, especially in the first two studies (33 and 57, respectively). In the current study, the average values of almost all genetic diversity indices were smaller in the Vietnamese collection than in the WA collection. This result was supported by STRUCTURE analysis (**Table 4**, **Fig. 4**), which assigned all VA accessions to the same subgroup at K = 2. These results indicate that the Vietnamese collection has a lower level of genetic diversity than the WA collection.

In genetic differentiation analysis (**Table 3**), we found significant differences between the Vietnamese group and all overseas groups except the African one. The F_{ST} was relatively very large between Vietnam and America (0.284), large between Vietnam and South Asia (0.163), and moderate between Vietnam and East Asia (0.078). These data indicate that the Vietnamese collection is relatively closely genetically related to the East Asian collection but is divergent in comparison with the South Asian and American collections. The result was confirmed by the Nei's genetic distance (**Table 3**) and PCoA results (**Fig. 5**); PCoA plot showed the differentiation of Vietnamese accessions from American and South Asian accessions, and a close distance between Vietnamese and East Asian accessions. Interestingly, Vietnamese accessions are genetically closer to the East Asian accessions (F_{ST} = 0.078; Nei's distance = 0.042) than to the Southeast Asian accessions (F_{ST} = 0.116; Nei's distance = 0.093), likely because all East Asian accessions originated from China (a neighbor of Vietnam), whereas no Southeast Asian accessions originated from two other neighbors, Laos and Cambodia. Another cause could be the dispersal events of amaranth from East Asia (China) to Vietnam, partly owing to the migrations of ancestors of

most contemporary Vietnamese ethnic groups from south of the Yangtze River of China (Hall and Patrinos 2012) and crop introduction during the long period of Chinese domination in Vietnam from 111 BC to 938 AD (Le *et al.* 1991). African accessions were closely related to Vietnamese, East Asian, and Southeast Asian accessions ($F_{ST} < 0.05$); this relationship (F_{ST}) was not significant, possibly because of the small size of the African collection ($n = 3$). The American collection had a very large genetic distance from most groups ($F_{ST} > 0.25$).

The STRUCTURE analysis assumed that the accessions could be grouped into two or four subpopulations (**Fig. 4**). At K = 2, all accessions were assigned to two subgroups (Q > 0.95) without admixture, indicating rare genetic communication between the subgroups. At K = 4, all accessions were allocated to four subgroups with one admixed group. Subgroup 1 included mostly accessions from East Asia and Vietnam, and contained 90.7% of all East Asian accessions; subgroup 2 included mostly accessions from Vietnam and subgroup 3 from South Asia; and subgroup 4 included 60% of accessions from America (**Table 4**). The results suggest association of the four subgroups with geographical areas and partly agree with the result of Wang and Park (2013), who claimed that the extent of genetic exchange within or among amaranth species from South and Southeast Asia is relatively low. On the other hand, Suresh *et al.* (2014) used 11 SSR markers for STRUCTURE analysis of 348 amaranth accessions from 33 species and found no significant correlation between geographic distance and genetic diversity. The positive relationship between amaranth accessions and their origins can be partly explained by the low extent of genetic exchange of *A. tricolor* accessions from different geographical areas. Our STRUCTURE analysis at K = 4 divided Vietnamese *A. tricolor* into two major types, one common type in East Asia (52 accessions; subgroup 1) and the other one unique type to Vietnam (61 accessions; subgroup 2); the latter type included only one accession from East Asia and one from Southeast Asia. It seems that genetic diversity in Vietnamese amaranths was established by dispersal events mainly from East Asia, adaptation to local environments through cross pollination, and the keeping of cross-pollinated seeds by farmers for the next season without selection.

In the current study, we used nine commercial cultivars commonly in China, India, Japan and Southern Asian areas (six cultivars from Evergreen Seeds Company, Anaheim, California, USA and three from Japanese companies). These cultivars have a low level of genetic diversity (**Table 2**) and are relatively closely genetically related to the accessions in the Vietnamese collection, as indicated by pairwise F_{ST} values and Nei's genetic distances (**Table 3**). The close relationship between these cultivars and Vietnamese accessions was confirmed by STRUCTURE analysis at K = 4 (**Table 4**, **Fig. 4**); eight commercial cultivars were assigned to subgroup 1, which included 52 accessions from Vietnam. However, no commercial cultivars were present in subgroup 2, which we considered unique type to Vietnam.

In Vietnam, we found that the variation among geographical groups was significant but small, about 2.69% (**Table 5**). The result is supported by the genetic differentiation analysis, in which the F_{ST} and Nei's genetic distance in pairwise comparisons between Vietnamese geographical groups ranged from small to moderate (**Table 6**). The STRUCTURE analysis grouped all VA accessions into eight subpopulations (Q > 7.5) and one admixed group, with accessions from different geographical areas present in each subgroup (**Fig. 6**, **Table 7**). The result suggests no relationship between geographic distance and genetic diversity and a high level of genetic exchange among *A. tricolor* accessions in Vietnam. In conclusion, SSR markers developed in the current study allowed us to evaluate the overall genetic diversity of the Vietnamese and overseas collections of *A. tricolor* and genetic relationships among geographic groups. The STRUCTURE and F_{ST} analyses showed a positive relationship between geographic distance and genetic diversity among most of the overseas groups and Vietnamese collection, but not among geographic groups within Vietnamese collection. The majority of Vietnamese accessions could be identified by using only the *matK* barcoding marker and included 120 *A. tricolor* accessions, with half of them belonging to the type unique to Vietnam. The genetic information from our research may be useful for effectively managing and exploiting the genetic resources in gene banks and for genetically improving amaranths.

Author Contribution Statement

TDS and TTTH provided the information of Vietnamese amaranth accessions and their leaves; NDC and YY carried out analysis, and wrote the manuscript; TDS, TTTH and RO participated in discussion of results and assisted in revising the manuscript; YY and RO conceived the study, designed the experiment and led the project.

Acknowledgments

We are grateful to the USDA National Plant Germplasm System for providing the seed materials. We also thank A. Shimono of the Toho University, and N. Nishi, R. Chen, S. Nakamura, and A. Fukasaku of the University of Tsukuba for technical assistance. This work was supported by a grant (PGRAsia Project) from the Ministry of Agriculture, Forestry and Fisheries of the Government of Japan.

Literature Cited

Achigan-Dako, E.G., O.E.D. Sogbohossou and P. Maundu (2014) Current knowledge on *Amaranthus* spp.: research avenues for improved nutritional value and yield in leafy amaranths in sub-Saharan Africa. Euphytica 197: 303–317.

Akubugwo, I.E., N.A. Obasi, G.C. Chinyere and A.E. Ugbogu (2007) Nutritional and chemical value of *Amaranthus hybridus* L. leaves from Afikpo, Nigeria. Afr. J. Biotechnol. 6: 2833–2839.

Assad, R., Z.A. Reshi, S. Jan and I. Rashid (2017) Biology of Amaranths. Bot. Rev. 83: 382–436.

Brenner, D., D. Baltensperger, P. Kulakow, J. Lehmann, R. Myers, M. Slabbert and B. Sleugh (2000) Genetic resources and breeding of Amaranthus. Plant Breed. Rev. 19: 227–285.

Bui, V.L., N.T. Do My, M. Jeanneau, S. Sadik, S. Tu, J. Vidal and T.V. Tran (1998) Rapid plant regeneration of a C_4 dicot species: *Amaranthus edulis*. Plant Sci. 132: 45–54.

Chase, M.W., R.S. Cowan, P.M. Hollingsworth, C. Van Den Berg, S. Madrinan, G. Petersen, O. Seberg, T. Jorgsensen, K.M. Cameron, M. Carine *et al.* (2007) A proposal for a standardised protocol to barcode all land plants. Taxon 56: 295–299.

Costea, M. and D.A. Demason (2001) Stem morphology and anatomy in *Amaranthus* L. (Amaranthaceae), Taxonomic significance. J. Torrey Bot. Soc. 128: 254–281.

Evanno, G., S. Regnaut and J. Goudet (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol. Ecol. 14: 2611–2620.

Excoffier, L. and H.E.L. Lischer (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol. Ecol. Resour. 10: 564–567.

Faircloth, B.C. (2008) msatcommander: detection of microsatellite repeat arrays and automated, locus-specific primer design. Mol. Ecol. Resour. 8: 92–94.

Grubben, G.J.H. and D.H. van Sloten (1981) Genetic resources of amaranths: A global plan of action. AGP:IBPGR/80/2. International Board for Plant Genetics Resources, Food and Agriculture Organizaion of the United Nations, Rome, Italy, p. 57.

Grubben, G.J.H. and O.A. Denton (2004) Plant resources of tropical Africa 2. Vegetables, PROTA Foundation, Netherlands, p. 668.

Guichoux, E., L. Lagache, S. Wagner, P. Chaumeil, P. Leger, O. Lepais, C. Lepoittevin, T. Malausa, E. Revardel, F. Salin *et al.* (2011) Current trends in microsatellite genotyping. Mol. Ecol. Resour. 11: 591–611.

Hall, G.H. and A.H. Patrinos (2012) Indigenous Peoples, Poverty, and Development. Cambridge Univ. Press, England, p. 406.

Hilu, K.W. and H.P. Liang (1997) The *matK* gene: Sequence variation and application in plant systematics. Am. J. Bot. 84: 830–839.

Hilu, K.W., T. Borsch, K. Muller, D.E. Soltis, P.S. Soltis, V. Savolainen, M.W. Chase, M.P. Powell, L.A. Alice, R. Evans *et al.* (2003) Angiosperm phylogeny based on matK sequence information. Am. J. Bot. 90: 1758–1776.

Jimenez-Aguilar, D.M. and M.A. Grusak (2017) Minerals, vitamin C, phenolics, flavonoids and antioxidant activity of *Amaranthus* leafy vegetables. J. Food Compost. Anal. 58: 33–39.

Khaing, A.A., K.T. Moe, J.W. Chung, H.J. Baek and Y.J. Park (2013) Genetic diversity and population structure of the selected core set in *Amaranthus* using SSR markers. Plant Breed. 132: 165–173.

Khandaker, L., M.B. Ali and S. Oba (2008) Total polyphenol and antioxidant activity of red amaranth (*Amaranthus tricolor* L.) as affected by different sunlight level. J. Japan. Soc. Hort. Sci 77: 395–401.

Kietlinski, K.D., F. Jimenez, E.N. Jellen, P.J. Maughan, S.M. Smith and D.B. Pratt (2014) Relationships between the weedy *Amaranthus hybridus* (Amaranthaceae) and the grain amaranths. Crop Sci. 54: 220–228.

Kumar, P., V.K. Gupta, A.K. Misra, D.R. Modi and B.K. Pandey (2009) Potential of molecular markers in plant biotechnology. Plant Omics 2: 141–162.

Kumar, S., G. Stecher and K. Tamura (2016) MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for bigger datasets. Mol. Biol. Evol. 33: 1870–1874.

Lahaye, R., M. Van Der Bank, D. Bogarin, J. Warner, F. Pupulin, G. Gigot, O. Maurin, S. Duthoit, T.G. Barraclough and V. Savolainen (2008) DNA barcoding the floras of biodiversity hotspots. Proc. Natl. Acad. Sci. USA 105: 2923–2928.

Le, P.H., Q.T. Vuong, V.H. Tan and L. Ninh (1991) Histoty of Vietnam, VietNam Education Publishing House Hanoi, pp. 162–167.

Lee, J.R., G.Y. Hong, A. Dixit, J.W. Chung, K.H. Ma, J.H. Lee, H.K. Kang, Y.H. Cho, J.G. Gwag and Y.J. Park (2008) Characterization of microsatellite loci developed for *Amaranthus hypochondriacus* and their cross-amplifications in wild species. Conserv. Genet. 9: 243–246.

Liu, K.J. and S.V. Muse (2005) PowerMarker: an integrated analysis environment for genetic marker analysis. Bioinformatics 21: 2128–2129.

Mallory, M.A., R.V. Hall, A.R. Mcnabb, D.B. Pratt, E.N. Jellen and P.J. Maughan (2008) Development and characterization of microsatellite markers for the grain amaranths. Crop Sci. 48: 1098–1106.

Mosyakin, S.L. and K.R. Robertson (1996) New infrageneric taxa and combinations in *Amaranthus* (Amaranthaceae). Ann. Bot. Fenn. 33: 275–281.

Muller, K.F., T. Borsch and K.W. Hilu (2006) Phylogenetic utility of rapidly evolving DNA at high taxonomical levels: Contrasting *matK*, *trnT-F*, and *rbcL* in basal angiosperms. Mol. Phylogenet. Evol. 41: 99–117.

Ooi, K., Y. Endo, J. Yokoyama and N. Murakami (1995) Useful primer designs to amplify DNA fragment of the plastid gene *matK* from angiosperm plants. J. Jap. Bot. 70: 328–333.

Peakall, R. and P.E. Smouse (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Mol. Ecol. Notes 6: 288–295.

Peakall, R. and P.E. Smouse (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. Bioinformatics 28: 2537–2539.

Pham, H.H. (1999) An illustrated flora of Vietnam I, Tre Publishing House, Ho Chi Minh City, p. 991.

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

Rastogi, A. and S. Shukla (2013) Amaranth: A new millennium crop of nutraceutical values. Crit. Rev. Food Sci. Nutr. 53: 109–125.

Sauer, J.D. (1967) The grain amaranths and their relatives: A revised taxonomic and geographic survey. Ann. Mo. Bot. Gard. 54: 103–137.

Sauer, J.D. (1993) Historical geography of crop plants: A select roster, CRC Press, Boca Raton, p. 320.

Shimizu, T. and K. Yano (2011) A post-labeling method for multiplexed and multicolored genotyping analysis of SSR, indel and SNP markers in single tube with bar-coded split tag (BStag). BMC Res. Notes 4: 161.

Shukla, S., A. Bhargava, A. Chatterjee, A. Srivastava and S.P. Singh (2006) Genotypic variability in vegetable amaranth (*Amaranthus tricolor* L for foliage yield and its contributing traits over successive cuttings and years. Euphytica 151: 103–110.

Shukla, S., A. Bhargava, A. Chatterjee, A.C. Pandey and B.K. Mishra (2010) Diversity in phenotypic and nutritional traits in vegetable amaranth (*Amaranthus tricolor*), a nutritionally underutilised crop. J. Sci. Food Agric. 90: 139–144.

Silva, P.I., A.M. Martins, E.G. Gouvea, M. Pessoa-Filho and M.E. Ferreira (2013) Development and validation of microsatellite markers for *Brachiaria ruziziensis* obtained by partial genome assembly of Illumina single-end reads. BMC Genomics 14: 17.

Singh, B., S. Pandey and J. Kumar (2013) A comparative study of Inter Simple Sequence Repeat (ISSR), Random Amplified Polymorphic DNA (RAPD) and Simple Sequence Repeat (SSR) loci in assessing genetic diversity in *Amaranthus*. Indian J. Genet. Plant Breed. 73: 411–418.

Srivastava, R. (2017) An updated review on phyto-pharmacological and pharmacognostical profile of *Amaranthus tricolor*: A herb of nutraceutical potentials. Pharma Innov. 6: 124–129.

Stetter, M.G. and K.J. Schmid (2017) Analysis of phylogenetic relationships and genome size evolution of the *Amaranthus* genus using GBS indicates the ancestors of an ancient crop. Mol. Phylogenet. Evol. 109: 80–92.

Suresh, S., J.W. Chung, G.T. Cho, J.S. Sung, J.H. Park, J.G. Gwag and H.J. Baek (2014) Analysis of molecular genetic diversity and population structure in *Amaranthus* germplasm using SSR markers. Plant Biosyst. 148: 635–644.

Taheri, S., T.L. Abdullah, M.R. Yusop, M.M. Hanafi, M. Sahebi, P. Azizi and R.R. Shamshiri (2018) Mining and development of novel SSR markers using next generation sequencing (NGS) data in Plants. Molecules 23: 20.

Varshney, R.K., A. Graner and M.E. Sorrells (2005) Genic microsatellite markers in plants: features and applications. Trends Biotechnol. 23: 48–55.

Vieira, M.L.C., L. Santini, A.L. Diniz and C.D. Munhoz (2016) Microsatellite markers: what they mean and why they are so useful. Genet. Mol. Biol. 39: 312–328.

Viljoen, E., D.A. Odeny, M.P.A. Coetzee, D.K. Berger and D.J.G. Rees (2018) Application of chloroplast phylogenomics to resolve species relationships within the plant genus *Amaranthus*. J. Mol. Evol. 86: 216–239.

Wang, X.Q. and J.Y. Park (2013) Comparison of genetic diversity among amaranth accessions from South and Southeast Asia using SSR markers. Korean J. Medicinal Crop Sci. 21: 220–228.

Waselkov, K.E., A.S. Boleda and K.M. Olsen (2018) A phylogeny of the genus *Amaranthus* (Amaranthaceae) based on several low-copy nuclear loci and chloroplast regions. Syst. Bot. 43: 439–458.

Yu, J., J.H. Xue and S.L. Zhou (2011) New universal *matK* primers for DNA barcoding angiosperms. J. Syst. Evol. 49: 176–181.