Genetic diversity analysis and marker-trait associations in *Amaranthus* species

Norain Jamalluddin¹, Festo J. Massawe¹, Sean Mayes¹,²,³, Wai Kuan Ho¹,³, Rachael C. Symonds⁴*

¹ Future Food Beacon, School of Biosciences, University of Nottingham Malaysia, Jalan Broga, Semenyih, Selangor, Malaysia, ² Plant and Crop Sciences, Biosciences, University of Nottingham, Sutton Bonington Campus, Loughborough, Leicestershire, United Kingdom, ³ Crops for the Future (UK) CIC, NIAB, Cambridge, United Kingdom, ⁴ Liverpool John Moores University, School of Biological and Environmental Sciences, Liverpool, Merseyside, United Kingdom

* R.C.Symonds@ljmu.ac.uk

Abstract

Amaranth (*Amaranthus* spp.) is a highly nutritious, underutilized vegetable and pseudocereal crop. It possesses diverse abiotic stress tolerance traits, is genetically diverse and highly phenotypically plastic, making it an ideal crop to thrive in a rapidly changing climate. Despite considerable genetic diversity there is a lack of detailed characterization of germplasm or population structures. The present study utilized the DArTSeq platform to determine the genetic relationships and population structure between 188 amaranth accessions from 18 agriculturally important vegetable, grain, and weedy species. A total of 74,303 SNP alleles were generated of which 63,821 were physically mapped to the genome of the grain species *A. hypochondriacus*. Population structure was inferred in two steps. First, all 188 amaranth accessions comprised of 18 species and second, only 120 *A. tricolor* accessions. After SNP filtering, a total of 8,688 SNPs were generated on 181 amaranth accessions of 16 species and 9,789 SNPs generated on 118 *A. tricolor* datasets. Both SNP datasets produced three major sub-populations (K = 3) and generate consistent taxonomic classification of the amaranth sub-genera (*Amaranthus*, *Amaranthus Acnida* and *Amaranthus albersia*), although the accessions were poorly demarcated by geographical origin and morphological traits. *A. tricolor* accessions were well discriminated from other amaranth species. A genome-wide association study (GWAS) of 10 qualitative traits revealed an association between specific phenotypes and genetic variants within the genome and identified 22 marker trait associations (MTAs) and 100 MTAs (P<0.01, P<0.001) on 16 amaranth species and 118 *A. tricolor* datasets, respectively. The release of SNP markers from this panel has produced invaluable preliminary genetic information for phenotyping and cultivar improvement in amaranth species.

Introduction

Climate predictions indicate that the agriculture sector in many parts of the world will be subjected to increasingly detrimental weather conditions such as droughts and elevated
temperatures, directly impacting global food supply chains. A strategy to mitigate climate-related agricultural losses is to diversify the food basket with a wide range of underutilized crop species with increased abiotic stress tolerance traits [1]. Amaranth (Amaranthus spp.), an ancient, nutrient-dense and climate-smart crop has high degree of genetic variation, environmental adaptability and phenotypic plasticity [2, 3]. Amaranth belongs to the Amaranthaceae family and is a C₄ dicotyledonous plant [4]. It consists of approximately 60–70 species grouped into three sub-generic; Amaranthus Albersia (vegetable amaranth), Amaranthus Amaranthus (cultivated grain amaranth) and Amaranthus Acnida (weedy amaranth) [5].

*Amaranthus tricolor* is a leafy vegetable amaranth species, widely cultivated in South Asia and Africa [6, 7], and is an excellent source of vitamins, protein, carotenoid, minerals and antioxidants, greater than other leafy vegetables such as lettuce and spinach [2, 8]. *A. tricolor* has the capacity to alter its physiological characteristics in response to environmental changes, for instance, increasing transpiration efficiency [9] and accumulation of compatible solutes such as proline, in response to drought stress [10]. It also had high genetic and phenotypic diversity which may provide an excellent opportunity for varietal development with increased drought tolerance characteristics [11–13].

Correct genotypic identification and preservation of genetic variation in amaranth is important to maintain ecotypes with desired traits useful for breeding programmes. The assembly of very high-quality grain amaranth, *Amaranthus hypochondriacus* (“Plainsman” cultivar) sequence genome by [14] allows anchoring of genotyping-by-sequencing (GBS) markers for all the SNP loci and allele sequences discovered, and GBS has proven to be the most efficient method to evaluate genetic diversity of grain amaranth as well as to validate the phylogeny of the genus [15–17]. This genome assembly was used as a reference genome for an annotation framework and gene discovery of MYB-like transcription factor genes that regulate the betalain red pigment pathway, which gives rise to stem and seed colour variations [18] through traditional bi-parental mapping [14] and now through genome-wide association studies (GWAS) [19]. More recently, this plainsman reference genome together with low-coverage PacBio reads and the contigs of amaranth draft genome [20] were used to assemble *A. hypochondriacus* (A.hyp_K_white), a landrace cultivated in India [21]. This assembly offers a better reference genome for the improvement of grain and vegetable amaranth crops in South Asia as it is genetically closer to most landraces and accessions originated from India and South Asia.

Nevertheless, vegetable species of amaranth have been less studied by molecular means than pseudo-cereal grain amaranths and weed species, especially when both are phylogenically related and the domestication events separating them have been revealed [15, 19, 22, 23]. Limited knowledge of the genetic diversity in these leafy vegetable amaranth species and the lack of availability of suitable molecular markers hamper breeding efforts. Cultivar development and improvement relies on access to a well characterised, genetically diverse pool of material and so a comprehensive knowledge of these genetic relationships is essential. To date, there is only one molecular study that exploits a large number of *A. tricolor* accessions using simple sequence repeat (SSR) and *matK* protein-coding chloroplast gene, which concluded that the genetic diversity in Vietnamese amaranths was established by dispersal events mainly from East Asia and adaptation to local environments [24]. While the amaranth marker studies have been useful for evolutionary and phylogenetic studies, further germplasm characterization and marker validation is needed.

GBS offers a number of potential advantages to SSR markers; it is more practical, inexpensive and has driven genotyping to be applied for non-model organisms [25, 26]. DArTSeq™ technology based on GBS methods is a platform developed by Diversity Arrays Technology Pty Ltd. (Canberra, Australia) for high-throughput genotyping via an intelligent selection of
genome fraction by targeting active genes and low copy DNA areas \cite{27, 28}. This present study is the first to utilize the DArTSeq platforms in amaranth to determine the genetic relationships and population structure between 188 amaranth accessions from 18 agronomically important vegetable, grain, and weedy species. This study also aimed to investigate the genetic relationship among a numerically larger group of \textit{A. tricolor} accessions. The development of SNP markers from this panel has allowed a GWAS analysis on morphological traits such as shape, size, and colour of the leaf, stem, and inflorescence. These traits are fast and easy to assess for direct use by farmers and are of great help to plant breeders when selecting potential parental lines \cite{12, 29}. This will facilitate understanding of the genetic bases and dissection of complex genes controlling economic traits such as drought tolerance and provides useful information on the degree of genetic variation and its correlations with agronomic traits.

**Materials and methods**

**Plant materials, growing conditions and morphological assessment**

A total of 188 amaranth accessions, comprising 18 species originating worldwide were used for genetic diversity analysis. Out of 188 accessions, 131 accessions were obtained from the World Vegetable Center Genebank, Taiwan (AVRDC), 52 accessions from the United State Department of Agriculture Genebank (USDA) and five commercial varieties were included as checks, of which three African varieties from East-West Seed, Thailand, were included and two local varieties from Serbajadi Gardening, Malaysia (Table 1).

A single plant of each accession was grown under shade-house conditions at University of Nottingham Malaysia (latitude 2.940°N, longitude 101.8740°E), with an average of 36°C daytime temperature, 28°C night temperature and 66% relative humidity. Plants were grown in a 16 x 12.5 x 14.5 cm plastic pot containing 2 kg black compost (Holland peat, Malaysia), irrigated daily to field capacity and at 3 weeks old, 3 g of 15N: 15P: 15K fertilizer was applied once to individual pots.

Ten qualitative traits including leaf, petiole and stem pigmentation, growth habit, branching index and, leaf shape and margin were recorded at 7 weeks post-emergence and terminal inflorescence color, shape and attitude were recorded when all accessions had fully set (at 11 weeks post emergence) using AVRDC descriptors (https://avrdc.org/seed/) (Table 1). Young leaf material was collected and snap frozen in liquid N\textsubscript{2} and stored at -80°C for DNA analysis.

**DNA extraction and DArTSeq genotyping**

Total genomic DNA of 188 amaranth accessions was isolated from young leaves using a Qiagen DNeasy plant DNA extraction kit (Qiagen, USA) and DNA quality and quantity was evaluated using a Nanodrop spectrophotometer (Thermo Scientific, USA). The DNA concentration was adjusted within the range of 50-100ng/μl. 2 μg of high molecular weight and good quality DNA per sample was sent to Diversity Arrays Technology Pty Ltd, Canberra, Australia for DArTSeq analysis.

In brief, DArTSeq technology relies on the combination of a complexity reduction method to enrich genomic representations, followed by next-generation sequencing by HiSeq2000 (Illumina, USA), as described by Kilian et al. \cite{27}. In this study, a combination of a rare cutting methylation-sensitive restriction enzyme (RE) \textit{PstI} with secondary frequently cutting RE \textit{MseI} were selected to optimize the locus coverage, reproducibility and polymorphisms. The \textit{PstI}-compatible adapter consists of the Illumina flow cell attachment sequence, sequencing primer and a ‘staggered’ of varying length barcode region. The reverse adapter consists of Illumina flow cell attachment region and \textit{MseI} overhang sequence. The ligated fragments with both a \textit{PstI} and \textit{MseI} adapter were amplified via polymerase chain reaction (PCR) with a programme
Table 1. List of 188 amaranth accessions and their morphological traits observed under shade-house conditions.

| Entry | Accessions ID | Species | Origin country | Germplasm | Growth habit | Branching index | Stem colour | Leaf colour | Petiole colour | Inflorescence colour | Leaf shape | Leaf margin | Terminal inflorescence shape | Terminal inflorescence attitude |
|-------|---------------|---------|----------------|------------|--------------|----------------|-------------|-------------|----------------|----------------------|----------------|-------------|-------------------------------|-------------------------------|
| 1     | AV ATR       | VI044435 | Atropurpurus   | Indonesia  | AVRDC       | Erect          | Along the stem | Pink        | Normal green | Green               | Green          | Lamelolate  | Entire                        | Short branches                 |
| 2     | AV GRA       | VI036225 | Graecizans     | Hungary    | AVRDC       | Erect          | Along the stem | Purple/Pink | Normal green | White               | Pink           | Lamelolate  | Entire                        | Other                        |
| 3     | AV GRA SIL   | VI044403 | Graecizans     | India      | AVRDC       | Erect          | Many branches  | Purple/Pink | Normal green | Pin/P Green         | Pink           | Cuneate     | Entire                        | Other                        |
| 4     | AV GRA ASC   | VI048388 | Graecizans     | India      | AVRDC       | Erect          | Many branches  | Purple/Pink | Normal green | Pin/P Green         | Green          | Lamelolate  | Other                         | Erect                        |
| 5     | AV MAN       | VI044427 | Mantegazzianus | USA        | AVRDC       | Erect          | Few branches   | Green        | Normal green | Green               | Lamelolate  | Entire      | Club-shaped                   | Erect                        |
| 6     | AV BLITO     | VI036227 | Bitoides       | Hungary    | AVRDC       | Prostrate      | Many branches  | Purple/Pink | Normal green | Green               | Lamelolate  | Entire      | Other                         | Erect                        |
| 7     | AV LEU       | VI044445 | Leucocarpus    | India      | AVRDC       | Erect          | Along the stem | Green        | Normal green | Green               | Green          | Elliptical  | Entire                        | Spike (dense)                |
| 8     | AV PAL       | VI044473 | Palmeri        | Senegal    | AVRDC       | Erect          | Few branches   | Purple/Pink | Normal green | Green               | Green          | Green       | Elliptical                    | Entire                        |
| 9     | AV RET 1     | VI048310 | Retroflexus    | Viet Nam   | AVRDC       | Erect          | Few branches   | Purple/Pink | Normal green | Margins/ Vein      | Red            | Ovariate    | Entire                        | Spike (dense)                |
| 10    | AV RET 2     | VI048311 | Retroflexus    | Viet Nam   | AVRDC       | Erect          | Along the stem | Green        | Normal green | Green               | Green          | Lamelolate  | Entire                        | Spike (dense)                |
| 11    | AV RET 3     | VI033461 | Retroflexus    | Venezuela  | AVRDC       | Erect          | Many branches  | Purple/Pink | Chlorotic stripe | Green       | Rhombic     | Entire                        | Undulate                    |
| 12    | AV RET 4     | VI048391 | Retroflexus    | Viet Nam   | AVRDC       | Erect          | Along the stem | Purple/Pink | Central spot  | Purple           | Red            | Rhombic     | Entire                        | Spike (dense)                |
| 13    | US RET 1     | Ames 26236 | Retroflexus  | China      | USDA        | Erect          | Along the stem | Pink         | Normal green | Green               | Green          | Elliptical  | Entire                        | Long branches                |
| 14    | AV SPI 11    | VI044410 | Spinousus      | Puerto Rico| AVRDC       | Erect          | Along the stem | Green        | Normal green | Green               | Green          | Lamelolate  | Entire                        | Short branches               |
| 15    | AV SPI 14    | VI040944 | Spinousus      | Thailand   | AVRDC       | Erect          | Many branches  | Green        | Normal green | Green               | Green          | Rhombic     | Entire                        | Short branches               |
| 16    | AV SPI 15    | VI048723 | Spinousus      | Thailand   | AVRDC       | Erect          | Along the stem | Green        | One stripe     | Green       | Green       | Rhombic                        | Undulate                    |
| 17    | AV SPI 16    | VI046123 | Spinousus      | Laos       | AVRDC       | Erect          | Many branches  | Purple/Pink | Normal green | Green               | Lamelolate  | Entire      | Spike (dense)                 | Erect                        |
| 18    | AV SP 1      | VI050233 | Sp             | Taiwan     | AVRDC       | Erect          | Many branches  | Green        | Others        | Green               | Green          | Rhombic     | Undulate                      | Other                        |
| 19    | AV SP 2      | VI046930 | Sp             | Thailand   | AVRDC       | Erect          | Along the stem | Purple/Pink | Normal green | Green               | Other         | Elliptical  | Entire                        | Short branches               |
| 20    | AV SP 3      | VI054799 | Sp             | Laos       | AVRDC       | Erect          | Along the stem | Green        | Normal green | Green               | Green          | Cuneate     | Undulate                      | Short branches               |
| 21    | AV SP 4      | VI033471 | Sp             | Malaysia   | AVRDC       | Erect          | Many branches  | Green        | Normal green | Green               | Green          | Lamelolate  | Entire                        | Other                        |
| 22    | AV SP 5      | VI044440 | Sp             | India      | AVRDC       | Erect          | Many branches  | Purple/Marginal veins | Purple      | Greens/Red     | Rhombic                        | Undulate                    | Other                    | Erect                        |
| 23    | AV SP 6      | VI056563 | Sp             | Bangladesh | AVRDC       | Erect          | Many branches  | Purple/Pink | Entire lamina  | Dark purple | Red           | Rhombic                       | Spike (dense)               |
| 24    | AV SP 7      | VI056560 | Sp             | Bangladesh | AVRDC       | Erect          | Many branches  | Green        | Dark green    | Green               | Green          | Rhombic     | Entire                        | Spike (dense)               |
| 25    | AV VIR 1     | VI049898 | Viridis        | Thailand   | AVRDC       | Prostrate      | Many branches  | Purple/Pink | Normal green | Green               | Pink           | Rhombic     | Entire                        | Short branches               |
| 26    | AV VIR 4     | VI049001 | Viridis        | Thailand   | AVRDC       | Erect          | Along the stem | Purple/Pink | Normal green | Purple               | Green          | Cuneate     | Entire                        | Short branches               |
| 27    | AV VIR 6     | VI046976 | Viridis        | Thailand   | AVRDC       | Erect          | Many branches  | Purple/Pink | Normal green | Green               | Pink           | Ovariate    | Entire                        | Short branches               |
| Entry | Accessions ID | Species | Origin country | Germplasm | Growth habit | Branching | Stem colour | Leaf colour | Petiole colour | Inflorescence | Leaf shape | Leaf margin | Terminal habit | Leaf size index | Inference source shape | Terminal inflorescence shape | Leaf lamina | Inference one cm area | Leaf lamina | Inference one cm area |
|-------|---------------|---------|----------------|-----------|-------------|-----------|-------------|-------------|---------------|--------------|-----------|-------------|----------------|----------------|--------------------------|--------------------------|-------------|------------------|-------------|------------------|
| 28    | AV VIR 9     | Viridis | Malaysia       | AVRDC     | Erect       | Few branches | Green       | Normal      | Green         | Long branches | Erect     | Pink         | Entire          | Entire         | Entire          | Entire          | Entire       | Entire           | Entire       | Entire           |
| 29    | AV VIR 12    | Viridis | Laos           | AVRDC     | Erect       | Along the stem | Purple/Pink | Normal      | Green         | Long branches | Erect     | Pink         | Entire          | Entire         | Entire          | Entire          | Entire       | Entire           | Entire       | Entire           |
| 30    | AV CRU 1     | Cruentus | Austria        | AVRDC     | Erect       | Along the stem | Purple/Pink | Normal      | Pink          | Short branches | Drooping  | Green        | Entire          | Entire         | Entire          | Entire          | Entire       | Entire           | Entire       | Entire           |
| 31    | AV CRU 2     | Cruentus | Germany        | AVRDC     | Erect       | Along the stem | Purple/Pink | Normal      | Pink          | Short branches | Drooping  | Green        | Entire          | Entire         | Entire          | Entire          | Entire       | Entire           | Entire       | Entire           |
| 32    | AV CRU 3     | Cruentus | Ethiopia       | AVRDC     | Erect       | Along the stem | Green       | Normal      | Green         | Short branches | Drooping  | Green        | Entire          | Entire         | Entire          | Entire          | Entire       | Entire           | Entire       | Entire           |
| 33    | AV CRU 4     | Cruentus | Austria        | AVRDC     | Erect       | Along the stem | Purple/Pink | Normal      | Pink          | Short branches | Drooping  | Green        | Entire          | Entire         | Entire          | Entire          | Entire       | Entire           | Entire       | Entire           |
| 34    | AV CRU 5     | Cruentus | Vietnam        | AVRDC     | Erect       | Along the stem | Green       | Normal      | Green         | Short branches | Drooping  | Green        | Entire          | Entire         | Entire          | Entire          | Entire       | Entire           | Entire       | Entire           |
| 35    | AV CRU 6     | Cruentus | Sudan          | AVRDC     | Erect       | Along the stem | Green       | Normal      | Purple         | Short branches | Drooping  | Green        | Entire          | Entire         | Entire          | Entire          | Entire       | Entire           | Entire       | Entire           |
| 36    | AV CRU 7     | Cruentus | Zimbabwe       | AVRDC     | Erect       | Along the stem | Green       | Normal      | Purple         | Short branches | Drooping  | Green        | Entire          | Entire         | Entire          | Entire          | Entire       | Entire           | Entire       | Entire           |
| 37    | AV CRU 8     | Cruentus | Tanzania       | AVRDC     | Erect       | Along the stem | Green       | Normal      | Green         | Short branches | Drooping  | Green        | Entire          | Entire         | Entire          | Entire          | Entire       | Entire           | Entire       | Entire           |
| 38    | AV CRU 9     | Cruentus | Cambodia       | AVRDC     | Erect       | Along the stem | Green       | Normal      | Pink          | Short branches | Drooping  | Green        | Entire          | Entire         | Entire          | Entire          | Entire       | Entire           | Entire       | Entire           |
| 39    | AV CRU 10    | Cruentus | Thailand       | AVRDC     | Erect       | Along the stem | Green       | Normal      | Purple         | Short branches | Drooping  | Green        | Entire          | Entire         | Entire          | Entire          | Entire       | Entire           | Entire       | Entire           |
| 40    | AV CRU 11    | Cruentus | Vietnam        | AVRDC     | Erect       | Along the stem | Green       | Normal      | Pink          | Short branches | Drooping  | Green        | Entire          | Entire         | Entire          | Entire          | Entire       | Entire           | Entire       | Entire           |
| 41    | AV CRU 12    | Cruentus | Indonesia      | AVRDC     | Erect       | Along the stem | Green       | Normal      | Pink          | Short branches | Drooping  | Green        | Entire          | Entire         | Entire          | Entire          | Entire       | Entire           | Entire       | Entire           |
| 42    | AV CRU 13    | Cruentus | Nepal          | AVRDC     | Erect       | Along the stem | Green       | Normal      | Purple         | Short branches | Drooping  | Green        | Entire          | Entire         | Entire          | Entire          | Entire       | Entire           | Entire       | Entire           |
| 43    | AV CRU 14    | Cruentus | Vietnam        | AVRDC     | Erect       | Along the stem | Green       | Normal      | Pink          | Short branches | Drooping  | Green        | Entire          | Entire         | Entire          | Entire          | Entire       | Entire           | Entire       | Entire           |
| 44    | AV CRU 15    | Cruentus | Laos           | AVRDC     | Erect       | Along the stem | Green       | Normal      | Purple         | Short branches | Drooping  | Green        | Entire          | Entire         | Entire          | Entire          | Entire       | Entire           | Entire       | Entire           |
| 45    | AV CRU 16    | Cruentus | Vietnam        | AVRDC     | Erect       | Along the stem | Green       | Normal      | Pink          | Short branches | Drooping  | Green        | Entire          | Entire         | Entire          | Entire          | Entire       | Entire           | Entire       | Entire           |
| 46    | AV CRU 17    | Cruentus | Vietnam        | AVRDC     | Erect       | Along the stem | Green       | Normal      | Purple         | Short branches | Drooping  | Green        | Entire          | Entire         | Entire          | Entire          | Entire       | Entire           | Entire       | Entire           |
| 47    | AV CRU 18    | Cruentus | Vietnam        | AVRDC     | Erect       | Along the stem | Green       | Normal      | Pink          | Short branches | Drooping  | Green        | Entire          | Entire         | Entire          | Entire          | Entire       | Entire           | Entire       | Entire           |
| 48    | AV CRU 19    | Cruentus | Vietnam        | AVRDC     | Erect       | Along the stem | Green       | Normal      | Purple         | Short branches | Drooping  | Green        | Entire          | Entire         | Entire          | Entire          | Entire       | Entire           | Entire       | Entire           |
| 49    | AV CRU 20    | Cruentus | Vietnam        | AVRDC     | Erect       | Along the stem | Green       | Normal      | Pink          | Short branches | Drooping  | Green        | Entire          | Entire         | Entire          | Entire          | Entire       | Entire           | Entire       | Entire           |
| 50    | AV CRU 21    | Cruentus | Vietnam        | AVRDC     | Erect       | Along the stem | Green       | Normal      | Purple         | Short branches | Drooping  | Green        | Entire          | Entire         | Entire          | Entire          | Entire       | Entire           | Entire       | Entire           |
| 51    | AV CRU 22    | Cruentus | Vietnam        | AVRDC     | Erect       | Along the stem | Green       | Normal      | Pink          | Short branches | Drooping  | Green        | Entire          | Entire         | Entire          | Entire          | Entire       | Entire           | Entire       | Entire           |
| 52    | AV CRU 23    | Cruentus | Vietnam        | AVRDC     | Erect       | Along the stem | Green       | Normal      | Purple         | Short branches | Drooping  | Green        | Entire          | Entire         | Entire          | Entire          | Entire       | Entire           | Entire       | Entire           |
| 53    | AV CRU 24    | Cruentus | Vietnam        | AVRDC     | Erect       | Along the stem | Green       | Normal      | Pink          | Short branches | Drooping  | Green        | Entire          | Entire         | Entire          | Entire          | Entire       | Entire           | Entire       | Entire           |
| 54    | AV CRU 25    | Cruentus | Vietnam        | AVRDC     | Erect       | Along the stem | Green       | Normal      | Purple         | Short branches | Drooping  | Green        | Entire          | Entire         | Entire          | Entire          | Entire       | Entire           | Entire       | Entire           |
| Entry | Accessions ID | Species | Origin country | Germplasm | Growth habit | Branching | Leaf colour | Petiole colour | Inflorescence | Leaf shape | Leaf margin | Terminal habit | Inference on community structure | Terminal inflorescence shape | Expression of traits | Expression on other traits |
|-------|---------------|---------|----------------|-----------|-------------|-----------|-------------|---------------|--------------|------------|-------------|----------------|-----------------------------|--------------------------|------------------------|--------------------------|
| 55    | AV I1114      | Blitum  | Malaysia       | AVRDC     | Erect       | Few branches | Purple | Dark | Elliptical | Entire | Spike (dense) | Short branches | Effaceous | Effaceous | Effaceous |
| 56    | AV I1111      | Blitum  | Thailand       | AVRDC     | Erect       | Few branches | Purple | White | Official | Entire | Spike (dense) | Short branches | Effaceous | Effaceous | Effaceous |
| 57    | AV I1110      | Blitum  | Vietnam        | AVRDC     | Erect       | Few branches | Purple | Red  | Elliptical | Entire | Spike (dense) | Short branches | Effaceous | Effaceous | Effaceous |
| 58    | AV I1112      | Blitum  | India          | AVRDC     | Erect       | Many branches | Purple | White | Official | Entire | Spike (dense) | Short branches | Effaceous | Effaceous | Effaceous |
| 59    | AV I1113      | Blitum  | Cambodia       | AVRDC     | Erect       | Many branches | Purple | Green | Ovatainate | Undulate | Spike (dense) | Short branches | Effaceous | Effaceous | Effaceous |
| 60    | AV I1114      | Blitum  | Korea          | AVRDC     | Prostrate   | Along the stem | Purple | Green | Green  | Central | Spike (dense) | Short branches | Effaceous | Effaceous | Effaceous |
| 61    | AV I1115      | Blitum  | Vietnam        | AVRDC     | Prostrate   | Central spot | Purple | Red  | Green  | Entire | Spike (dense) | Short branches | Effaceous | Effaceous | Effaceous |
| 62    | AV DUB 1      | Dubius  | Viet Nam       | AVRDC     | Erect       | Along the stem | Purple | Red  | Ovatainate | Entire | Spike (dense) | Short branches | Effaceous | Effaceous | Effaceous |
| 63    | AV DUB 2      | Dubius  | Viet Nam       | AVRDC     | Erect       | Along the stem | Purple | Red  | Ovatainate | Entire | Spike (dense) | Short branches | Effaceous | Effaceous | Effaceous |
| 64    | AV DUB 3      | Dubius  | Thailand       | AVRDC     | Erect       | Along the stem | Purple | Red  | Ovatainate | Entire | Spike (dense) | Short branches | Effaceous | Effaceous | Effaceous |
| 65    | AV DUB 4      | Dubius  | Tanzania       | AVRDC     | Erect       | Along the stem | Purple | Red  | Ovatainate | Entire | Spike (dense) | Short branches | Effaceous | Effaceous | Effaceous |
| 66    | AV DUB 5      | Dubius  | Tanzania       | AVRDC     | Erect       | Along the stem | Purple | Red  | Ovatainate | Entire | Spike (dense) | Short branches | Effaceous | Effaceous | Effaceous |
| 67    | AV DUB 6      | Dubius  | Tanzania       | AVRDC     | Erect       | Along the stem | Purple | Red  | Ovatainate | Entire | Spike (dense) | Short branches | Effaceous | Effaceous | Effaceous |
| 68    | AV DUB 7      | Dubius  | Tanzania       | AVRDC     | Erect       | Along the stem | Purple | Red  | Ovatainate | Entire | Spike (dense) | Short branches | Effaceous | Effaceous | Effaceous |
| 69    | AV DUB 8      | Dubius  | Tanzania       | AVRDC     | Erect       | Along the stem | Purple | Red  | Ovatainate | Entire | Spike (dense) | Short branches | Effaceous | Effaceous | Effaceous |
| 70    | AV DUB 9      | Dubius  | Tanzania       | AVRDC     | Erect       | Along the stem | Purple | Red  | Ovatainate | Entire | Spike (dense) | Short branches | Effaceous | Effaceous | Effaceous |
| 71    | AV DUB 10     | Dubius  | Tanzania       | AVRDC     | Erect       | Along the stem | Purple | Red  | Ovatainate | Entire | Spike (dense) | Short branches | Effaceous | Effaceous | Effaceous |
| 72    | AV DUB 11     | Dubius  | Tanzania       | AVRDC     | Erect       | Along the stem | Purple | Red  | Ovatainate | Entire | Spike (dense) | Short branches | Effaceous | Effaceous | Effaceous |
| 73    | AV DUB 12     | Dubius  | Tanzania       | AVRDC     | Erect       | Along the stem | Purple | Red  | Ovatainate | Entire | Spike (dense) | Short branches | Effaceous | Effaceous | Effaceous |
| 74    | AV DUB 13     | Dubius  | Tanzania       | AVRDC     | Erect       | Along the stem | Purple | Red  | Ovatainate | Entire | Spike (dense) | Short branches | Effaceous | Effaceous | Effaceous |
| 75    | AV DUB 14     | Dubius  | Tanzania       | AVRDC     | Erect       | Along the stem | Purple | Red  | Ovatainate | Entire | Spike (dense) | Short branches | Effaceous | Effaceous | Effaceous |
| 76    | AV DUB 15     | Dubius  | Tanzania       | AVRDC     | Erect       | Along the stem | Purple | Red  | Ovatainate | Entire | Spike (dense) | Short branches | Effaceous | Effaceous | Effaceous |
| 77    | AV DUB 16     | Dubius  | Tanzania       | AVRDC     | Erect       | Along the stem | Purple | Red  | Ovatainate | Entire | Spike (dense) | Short branches | Effaceous | Effaceous | Effaceous |
| 78    | AV DUB 17     | Dubius  | Tanzania       | AVRDC     | Erect       | Along the stem | Purple | Red  | Ovatainate | Entire | Spike (dense) | Short branches | Effaceous | Effaceous | Effaceous |
| 79    | AV DUB 18     | Dubius  | Tanzania       | AVRDC     | Erect       | Along the stem | Purple | Red  | Ovatainate | Entire | Spike (dense) | Short branches | Effaceous | Effaceous | Effaceous |
| 80    | AV DUB 19     | Dubius  | Tanzania       | AVRDC     | Erect       | Along the stem | Purple | Red  | Ovatainate | Entire | Spike (dense) | Short branches | Effaceous | Effaceous | Effaceous |
| 81    | AV DUB 20     | Dubius  | Tanzania       | AVRDC     | Erect       | Along the stem | Purple | Red  | Ovatainate | Entire | Spike (dense) | Short branches | Effaceous | Effaceous | Effaceous |
| Entry | ID | Species | Origin | Country | Growth habit | Leaf colour | Leaf shape | Leaf margin | Terminal inflorescence | Terminal inflorescence shape | Prickles | Spike | Spike (dense) |
|-------|----|---------|--------|---------|-------------|-------------|------------|-------------|------------------------|-------------------------------|----------|-------|-------------|
| 82    | AV TRI 15 | Zeya | Malaysia | E-WEST | Erect | Along the stem | Green | Ovatainate | Entire | Spike | Spike (dense) | Other | Erect | Erect |
| 83    | AV TRI 16 | Zeya | Malaysia | E-WEST | Erect | Along the stem | Green | Ovatainate | Entire | Spike | Spike (dense) | Other | Erect | Erect |
| 84    | AV TRI 17 | Zeya | Malaysia | E-WEST | Erect | Along the stem | Green | Ovatainate | Entire | Spike | Spike (dense) | Other | Erect | Erect |
| 85    | AV TRI 18 | Zeya | Malaysia | E-WEST | Erect | Along the stem | Green | Ovatainate | Entire | Spike | Spike (dense) | Other | Erect | Erect |
| 86    | AV TRI 19 | Zeya | Malaysia | E-WEST | Erect | Along the stem | Green | Ovatainate | Entire | Spike | Spike (dense) | Other | Erect | Erect |
| 87    | AV TRI 20 | Zeya | Malaysia | E-WEST | Erect | Along the stem | Green | Ovatainate | Entire | Spike | Spike (dense) | Other | Erect | Erect |
| 88    | AV TRI 21 | Zeya | Malaysia | E-WEST | Erect | Along the stem | Green | Ovatainate | Entire | Spike | Spike (dense) | Other | Erect | Erect |
| 89    | AV TRI 22 | Zeya | Malaysia | E-WEST | Erect | Along the stem | Green | Ovatainate | Entire | Spike | Spike (dense) | Other | Erect | Erect |
| 90    | AV TRI 23 | Zeya | Malaysia | E-WEST | Erect | Along the stem | Green | Ovatainate | Entire | Spike | Spike (dense) | Other | Erect | Erect |
| 91    | AV TRI 24 | Zeya | Malaysia | E-WEST | Erect | Along the stem | Green | Ovatainate | Entire | Spike | Spike (dense) | Other | Erect | Erect |
| 92    | AV TRI 25 | Zeya | Malaysia | E-WEST | Erect | Along the stem | Green | Ovatainate | Entire | Spike | Spike (dense) | Other | Erect | Erect |
| 93    | AV TRI 26 | Zeya | Malaysia | E-WEST | Erect | Along the stem | Green | Ovatainate | Entire | Spike | Spike (dense) | Other | Erect | Erect |
| 94    | AV TRI 27 | Zeya | Malaysia | E-WEST | Erect | Along the stem | Green | Ovatainate | Entire | Spike | Spike (dense) | Other | Erect | Erect |
| 95    | AV TRI 28 | Zeya | Malaysia | E-WEST | Erect | Along the stem | Green | Ovatainate | Entire | Spike | Spike (dense) | Other | Erect | Erect |
| 96    | AV TRI 29 | Zeya | Malaysia | E-WEST | Erect | Along the stem | Green | Ovatainate | Entire | Spike | Spike (dense) | Other | Erect | Erect |
| 97    | AV TRI 30 | Zeya | Malaysia | E-WEST | Erect | Along the stem | Green | Ovatainate | Entire | Spike | Spike (dense) | Other | Erect | Erect |
| 98    | AV TRI 31 | Zeya | Malaysia | E-WEST | Erect | Along the stem | Green | Ovatainate | Entire | Spike | Spike (dense) | Other | Erect | Erect |
| 99    | AV TRI 32 | Zeya | Malaysia | E-WEST | Erect | Along the stem | Green | Ovatainate | Entire | Spike | Spike (dense) | Other | Erect | Erect |
| 100   | AV TRI 33 | Zeya | Malaysia | E-WEST | Erect | Along the stem | Green | Ovatainate | Entire | Spike | Spike (dense) | Other | Erect | Erect |
| 101   | AV TRI 34 | Zeya | Malaysia | E-WEST | Erect | Along the stem | Green | Ovatainate | Entire | Spike | Spike (dense) | Other | Erect | Erect |
| 102   | AV TRI 35 | Zeya | Malaysia | E-WEST | Erect | Along the stem | Green | Ovatainate | Entire | Spike | Spike (dense) | Other | Erect | Erect |
| 103   | AV TRI 36 | Zeya | Malaysia | E-WEST | Erect | Along the stem | Green | Ovatainate | Entire | Spike | Spike (dense) | Other | Erect | Erect |
| 104   | AV TRI 37 | Zeya | Malaysia | E-WEST | Erect | Along the stem | Green | Ovatainate | Entire | Spike | Spike (dense) | Other | Erect | Erect |
| 105   | AV TRI 38 | Zeya | Malaysia | E-WEST | Erect | Along the stem | Green | Ovatainate | Entire | Spike | Spike (dense) | Other | Erect | Erect |
| 106   | AV TRI 39 | Zeya | Malaysia | E-WEST | Erect | Along the stem | Green | Ovatainate | Entire | Spike | Spike (dense) | Other | Erect | Erect |
| 107   | AV TRI 40 | Zeya | Malaysia | E-WEST | Erect | Along the stem | Green | Ovatainate | Entire | Spike | Spike (dense) | Other | Erect | Erect |
| 108   | AV TRI 41 | Zeya | Malaysia | E-WEST | Erect | Along the stem | Green | Ovatainate | Entire | Spike | Spike (dense) | Other | Erect | Erect |

(Continued)
| Entry | Accessions | ID    | Species | Origin country | Germplasm | Growth habit | Branching index | Stem colour | Leaf colour | Petiole colour | Inflorescence colour | Leaf shape | Leaf margin | Terminal inflorescence shape | Terminal inflorescence attitude |
|-------|------------|-------|---------|---------------|-----------|--------------|-----------------|-------------|-------------|----------------|------------------------|------------|-------------|----------------------------|-------------------------------|
| 109   | AV TRI 41  | V1044430| Tricolor | Papua New Guinea| AVRDC | Erect | Along the stem | Green | Spotted purple | Purple | Green | Lanceolate | Undulate | Other | Erect |
| 110   | AV TRI 42  | V1044407| Tricolor | Papua New Guinea| AVRDC | Erect | Along the stem | Green | Spotted purple | Purple | Green | Elliptical | Entire | Other | Other |
| 111   | AV TRI 43  | V1048301| Tricolor | Bangladesh | AVRDC | Erect | Along the stem | Purple | Dark purple | Red | Ovate | Undulate | Short branches | Drooping |
| 112   | AV TRI 44  | V1048286| Tricolor | Bangladesh | AVRDC | Erect | Along the stem | Purple | Dark purple | Red | Ovate | Undulate | Short branches | Erect |
| 113   | AV TRI 45  | V1048021| Tricolor | Bangladesh | AVRDC | Erect | Few branches | Purple | Purple | Pink | Green | Other | Entire | Other | Short branches | Drooping |
| 114   | AV TRI 46  | V1047929| Tricolor | Bangladesh | AVRDC | Erect | Along the stem | Purple | Dark purple | Red | Rhombic | Entire | Short branches | Drooping |
| 115   | AV TRI 47  | V1047682| Tricolor | Bangladesh | AVRDC | Erect | Along the stem | Pink | Pink | Pink | Rhombic | Undulate | Spike (dense) | Drooping |
| 116   | AV TRI 48  | V1047681| Tricolor | Bangladesh | AVRDC | Erect | Few branches | Green | Normal green | Green | Green | Ovate | Undulate | Spike (dense) | Erect |
| 117   | AV TRI 49  | V1047504| Tricolor | Bangladesh | AVRDC | Erect | Few branches | Purple | Dark purple | Red | Rhombic | Entire | Spike (dense) | Erect |
| 118   | AV TRI 50  | V1047501| Tricolor | Bangladesh | AVRDC | Erect | Few branches | Purple | Green | Red | Ovate | Undulate | Short branches | Erect |
| 119   | AV TRI 51  | V1035270| Tricolor | Cambodia | AVRDC | Erect | Few branches | Green | Normal green | Green | Red | Rhombic | Entire | Short branches | Drooping |
| 120   | AV TRI 52  | V1036168| Tricolor | Cambodia | AVRDC | Erect | Few branches | Purple | Normal green | Green | Green | Other | Undulate | Short branches | Erect |
| 121   | AV TRI 53  | V1042979| Tricolor | Indonesia | AVRDC | Erect | Few branches | Green | Normal green | Green | Green | Ovate | Undulate | Spike (dense) | Drooping |
| 122   | AV TRI 54  | V1042978| Tricolor | Indonesia | AVRDC | Erect | Few branches | Green | Normal green | Green | Green | Ovate | Undulate | Spike (dense) | Erect |
| 123   | AV TRI 55  | V1039413| Tricolor | India | AVRDC | Erect | Along the stem | Green | Normal green | Green | Green | Rhombic | Undulate | Other | Erect |
| 124   | AV TRI 56  | V1038498| Tricolor | India | AVRDC | Erect | Along the stem | Purple | Pink | Pink | Lanceolate | Undulate | Other | Erect |
| 125   | AV TRI 57  | V1044426| Tricolor | Malaysia | AVRDC | Erect | Few branches | Green | Normal green | Green | Green | Rhombic | Undulate | Other | Drooping |
| 126   | AV TRI 58  | V1035139| Tricolor | Malaysia | AVRDC | Erect | Few branches | Purple | Central spot | Purple | Green | Rhombic | Entire | Other | Drooping |
| 127   | AV TRI 59  | V1035062| Tricolor | Malaysia | AVRDC | Erect | Along the stem | Green | Chlorotic stripe | Green | Green | Lanceolate | Entire | Other | Erect |
| 128   | AV TRI 60  | V1033490| Tricolor | Malaysia | AVRDC | Erect | Along the stem | Purple | Central spot | Purple | Green | Rhombic | Undulate | Other | Erect |
| 129   | AV TRI 61  | V1033480| Tricolor | Malaysia | AVRDC | Erect | Along the stem | Green | Normal green | Green | Green | Ovate | Undulate | Other | Erect |
| 130   | AV TRI 62  | V1033474| Tricolor | Malaysia | AVRDC | Erect | Along the stem | Green | Others | Green | Green | Other | Undulate | Other | Erect |
| 131   | AV TRI 63  | V1033473| Tricolor | Malaysia | AVRDC | Erect | Along the stem | Pink | Central spot | Pink | Green | Rhombic | Undulate | Short branches | Drooping |
| 132   | AV TRI 64  | V1040005| Tricolor | Thailand | AVRDC | Erect | Along the stem | Green | Chlorotic stripe | Green | Green | Other | Undulate | Other | Drooping |
| 133   | AV TRI 65  | V1046379-A | Tricolor | USA | AVRDC | Erect | Few branches | Green | Normal green | Green | Green | Rhombic | Undulate | Other | Erect |
| 134   | AV TRI 66  | V1047326-A | Tricolor | Viet Nam | AVRDC | Erect | Along the stem | Purple | Pink | Pink | Rhombic | Entire | Other | Other | Erect |
| 135   | AV TRI 67  | V1047387| Tricolor | Viet Nam | AVRDC | Erect | Along the stem | Green | Normal green | Green | Green | Lanceolate | Entire | Spike (dense) | Erect |
| Entry | Accession ID | Country | Origin country | Germplasm | Growth habit | Leaf shape | Terminal inflorescence shape | Petiole colour | Stem colour | Leaf colour | Inflorescence colour | Terminal inflorescence attitude |
|-------|--------------|---------|----------------|-----------|-------------|------------|-----------------------------|----------------|-------------|-------------|---------------------|-----------------------------|
| 136   | AV TRI 68    | Taiwan  | AVTRI 68       | Taiwan    | Erect       | Ovate      | Spike (dense)               | Green          | Green       | Green       | Normal              | Entire                     |
| 137   | AV TRI 69    | Taiwan  | AVTRI 69       | Taiwan    | Erect       | Ovate      | Spike (dense)               | Green          | Green       | Red         | Normal              | Entire                     |
| 138   | US TRI 1     | Bangladesh | US TRI 1 | Bangladesh | Erect       | Ovate      | Spike (dense)               | Green          | Red         | Red         | Normal              | Entire                     |
| 139   | US TRI 2     | Brazil   | US TRI 2       | Brazil    | Erect       | Ovate      | Spike (dense)               | Green          | Red         | Red         | Normal              | Entire                     |
| 140   | US TRI 3     | China    | US TRI 3       | China     | Erect       | Ovate      | Spike (dense)               | Green          | Red         | Red         | Normal              | Entire                     |
| 141   | US TRI 4     | China    | US TRI 4       | China     | Erect       | Ovate      | Spike (dense)               | Green          | Red         | Red         | Normal              | Entire                     |
| 142   | US TRI 5     | China    | US TRI 5       | China     | Erect       | Ovate      | Spike (dense)               | Green          | Red         | Red         | Normal              | Entire                     |
| 143   | US TRI 6     | China    | US TRI 6       | China     | Erect       | Ovate      | Spike (dense)               | Green          | Red         | Red         | Normal              | Entire                     |
| 144   | US TRI 7     | China    | US TRI 7       | China     | Erect       | Ovate      | Spike (dense)               | Green          | Red         | Red         | Normal              | Entire                     |
| 145   | US TRI 8     | China    | US TRI 8       | China     | Erect       | Ovate      | Spike (dense)               | Green          | Red         | Red         | Normal              | Entire                     |
| 146   | US TRI 9     | China    | US TRI 9       | China     | Erect       | Ovate      | Spike (dense)               | Green          | Red         | Red         | Normal              | Entire                     |
| 147   | US TRI 10    | China    | US TRI 10      | China     | Erect       | Ovate      | Spike (dense)               | Green          | Red         | Red         | Normal              | Entire                     |
| 148   | US TRI 11    | China    | US TRI 11      | China     | Erect       | Ovate      | Spike (dense)               | Green          | Red         | Red         | Normal              | Entire                     |
| 149   | US TRI 12    | Indonesia| US TRI 12      | Indonesia | Erect       | Ovate      | Spike (dense)               | Green          | Red         | Red         | Normal              | Entire                     |
| 150   | US TRI 13    | Indonesia| US TRI 13      | Indonesia | Erect       | Ovate      | Spike (dense)               | Green          | Red         | Red         | Normal              | Entire                     |
| 151   | US TRI 14    | Indonesia| US TRI 14      | Indonesia | Erect       | Ovate      | Spike (dense)               | Green          | Red         | Red         | Normal              | Entire                     |
| 152   | US TRI 15    | Indonesia| US TRI 15      | Indonesia | Erect       | Ovate      | Spike (dense)               | Green          | Red         | Red         | Normal              | Entire                     |
| 153   | US TRI 16    | Indonesia| US TRI 16      | Indonesia | Erect       | Ovate      | Spike (dense)               | Green          | Red         | Red         | Normal              | Entire                     |
| 154   | US TRI 17    | Indonesia| US TRI 17      | Indonesia | Erect       | Ovate      | Spike (dense)               | Green          | Red         | Red         | Normal              | Entire                     |
| 155   | US TRI 18    | Indonesia| US TRI 18      | Indonesia | Erect       | Ovate      | Spike (dense)               | Green          | Red         | Red         | Normal              | Entire                     |
| 156   | US TRI 19    | Indonesia| US TRI 19      | Indonesia | Erect       | Ovate      | Spike (dense)               | Green          | Red         | Red         | Normal              | Entire                     |
| 157   | US TRI 20    | Indonesia| US TRI 20      | Indonesia | Erect       | Ovate      | Spike (dense)               | Green          | Red         | Red         | Normal              | Entire                     |
| 158   | US TRI 21    | Indonesia| US TRI 21      | Indonesia | Erect       | Ovate      | Spike (dense)               | Green          | Red         | Red         | Normal              | Entire                     |
| 159   | US TRI 22    | Indonesia| US TRI 22      | Indonesia | Erect       | Ovate      | Spike (dense)               | Green          | Red         | Red         | Normal              | Entire                     |
| 160   | US TRI 23    | Indonesia| US TRI 23      | Indonesia | Erect       | Ovate      | Spike (dense)               | Green          | Red         | Red         | Normal              | Entire                     |
| 161   | US TRI 24    | Indonesia| US TRI 24      | Indonesia | Erect       | Ovate      | Spike (dense)               | Green          | Red         | Red         | Normal              | Entire                     |
| 162   | US TRI 25    | Indonesia| US TRI 25      | Indonesia | Erect       | Ovate      | Spike (dense)               | Green          | Red         | Red         | Normal              | Entire                     |
| Entry Accession | ID | Species | Origin country | Germplasm Growth habit | Growth habit traits | Branching index | Leaf colour | Petiole colour | Inflorescence colour | Leaf shape | Leaf margin | Terminal inflorescence shape | Terminal inflorescence attitude |
|-----------------|----|---------|----------------|------------------------|---------------------|-----------------|-------------|----------------|----------------------|------------|-------------|-------------------------|-------------------------------|
| 163             | US TRI 26 | Tricolor | Zaire | USDA | Erect | Along the stem | Green | Green | Green | Ovatainate | Entire | Other | Short branches | Erect |
| 164             | US TRI 27 | Tricolor | China | USDA | Erect | Along the stem | Pink | Basal area | Green | Red | Lamina | Entire | Other | Short branches | Erect |
| 165             | Local Red | Tricolor | Malaysia | USDA | Erect | Along the stem | Green | Central spot | Green | Green | Ovatainate | Entire | Short branches | Erect |
| 166             | US TRI 29 | Tricolor | China | USDA | Erect | Along the stem | Purple | Entire lamina | Purple | Red | Lamina | Undulate | Short branches | Erect |
| 167             | US TRI 30 | Tricolor | Hong Kong | USDA | Erect | Along the stem | Purple | Central spot | Pink | Green | Lanceolate | Entire | Long branches | Erect |
| 168             | US TRI 31 | Tricolor | Hong Kong | USDA | Erect | Along the stem | Pink | Central spot | Green | Green | Other | Short branches | Erect |
| 169             | US TRI 32 | Tricolor | Malaysia | USDA | Erect | Along the stem | Green | Basal area | Green | Green | Entire | Other | Other | Short branches | Erect |
| 170             | US TRI 33 | Tricolor | Malaysia | USDA | Erect | Along the stem | Purple | Entire lamina | Purple | Red | Lamina | Undulate | Short branches | Erect |
| 171             | US TRI 34 | Tricolor | Malaysia | USDA | Erect | Along the stem | Purple | Central spot | Pink | Green | Lanceolate | Entire | Other | Short branches | Erect |
| 172             | US TRI 35 | Tricolor | Malaysia | USDA | Erect | Along the stem | Purple | Central spot | Pink | Green | Lanceolate | Entire | Other | Short branches | Erect |
| 173             | US TRI 36 | Tricolor | Malaysia | USDA | Erect | Along the stem | Purple | Central spot | Pink | Green | Lanceolate | Entire | Other | Short branches | Erect |
| 174             | US TRI 37 | Tricolor | Malaysia | USDA | Erect | Along the stem | Purple | Central spot | Pink | Green | Lanceolate | Entire | Other | Short branches | Erect |
| 175             | US TRI 38 | Tricolor | Malaysia | USDA | Erect | Along the stem | Purple | Central spot | Pink | Green | Lanceolate | Entire | Other | Short branches | Erect |
| 176             | US TRI 39 | Tricolor | Malaysia | USDA | Erect | Along the stem | Purple | Central spot | Pink | Green | Lanceolate | Entire | Other | Short branches | Erect |
| 177             | US TRI 40 | Tricolor | Malaysia | USDA | Erect | Along the stem | Purple | Central spot | Pink | Green | Lanceolate | Entire | Other | Short branches | Erect |
| 178             | US TRI 41 | Tricolor | Malaysia | USDA | Erect | Along the stem | Purple | Central spot | Pink | Green | Lanceolate | Entire | Other | Short branches | Erect |
| 179             | US TRI 42 | Tricolor | Malaysia | USDA | Erect | Along the stem | Purple | Central spot | Pink | Green | Lanceolate | Entire | Other | Short branches | Erect |
| 180             | US TRI 43 | Tricolor | Malaysia | USDA | Erect | Along the stem | Purple | Central spot | Pink | Green | Lanceolate | Entire | Other | Short branches | Erect |
| 181             | US TRI 44 | Tricolor | Malaysia | USDA | Erect | Along the stem | Purple | Central spot | Pink | Green | Lanceolate | Entire | Other | Short branches | Erect |
| 182             | US TRI 45 | Tricolor | Malaysia | USDA | Erect | Along the stem | Purple | Central spot | Pink | Green | Lanceolate | Entire | Other | Short branches | Erect |
| 183             | US TRI 46 | Tricolor | Malaysia | USDA | Erect | Along the stem | Purple | Central spot | Pink | Green | Lanceolate | Entire | Other | Short branches | Erect |
| 184             | US TRI 47 | Tricolor | Malaysia | USDA | Erect | Along the stem | Purple | Central spot | Pink | Green | Lanceolate | Entire | Other | Short branches | Erect |
| 185             | US TRI 48 | Tricolor | Malaysia | USDA | Erect | Along the stem | Purple | Central spot | Pink | Green | Lanceolate | Entire | Other | Short branches | Erect |
| 186             | US TRI 49 | Tricolor | Malaysia | USDA | Erect | Along the stem | Purple | Central spot | Pink | Green | Lanceolate | Entire | Other | Short branches | Erect |
| 187             | US TRI 50 | Tricolor | Malaysia | USDA | Erect | Along the stem | Purple | Central spot | Pink | Green | Lanceolate | Entire | Other | Short branches | Erect |
| 188             | US TRI 51 | Tricolor | Malaysia | USDA | Erect | Along the stem | Purple | Central spot | Pink | Green | Lanceolate | Entire | Other | Short branches | Erect |
set to an initial denaturation step of 94°C for 1 min, followed by 30 cycles of denaturation at 94°C for 20 s, annealing at 58°C for 30 s and extension at 72°C for 45 s, before a final extension at 72°C for 7 min. Equimolar amounts of PCR products from each sample were combined followed by a single end sequencing of 77 cycles on an Illumina Hiseq2500. Twenty-four DNA samples were also genotyped in two technical replications to obtain the reproducibility of the marker data. The full SNP dataset is shown in S1 Table.

Data analysis

SNP filtering. The SNP data generated from DArTSeq technology were first physically mapped to *Amaranthus hypochondriacus* genome v2.1 [14] using CLC Genomic Workbench v8 (Qiagen), based on match of aligned sequence tags against the reference genome, with 80% length and similarity fraction [29]. To investigate species-specific SNPs among 12 amaranth species (not including species with one representative), the amaranth species were manually examined for unique SNPs presence in the mapped SNP markers. Six species with the highest species-specific SNPs were subjected to a Venn diagram to visualize the SNP loci shared among the species. The Venn diagram of overlapping SNP loci was generated using the online program Van de Peer Lab (http://bioinformatics.psb.ugent.be/). Genetic diversity and population structure was carried out in two steps. First, all 188 amaranth accessions consisting of 18 species were analyzed together and second, a subset of 120 *A. tricolor* accessions were analyzed separately, aiming to explore the genetic distances and population structure among the *A. tricolor* populations, which were of primary interest. In each dataset, the mapped SNP markers were trimmed by removing SNPs with <97% reproducibility, <70% call rate and <0.05 polymorphic information content (PIC) and SNPs located on minor contigs that were not have been annotated. Individual accessions with >30% missing data and SNP loci with >30% missing data were removed. The most informative SNPs with minimum allele frequency (MAF) >0.05 imputed using TASSEL v5.2.52 software [30] were selected for further analysis.

Population structure was constructed using the structure-like population genetic analyses using R package LEA [31–33]. The number of populations was determined using cross-entropy criterion, based on the predictions of a fraction of masked genotypes (matrix completion) and on the cross-validation approach, with runs of eight values of K (K = 1:8). A distance matrix was generated using TASSEL v5.2.52 software which was used to conduct principal coordinate analysis (PCoA) and a phylogenetic tree based on UPGMA distance.

Genome-wide association study of morphological traits. GWAS was conducted on the observed ten morphological traits on the same SNP datasets used for genetic diversity analysis. A mixed linear model (MLM) was generated to determine the associations by using the Q-matrix from population structure analysis (R package LEA) and kinship (K) from centered IBS method via TASSEL v5.2.52 and marker trait association (MTA) was determined at P≤0.01 and P≤0.001. The Manhattan plots of–log(p-values) and the quantile-quantile plots (Q-Q) of expected vs observed p-values for SNP based genotype-phenotype associations were generated using TASSEL v5.2.52. The most significant flanking sequences of SNPs associated with the traits (P≤0.001) were queried against JBrowse Phytozome v13 database to obtain the putative biological functions.

Results

SNP marker discovery

DArTSeq generated 74,306 polymorphic SNP reads from the 188 amaranth accessions of 18 species (S1 Table). Of these reads, 63,821 SNPs could be physically mapped to the *Amaranthus hypochondriacus* genome with an averaged of 100% reproducibility (max = 100%, min = 93%,
median = 100%), 74% call rate (max = 100%, min = 19%, median = 74%) and 0.14 PIC (max = 0.50, min = 0, median = 0.09). The majority of the SNPs were an A/G or C/T transition mutation (62%) while the other 38% were A/C, A/T, C/G, and G/T transversion mutation. A Venn diagram of the six largest sets of amaranth species showed species-specific SNP loci, with A. thunbergii showed the highest number of unique SNPs (26,629), followed by A. spinosus (1,008), A. gracizans (1,067), A. tricolor (820), A. hypochondriacus (437) and A. hybridus (296) (Fig 1). There were only 1,394 polymorphic SNP shared by all six species group.

Genetic diversity and population structure of two amaranth sets

First, for all 18 amaranth species, individual genotypes with >30% missing SNP data including A. atropurpurus (AV-ATR), A. blitoides (AV-BLITO), and A. spinosus (AV-SPI 1, AV-SPI 5 and AV-SPI 6), A. retroflexus (US-RET 1) and A. hybridus (AV-HYB 3) were removed and a total of 8,668 SNPs remained for 16 amaranth species, comprised of 181 accessions, with an averaged of 100% reproducibility, 91% call rate, 0.28 PIC, 0.16 MAF, and 6.97% averaged missing value in SNP loci and 5% averaged missing value at the individual-level. Second, for 120 A. tricolor accessions, two individual accessions (AV-TRI 20 and AV-TRI 28) which contribute to 30% of the missing values were removed and a total of 9,789 SNPs remained for 118 A. tricolor accessions, with and averaged of 100% reproducibility, 78% call rate, 0.20 PIC, 0.07 MAF, and 2% averaged missing values in SNP loci and at individual-level. Both SNP datasets (from

Fig 1. Venn diagram showing the presence, average and overlap of SNPs in the six largest amaranth species sets.

https://doi.org/10.1371/journal.pone.0267752.g001
16 amaranth species [181 accessions] and 118 *A. tricolor* accessions) shared 1346 SNPs identical markers.

Population structure analysis demonstrated that the K-values of the 16 amaranth species dataset and the *A. tricolor* subset were K = 3 respectively, based on minimal cross-entropy (S1 Fig) and the Q-matrix is displayed in a bar plot representation (Fig 2A and 2B). Each vertical bar represents a single accession, and the length of each bar represents the proportion contributed by each sub-population (admixture) and the grouping of the populations are illustrated in UPGMA phylogenetic tree (Fig 3A and 3B). The PCoA demonstrates the genetic divergence of both marker datasets was consistent with the output of the population structure (Fig 4A and 4B).

The 16 amaranth species were grouped into three populations. The majority of *A. tricolor* accessions belonged to Pop 1, with the exception of six *A. tricolor* accessions which originated from Bangladesh and belonged to POP 2 (brown-colored accessions) while the two out-grouped *A. tricolor* accessions (AV-TRI 20 and AV-TRI 28) were separated into POP 3 (yellow-dotted colour accessions). The two grain-types amaranth species (*A. hypochondriacus* and *A. cruentus*) belonged to Pop 3 together with their putative progenitor (*A. hybridus*), with the exception of one *A. cruentus* accession (AV-CRU 5) which belongs to Pop 1. Other cultivated vegetable-type species such as *A. blitum*, *A. graccizan*, *A. sp* and *A. thunbergii* were closely related to *A. tricolor* in Pop 1 (red-colored accessions), although several accessions belonged to Pop 2. The weed-type species such as *A. retroflexus* and *A. viridis* were diverse between the three populations. The PCoA demonstrated that Pop 1 clustered tightly together depicting that

---

(A) Ancestry matrix of 16 species (K=3)

(B) Ancestry matrix of 118 accessions (K=3)

Fig 2. Population structure of (A) 16 amaranth species and (B) 118 *A. tricolor* accessions at K = 3, respectively. Each vertical bar represents a single accession and the length of each bar represents the proportion contributed by each sub-population. The group membership for each population structure is similar to the UPGMA dendogram.

https://doi.org/10.1371/journal.pone.0267752.g002
Genetic diversity analysis in *Amaranthus* species
little diversity may exist within the populations and closer to Pop 2 which may explain the
inter-specific admixtures. Meanwhile Pop 2 and Pop 3 showed some dispersal and diversity
within the populations.

The *A. tricolor* subset demonstrated that accessions were divided into three sub-popula-
tions. Sub-pop 1 was made up of 105 accessions from 12 countries of origin, Sub-Pop 2 com-
prised of seven accessions, of which three accessions were from Papua New Guinea and four
accessions from USA, and Sub-pop 3 consists of six Bangladeshi accessions with distinct mor-
phological traits (had branches along the stem, purple-pink stem color, purple leaf and petiole
color, red-green inflorescence color and erect terminal inflorescence attitude) (Table 1). In
comparison with the 16 amaranth species population structure, *A. tricolor* accessions that
belong to Sub-pop 2 grouped together with the rest of *A. tricolor* accessions in Pop 1 (brown-
colored accessions). Meanwhile, the six distinct Bangladeshi *A. tricolor* accessions of Sub-pop
3 remained separated from the rest of *A. tricolor* accessions similar to Pop 2. The PCoA dis-
played a clear division between the sub-populations and the overall population statistic calcu-
lated using a Monte-Carlo test revealed that there is an overall significant difference between
the sub-populations (*P* = 0.002).

**SNP associations for morphological traits**

GWAS identified 22 significant and “suggestive” MTAs on 16 chromosomes of 16 amaranth
species that underline four morphological traits observed in branching index, inflorescence
color, leaf shape (P ≤ 0.01, P ≤ 0.001) and leaf pigmentation (P ≤ 0.01) (Table 2; S2 Table). At
P ≤ 0.001, four SNP markers were associated with branching index, six SNP markers associated
with inflorescence color and two SNP markers associated with leaf shape. Meanwhile, 100 sig-
nificant MTAs were generated from 118 *A. tricolor*, distributed among 16 chromosomes that
underline four morphological traits observed in inflorescence color, and leaf, petiole and stem pigmentations (P < 0.01, P < 0.001) (Table 3, S2 Table). At P < 0.001, forty-five SNP markers were associated with leaf pigmentation, eight SNP markers associated with petiole pigmentation, four SNP markers associated with inflorescence color and two SNP markers associated with stem pigmentation.

Furthermore, the mapping of this amaranth panel with the reference genome, *A. hypochondriacus* [14] identified twelve putative candidate genes with functional protein. These markers had low phenotypic variation (< 20%) evaluated on all respective traits. The Manhattan plots of –log(p) > 3 and the Q-Q plots of these traits are presented in Figs 5 and 6.

**Discussion**

The evaluation of molecular markers and morphological traits was carried out on single plants to retain homogeneity of germplasm, as morphological variations were observed among amaranth plants within one collection. The evaluation of single plants is necessary as amaranth has high phenotypic plasticity which appears to be heterogamous in field plantings and thus adapts easily to the environmental changes, even though selection within cultivar/landrace has the possibility to be infertile [34]. The capacity of amaranth to have wide genetic variability provides new prospects in the development of new crop varieties. Therefore, the construction of

| Trait         | SNP Allele ID | Chr | SNP position (bp) | P    | Marker R² | Transcript name | Protein homologs          |
|---------------|---------------|-----|-------------------|------|-----------|-----------------|---------------------------|
| Branching index | 33457123[F0–18:G>A; A–18: G>A] | 3   | 1229325           | 0.000 | 0.09588   | -               |                           |
|               | 33427832[F0–36:A>T; T–36:A>T] | 14  | 1.1E+07           | 0.001 | 0.0854    | -               |                           |
|               | 33440918[F0–66:G>T; T–66:G>T] | 5   | 8486270           | 0.001 | 0.08314   | -               |                           |
|               | 33416254[F0–26:C>T; T–26:C>T] | 1   | 2988352           | 0.001 | 0.08504   | -               |                           |
| Inflorescence color | 33414045[F0–22:A>G; A–22: A>G] | 14  | 2E+07             | 0.000 | 0.13234   | -               |                           |
|               | 33416574[F0–45:G>A; A–45: G>A] | 11  | 1.4E+07           | 0.000 | 0.14652   | -               |                           |
|               | 33420897[F0–6:A>G; G–6:A>G] | 12  | 3726457           | 0.000 | 0.10482   | AH018341-RA FTIP1: FT-interacting protein 1 (A. thaliana) |
| Leaf shape    | 33430487[F0–18:C>T; T–18:C>T] | 5   | 3281805           | 0.000 | 0.18777   | -               |                           |
|               | 33423745[F0–8:C>G; C–8:C>G] | 2   | 2.2E+07           | 0.001 | 0.08268   | -               |                           |
|               | 33435567[F0–16:T>C; C–16:T>C] | 16  | 7310907           | 0.001 | 0.08184   | -               |                           |
|               | 33459037[F0–16:T>C; C–16:T>C] | 6   | 4618929           | 0.001 | 0.08407   | -               |                           |
|               | 33443436[F0–14:G>A; A–14: G>A] | 6   | 4371684           | 0.001 | 0.08261   | -               |                           |

https://doi.org/10.1371/journal.pone.0267752.t002
Table 3. 58 MTAs (P < 0.001) of four morphological traits; inflorescence color, and leaf, petiole and stem pigmentation in 118 A. tricolor accessions.

| Trait                  | SNP Allele ID | Chr SNP position (bp) | P     | Marker | Protein homologs                      |
|------------------------|---------------|-----------------------|-------|--------|----------------------------------------|
| **Inflorescence color**| 3342103|F|G-23:A 15 4478322 0.000 0.16 | AH022173 Similar to MIEL1: E3 ubiquitin-protein ligase MIEL1 (A. thaliana) |
|                       | 3345567|F|G-23:A 16 7310907 0.000 0.14 | |
|                       | 3341674|F|A-45:G 11 1360378 0.001 0.13 | |
|                       | 3343943|F|G-66:G 3 1842058 0.001 0.12 | AH006016 Similar to MOT1: MOLYBDATE TRANSPORTER 1 (A. thaliana) |
| **Leaf pigmentation** | 3343818|F|G-31:A 14 1931154 0.000 0.23 | |
|                       | 3343686|F|C-13:T 10 12170324 0.000 0.17 | |
|                       | 3342765|F|C-13:T 2 19932945 0.000 0.21 | |
|                       | 3344037|F|G-5:G 2 32078565 0.000 0.20 | |

**Petiole pigmentation**

| Trait                  | SNP Allele ID | Chr SNP position (bp) | P     | Marker | Protein homologs                      |
|------------------------|---------------|-----------------------|-------|--------|----------------------------------------|
|                       | 3342464|F|A-45:G 7 19047239 0.000 0.21 | AH011978 Similar to DMS3: Protein DEFECTIVE IN MERISTEM SILENCING 3 (A. thaliana) |
|                       | 3345692|F|G-8:G 4 14612733 0.001 0.15 | |
|                       | 3343221|F|A-41:G 10 21992996 0.000 0.13 | |
|                       | 3347528|F|G-10:G 4 20557141 0.000 0.13 | |
|                       | 3345993|F|C-25:T 4 2054681 0.000 0.13 | |
|                       | 3344525|F|G-6:G 10 22045821 0.000 0.13 | |
|                       | 3344339|F|C-18:T 8 16678507 0.000 0.13 | |
|                       | 3343295|F|G-20:G 16 6291635 0.000 0.17 | |
|                       | 3341539|F|C-51:C 4 27753448 0.000 0.12 | |
|                       | 3345928|F|C-59:G 4 10349586 0.000 0.12 | |
|                       | 3342908|F|A-36:C 2 2084422 0.000 0.13 | |
|                       | 3347404|F|C-39:T 13 2722366 0.000 0.13 | |

**Stem pigmentation**

| Trait                  | SNP Allele ID | Chr SNP position (bp) | P     | Marker | Protein homologs                      |
|------------------------|---------------|-----------------------|-------|--------|----------------------------------------|
|                       | 3342499|F|G-57:T 3 8553631 0.000 0.20 | AH015492 Similar to A1g15820: Zinc finger CCCH domain-containing protein 15 (A. thaliana) |

https://doi.org/10.1371/journal.pone.0267752.1003
population structure in amaranth through a combination of morphological and molecular data is needed in order to develop a framework for future breeding programmes.

GBS data can have a high proportion of missing values [16] and the number of SNPs retained for the analysis depends on the quality control method [35]. In this study, a large number of SNP markers (74,306 SNP) were generated through the DArTSeq method, a non-reference based approach (de novo) using the PstI and MseI endonucleases in the library preparation step. After aligning the sequence tags against the very high quality and full length macromolecules of the A. hypochondriacus reference genome for SNP locations [14], the DArTSeq was able to generate relatively large numbers of SNP marker which could be mapped to the A. hypochondriacus genome (63,821 SNP) and suggests that DArTSeq as a technique should provide for full genome coverage. The number of SNP loci discovered in this study compared favorably with previous GBS studies generated in amaranth species that used ApeKI single enzyme cutting combined with deep reference-based assembly methodsm [17] as well as studies that used two library preparations via reference-based and non-reference based assembly methods [15, 16]. After filtration, the range of polymorphic SNP markers used in this study was comparable with other findings, such as 3,974 DArTSeq SNPs successfully used for population structure of 67 wild Galapagos tomato accessions (Solanum cheesmaniae and S. galapagensis) [36] and 3,956 DArTSeq SNPs used in 80 macadamia accessions (Macadamia integrifolia, M. tetraphylla and hybrids) [37].

Population structure analysis on 16 amaranth species generates consistent taxonomic classification of amaranth sub-genera which was previously defined using seeds, inflorescence and floral characteristics [7, 38]. Three amaranth sub-genera Amaranthus Amaranthus, Amaranthus Acnida and Amaranthus Albersia were well defined in this study, consistent with other
GBS findings by [15]. Subgenus *Amaranthus*, comprised of grain amaranth (*A. hypochondriacus* and *A. cruentus*) and its weed progenitor (*A. hybridus*) were distinguished in Pop 3. Subgenus *Albersia*, which comprised of vegetable amaranth including *A. tricolor* were distinguished in Pop 1 and Pop 2, together with six out of seven *A. blitum* accessions, three *A. graecizans* accessions and four of six *A. viridis* accessions. Meanwhile, species belonging to subgenus *Acnida*, which comprised of weedy amaranth, *A. spinosus* and *A. palmeri* were diverse between the three sub-populations. Another important finding was *A. hybridus* that belonged to subgenus *Amaranthus* was split into sub-genus *Albersia*. *A. hybridus* is the direct ancestor of cultivated grain amaranth species [39, 40], and the split of accessions identity could be due to inter-varietal hybridization. Weedy amaranth, *A. spinosus* is a cross-pollinated and subsequent gene flow between populations may occur more rapidly than the primarily self-pollinated amaranth species [40]. Lee et al. [47] also have stated that varying amounts of outcrossing and frequent interspecific and inter-varietal hybridization have occurred in amaranth accessions even though it is self-pollinated. Therefore, this could explain the admixture between amaranth species. Besides, this study found that weedy amaranth possessed more unique SNPs per accession than grain amaranth evidently perhaps suggesting that weedy species have had far less selection pressure than the cultivated grain species, which is useful from a breeding perspective.

There is also genetic differentiation between grain and vegetable amaranth in this study, which has also been observed in many molecular markers studies, including AFLP [41], SSR [23, 42, 43] and GBS [15], although those studies incorporated far fewer *A. tricolor* accessions. This genetic analysis has not only revealed duplicates and genetically closely related
individuals, but also allowed categorization of accessions into the correct species. In this study, two *A. tricolor* accessions (AV-TRI 20 and AV-TRI 28) from Asia deviated from the *A. tricolor* clade and were grouped together with sub-genera *Amaranthus*, which mainly belonged to grain and weed amaranths. There are two assumptions for this finding, either the two amaranths were incorrectly identified as *A. tricolor* [17] or were originally a landrace that was grown in a region where grain amaranth was traditionally cultivated over a long time through seeds exchange [44–46]. In a previous study, GBS accurately identified *A. caudatus* accession PI 490752, characterized as *A. hypochondriacus* by 11 SSR markers [39], but it should be assigned into the *A. caudatus* group [17]. Therefore, re-analysis should be carried out for these two *A. tricolor* accessions, with addition of larger morphological dataset, which could correct the possible misclassification. The occurrence of admixed/hybrid genotypes may indicate frequent hybridization or introgression events. An experiment based on SSR markers by [23] revealed that *A. tricolor* accessions did not correlate between groups which may imply that *A. tricolor* had larger genetic variation. There was also uncertainty in positioning phylogeny of *A. tricolor* accessions among amaranth species, although *A. tricolor* accessions were grouped together in a clade [15]. *A. tricolor* had by far the largest estimated genome size (782.7Mbp) among 35 amaranth species, and this suggests that polyploidization likely influenced the genome size of this species [15].

In this study, the species groupings were independent of the accession’s geographical origin, contradicting previous GBS findings [15–17]. In previous studies, geographical patterns demonstrate that comprehensive origin sampling can assist in understanding the evolution of the species as shown by a strong split of geographic pattern in *A. hybridus* between accessions from Central and South America, which later supports the hypothesis that two different lineages were the ancestors of the grain amaranth [15]. In this study, the genetic differentiation between species and geographical origin was weak, although a strong split of geographical pattern was observed in *A. hybridus* where accessions from America and Africa were divided into two clusters, which may explain the genetic differentiation of *hybridus* complex [23]. This is probably due to the cosmopolitan nature of the genus, or the results of human activities such as breeding and resource exchange [47]. While the current study used a different restriction endonuclease frequent cutter for construction of the genomic representations sequenced, the biased number of accession per species could contribute to the lack discrimination of geographical origin and species level. This was also observed in 3,431 DArTSeq SNPs used to conduct genetic diversity in 89 safflower accessions (*Carthamus tinctorius* L.), in which the SNPs showed weak correlation between safflower diversity pattern and origins, when compared with a larger SNP dataset [48]. However, for a large set of 118 *A. tricolor* accessions, genetic differentiation of Bangladeshi accessions was clear as they clustered together and had distinct morphological characters.

The closely related *A. hypochondriacus* genome was used as the genome reference for association mapping as no *A. tricolor* genome is available to date. The utilities of the reference quality genome were demonstrated in two ways, i.e. chromosomal evolution and mapping of genetic locus responsible for stem color, hence ample support to clarify the scientific understanding of a useful agricultural trait in amaranth. The highly significant MTA found in morphological traits in this study illustrate how this DArTSeq data can provide high resolution genome coverage for mapping opportunities. However, the most significant associations detected in the MLM model had a lower threshold (−log(p-value)<4, although the mixed model was superior, it still could be lead to at least one false negative and false positive [49]. This could be due to the use of different amaranth species (*A. hypochondriacus*) as a reference genome instead of the *A. tricolor* genome. The difficulty of working with plant genomes is that they are highly repetitive and feature extensive structural variation between members of the
same species, mostly attributed to their active transposons [50] and chromosomal rearrangements. For example, in the well-studied species Arabidopsis thaliana, natural accessions are missing 15% of the reference genome, indicating a similar fraction would be absent from the reference, but present in other accessions [51]. Moreover, although A. thaliana has a small (140 Mb) and not very repetitive genome compared to many other plants, SNPs may be assigned to incorrect positions due to sequence similarity shared between unlinked loci [52]. Therefore, more extensive structural variation would be expected in a larger A. tricolor genome, which contain a higher proportion of repeats and has undergone ancient and recent rounds of polyploidization [15].

Conclusions

The findings in this study demonstrated that the DArTSeq SNP data generated from 181 amaranth accessions comprised of 16 species was capable of differentiating vegetable amaranth, A. tricolor from grain and wild amaranth species. The species groupings were independent of accessions’ geographical origin. This is likely a result of germplasm origin being registered as where the seeds were donated from, which may not be the actual origin of the accession or movement of germplasm in recent historical time. For a larger A. tricolor data set, there was likelihood that a good differentiation of A. tricolor could be achieved based on a combined analysis of molecular markers, geographical origin and morphological traits. GWAS used to conduct a pilot genome association for 10 morphological traits demonstrates the potential effectiveness of the amaranth diversity panel for trait dissection. The high degree of morphological variation observed in amaranth may be beneficial in terms of its adaptive capabilities in different climatic conditions.

Supporting information

S1 Table. DArTSeq SNP reads from the 188 amaranth accessions for18 species.
(XLSX)

S2 Table. 10 morphological traits of A. tricolor subset observed under shade-house conditions.
(XLSX)

S1 Fig. Cross-entropy plot for (a) first population structure: 181 amaranth accessions of 16 species and (b) second population structure: 118 A. tricolor accessions. A range of K = 1:8 was tested and K = 3 was chosen as the cross-entropy curve exhibits a plateau in both datasets.
(TIF)

Acknowledgments

We thank the School of Biosciences, University of Nottingham Malaysia and Crops for the Future for the technical support. We are grateful to Dr David Brenner and the USDA for the donation of 179 accessions from the USDA amaranth collection.

Author Contributions

Conceptualization: Norain Jamalluddin, Festo J. Massawe, Sean Mayes, Rachael C. Symonds.

Data curation: Norain Jamalluddin.

Formal analysis: Norain Jamalluddin, Wai Kuan Ho, Rachael C. Symonds.

Funding acquisition: Festo J. Massawe, Sean Mayes, Rachael C. Symonds.
Investigation: Norain Jamalluddin, Festo J. Massawe, Sean Mayes, Wai Kuan Ho, Rachael C. Symonds.

Methodology: Norain Jamalluddin, Sean Mayes, Wai Kuan Ho.

Project administration: Festo J. Massawe, Rachael C. Symonds.

Supervision: Festo J. Massawe, Sean Mayes, Rachael C. Symonds.

Validation: Rachael C. Symonds.

Writing – original draft: Norain Jamalluddin.

Writing – review & editing: Festo J. Massawe, Sean Mayes, Wai Kuan Ho, Rachael C. Symonds.

References

1. Mayes S, Massawe FJ, Alderson PG, Roberts JA, Azam-Alli SN, Hermann M. The potential for underutilized crops to improve security of food production. J Exp Bot. 2011; 63:1075–1079. https://doi.org/10.1093/jxb/err396 PMID: 22131158

2. Rastogi A, Shukla S. Amaranth: A new millennium crop of nutraceutical values. Crit Rev Food Sci Nutr. 2013; 53:109–125. https://doi.org/10.1080/10408398.2010.517876 PMID: 23072528

3. Jamalluddin N, Symonds RC, Mayes S, Ho WK, Massawe F. Chapter 6: Diversifying crops for food and nutrition security: a case of vegetable amaranth, an ancient climate-smart crop. In: Galanakis CM. Food Security and Nutrition. Elsevier; 2021. pp. 125–146.

4. Kauffman CS, Weber LE. Grain amaranth. In: Janick J, Simon JE (ed). Advances in new crops. Timber Press, Portland; 1990. pp. 127–139.

5. Mosyakin SL, Robertson KR. New infrageneric taxa and combinations in Amaranthus (Amaranthaceae). Ann Bot Fenn. 1996; 33:275–282.

6. Grubben GJH, van Sloten DH. Genetic resources of Amaranths: A global plan of action. ACP:IBPGR/80/2. International Board for Plant Genetic Resources, Food and Agriculture Organization of the United Nations; 1981. Rome, Italy. pp. 57.

7. Achigan-Dako EG, Sogbohossou OED, Maundu P. Current knowledge on Amaranthus spp.: Research avenues for improved nutritional value and yield in leafy amaranths in sub-Saharan Africa. Euphytica. 2014; 197:303–317.

8. Jiménez-Aguilar DM, Grusak MA. Minerals, vitamin C, phenolics, flavonoids and antioxidant activity of Amaranthus leafy vegetables. J Food Compos Anal. 2017; 58:33–39.

9. Jamalluddin N, Massawe F, Symonds RC. Transpiration efficiency of amaranth (Amaranthus sp.) in response to drought stress. J Hortic Sci Biotechnol. 2018; 94:448–459.

10. Sarker U, Islam MT, Oba S. Salinity stress accelerates nutrients, dietary fiber, minerals, phytochemicals and antioxidant activity in Amaranthus tricolor leaves. PLoS ONE; 2018. 13(11):e0206388. https://doi.org/10.1371/journal.pone.0206388 PMID: 30383779

11. Alemayehu FR, Bendevis MA, Jacobsen SE. The potential for utilizing the seed crop Amaranth (Amaranthus spp.) in East Africa as an alternative crop to support food security and climate change mitigation. J Agron Crop Sci. 2014; 201(5):321–329.

12. Sarker U, Islam MT, Rabbanig OB, Oba S. Genotypic variability for nutrient, antioxidant, yield and yield contributing traits in vegetable amaranth. J Food Agric Environ. 2014; 12(3&4):168–174.

13. Sogbohossou EOD, Achigan-Dako EG. Phenetic differentiation and use-type delimitation in Amaranthus spp. from worldwide origins. Sci Hortic. 2014; 178:31–42.

14. Lightfoot DJ, Jarvis DE, Ramara J, Lee R, Jellen EN, Maughan PJ. Single-molecule sequencing and Hi-C-based proximity-guided assembly of amaranth (Amaranthus hypochondriacus) chromosomes provide insights into genome evolution. BMC Biol. 2017; 15:74. https://doi.org/10.1186/s12915-017-0412-4 PMID: 28654926

15. Stetter M.G, Schimid K. Analysis of phylogenetic relationships and genome size evolution of the Amaranthus genus using GBS indicates the ancestors of an ancient crop. Mol Phylogenet Evol. 2017; 109:80–92. https://doi.org/10.1016/j.ympev.2016.12.029 PMID: 28057554
16. Stetter MG, Müller T, Schmid KJ. Genomic and phenotypic evidence for an incomplete domestication of South American grain amaranth (Amaranthus caudatus). Mol Ecol. 2017; 26(3): 871–886. https://doi.org/10.1111/mec.13974 PMID: 2809043

17. Wu X, Blair MW. Diversity in grain amaranths and relatives distinguished by genotyping by sequencing (GBS). Front Plant Sci. 2017; 8: 1960. https://doi.org/10.3389/fpls.2017.01960 PMID: 29204149

18. Gates DJ, Strickler SR, Mueller LA, Olson BJ, Smith SD. Diversification of R2R3-MYB transcription factors in the tomato family Solanaceae. J Mol Evol. 2016; 83(1–2):26–37. https://doi.org/10.1007/s00239-016-9750-z PMID: 27364496

19. Stetter MG, Vidal-Villarejo M, Schmid KJ. Parallel seed color adaptation during multiple domestication attempts of an ancient new world grain. Mol Biol Evol. 2020; 37:1407–1419. https://doi.org/10.1093/molbev/msz304 PMID: 3186092

20. Sunil M, Hariharan AK, Nayak S, Gupta S, Nambisan SR, Gupta RP, et al. The draft genome and transcriptome of Amaranthus hypochondriacus: A C4 dicot producing high-lysine edible pseudo-cereal. DNA Res. 2014; 21:585–602. https://doi.org/10.1093/dnares/dsu021 PMID: 25070797

21. Deb S, Jayaprasad S, Ravi S, Rao KR, Whadga S, Hariharan N, et al. Classification of grain amaranths using chromosome-level genome assembly of Ramdana, A. hypochondriacus. Front Plant Sci. 2020; 579529. https://doi.org/10.3389/fpls.2020.579529 PMID: 33262776

22. Mallory MA, Hall RV, McNabb AP, Pratt DB, Jellen EN, Maughan PJ. Development and characterization of microsatellite markers for the grain amaranths. Crop Sci. 2008; 48:1098–1106.

23. Khaing AA, Moe KT, Chung JW, Baek HJ, Park, YJ. Genetic diversity and population structure of the South American grain amaranth (Amaranthus hypochondriacus) with RAPD marker analysis. Front Plant Sci. 2020; 11:539857. https://doi.org/10.3389/fpls.2020.539857 PMID: 31988628

24. Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES, et al. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. PLoS ONE. 2011; 6(5):e19379. https://doi.org/10.1371/journal.pone.001758.x PMID: 21573248

25. Andrews KR, Good JM, Miller MR, Luikart G, Hohenlohe PA. Harnessing the power of RADseq for ecological and evolutionary genomics. Nat Rev Genet. 2016; 17:81–92. https://doi.org/10.1038/nrg.2015.28 PMID: 26729255

26. Andrews KR, Good JM, Miller MR, Luikart G, Hohenlohe PA. Harnessing the power of RADseq for ecological and evolutionary genomics. Nat Rev Genet. 2016; 17:81–92. https://doi.org/10.1038/nrg.2015.28 PMID: 26729255

27. Kilian A, Wenzl P, Hutten E, Carling J, Xia L, Blois H, et al. Diversity arrays technology: a generic genome profiling technology on open platforms. Methods Mol Biol. 2012; 888:67–89. https://doi.org/10.1007/978-1-61779-870-2_5 PMID: 22665276

28. Li HS, Bhavani J, Vikram S, Sehgal P, Huerta-Espino D, Kilian A, et al. A high density GBS map of bread wheat and its application for dissecting complex disease resistance traits. BMC Genomics. 2015; 16:216. https://doi.org/10.1186/s12864-015-1424-5 PMID: 25887001

29. Ho WK, Chai HH, Kendabie P, Ahmad NS, Jani J, Massawe F, et al. Integrating genetic maps in bambara groundnut [Vigna subterranea (L) Verdc.] and their syntentic relationships among closely related legumes. BMC Genomics. 2017; 18:192. https://doi.org/10.1186/s12864-016-3393-8 PMID: 28219341

30. Kilian A, Wenzl P, Hutten E, Carling J, Xia L, Blois H, et al. Diversity arrays technology: a generic genome profiling technology on open platforms. Methods Mol Biol. 2012; 888:67–89. https://doi.org/10.1007/978-1-61779-870-2_5 PMID: 22665276

31. Li HS, Bhavani J, Vikram S, Sehgal P, Huerta-Espino D, Kilian A, et al. A high density GBS map of bread wheat and its application for dissecting complex disease resistance traits. BMC Genomics. 2015; 16:216. https://doi.org/10.1186/s12864-015-1424-5 PMID: 25887001

32. Ho WK, Chai HH, Kendabie P, Ahmad NS, Jani J, Massawe F, et al. Integrating genetic maps in bambara groundnut [Vigna subterranea (L) Verdc.] and their syntentic relationships among closely related legumes. BMC Genomics. 2017; 18:192. https://doi.org/10.1186/s12864-016-3393-8 PMID: 28219341

33. Kilian A, Wenzl P, Hutten E, Carling J, Xia L, Blois H, et al. Diversity arrays technology: a generic genome profiling technology on open platforms. Methods Mol Biol. 2012; 888:67–89. https://doi.org/10.1007/978-1-61779-870-2_5 PMID: 22665276

34. Falush D, Stephens M, Pritchard JK. Inference of population structure using multilocus genotype data: Dominant markers and null alleles. Mol Ecol Notes. 2007; 7(4):574–578. https://doi.org/10.1111/j.1471-2229.2007.01758.x PMID: 18784791

35. François O. Running structure-like population genetic analyses with R, R tutorials in population genet- ics. U Grenoble -Alpes. 2016; 1–9.

36. Guillen-Portal FR, Baltensperger DD, Nelson LA. Plant population influence on yield and agronomic traits in ‘Plainsman’ grain amaranth. In: Janick J (ed.) Perspectives on new crops and new uses. ASHS Press, Alexandria, VA; 1999. pp 190–193.

37. Marees AT, de Kluiver H, Stringer S, Vorspan F, Curis E, Marie-Claire C, et al. A tutorial on conducting genome-wide association studies: Quality control and statistical analysis. Int J Methods Psychiat Res. 2018; 27(2):e1608. https://doi.org/10.1002/mpr.1608 PMID: 29484742

38. Pailles Y, Ho S, Pires IS, Tester M, Negrão S, Schmölckel SM. Genetic diversity and population structure of two tomato species from the Galápagos Islands. Front Plant Sci. 2017; 8:138. https://doi.org/10.3389/fpls.2017.00138 PMID: 28261227
37. Alam M, Neal J, O’Connor K, Kilian A, Topp B. Ultra-high-throughput DArTseq-based silicoDArT and SNP markers for genomic studies in macadamia. PloS ONE. 2018; 13(8): e0203465. https://doi.org/10.1371/journal.pone.0203465 PMID: 30169500

38. Das S. Systematics and taxonomic delimitation of vegetable, grain and weed amaranths: A morphological and biochemical approach. Genet Resour Crop Evol. 2012; 59:289–303.

39. Kietlinski KD, Jimenez F, Jellen EN, Maughan PJ, Smith SM, Pratt DB. Relationships between the weedy (Amaranthaceae) and the grain amaranths. Crop Sci. 2014; 54(1):220–228.

40. Stetter MG, Zeitzler L, Steinhaus A, Kroener K, Biljecki M, Schmid KJ. Crossing methods and cultivation conditions for rapid production of segregating populations in three grain amaranth species. Front Plant Sci. 2016; 7:816. https://doi.org/10.3389/fpls.2016.00816 PMID: 27375666

41. Costea M, Weaver SE, Tardif SJ. The biology of Canadian weeds. 130. Amaranthus retroflexus L., A. powellii S. Watson and A. hybridus L. Can J Plant Sci. 2004; 84(2):631–668.

42. Oo WH, Park YJ. Analysis of the genetic diversity and population structure of amaranth accessions from South America using 14 SSR markers. Korean J Crop Sci. 2013; 58(4): 336–346.

43. Suresh S, Chung JW, Cho GT, Sung JS, Park JH, Gwag JG, et al. Analysis of molecular genetic diversity and population structure in Amaranthus germplasm using SSR markers. Plant Biosyst. 2014; 148(4): 635–644.

44. Brenner D, Baltensperger D, Kulakow P, Lehmann J., Myers R, Slabbert M, et al. Genetic resources and breeding of Amaranthus. Plant Breed Rev. 2010; 19:227–285.

45. Jimenez FR, Maughan PJ, Alvarez A, Kietlinski KD, Smith SM, Pratt DB, et al. Assessment of genetic diversity in Peruvian Amaranth (Amaranthus caudatus and A. hybridus) germplasm using single nucleotide polymorphism markers. Crop Sci. 2013; 53(2): 532–541.

46. Das S. Infrageneric classification of amaranths. In: Das S. (Eds.) Amaranthus: A promising crop of future. Springer, Singapore; 2016. pp. 49–56.

47. Lee JR, Hong GY, Dixit A, Chung JW, Ma KH, et al. Characterization of microsatellite loci developed for Amaranthus hypochondriacus and their cross-amplifications in wild species. Conserv Genet. 2008; 9:243–246.

48. Hassani SMR, Talebi R, Pourdad SS, Najl AM, Fayaz F. In-depth genome diversity, population structure and linkage disequilibrium analysis of worldwide diverse safflower (Carthamus tinctorius L.) accessions using NGS data generated by DArTseq technology. Mol Biol Rep. 2020; 47(3):2123–2135. https://doi.org/10.1007/s11033-020-05312-x PMID: 32062796

49. Voichek Y, Weigel D. Identifying genetic variants underlying phenotypic variation in plants without complete genomes. Nat Genet. 2020; 52:534–540. https://doi.org/10.1038/s41588-020-0612-7 PMID: 32284578

50. Bennetzen JL. Transposable element contributions to plant gene and genome evolution. Plant Mol Biol. 2000; 42:251–269. PMID: 10688140

51. 1001 Genomes Consortium. 1,135 genomes reveal the global pattern of polymorphism in Arabidopsis thaliana. Cell 2016; 166(2):481–491. https://doi.org/10.1016/j.cell.2016.05.063 PMID: 27293186

52. Long Q, Rabanal FA, Meng D, Huber CD, Farlow A, Platzer A, et al. Massive genomic variation and strong selection in Arabidopsis thaliana lines from Sweden. Nat Genet. 2013; 45(8):884–890. https://doi.org/10.1038/ng.2678 PMID: 23793030