Study on Genetic Diversity of Leafy Amaranth (*Amaranthus tricolor* L.) in Vietnam

August 2019

NGUYEN DUC CHINH

Study on Genetic Diversity of Leafy Amaranth (*Amaranthus tricolor* L.) in Vietnam

A Dissertation Submitted to
the Graduate School of Life and Environmental Sciences,
the University of Tsukuba
in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy in Agricultural Science

NGUYEN DUC CHINH

Table of contents

List of tables

CHAPTER 2

CHAPTER 3

List of figures

CHAPTER 2

CHAPTER 3

CHAPTER 4

CHAPTER 1: General Introduction

Climate change is considered one of the greatest environmental threats to sustainable agricultural production and food security in many parts of the world, particularly under increasing pressures from a growing population (FAO 2008). Poor developing countries are at greatest risk from climate change effects due to their limited capacity to adapt in terms of resources and technology and the fact that their economies and the incomes of their people heavily rely on the agricultural sector, which is very sensitive to climate variance (Mertz *et al.* 2009). Among these, Vietnam has been listed as one of the countries that is most seriously affected by climate change (Dasgupta *et al.* 2007).

The production of crops and livestock is directly impacted by changes in temperature and the frequency and severity of extreme events such as droughts and floods (Cumhur and Malcolm 2008). In Vietnam, the annual average temperature increased by approximately 0.5°C–0.7°C and the sea level rose by approximately 20 cm between 1958 and 2007 (FAO 2011). This increase in temperature alongside greater rainfall variation has resulted in more heatwaves and droughts occurring during the summer, which has reduced crop productivity, resulting in estimated financial losses to the agricultural sector of US$54.9 million per year between 1995 and 2007 (Ho 2018). Furthermore, three climate change scenarios to 2050 have predicted that the average annual temperature and sea level will increase in all agroecological areas in Vietnam, with the total area that is affected by salinity intrusion (>4 g/L NaCl) during the dry season increasing by 420,000 ha (Tran *et al.* 2016). Consequently, the National Strategy on Climate Change has emphasized that the Vietnamese Government should invest more resources in national

breeding programs to develop new crop cultivars with high resistance and tolerance to drought and salinity (Ho 2018).

Agricultural production in Vietnam is strongly reliant on chemical pesticides, so ensuring vegetable safety remains a major issue (Hoi *et al.* 2016). In a survey of consumers and buyers in Hanoi, Wang *et al.* (2012) found that 88.5% of those surveyed were worried about the quality of vegetables due to the high levels of pesticide residues. Furthermore, it has been estimated that the high levels of agrochemical residues on fruit and vegetable products costs the economy US$700 million per year in terms of domestic human health and lost export opportunities (Hoi *et al.* 2013). These high levels of agrochemical residues likely result from the fact that Vietnam is an agricultural country that is dominated by millions of small-scale farmers and the Government has only a limited capacity to put controls in place at the farm level (Hoi *et al.* 2016). In addition, Vietnamese vegetable farmers often apply pesticides at higher rates than are recommended on the label (Nguyen *et al.* 2018) and also violate the recommended preharvest interval (Lan *et al.* 2014). Therefore, the development of new cultivars with high resistance to disease will be an effective strategy for reducing pesticide use in vegetable production and ensuring vegetable safety.

1.1. Introduction to the genus *Amaranthus*

The genus *Amaranthus* L. (2n=32 or 34) contains approximately 60–75 species (Sauer 1967, Mosyakin and Robertson 1996, Brenner *et al.* 2000, Costea and DeMason 2001, Waselkov *et al.* 2018), among which there are at least 17 edible leafy species and 3 grain species (grown for their seeds) (Grubben and Denton 2004). Although amaranth species are widely distributed, most occur in tropical and warm temperate regions (Sauer 1967, Suresh *et al.* 2014), with approximately 55 species being native to the Americas

and the remainder having originated in Europe, Asia, Africa, and Australia (Sauer 1967, Waselkov *et al.* 2018). The first record of amaranth grain being used as a staple food was by the Aztec civilization in Central Mexico (Sauer 1993, Brenner *et al.* 2000), while amaranth leaves and stems are used as vegetables in South Asian and African countries (Grubben and Denton 2004, Rastogi and Shukla 2013, Achigan-Dako *et al.* 2014), competing with spinach leaves in terms of their protein content (Bui *et al.* 1998). Cultivated amaranths have also been used as forage and ornamental crops, while other species occur as weeds (Brenner *et al*., 2000). In the last few decades, amaranths have been regarded as new millennium crops globally due to their excellent nutritional values, high adaptability to severe conditions, and lack of major diseases (Shukla *et al.* 2006a, Rastogi and Shukla 2013, Achigan-Dako *et al.* 2014).

1.2. Origin, domestication, and distribution of amaranths

Amaranths are believed to have originated in South and Central America, where the highest number of native species of the genus *Amaranthus* are found and the domestication of grain amaranths occurred. The archeological record shows that domesticated amaranth seeds occurred over a wide geographical range, from Argentina to the southern United States (Sauer 1967). The earliest archeological records are of 6,000-year-old seeds of *A. cruentus* L. and 1,500-year-old seeds of *A. hypochondriacus* L., both of which were found in the same cave in Tehuacán, Puebla, Mexico, although the domestication of these plants may have occurred much earlier than this (Sauer 1969). While there is some debate around the domestication of grain amaranths, most studies agree that they originated from the two closely related wild species *A. hybridus* L. and *A. powellii* S. Wats. in separate areas, resulting in the production of *A. caudatus* L. in the Andes and *A. cruentus* and *A. hypochondriacus* in Central America (Sauer 1967, Sauer

1969, Xu and Sun 2001, Mallory *et al.* 2008, Maughan *et al.* 2011). In a recent analysis of the phylogenetic relationships among 94 amaranth accessions representing 35 species through genotyping by sequencing and genome size measurement, Stetter and Schmid (2017) confirmed that two different *A. hybridus* lineages in different areas were the ancestors of the three grain amaranths, with a possible contribution by *A. quitensis* Kunth in the domestication of *A. caudatus*.

Amaranths were widely cultivated as a major food and ceremonial grain at the time of the Spanish Conquest in Mexico, being as important as corn and beans in the Aztec Empire (Sauer 1950b, a). However, after the arrival of the Spanish conquistadors in early 1500 AD, the Spanish attempted to suppress Aztec culture and religion, resulting in *Amaranthus* almost disappearing as a crop in America. Consequently, the production and use of grain amaranths only survived in scattered parts of Mexico and Central and South America.

After the Spanish invasion of Latin America, amaranths spread to Europe and later Africa and Asia (Brenner *et al.* 2000). There was also some contact between American and Asian peoples in pre-Colombian times (Sauer 1950a), which resulted in Mexican and Peruvian races of amaranths being distributed in Southeast Asia by primitive hill people from Szechuan through to Tibet and India (Sauer 1950b). The study of Waselkov *et al.* (2018), which was based on several low-copy nuclear loci and chloroplast regions of 58 amaranth species, lent further support to the proposal that there were at least two long dispersal events of amaranths from South America to Eurasia, Africa, and Australia.

By the mid-20th century, grain amaranths had declined to a minor relic in most of their ancient homelands, with their cultivation only actively expanding in India. However, since the 1970s, there has been a widespread interest in increasing the use of this crop due

to its excellent protein content, the hardiness of the plants, and its adaptability to cultivation with hand tools (Sauer 1993).

Two main *Amaranthus* species have also been semi-domesticated as vegetable crops in Europe, Asia, and Africa. These include *A. tricolor* L., which originated in South and Southeast Asia, and *A. blitum* L., which originated in Eurasia. In addition, several other species have been harvested in the wild or cultivated for use as vegetables, including *A. dubius* Mart. ex Thell., *A. cruentus*, *A. thunbergii* Moq., *A. graecizans* L., *A. hypochondriacus*, and *A. viridis* L. (Grubben and Denton 2004).

1.3. Taxonomy and classification of the genus *Amaranthus*

The genus *Amaranthus* was first established by Linnaeus in 1753, at which time it consisted of 11 species divided into two groups: "Triandri" (six species) and "Pentandri" (five species), which correspond to the subgenera *Albersia* and *Amaranthus*, respectively, that are recognized today. From 1755 to 1771, Linnaeus added a further 12 species to the genus, giving a total of 23 species (Iamonico 2015). Various *Amaranthus* species were initially recognized as belonging to other genera, particularly the dioecious species and the monoecious species with dehiscent or indehiscent fruits (Waselkov *et al.* 2018). For instance, *Acnida*, which mostly included dioecious plants, was widely recognized as a separate genus by most authors until 1955, when Sauer (1955) provided evidence that it belonged to *Amaranthus*. However, he did not formally nomenclatural combination of these groups at either the subgeneric or sectional levels. Furthermore, although Aellen (1959) proposed *Amaranthus* subgen. *Acnida* (L.), this combination was not valid because he did not cite the basionym, although the combination was later validated by Robertson (1981). Similarly, *Albersia* was established as a separate genus by Kunth (1838) to accommodate the species of *Amaranthus* sensu lato with indehiscent utricles, with the

first combination at the subgeneric level appearing to have been *Amaranthus* subgen. *Albersia* (Kunth) Gren. & Godr. (Mosyakin and Robertson 1996).

These genera are presently recognized as three subgenera in the genus *Amaranthus* by most authors (Mosyakin and Robertson 1996, Achigan-Dako *et al.* 2014, Iamonico 2015, Stetter and Schmid 2017, Waselkov *et al.* 2018): subgenus *Amaranthus*, subgenus *Acnida* (L.) Aellen ex K.R.Robertson, and subgenus *Albersia*. These subgenera can be determined on the basis of characteristics related to the reproductive morphology (dioecious or monoecious), tepal number (2–3 or 5), type of inflorescence (axillary or terminal), and type of fruit (dehiscent or indehiscent) (Mosyakin and Robertson 1996). However, while keys have been provided for the infrageneric taxa in subgenera *Amaranthus* and *Acnida* (Mosyakin and Robertson 1996), they have not yet been adequately developed for *Albersia*, which is currently delimited through exclusion of the other two subgenera rather than a description of its crucial morphological characteristics. Furthermore, this subgenus still includes groups that are too polymorphic and geographically widespread to be natural. The best taxonomic solution for the subgenus *Albersia* seems to be the recognition of several narrower sections and subsections based on the whole of its morphological characteristics. However, the taxonomic relationships within this subgenus require additional investigation.

Amaranth species can also be divided into three categories based on their uses (Das 2012): (1) vegetable amaranths, such as *A. tricolor* var. *tricolor* and *A. tricolor* var. *tristis*; (2) grain amaranths, which include *A. hypochondriacus*, *A. caudatus*, and *A. cruentus*; and (3) weed amaranths, such as *A. spinosus* L., *A. viridis*, *A. retroflexus* L., *A. graecizans*, *A. dubius*, and *A. hybridus*.

1.4. Use and cultivation of amaranths

Amaranths have high genetic diversity and variability in form and have been cultivated for use as a pseudocereals, vegetables, medicines, dyes, forage, and ornamental crops (Sauer 1950a, Sauer 1967, Rastogi and Shukla 2013, Achigan-Dako *et al.* 2014, Assad *et al.* 2017). Amaranths can grow on a wide range of soil types with varying moisture levels, but grow particularly well in loam, sandy-loam, or silty-loam soils with good water holding capacities and pHs between 4.5 and 8.0 (Assad *et al.* 2017).

Amaranthus caudatus, *A. hypocondriacus*, and *A. cruentus* are grown for their grain in scattered parts of Central and South America and Asia (particularly India). These species are usually planted in dispersed patches near homes, in gardens, or mixed with corn and vegetables (Sauer 1967, Brenner *et al.* 2000). A traditional use of amaranth grains is to mix popped amaranth with honey to make a type of snack bar or cake. Amaranth grains are also ground into flour to make cake and are sometimes mixed with rice before cookingIn the United States, the commercial production of grain amaranths started in the late 1970s. However, while these crops are widely adapted to the local conditions and can be grown using conventional equipment that is used for other grain crops, they are only produced over a very small area (800–1200 ha) (Brenner *et al.* 2000). The main commercial products are crackers, cookies, and amaranth flour for the production of baked goods (Rastogi and Shukla 2013). (Rastogi and Shukla 2013).

Vegetable amaranths are widely grown in Africa and Asia, where the leaves and young stems are used as a frying vegetable or in soup. In Asian countries, they are also occasionally eaten raw in salads (Grubben and Denton 2004). Amaranth leaves taste the same as spinach. Before use, it is recommended that the leaves are boiled for 5 min to remove any antinutritional compounds. The cooked mass can then be eaten separately or

with other food (Rastogi and Shukla 2013). Vegetable amaranths are usually grown commercially as sole crops but are also found intercropped with other food crops in home gardens.

The amaranth species that are most widely cultivated for use as vegetables are *A. tricolor*, *A. blitum*, *A. cruentus*, and *A. dubius*. *Amaranthus tricolor* is one of the major leaf vegetables and the most important amaranth species in South and Southeast Asia, where it is mainly grown in small plots. In addition, *A. tricolor* is grown on a limited scale in home gardens or as a commercial vegetable in Africa, but is of little economic significance in this country. *Amaranthus blitum* probably originated in India, where it is still an important vegetable, and is also important in East Africa, where it is frequently collected as a wild plant and is a popular vegetable in home gardens, although its economic value as a market vegetable is limited. *Amaranthus cruentus* was domesticated as a pseudocereal (grain amaranth) in Central America but is now a popular traditional vegetable in humid lowland countries in Africa and is also grown as a leaf vegetable throughout Southeast Asia (albeit to a lesser extent than *A. tricolor*) where it is cultivated in small scattered plots and dispersed in small street markets (Grubben and Denton 2004). Finally, *A. dubius* is most widely used in the coastal parts of Kenya and Tanzania that are adjacent to lowlands below 1,500 m above sea level (Achigan-Dako *et al.* 2014). Vegetable amaranths in general are recommended as healthy foods and have medicinal properties for young children, lactating mothers, and patients with fever, hemorrhage, anemia, or kidney complaints. In addition, *A. blitum* is used as a medicine for lung disorders (Grubben and Denton 2004).

Amaranths have also been grown for use as dyes since ancient times in both the Old and New World. The amaranth species that were cultivated as dye plants by the Pueblo

people included *A. caudatus*, *A. cruentus*, *A. hybridus*, and *A. palmeri* S. Wats. (Sauer 1950a). Furthermore, in the last decade, the red-violet betacyanin pigments that are produced by amaranth plants have been considered a potential alternative source of natural pigments for use as colorants in foods and medicines (Cai *et al.* 2005). Amaranths are also used as forage crops in both the tropics and temperate zones, while *A. tricolor* and *A. caudatus* are used as ornamental plants throughout the world (Brenner *et al.* 2000).

During the 1980s, several advanced amaranth lines were introduced into China from the Rodale Research Center (RRC), Pennsylvania, USA. After several years of line purification and selection, versions of lines K112 and R104 (*A. hypochondriacus*) were obtained, which have since become the two most widespread cultivars in China (Yue *et al.* 1993). The annual cultivated area of grain amaranths in China is approximately 60,000 ha, making China the biggest producer in the world (Yue *et al.* 1993, Wu *et al.* 1995). This production is currently driven by the demand for animal feed (Wu *et al.* 2000).

1.5. Potential of the genus *Amaranthus*

Globally, amaranths are considered new millennium crops due to their excellent nutritional value, high adaptability to severe conditions, and lack of major diseases (Shukla *et al.* 2006a, Rastogi and Shukla 2013, Achigan-Dako *et al.* 2014, Assad *et al.* 2017).

Amaranth grains contain more dietary fiber and 5–20 times the amounts of calcium and iron than other types of grains. Furthermore, they have a higher protein content than other grains (12.5%–21.0%) (Venskutonis and Kraujalis 2013), which includes large amounts of the essential amino acids lysine (0.73%–0.84% of the total protein content) and trytophan (0.18%–0.28%) (Bressani *et al.* 1987), as well as higher amounts of the sulfur-containing amino acids methionine and cysteine than legumes (Venskutonis and

Kraujalis 2013). Thus, amaranth grains have the potential to improve the world food situation by providing an alternative source of protein (Assad *et al.* 2017).

In terms of the vegetative parts, 1 cup of cooked amaranth leaf makes a large contribution toward the daily requirement for vitamin C (90%), vitamin A (73%), calcium (28%), and iron (17%) (Achigan-Dako *et al.* 2014). In addition, Shukla *et al.* (2006a, 2006b) showed that the leaves of *A. tricolor* contain high concentrations of protein (1.24 ± 0.03 mg/100 mg), carotenoids (0.83 ± 0.02 mg/g), and minerals such as calcium (1.7 ± 0.04 g/100 g), iron (1233.8 ± 50.02 mg/kg), and zinc (791.7 ± 28.98 mg/kg), as well as significant amounts of ascorbic acid (112.33 ± 5.00 mg/100 g) and fiber (8.39% ± 0.10%). Schonfeldt and Pretorius (2011) further showed that the raw leaves of Amaranthus (*A. tricolor*), pumpkin (*Cucubita maxima* L.), and cat's whiskers (*Cleome gynandra* L.) contain higher concentrations of iron, zinc, and magnesium than those of cowpea (*Vigna unguiculata* (L.) Walp.) and wild jute (*Corchorus olitorius* L.), while Fasuyi and Akindahunsi (2009) found that *A. cruentus* leaves contain high levels of the essential amino acids lysine and methionine.

Amaranths are extremely adaptable to challenging growing conditions, such as heat stress, drought, and disease, and consequently grow well in agriculturally marginal lands (Shukla *et al.* 2006a, Rastogi and Shukla 2013). They are also tolerant to a variety of other unfavorable abiotic conditions, including high salinity, acidity, or alkalinity, making them uniquely suited for subsistence agriculture (Maughan *et al*. 2011). These plants are cultivated in all agroecological zones of West Africa, from the coastal sector in the Guinean zone to the dry forests and herbaceous savannahs in the Sudanian zone (Achigan-Dako *et al.* 2014). Furthermore, it has been shown that drought stress does not affect the water use efficiency of vegetable amaranths (Liu and Stutzel 2004) due to their C4

photosynthetic pathway, deep and extensive root system, ability to become dormant under extreme drought conditions (Assad *et al.* 2017), and ability to shut down transpiration through wilting and then recovering easily once moisture is available again (Myers 1996). In addition, amaranths are generally considered to be tolerant of nematodes, even being recommended as rotation crops to reduce nematode populations for subsequent crops, and are able to sustain their yield even when insects chew on their leaves (O'Brien and Price. 1983).

1.6. Amaranth genetic resources and breeding programs

Plant genetic resource conservation is extremely important for preserving the germplasm diversity and providing native genetic material with a large number of desirable traits (e.g., high quality, disease resistance, abiotic stress tolerance) for use in breeding programs. Plants in the genus *Amaranthus* have been collected and maintained as *ex situ* germplasm collections in many parts of the world, resulting in approximately 28,300 amaranth accessions globally (Ebert 2013), most of which are grain landrace types (Brenner *et al.* 2000). The National Bureau of Plant Genetic Resources in India has one of the best qualitative collections of amaranth germplasm in the world, with more than 5,700 accessions representing approximately 20 species. This is followed by the North Central Regional Plant Introduction Station (USA020) in the United States, which maintains approximately 3,300 accessions representing half of the species in the genus, the EMBRAPA Recursos Geneticos e Biotecnologia (CENARGEN) in Brazil with more than 2300 accessions, the Universidad Nacional San Antonio Abad del Cusco (UNSAAC/CICA) in Peru with approximately 1,600 accessions, the Institute of Crop Germplasm Resources-Chinese Academy of Agricultural Sciences (ICGR-CAAS) with

1,459 accessions (Ebert 2013), and the World Vegetable Center (AVRDC) with 520 accessions representing 18 species (Achigan-Dako *et al.* 2014).

Although amaranths are ancestral crops, they did not gain the attention of researchers until the early 1970s, when scientists at RRC revealed that grain amaranths have a high content and balanced profile of proteins. Research interest then grew in the 1980s, forming the basis for continued development of the crop in the 1990s (Brenner *et al.* 2000). RRC breeders initially selected numerous lines from Mexican germplasm but subsequently made crosses among the various grain amaranths from 1978 to 1990. These breeders treated amaranths as self-pollinating crops and developed a pedigree method in which selections were made in the F3 generation and uniformity was generally achieved by the F6 generation (Brenner *et al.* 2000). During the 1980s and 1990s, breeders in the United States led research efforts to improve grain amaranths, with additional research attention in China, India, and Mexico, as well as by small numbers of scientists in other countries (Myers 1996). The early breeding aims for grain amaranths at RRC emphasized a number of desirable traits, including increased seed size, synchronous dry down of the plant and seed head, vigorous seedling growth, resistance to lodging, reduction in seed shattering, improved pest tolerance, and increased seed protein and functional traits. This was followed by the proposal of three international lists of breeding purposes for grain amaranths by Williams and Brenner in 1995, which focused on enhanced food quality, reduced plant height, and increased yield (Brenner *et al.* 2000). In general, the breeding aims for grain amaranths are increased yield, enhanced pest resistance, and improved harvestability (Myers 1996).

Initial breeding efforts for vegetable amaranths have mostly been recorded in Asian and African countries since the late 1990s. Several Indian breeders have investigated the

genetic variability among the available accessions and the correlation among foliage yield (Reddy and Varalakshmi 1998, Shukla and Singh 2000, Shukla *et al.* 2004, Shukla *et al*. 2006a, Shukla *et al.* 2009). In addition, some productive commercial cultivars of *A. tricolor* were developed at Tamil Nadu Agricultural University, Coimbatore, India (Grubben and Denton 2004). Breeding work on *A. tricolor* has also been carried out in Indonesia, Thailand, and Taiwan, and some commercial cultivars are now available in South and Southeast Asian countries (e.g., from the East-West Seed Company in Indonesia and Thailand). Breeding work on *A. dubius* as a leafy vegetable has also been carried out at Tamil Nadu Agricultural University, and several commercial cultivars are now available. The breeding of *A. cruentus* as a leafy vegetable has been limited to a selection of local cultivars or landraces in a small number of countries (Benin, Nigeria) (Grubben and Denton 2004). Finally, the value of producing vegetable amaranths as a source of fresh greens in the northeastern temperate climate of the United States has also been evaluated (Schweig and Brown 2018).

1.7. Amaranth genetic resources and uses in Vietnam

Amaranths are one of the most popular summer vegetables in Vietnam, with its young leaves and shoots being used in stir-fries, soups, and boiled dishes. This vegetable can grow well in the summer season owing to its strong root system and is often cultivated in scattered plots in home gardens or small plots in fields from February to July. The crop can be harvested from 25 to 30 days after sowing and has a yield of 4–7 tons/ha (Nguyen and Tran 1996). Most of the amaranth germplasms that are grown by Vietnamese farmers are landraces that exhibit some desirable phenotypes. However, some amaranth cultivars (green leaf, red round leaf, and red narrow leaf type) are provided by commercial seed companies with limited information about their origin.

Seven amaranth species have been recorded in Vietnam: *A. tricolor*, *A. caudatus*, *A. viridis*, *A. spinosus*, *A. hybridus*, *A. lividus* L. (synonym: *A. blitum*), and *A. retroflexus* (Pham 1999). The Plant Resources Center (PRC) in Vietnam maintains approximately 280 amaranth accessions that were collected from eight ecological areas of Vietnam and are expected to contain a wide amount of genetic variation and desirable traits for crop improvement (PRC 2018). However, the Vietnamese *Amaranthus* (VA) collection has not been taxonomically classified or evaluated for its genetic diversity.

Although amaranths have been a popular vegetable in Vietnam for many years and is widely considered a promising crop for the next century due to its excellent nutritional values and high tolerance to biotic and abiotic stresses, there has been a lack of studies investigating its genetic diversity and no breeding programs to improve its genetics. Thus, it is important that the genetic diversity of amaranth is evaluated and promoted for further breeding efforts, as well as for the management and exploitation of plant genetic resources from gene banks.

1.8. Objectives

The aims of this thesis were to identify and evaluate the genetic diversity of the *A. tricolor* resources in Vietnam in comparison with those in other parts of the world in order to explore the genetic potential of this crop for further effective exploitation, conservation and improvement. The four main components of this study are presented in separate chapters. In Chapter 2, the amaranth resources in Vietnam are analyzed using the maturase K (*matK*) marker to classify the species. This allowed the identification of leafy amaranth accessions (particularly *A. tricolor*) from the collection that would be suitable for further analysis and genetic improvement. In Chapter 3, the genetic diversity of Vietnamese leafy amaranth (*A. tricolor*) is compared with that of *A. tricolor* collected

from other parts of the world. This analysis was performed by developing simple sequence repeat (SSR) markers for *A. tricolor* based on PacBio sequence data obtained from the genomic DNA of the cultivar ‘Biam’ and using these to evaluate the genetic diversity and structure of Vietnamese and overseas accessions of this species. This information allowed the genetic diversity indices, genetic distancnd population structure of the germplasm to be assessed, which will help breeders to effectively exploit genetic resources in their breeding programs and will also be valuable for effectively managing and preserving the genetic resources in gene banks. In Chapter 4, the morphological variation of Vietnamese *A. tricolor* accessions is evaluated, the results of which will be useful for prioritizing suitable traits in improvement programs. Finally, Chapter 5 provides a comprehensive discussion of the results of the above studies and summarizes the main conclusions.

CHAPTER 2: Classification of amaranth germplasm accessions in Vietnam using the *matK* marker

2.1. Introduction

The evaluation of genetic diversity is a key prerequisite for crop breeding programs and is also useful for the management and exploitation of plant genetic resources from gene banks (Shukla *et al.* 2009). However, *Amaranthus* has been widely considered a "difficult" genus for the determination of taxonomy and analysis of genetic variation based on morphology alone due to the overall similarity among many species, the small sizes of the diagnostic parts, and the presence of intermediate forms (Costea and DeMason 2001, Achigan-Dako *et al.* 2014, Assad *et al.* 2017). Therefore, to overcome these difficulties, the genetic diversity of amaranths has recently been investigated using chloroplast sequences (Viljoen *et al.* 2018, Waselkov *et al.* 2018).

The *matK* gene is one of the most variable protein-coding chloroplast genes in angiosperms (Hilu *et al.* 2003, Yu *et al.* 2011) and has a very high evolutionary rate, making it suitable for phylogenetic analyses at taxonomic levels ranging from order to species (Hilu and Liang 1997, Hilu *et al.* 2003, Muller *et al.* 2006, Chase *et al.* 2007, Lahaye *et al.* 2008). An evaluation of eight potential DNA barcodes for biodiversity inventories using >1,600 samples collected in southern Africa and >1,000 species of Mesoamerican orchids suggested that *matK* is a universal barcode for flowering plants (Lahaye *et al.* 2008). Furthermore, *matK* primers showed high amplification (93.1%) and sequencing (92.6%) rates in 58 species from 47 families of angiosperm plants (Yu *et al.* 2011).

In a phylogenetic analysis of the family Amaranthaceae based on *matK* DNA sequence data, Ogundipe and Chase (2009) found that there was high bootstrap support

(99%) for the genus *Amaranthus* being a monophyletic group, as well as high congruence between the *matK* phylogenies of amaranth species and various multigene/multigenome phylogenies of angiosperms. In addition, Viljoen *et al.* (2018) undertook a phylogenetic evaluation of 59 *Amaranthus* accessions representing 13 known species and 14 unknown species by comparing the barcoding genes *matK*, ribulose-bisphosphate carboxylase (*rbcL*), and the internal transcribed spacer (ITS) with the whole chloroplast sequence, which showed that leafy amaranths (*A. tricolor*, *A. blitum*, *A. viridis*, and *A. graecizans*) formed a stable sister lineage, while the weed amaranths (*A. retroflexus* and *A. powellii*) shared a common ancestor with the so-called "hybridus complex," which includes the three grain amaranths (*A. hypochondriacus*, *A. caudatus*, and *A. cruentus*) and their putative ancestors (*A. hybridus* and *A. quitensis*). Furthermore, the same authors found that the accessions of leafy amaranths and weed amaranths grouped together at the species level with high bootstrap values, whereas there was a poor resolution of species in the hybridus complex. Another recent study on the phylogenetic relationship among 58 amaranth species based on ITS, three low-copy nuclear genes [A36, glycerol-3-phosphate dehydrogenase (G3PDH), and *Waxy*], and the sequence data from two chloroplast regions (*trnL*5′*–trnL*3′ and *matK*/*trnK*) indicated the presence of three major clades that were roughly similar to the three subgenera of *Amaranthus* that were morphologically classified by Mosyakin and Robertson (1996), Costea and DeMason (2001), and Bayón (2015).

Therefore, in this study, the chloroplast *matK* marker was used to classify leafy amaranth species from Vietnam.

2.2. Materials and Methods

2.2.1. Plant materials and DNA isolation

A total of 272 VA accessions that were held at the PRC in Vietnam (Supplemental Table 1) and 27 world *Amaranthus* (WA) accessions that had been collected from other countries (Supplemental Table 2) were used in this study. Among the VA accessions, 268 represented unknown species that had been collected in eight ecological areas of Vietnam (Figure 2.1) and 4 were imported from the World Vegetable Center (WorldVeg). Among the WA accessions, 11 were provided by the United States Department of Agriculture (USDA) National Plant Germplasm System and 16 were preserved at the University of Tsukuba, Japan. The WA collection included 26 accessions of *A. tricolor* and 1 accession of *A. hypochondriacus*, which was used as an outgroup species in this study. DNA was extracted from the first true leaf of each seedling 14 days after sowing using a DNeasy® Plant Mini Kit (Qiagen, Hilden, Germany) according to the manufacturer's protocol with minor modifications.

2.2.2. Phylogenetic analysis

Phylogenetic analysis was performed using the *matK* marker developed by Ooi *et al.* (1995). The *matK* gene was amplified from the 272 VA and 27 WA accessions using the primer pair *matK*-F (5′-CTATATCCACTTATCTTTCAGGAGT-3′) and *matK*-R (5′-AAAGTTCTAGCACAAGAAAGTCGA-3′). Each sample was amplified using Ampdirect Plus PCR Mastermix (Ampdirect Plus Buffer, Blend Tag Plus DNA Polymerase, and 3 µM forward and reverse primers) in a final volume of 10 µl. Polymerase chain reaction (PCR) was then performed in a thermal cycler under the following conditions: denaturation at 94°C for 2 min, 40 cycles at 94°C for 30 s, 55°C for

30 s, 72°C for 1 min, and a final extension at 72°C for 10 min. The PCR products were cleaned using ExoSAP-IT™ Express PCR Product Cleanup and stained with the BigDye® Terminator v3.1, and the DNA precipitate was dissolved in 10 µl of Hi-Di™ formamide. The DNA was then sequenced in an Applied Biosystems DNA Sequencing analyzer at the Plant Breeding Laboratory of the University of Tsukuba.

The sequencing data were analyzed with Sequencing Analysis Software Version 5.1.1 and the GeneStudio software. The sequences obtained for the 272 VA and 27 WA accessions were aligned using the ClustalW algorithm in Molecular Evolutionary Genetics Analysis Version 7.0 (MEGA7) (Kumar *et al.* 2016) with default settings, and the alignment sequences were used to produce a dendrogram based on the maximum likelihood algorithm in MEGA7. The *matK* sequences of 45 Germplasm Resources Information Network (GRIN) amaranth accessions of known species that had been deposited in GenBank (KX079543–KX079587; Viljoen *et al.* 2018) were used as references.

2.3. Results

Alignment of the *matK* gene sequences from the 272 VA, 27 WA, and 45 GRIN accessions gave a total length of 434 bp per accession and contained 96.8% constant, 3.2% variable, and 2.5% parsimoniously informative sites (Table 2.1). The phylogenetic relationships that were inferred for the 344 amaranth accessions based on all of the nucleotide sites are shown in Figure 2.2. Leafy amaranths (*A. tricolor*, *A. blitum*, and *A. viridis*) and weed amaranths (*A. spinosus*, *A. retroflexus*, and *A. powellii*) clustered with accessions from the same species, with the exception of one *A. powellii* accession (GRIN28) that was grouped with the hybridus complex clade. The three grain amaranths (*A. hypochondriacus*, *A. caudatus*, and *A. cruentus*) and their putative ancestors (*A. hybridus* and *A. quitensis*) also clustered together.

All GRIN and 25 WA accessions of *A. tricolor* were placed in the same clade as 120 VA accessions (unknown species) with high bootstrap support (85%). The *A. tricolor* clade was clearly divided into two subclades, of which contained three accessions (WA012, WA016, and GRIN41 = WA129). Two GRIN accessions of *A. blitum*, which is another vegetable amaranth species that is native to Asia, clustered with 18 VA and 1 WA (WA008) accessions with 61% bootstrap support. All GRIN accessions of *A. viridis* were placed in a single clade with high bootstrap support (83%), together with one VA accession of *A. viridis* (VA226) and five VA unknown-species accessions. We also determined the positions of three weed amaranth species (*A. retroflexus*, bootstrap support of 83%; *A. powellii*, 62%; *A. spinosus*, 74%). However, none of the VA accession fell into these clades.

Two of the four GRIN accessions of *A. dubius*, which is a wild relative of the grain amaranths, clustered with 41 of the VA accessions (62% bootstrap support), whereas the

remaining two were placed in the hybridus complex clade (48% bootstrap support), which also contained the three grain amaranth species and their putative ancestors. The hybridus complex clade also included five GRIN accessions of *A. caudatus*, five GRIN accessions of *A. hybridus*, three GRIN accessions of *A. cruentus*, six GRIN accession of *A. hypochondriacus*, two GRIN accessions of *A. quitensis*, one WA accession of *A. hypochondriacus* (WA10), one VA accession of *A. hypochondriacus* (VA228), one VA accession of *A. dubius* (VA227), one VA accession of *A. caudatus* (VA179), and 84 VA accessions of unknown species.

Table 2.1. Nucleotide sequence analysis of the *matK* marker in 272 *Amaranthus* accessions from Vietnam and 27 amaranth accessions from over the world

Marker	Length (bp)	Constant sites (%)	Variable sites (%)	Parsimoniously informative sites (%)
matK	434	96.8	3.2	2.5

Table 2.2. Identification of Vietnamese *Amaranthus* (VA) accessions using the maximum likelihood method based on the Tamura-Nei model in MEGA7

Species	Number of VA accessions	Percentage (%)
A. tricolor L.	120	44.1
A. dubius Mart. ex Thell	41	15.1
A. blitum L.	18	6.6
A. viridis L.	6	2.2
Hybridus complex	87	32.0
Total accessions	272	100

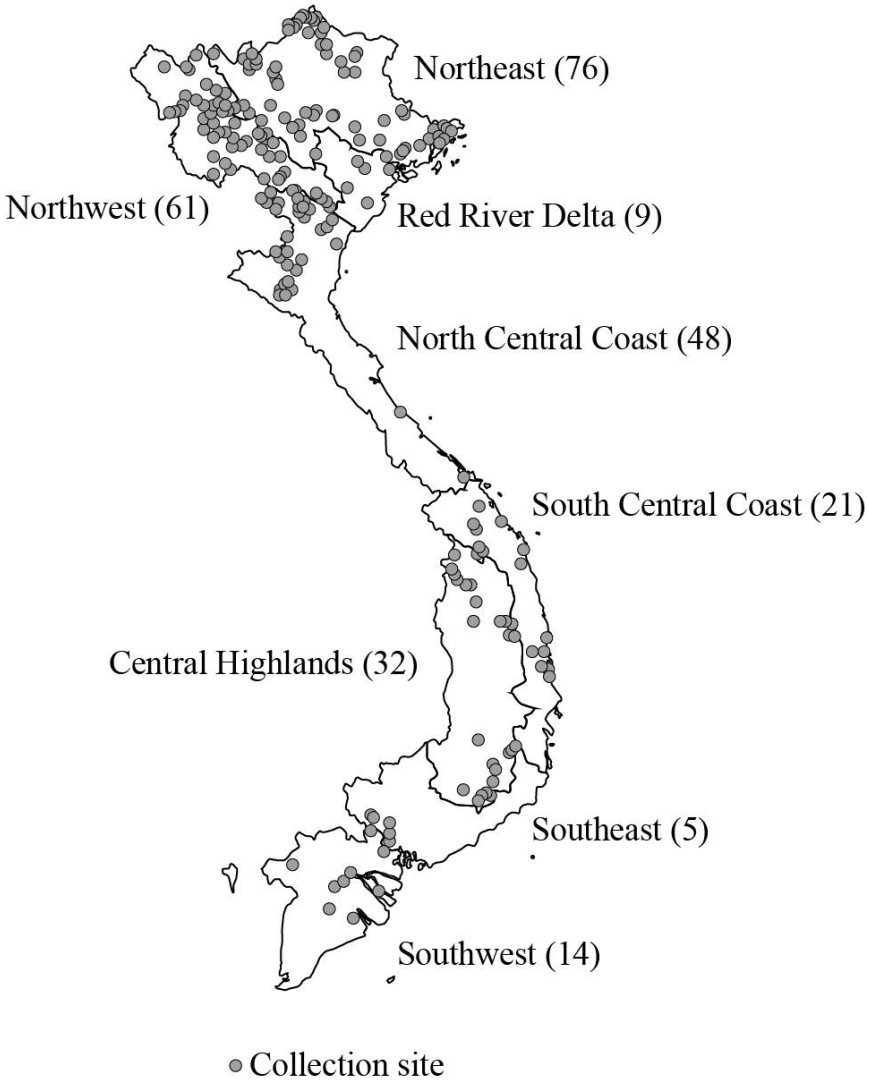


Figure 2.1. Map of Vietnam showing the collection sites of the *Amaranthus* accessions used in this study

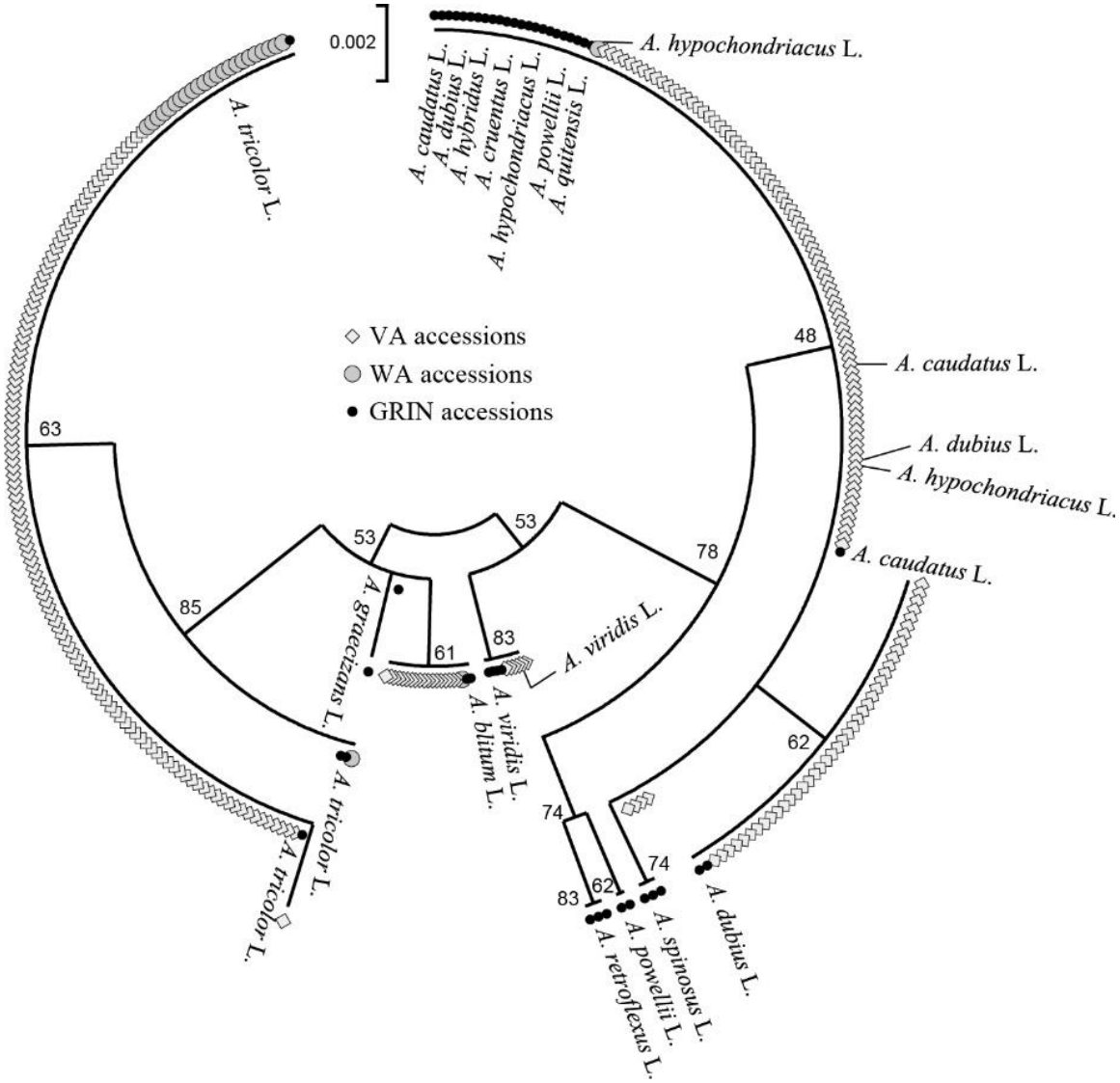


Figure 2.2. Phylogenetic dendrogram of the 344 *Amaranthus* accessions analyzed in this study as inferred using the maximum likelihood method based on the Tamura-Nei model in MEGA7 with 1,000 bootstrap replicates. VA, Vietnamese *Amaranthus*; WA, world *Amaranthus*; GRIN, Germplasm Resources Information Network

2.4. Discussion

Alignment of the chloroplast *matK* marker sequences (434 bp) from 272 VA and 27 WA accessions indicated that 96.8% of the sites were constant while the remaining 3.2% were variable (Table 2.1). The length of the aligned sequence was much shorter than has been observed in previous studies on the genetic diversity of amaranths using the *matK* marker (Viljoen *et al.* 2018, Waselkov *et al.* 2018). However, the percentage of parsimoniously informative sites (approximately 2.5%) was higher than was detected by Viljoen *et al.* (2018) (approximately 2.15%).

A phylogenetic tree was constructed by combining these sequences with those of 45 GRIN amaranth accessions that were previously analyzed by Viljoen *et al.* (2018) using whole chloroplast genome sequences. The interrelationships among the GRIN accessions, which included leafy amaranths (*A. tricolor*, *A. blitum*, *A. viridis*, and *A. graecizans*), weed amaranths (*A. spinosus*, *A. retroflexus*, and *A. powellii*), and *A. dubius*, were very similar between the dendrogram produced in the present study and that presented by Viljoen *et al.* (2018: fig.4). Most of the WA accessions of *A. tricolor* were assigned to the same group as the GRIN *A. tricolor* accessions, with the exception of WA008, which was assigned to the same clade as the *A. blitum* group, and this relationship was confirmed in another study (our unpublished data). The present phylogenetic analysis classified the species of 68% of the Vietnamese accessions, including 120 accessions of *A. tricolor* with a high bootstrap support (85%), 41 of *A. dubius*, 18 of *A. blitum*, and 6 of *A. viridis*. *Amaranthus tricolor* accounted for the largest percentage of accessions (44%) in the Vietnamese collection (Table 2.2), which agrees with a widely held view that *A. tricolor* is one of the most economically important vegetable amaranths in Vietnam. In this study, we could identify *A. dubius* and 3 (*A.*

tricolor and *A. blitum, A. viridis*) out of 7 amaranth species recorded in Vietnam by Pham (1999), which included *A. tricolor*, *A. caudatus*, *A. viridis*, *A. spinosus*, *A. hybridus*, *A. lividus* L. (synonyms *A. blitum*), and *A. retroflexus*.

In conclusion, the genetic diversity of 272 accessions that have been preserved in the PRC in Vietnam were successfully identified using the chloroplast DNA marker *matK*. Alignment of the *matK* gene sequences from the VA and WA accessions gave a total length of 434 bp per accession, 2.5% of which represented parsimoniously informative sites. The majority of the VA accessions (68%) could be identified to species using only the *matK* barcoding marker, among which 120 were identified as *A. tricolor*, 41 as *A. dubius*, 18 as *A. blitum*, and 6 as *A. viridis*. The remaining accessions were classified to grain amaranth species, such as *A. hypochondriacus*, *A. caudatus*, and *A. cruentus*. The genetic information obtained from this research may be useful for effectively managing and exploiting the genetic resources in gene banks and for genetically improving amaranths.

CHAPTER 3: Genetic diversity of *A. tricolor* in Vietnam by using SSR markers in comparison with worldwide amaranths

3.1. Introduction

Amaranthus tricolor is a popular summer vegetable in several developing countries in Asia and Africa owing to its excellent nutritional value, high adaptability to severe conditions, and an absence of major diseases (Andini *et al.* 2013, Achigan-Dako *et al.* 2014). However, only limited information is available on its genetic diversity and little effort has been made toward genetically improving this crop to increase its potential yield and quality (Shukla *et al.* 2009). The amaranths that have been collected from eight ecological areas of Vietnam are expected to contain a high level of genetic variation and to offer desirable traits for crop breeding programs. Therefore, their identification and evaluation will contribute to the improvement of leafy amaranths.

Over the last decades, SSRs have been considered the most powerful molecular markers for genetic analyses due to (1) their versatility, (2) high polymorphism provided by a large number of alleles per locus, (3) co-dominance making them suitable for direct assessments of heterozygosity, (4) small amounts of DNA required, and (5) high transferability to related species (Varshney *et al.* 2005, Kumar *et al.* 2009, Guichoux *et al.* 2011, Vieira *et al.* 2016, Taheri *et al.* 2018). However, the development of SSRs is labor-intensive, economically costly, and time-consuming (Taheri *et al.* 2018). The first study on the development of SSR markers for amaranths was carried out by Mallory *et al.* (2008), who produced 179 polymorphic SSR markers based on the genomic DNA of *A. hypochondriacus*. These markers amplified a total 731 alleles in 35 accessions representing 6 species (*A. hypochondriacus*, *A. caudatus*, *A. cruentus*, *A. hybridus*, and *A. powellii*), with an average of 4 alleles per locus and an expected heterozygosity (H_e) of

0.62. In the same year, Lee *et al.* (2008) isolated and characterized 12 polymorphic SSR markers for *A. hypochondriacus*, which detected a total of 92 alleles across 20 accessions, with an average of 7.7 alleles per locus and observed heterozygosity (H_o) and H_e values ranging from 0.00 to 0.95 and 0.49 to 0.92, respectively. Later studies on the genetic diversity of amaranths by Khaing *et al.* (2013), Singh *et al.* (2013), Wang and Park (2013), Kietlinski *et al.* (2014) and Suresh *et al.* (2014) were based only on 11–14 SSR markers, which were developed by Mallory *et al.* (2008) and Lee *et al.* (2008).

In the present study, eight accessions of *A. tricolor* were initially screened using the 179 SSR markers developed by Mallory *et al.* (2008) and 12 SSR markers developed by Lee *et al.* (2008). However, all of the tested markers were found to be monomorphic among these accessions. The advent of high-throughput sequencing technology along with bioinformatics tools has facilitated the rapid discovery of thousands of potentially amplifiable microsatellite regions at low cost, even for species with limited genomic information (Guichoux *et al.* 2011, Zalapa *et al.* 2012, Silva *et al.* 2013, Taheri *et al.* 2018). Therefore, in the present study, SSR markers were developed for *A. tricolor* based on PacBio sequence data obtained from the genomic DNA of the cultivar 'Biam' to evaluate the genetic diversity of *A. tricolor* in Vietnam and other countries.

3.2. Materials and Methods

3.2.1. Plant materials and DNA isolation

A total of 120 VA accessions that were identified as *A. tricolor* using the *matK* marker (see Chapter 2 and Supplemental Table 1) and 181 WA accessions collected from other countries (Supplemental Table 2) were used in this study. Among the WA accessions, 166 were provided by the USDA National Plant Germplasm System and 15 were preserved at the University of Tsukuba. The WA collection included 175 accessions of *A. tricolor* and 6 accessions of *A. hypochondriacus*, which was used as an outgroup species in this study. DNA was extracted from the first true leaf of each seedling 14 days after sowing using a DNeasy Plant Mini Kit (Qiagen, Hilden, Germany) according to the manufacturer's protocol with minor modifications.

3.2.2. Development of SSR primers for leafy amaranths

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

A total of 817 novel SSR markers were randomly selected from a pool of 27,690 markers, and amplification was assessed using 'Biam' DNA. The forward primers were 5′-labeled with the fluorescent dyes 6-carboxyfluorescein (6-FAM), VIC, NED, or PET (Shimizu and Yano 2011). PCR was carried out using a KAPA2GT Fast PCR Kit in a

final volume of 10 µl containing 1 µl of genomic DNA, 2 µl of 5X KAPA2G Buffer A, 0.2 µl of 10 mM deoxynucleoside triphosphates (dNTPs), 0.2 µl of 25 mM $MgCl_2$, 0.02 µl of 5 U/µl KAPA2G FAST DNA polymerase, 6.08 µl of Milli-Q water, and 0.5 µl of BStag primer mix. The reactions were performed in a thermal cycler (Applied Biosystems) under the following conditions: 94°C for 3 min; 30 cycles at 94°C for 20 s, 54°C for 30 s, and 62°C for 30 s; 3 cycles at 94°C for 20 s, 49°C for 10 s, and 72°C for 5 s; and a final extension at 72°C for 10 min. Amplification was assessed by electrophoresis in 8% polyacrylamide gels.

The polymorphism of the amplified SSR markers was assessed in eight WA accessions of *A. tricolor* (WA001, WA002, WA003, WA004, WA011, WA012, WA017, and WA022). Each sample was amplified using Ampdirect Plus PCR Mastermix (Ampdirect Plus Buffer, KAPA2G Fast DNA Polymerase, 0.5 µl of BStag primer mix) in a final volume of 5 µl, and the PCR reactions were performed in a thermal cycler (Applied Biosystems) under the conditions outlined above. The sizes of the amplified fragments were then estimated using an automated DNA analyzer (3130xl) with a GeneScan™ 600 LIZ® Size Standard and GeneMapper v. 4.0 software (Applied Biosystems).

3.2.3. Analysis of the genetic diversity and population structure using SSR markers

The genetic diversity of 301 accessions, including 120 VA accessions of *A. tricolor* (the species was determined using the *matK* marker) and 181 WA accessions (175 *A. tricolor* and 6 *A. hypochondriacus;* 25 *A. tricolor* and 1 *A. hypochondriacus* used in the matK test), was analysed by using 30 developed SSR markers (Table 3.1). PCR was performed in a C1000 Thermal Cycler (Bio-Rad Laboratories, Inc., Hercules, CA, USA) as outlined in section 3.2.2. The sizes of the amplified fragments were estimated by using an automated

DNA analyzer with the GeneScan-500LIZ size standard (Applied Biosystems). Fragment length was determined with GeneMapper software (Applied Biosystems).

The GenAlEx package v. 6.503 was used to calculate the number of alleles (Na), the number of effective alleles (Ne), observed heterozygosity (H_o), expected heterozygosity (H_e), information index (I), Nei's genetic distance, and the Principal Coordinates Analysis (PCoA) (Peakall and Smouse 2006, 2012). The major allele frequency and the polymorphic information content (PIC) were determined and a phylogenetic tree was constructed using PowerMarker v. 3.25 (Liu and Muse 2005). In addition, an analysis of molecular variance (AMOVA) was performed and pairwise population fixation index (F_{ST}) values were calculated using Arlequin 3.5.2.2 software (Excoffier and Lischer 2010).

The population structure was inferred using the STRUCTURE 2.3.4 software package (Pritchard *et al.* 2000). The number of subpopulations (*K*) was set at 1–9, with 10 replications per *K* value. The *K* value was determined by running an admixture and correlated allele frequencies model. Each run started with 1,000,000 burn-ins, followed by 100,000 Markov chain Monte Carlo iterations. The *ad hoc* statistic delta *K* (ΔK) was then calculated to detect populations using the online program STRUCTURE HARVESTER (Evanno *et al.* 2005).

3.3. Results

3.3.1. Genetic diversity of A. tricolor

Nine out of the 30 selected SSR markers performed poorly in GeneMapper, and one accession (VA276) was found to have a high missing rate (47.6%), resulting in its exclusion from the following analysis. The analysis indicated that the remaining 294 accessions of *A. tricolor* were supported as a monophyletic group, with 6 accessions of *A. hypochondriacus* formed an outgroup (Figure 3.1). The *A. tricolor* group was clearly divided into two subgroups, one of which contained 15 WA accessions, including WA012, WA016, and WA129, in agreement with the result obtained using the *matK* barcode.

The statistics for the 21 SSR markers that were used to genotype the 294 accessions of *A. tricolor* are summarized in Table 3.2. These markers amplified a total of 153 alleles, ranging from 2 (TAM002, TAM016, and TAM021) to 20 (TAM006) per marker, with a mean allelic richness (Na) of 7.29, and Ne of 2.28. The average values were major allele frequency, 0.71; H_o, 0.14; H_e, 0.38; and PIC, 0.35 (Table 3.2). The number of rare alleles (frequency <5%) was 102 (66.7%).

The genetic diversity indices for the VA and WA collections are shown in Table 3.3. The average number of alleles per marker was considerably lower in the VA collection (4.52) than in the WA collection (6.24). Similarly, the average values of almost all other indices (Na, Ne, H_e, I, and PIC) were also lower in the Vietnamese collection than in the WA one. The only exception was H_o (0.14 in both collections), indicating that most of the samples in the whole collection were highly homozygous. Among the overseas geographical groups, the average number of alleles ranged from 2.0 (Africa) to 4.90 (South Asia). The African collection also had the smallest values of Ne (1.79), I (0.48), H_e (0.28), and PIC (0.24); except for Ho. The American collection had the largest values

for I (0.90), H_e (0.52), and PIC (0.46). The average values of all indices were higher in the VA collection than in the commercial cultivars, which had the smallest or second smallest values for all indices except for H_e (Table 3.3). Among Vietnamese geographical groups, the average number of alleles ranged from 2.05 (Southeast) to 3.62 (Northeast). The average values of Na, Ne, I, H_e, and PIC were the largest in the Northeast collection. The Southeast and South Central Coast collections had the lowest and the second lowest for all the indices, except for Ho.

There was generally significant genetic differentiation (F_{ST}) among the WA geographical groups and the VA collection. The exceptions to this were the African group compared with all other groups and the American group compared with the Southeast Asian group. The pairwise F_{ST} ranged from 0.015 (Africa vs. East Asia) to 0.315 (America vs. East Asia) (Table 3.4). In the comparisons between the Vietnamese and overseas groups, F_{ST} was 0.284 for America, 0.163 for South Asia, 0.116 for Southeast Asia, 0.100 for commercial cultivars, 0.078 for East Asia, and 0.029 for Africa. Similar to the pairwise F_{ST}, Nei's genetic distance between Vietnam and other groups was also the highest for America, and the lowest for Africa and East Asia.

3.3.2. Population structure of A. tricolor

In the STRUCTURE analysis, the highest likelihood value of ΔK was observed at $K = 2$ with a secondary peak at $K = 4$ (Figure 3.2), indicating that the accessions could be grouped into two or four subpopulations.

At $K = 2$, all 294 accessions of *A. tricolor* were allocated to two subgroups, with membership probabilities (Q) of >0.95 without admixture. Most accessions (279) were assigned to subgroup 1, while only 15 (5.1%) were assigned to subgroup 2, all of which

were from overseas (9 from America, 5 from South Asia, and 1 from Southeast Asia; Figure. 3.3 and Table 3.5).

At $K = 4$, Q was >0.75 and 115 accessions (87.8% from East Asia and Vietnam) were assigned to subgroup 1, which included 90.7% of all East Asian accessions; 61 accessions were allocated to subgroup 2, most of which (96.7%) were from Vietnam; 83 accessions were assigned to subgroup 3, most of which (92.8%) were from South Asia; and 15 accessions were assigned to subgroup 4, which was the same as subgroup 2 at $K = 2$. Eight of the nine commercial cultivars were allocated to subgroup 1, while the ninth was assigned to an admixed group (Figure. 3.3 and Table 3.5). The result of the STRUCTURE analysis was supported by the PCoA (Figure 3.4), which assigned most accessions to four major subgroups; subgroup 1 mostly included accessions from East Asia and Vietnam; subgroup 2 from Vietnam; subgroup 3 from South Asia; subgroup 4 from America and South Asia. The first and second factors of the PCoA explained approximately 14.63% and 9.86%, respectively, of the variation in the genetic distance matrix.

3.3.3. Genetic diversity and population structure of the Vietnamese collection of **A. tricolor**

The AMOVA indicated that there were significant differences among groups at all hierarchical levels (Table 3.6). The percentage variation was highest among individuals within geographical groups (approximately 55.54%), followed by among loci within individuals (35.99%), while there were only low levels of variation between the VA and WA collections (5.78%) and among the geographical groups (2.69%).

The pairwise F_{ST} values between the WA collection and each Vietnamese group were significant and ranged from 0.065 (Central Highlands) to 0.157 (Southeast) (Table 3.7). By contrast, the pairwise F_{ST} values among Vietnamese groups ranged from 0.020 (Southeast

vs. North Central Coast) to 0.127 (Southeast vs. South Central Coast) and was only significant for some comparisons. Similar to the pairwise F_{ST}, Nei's genetic distance tended to be higher between the WA collection and each Vietnamese geographical group than among the Vietnamese groups (Table 3.7).

The STRUCTURE analysis of the 119 VA accessions of *A. tricolor* using 21 SSR markers showed that the highest likelihood value of ΔK was obtained at $K = 8$ (Figure 3. 5), indicating that the accessions could be grouped into eight subpopulations with a *Q* value of 0.75. In total, 13 accessions were assigned to subgroup 1, 5 to subgroup 2, 12 to subgroup 3, 13 to subgroup 4, 10 to subgroup 5, 9 to subgroup 6, 19 to subgroup 7, and 9 to subgroup 8. The admixed group had the largest number of accessions (29) and each subgroup included accessions from several geographical groups (Figure 3. 6 and Table 3.8). All accessions in subgroups 2, 4, and 7 were present in subgroup 1 at K=4 STRUCTURE analysis using whole accessions; and all of the accessions in subgroups 5 and 8 and most of the accessions in subgroups 1 (84.6%), 3 (75.0%), and 6 (78%) were present in subgroup 2 at K=4 STRUCTURE analysis using whole accessions.

Table 3.1. Characteristics of the 30 simple sequence repeat (SSR) markers that were developed from the sequence data of *Amaranthus tricolor* 'Biam'

Marker	Forward primer sequence	Reverse primer sequence	Repeat motif	Size (bp)
TAM001	ACCTCTCTCAGAAAGCCGTG	AAGCGAACCAGAAAGTGCAG	(AAAT)8	319
TAM002	ATGTGGTTAAGGCCTGAGGG	CCAACACCATCATAGAGTCCAC	(AAAT)6	204
TAM003	GGGAGGTTTAAGGGAGTGGG	ACCCTTCTCTTCAACCCACC	(AAT)6	209
TAM004	AAATAGTGGAGGGCTCGGAC	CTCGTAACGGGTGCAACTTC	(AAT)7	122
TAM005	AATGGGTTAGGGCGGGTATC	CAGCTTAACACTGCGGCTAC	(AT)8	213
TAM006	TGCATGCCAAGTCCGAATTC	GCTCAGCTTGTTGTCGACTC	(AAT)7	408
TAM007	ATCTGACCGTGGCTGATACC	ACGATGAGTATGGGCCCTTC	(AT)7	395
TAM008	TCTAGCAGTCACCGGTTAGC	CAGGTTTACGCTGGCAGAAG	(AAT)6	399
TAM009	CACGTACACTGTAGCTTGCG	TTGATGGTGCTCCCTACGTC	(ATC)6	390
TAM010	TTGTTTGGCACTGTATCGCC	GGGATAGGTAGTTGGAGGCG	(AAT)6	229
TAM011	AACGGCGCTATAAGTTGCAG	GTGCAAGCAGAGGGAAACAG	(ATC)6	288
TAM012	GGCTCGCCCATTTACTCATG	ACCTAGATCTGCGTTCGGTC	(AAT)6	212
TAM013	CCGAGGGAAACACGTCTTTG	CCGCAGCAAATAAGGGAAGG	(AT)7	380
TAM014	AAATGCGAGCAGTTCAGTCG	CACATGTGCAGAGGTTTGGG	(AAT)8	399
TAM015	TAAAGCCCGCATGATTCGTG	CGTTTGGGAATACGAGCACC	(AAG)7	192
TAM016	AACCGTGGGAAATGCTTTGG	CACGTTCGTTACTTCCCAGC	(AT)7	417
TAM017	GTGATTCCGCGTTCCATCTG	CAGAGTCCAAACACCGATGC	(AT)8	316
TAM018	TTTATGCGCCAAACTCCGAC	AGAGGGAAGATGCCATGTCC	(AT)9	402
TAM019	GACGCTTGGCTCTTGTTCTC	GGGTGCTCAGCTTGTTTCTC	(AAT)7	314
TAM020	GCAACGTGACCCAGAACATC	TCCGTCTTCTAACCATGGCC	(AAC)7	365
TAM021	TGGGTATAGTTCGCCACACC	ACACCCTTTCCTCCCTTTCC	(ATC)7	430
TAM022	TCGCACTGATTCAACGGTTG	AGCCCATCTCCCATAAGCAG	(AC)8	447
TAM023	GCGCACACCATCTTGTAGTG	GGCGTCCTCCTAGCATAGTC	(ATC)6	264
TAM024	TGCGTCTTTGTGATTGTCCG	ACGCGACAAGATATCCTCCC	(ATC)6	450
TAM025	CCGATTAAAGAATGGCGGCC	ATGTCCGTATCCTGTGTCCG	(AAG)8	408
TAM026	CAAGGCCAACACCAGCTATG	AGTATGAGGGTACGTGCTGC	(AGG)7	362
TAM027	ACTCACCCACACCATCTTCC	AATTGTTAGCACGAGTCGCC	(ATC)8	341
TAM028	CCCATGTGCATCAATCGACC	CACCTCCTTCAACAATGGCC	(AGG)6	262
TAM029	CTCCATCTCTCCTTGTCCCG	TTGCGAGGCCTTTAAAGCAG	(AT)8	435
TAM030	TACCGTCACATCCGTCGTAC	AGTTGCGCAAAGAGTTCGTC	(AAT)9	440

Table 3.2. Statistics for the 21 simple sequence repeat (SSR) markers that were used for genotyping 294 accessions of *Amaranthus tricolor*

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

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

Table 3.3. Genetic diversity indices of the world *Amaranthus* (WA) and Vietnamese *Amaranthus* (VA) collections of *A. tricolor* showing variations among the geographical groups

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

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

Table 3.4. Pairwise fixation index (F_{ST}; below the diagonal) and Nei's genetic distance (above the diagonal) among geographical groups from other countries and the Vietnamese collection of *Amaranthus tricolor*

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

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

Table 3.5. Geographical groupings of *Amaranthus tricolor* accessions as assumed by STRUCTURE

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

[a] Membership probabilities (*Q*) > 0.95.
[b] Membership probabilities (*Q*) > 0.75.

Table 3.6. Hierarchical analysis of molecular variance (AMOVA) among the eight Vietnamese geographical groups and the world *Amaranthus* (WA) collection for *A. tricolor*

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

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

Table 3.7. Pairwise fixation index (F_{ST}; below the diagonal) and Nei's genetic distance (above the diagonal) values among eight Vietnamese geographical groups and the world *Amaranthus* (WA) collection for *A. tricolor*

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

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

Table 3.8. Assignment of Vietnamese *Amaranthus tricolor* accessions to subgroups by the STRUCTURE analysis

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

All subgroups had membership probabilities (Q) > 0.75.

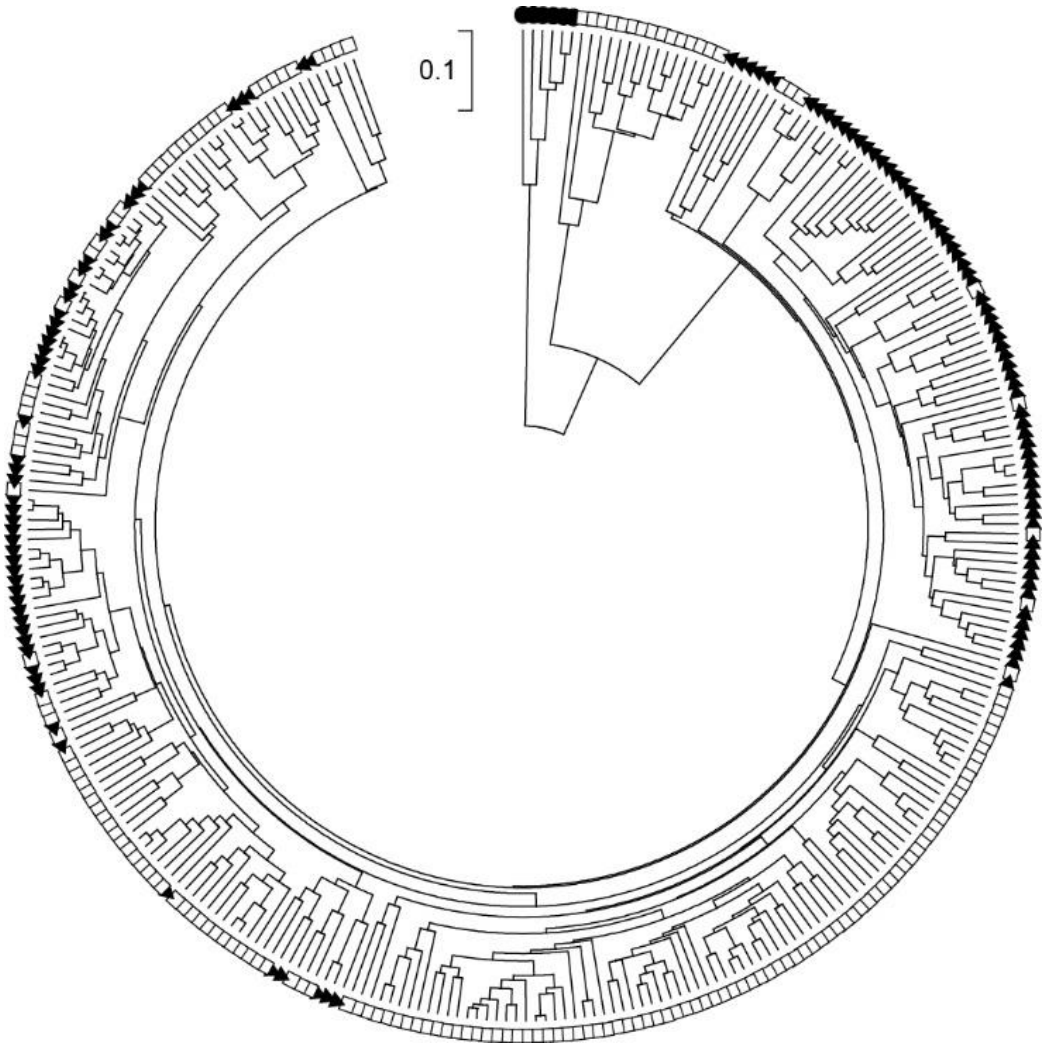


Figure 3. 1. Phylogenetic tree of 300 *Amaranthus* accessions constructed using the neighbor-joining method based on the Nei (1983) distance in PowerMarker. Open squares, world *Amaranthus* (WA) accessions of *A. tricolor*; black triangles, Vietnamese *Amaranthus* (VA) accessions of *A. tricolor*; black circles, *A. hypochondriacus* accessions (outgroup)

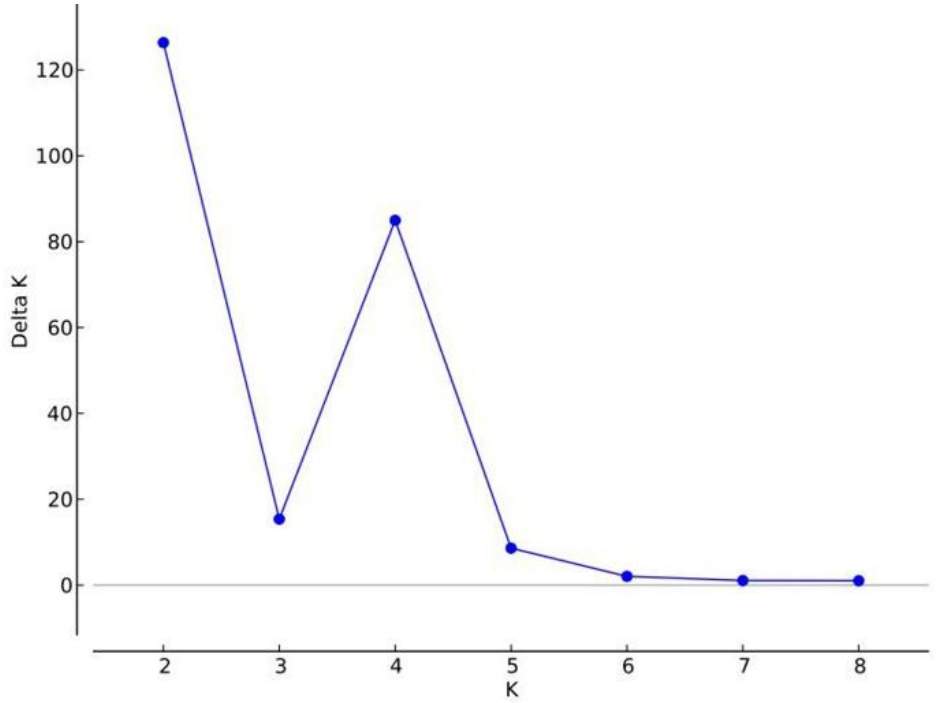


Figure 3.2. Δ*K* from the STRUCTURE analysis of 294 accessions of *Amaranthus tricolor* using the Evanno method

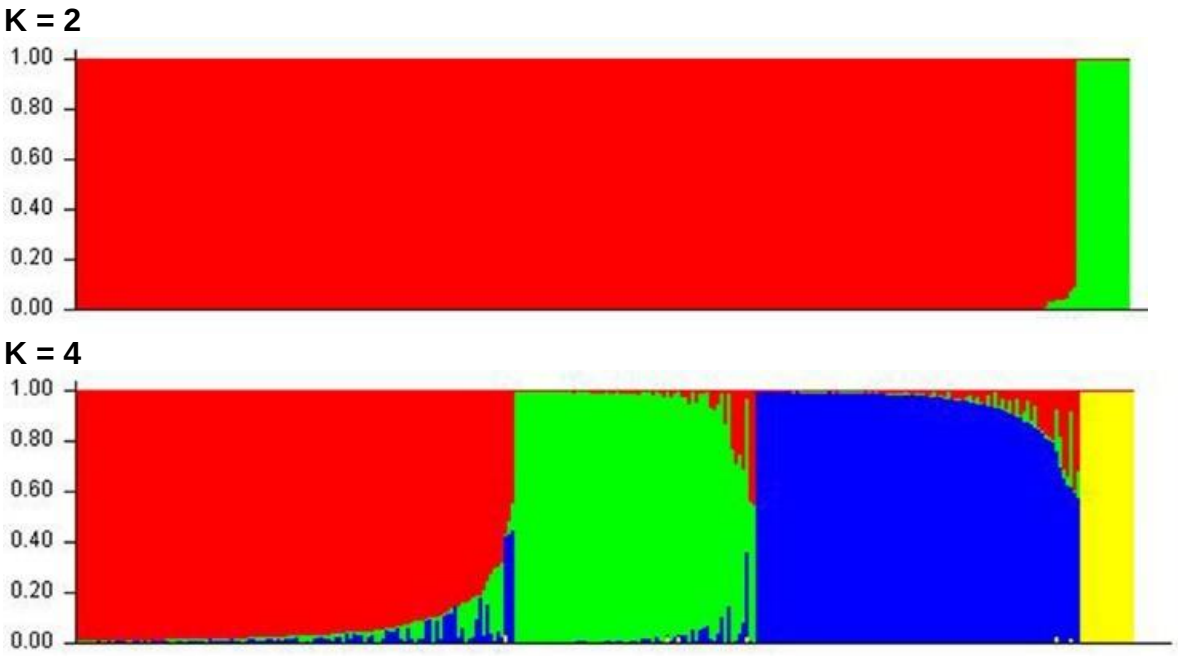


Figure 3.3. Population structure of 294 accessions of *Amaranthus tricolor* determined for different values of *K* using 21 simple sequence repeat (SSR) markers

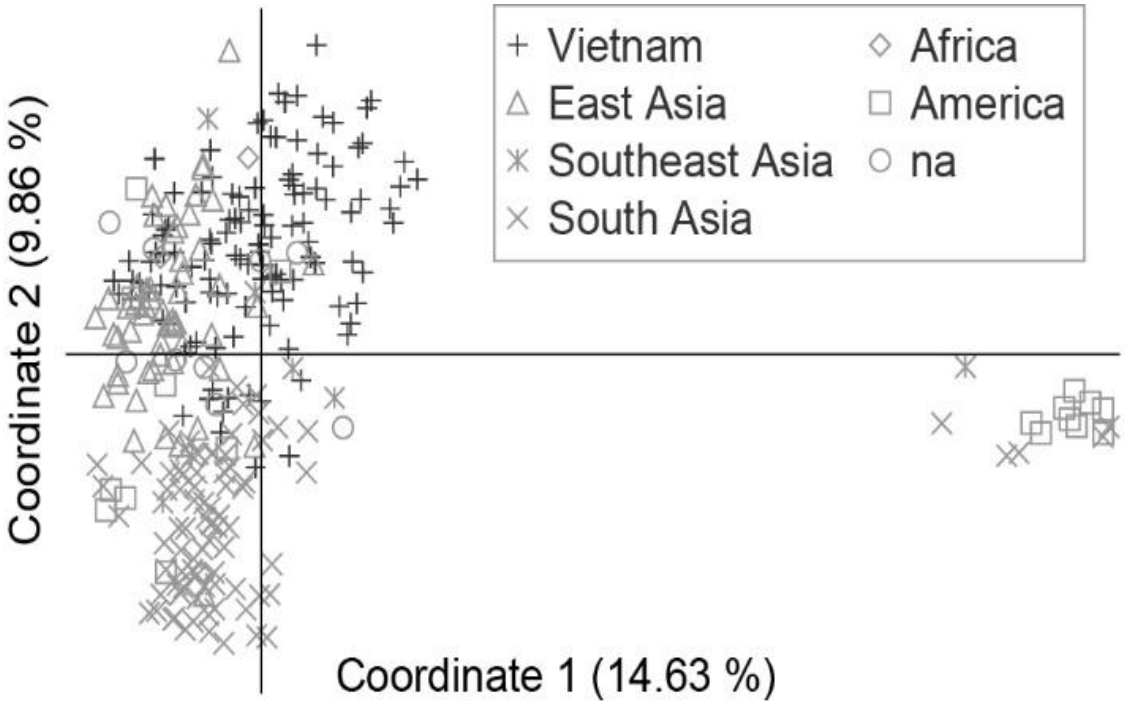


Figure 3.4. Principal coordinates analysis (PCoA) of 294 accessions of *Amaranthus tricolor* showing their areas of origin

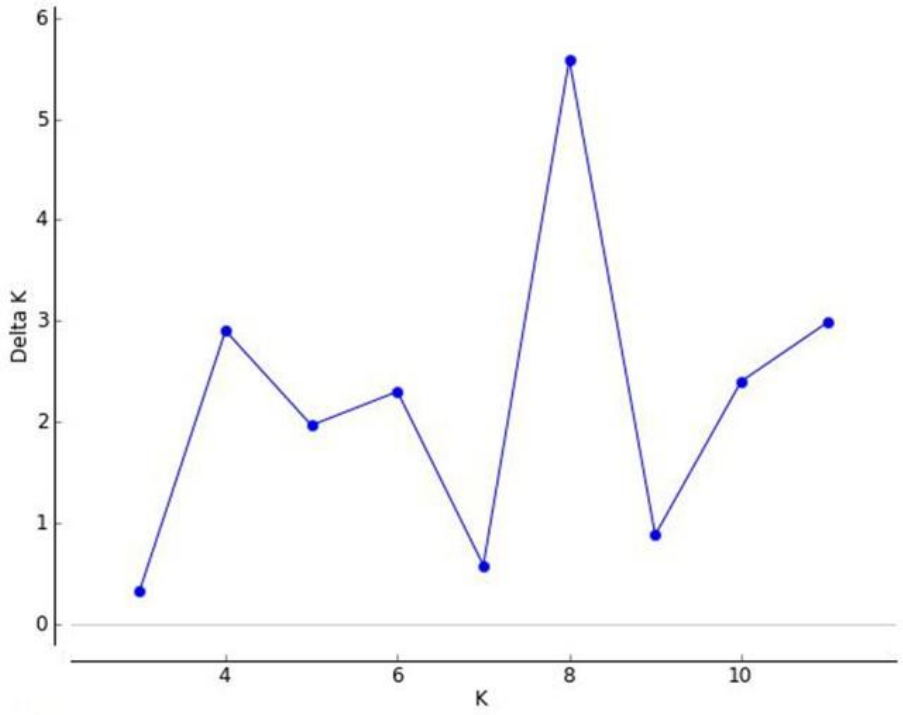


Figure 3.5. Δ*K* from the STRUCTURE analysis of 119 Vietnamese *Amaranthus* (VA) accessions of *A. tricolor* using the Evanno method

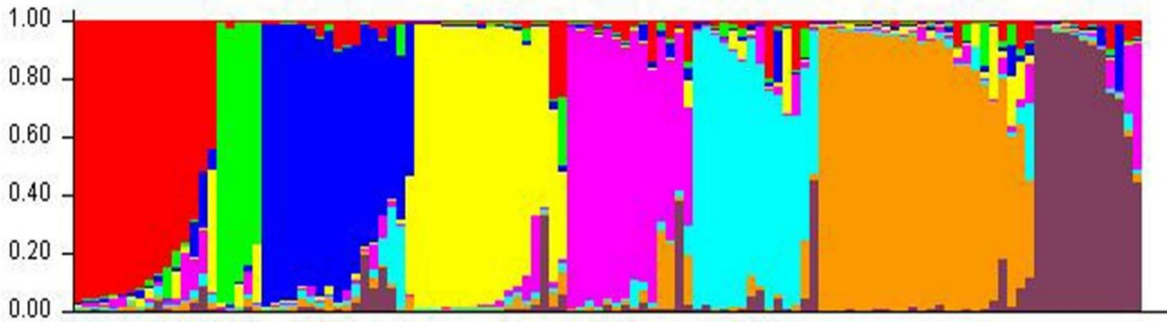


Figure 3. 2. Population structure of 119 Vietnamese *Amaranthus* (VA) accessions of *A. tricolor* determined at $K = 8$ using 21 simple sequence repeat (SSR) markers

3.4. Discussion

In this study, 21 SSR markers were developed from the PacBio sequence data for *A. tricolor* 'Biam' and were used to genetically analyze 294 accessions of *A. tricolor* from Vietnam and overseas. These markers successfully amplified single alleles with relatively high levels of polymorphism and so would be useful in studies on the genetic diversity of leafy amaranths. The 21 markers amplified a total of 153 loci in 294 accessions, with an average of 7.29 alleles per locus, Ho of 0.14, He of 0.38, and PIC of 0.35. The average allele richness in our study was larger than in Mallory *et al.* (2008) and Wang and Park (2013) (4.0 and 4.79, respectively), but smaller than in Khaing *et al.* (2013) and Suresh *et al.* (2014) (12.9 and 11.1, respectively). These differences may have resulted from the use of different SSR markers, as well as differences in the sample sizes, which were smaller in the former two studies and much larger in the latter two studies than in the present study. The average genetic diversity indices that were calculated in the present study were smaller than have been reported in most previous studies, such as Suresh *et al.* (2014) (H_o, 0.29; H_e, 0.70; PIC, 0.66), Khaing *et al.* (2013) (PIC, 0.71), Lee *et al.* (2008) (H_o, 0.28; H_e, 0.74), and Mallory et al. (2008) (H_e, 0.71). This may have been due to the much larger numbers of species that were used in all these studies, but particularly the first two (33 and 57, respectively), compared with the present study. In this study, the average values of nearly all of the genetic diversity indices were smaller for the VA collection than for the WA collection, and this result was supported by the STRUCTURE analysis (Table 3.5; Figure 3.3), which assigned all of the VA accessions to the same subgroup at $K = 2$. Thus, it appears that the VA collection of *A. tricolor* has a lower level of genetic diversity than the WA collection.

The genetic differentiation analysis showed that there were significant differences between the Vietnamese group and all overseas groups except that from Africa (Table 3.4). The F_{ST} was very large between Vietnam and America (0.284), large between Vietnam and South Asia (0.163), and moderate between Vietnam and East Asia (0.078), indicating that the VA collection is relatively closely genetically related to the East Asian collection but is divergent from the South Asian and American collections. This finding was confirmed by Nei's genetic distance (Table 3.4) and the PCoA results (Figure 3.4), which showed that the VA accessions were differentiated from the American and South Asian accessions but close to the East Asian accessions. Interestingly, the VA accessions are genetically closer to the East Asian accessions (F_{ST}= 0.078; Nei's distance = 0.042) than the Southeast Asian accessions (F_{ST}= 0.116; Nei's distance = 0.093). This may be because all East Asian accessions originated from China (a neighbor of Vietnam), whereas none of the Southeast Asian accessions came from Laos or Cambodia (other neighbors of Vietnam). Another possibility is that amaranths dispersed from East Asia (China) to Vietnam partly through the migrations of the ancestors of most of today's Vietnamese ethnic groups from south of the Yangtze River in China (Hall and Patrinos 2012) and the introduction of this crop during the long period of Chinese domination in Vietnam from 111 BC to 938 AD (Le *et al.* 1991). The African accessions were closely related to the Vietnamese, East Asian, and Southeast Asian accessions ($F_{ST} < 0.05$), but this relationship was not significant, partly due to the small size of the African collection (n = 3). Finally, the American collection had a very large genetic distance from most groups ($F_{ST} > 0.25$).

The STRUCTURE analysis indicated that the accessions could be grouped into two or four subpopulations (Figure 3.3). At K = 2, all accessions were assigned to two

subgroups ($Q > 0.95$) without admixture, indicating rare genetic communication between the subgroups. At $K = 4$, all accessions were allocated to four subgroups with one admixed group. Among these, subgroup 1 mostly included accessions from East Asia and Vietnam, and contained 90.7% of all East Asian accessions; subgroup 2 mostly included accessions from Vietnam; subgroup 3 mostly included accessions from South Asia; and subgroup 4 included 60% of the accessions from America (Table 3.5). These results demonstrate that the four subgroups are associated with different geographical areas, which partly supports the result of Wang and Park (2013), who claimed that there is relatively little genetic exchange within or among amaranth species from South and Southeast Asia. By contrast, in a STRUCTURE analysis of 348 amaranth accessions from 33 species using 11 SSR markers, Suresh *et al.* (2014) found no significant correlation between geographical distance and genetic diversity. The relationship that was observed between amaranth accessions and their origins in the present study can partly be explained by the low level of genetic exchange of *A. tricolor* among these geographical areas. The results of the STRUCTURE analysis at $K = 4$ also showed that Vietnamese *A. tricolor* could be divided into two major types: a common type in East Asia (52 accessions belonging to subgroup 1) and a unique type in Vietnam (61 accessions belonging to subgroup 2, which contained only one accession from each of East Asia and Southeast Asia). Thus, it appears that the genetic diversity of Vietnamese *A. tricolor* resulted from dispersal events mainly from East Asia and adaptation to the local environments through on-site cross pollination, as well as the retention of cross-pollinated-seeds until the next season without artificial selection.

Nine commercial cultivars that are commonly in China, India, Japan, and Southern Asia were also included in this study (six of these cultivars were obtained from Evergreen

Seeds Company, Anaheim, California, USA and three were sourced from Japanese companies). These cultivars were found to have a lower level of genetic diversity (Table 3.3) and a relatively close genetic distance (Table 3.4) to the VA collection. The close relationship between these cultivars and the VA accessions was confirmed by the STRUCTURE analysis (Table 3.5; Figure 3.3), which assigned eight of the commercial cultivars to subgroup 1 (52 accessions from Vietnam) at $K = 4$. However, no commercial cultivars were assigned to subgroup 2, which is considered a unique type in Vietnam.

The variation among Vietnamese geographical groups was significant but small at approximately 2.69% (Table 3.6). This observation was supported by the genetic differentiation analysis, in which the F_{ST} and Nei's genetic distance values for pairwise comparisons between the Vietnamese geographical groups ranged from low to moderate (Table 3.7). The STRUCTURE analysis grouped all VA accessions into eight subpopulations ($Q > 0.75$) and one admixed group, with accessions from different geographical areas being represented in each subgroup (Figure 3.6 and Table 3.8). This result suggests that there is no relationship between geographical distance and genetic diversity among *A. tricolor* accessions in Vietnam and indicates that there is a high level of genetic exchange.

In conclusion, the SSR markers that were developed in this study allowed the overall genetic diversity of the VA and WA collections of *A. tricolor* and the genetic relationships among geographical groups to be assessed. The STRUCTURE, PCoA, and F_{ST} analysis indicated that there is a positive relationship between geographical distance and genetic diversity among the VA collection and most of the overseas groups but not among the Vietnamese geographical groups. Although the genetic diversity of the VA collection was lower than that of the WA collection, Vietnamese *A. tricolor* could be

divided into two major types: a common type in East Asia and a unique type in Vietnam. The genetic information obtained from this research may be useful for effectively managing and exploiting the genetic resources in gene banks and for genetically improving *A. tricolor*.

CHAPTER 4: Analysis of the morphological diversity of Vietnamese *A. tricolor*

4.1. Introduction

Information about the genetic diversity of plant species is essential for the effective utilization of genetic resources and the development of breeding programs (Thapa and Blair 2018). In particular, such information allows the identification of diverse parental combinations that will create segregating progenies with maximum genetic variability and facilitates the introgression of desirable genes from diverse germplasm into the available genetic base (Shukla *et al.* 2009). Thus, the extent of variability in plant materials will determine the limitation of selection in a given improvement program (Shukla *et al.* 2006a).

Although the use of molecular techniques in plant genetic diversity analyses has brought numerous successes in recent decades, morphology is still widely used to evaluate the variability of germplasm accessions (Thapa and Blair 2018). The morphological method is based on visual characteristics, such as leaf color, seed shape, growth habit, and pigmentation, and so does not require expensive technology. However, this approach does need a large area of land for field experiments, which may make it more expensive than molecular techniques in developed countries and equally expensive in developing countries where the labor and land requirements are more readily available (Govindaraj *et al.* 2015). Morphological data also play an important role as a "reality check" for molecular results (Wiens 2004). For instance, in an assessment of the genetic variability of 29 distinct and promising strains of the vegetable amaranth *A. tricolor* using morphological traits over a 2-year period, Shukla *et al.* (2006a) observed small differences between the phenotypic and genotypic coefficients of variation and high levels of heritability and expected genetic advances for many of the morphological traits,

suggesting that all of the observed traits were under genotypic control. Thus, the foliage yield and quality traits of *A. tricolor* could be enhanced through the selection of plant types based on their component traits.

Amaranthus tricolor has been an important summer vegetable in Vietnam and other Southeast Asian countries for many years. However, most of the *A. tricolor* germplasm that is grown by farmers comprises landraces that exhibit some desirable phenotypes, and this crop has not been the subject of any breeding program. The PRC in Vietnam maintains approximately 280 amaranth accessions that were collected from eight ecological areas of Vietnam and are expected to contain a large amount of genetic variation and desirable traits for crop improvement. However, the VA collection has not been taxonomically classified or evaluated for its genetic diversity.

In this study, the genetic diversity of VA accessions of *A. tricolor*, as determined by the *matK* analysis in Chapter 2, was evaluated using morphological data.

4.2. Materials and Methods

4.2.1. Materials

The morphological traits of 280 VA accessions that had been preserved at the PRC in Vietnam were evaluated from 2015 to 2017 (n = 100 in 2015, 77 in 2016, and 103 in 2017). These accessions belonged to unknown species and had been collected in eight ecological areas of Vietnam (Fig. 1.1). In addition, four accessions were imported from WorldVeg. Among these accessions, 120 (Supplemental Table 1) were identified as *A. tricolor* using the *matK* marker (see Chapter 2) and so were used in this analysis. Among the 120 accessions, 5 (VA017, VA131, VA135, VA243, and VA276) had a shortage of morphological data owing to a lack of flowers and so were not included in this analysis. The results of a STRUCTURE analysis at K = 4 showed that the remaining 115 accessions could be divided into two major types: a common type in East Asia (48 accessions belonging to subgroup 1, named "SubWA1") and a unique type in Vietnam (59 accessions belonging to subgroup 2, named "SubWA2"). In addition, 2 accessions were placed in subgroup 3 ("SubWA3") and 6 accessions were assigned to the admixed group

4.2.2. Experimental site

This experiment was conducted in the experimental field of the PRC, which is located northwest of the Red River Delta in Hanoi, Vietnam (20°53′ to 21°23′N, 105°44′ to 106°02′E). The elevation of the land in this field ranges from 5 to 20 m, and this region has an average annual temperature and total rainfall of 24°C and 1676 mm, respectively (Luo *et al.* 2018). Seeds from each accession were germinated in plastic trays (6 × 12 holes; 4 cm diameter, 4.5 cm deep) containing a growing medium without fertilizer. Three weeks after sowing, 10 healthy plants per accession were selected and transplanted into

the PRC field at a density of 2.08 plants/m^2 with a fertilizer rate of 25 manure + 100 N + 70 P_2O_5 + 80 K_2O.

4.2.3. Methods

A total of 38 morphological characteristics were described according to the International Union for the Protection of New Varieties of Plants (UPOV) descriptors (UPOV 2008) with minor modifications (Table 4.1). Each trait was measured in five replicate plants at six different stages of growth: 1) 3–6 days after germination; 2) once plants had six to eight true leaves; 3) at the beginning of emergence of the inflorescence; 4) when 50% of plants were in full blossom; 5) at the mature stage; and 6) the dried seeds. A total of 12 quantitative and 26 qualitative morphological traits were assessed. However, the accessions were grown separately from 2015 to 2017 and thus may have been subjected to different environmental conditions, which can strongly affect quantitative traits. Therefore, only the 26 qualitative traits were included in this analysis to avoid any potential errors. Data analysis was performed using Microsoft Excel.

Table 4.1. List of the morphological traits that were evaluated in the Vietnamese *Amaranthus* (VA) collection

Score	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
Stage 1: 3–6 days after germination										
Betalain coloration of cotyledon		Absent								Present
Intensity of betalain coloration of hypocotyl		Absent		Weak		Medium		Strong		
Stage 2: When the plants had six to eight true leaves										
Leaf length (L)	Quantitative evaluation (cm)									
Leaf width (D)	Quantitative evaluation (cm)									
L/D ratio	Ratio of length to width									
Position of broadest leaves		In middle or slightly toward base			Moderately toward base			Strongly toward base		
Prominence of leaf veins				Weak		Medium		Strong		
Main color on upper side of leaves		Light green		Green		Dark green		Red		Purple
Distribution of secondary color on upper side of leaves				Colored basal area		Central blotch		Colored margin and veins		
Color on lower side of leaves				Green		Red		Purple		
Stage 3: At the beginning of emergence of the inflorescence										
Margin of fully expanded leaves		Entire								Sinuate
Stage 4: When 50% of plants were in full blossom										
Beginning of inflorescence emergence	Date (day/month)									
Time of flowering	Date (day/month)									
Stem color				Green	Yellow	Pink	Red	Purple		
Color of stem stripes		Light green		Red		Red-purple		Purple		
Betalain coloration of petioles		Absent								Present
Intensity of betalain coloration of petioles		Very weak		Weak		Medium		Strong		Very strong
Color of leaves		Light green		Green		Dark green		Red		Purple
Presence of blotch on leaf blade		Absent								Present
Color of inflorescences				Yellow	Green	Pink	Red	Purple	Brown	
Compactness of inflorescence				Compact		Intermediate		openOpen		
Density of glomerules on inflorescence				Sparse		Medium		Dense		
Type of inflorescence		Amarantiform								Glomerulate
Number of female flowers per glomerule	Quantitative evaluation (%)									

Score	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
Length of bract relative to utricle				Shorter		Equal		Longer		
Growth habit of inflorescence		Determinate								Indeterminate
Attitude of inflorescence		Upright		Weakly recurved		Moderately recurved		Strongly recurved		
Stage 5: At maturity										
Time of maturity	Date (day/month)									
Plant length	Quantitative evaluation (cm)									
Betalain coloration of stem base		Absent								Present
Number of lateral branches of main stem	Quantitative evaluation (number per plant)									
Stem diameter	Quantitative evaluation of main stem just above the ground (mm)									
Insect resistance		Very weak		Weak		Medium		strongStrong		Very strong
Disease resistance		Very weak		Weak		Medium		Strong		Very strong
Stage 6: Dried seeds										
Seed color		White		Yellow		Pink		Brown		Black
Seed type		Flint								Floury
Seed shape		Ellipsoid								Discoid
Seed weight per 1000 seeds	Quantitative evaluation									

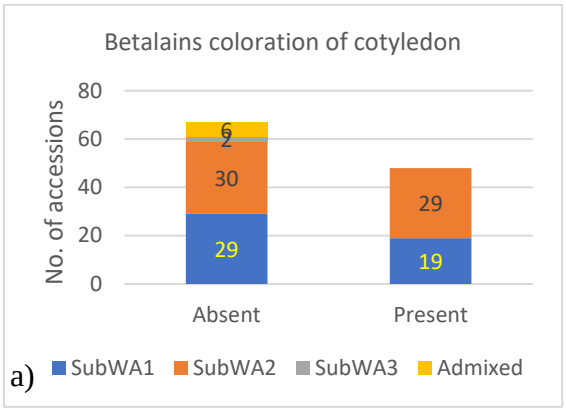
4.3. Results

In general, most of the morphological traits assessed were diverse in both the common type in East Asia and the unique type in Vietnam. However, the margin of the leaves, seed color, and seed shape were exceptions to this (Figure 4.1w, x, and z). Numerous accessions of both types exhibited the same characteristics for nearly all of the traits with the exception of the main color on the upper side of the leaves (Figure 4.1e), stem color (Figure 4.1h), color of the stem stripes (Figure 4.1i), color of the leaves (Figure 4.1k), and color of the inflorescences (Figure 4.1 m), which were green (or light green) for the majority of the common-type accessions and purple for most of the unique-type accessions. There was also some variation in the intensity of betalain coloration on the petioles (Figure 4.1j) and the growth habit of the inflorescences (Figure 4.1r).

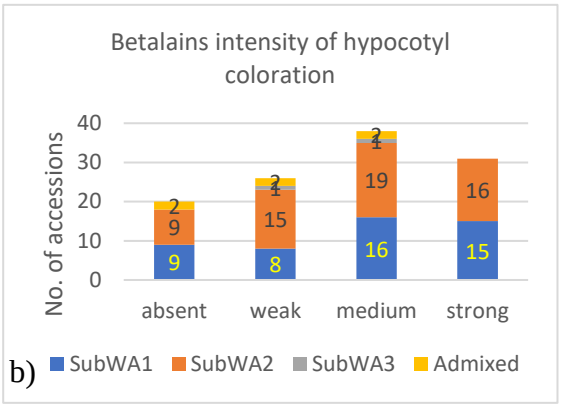
The majority of accessions of the common and unique types had a strong prominence of leaf veins (70.8% and 89.8%, respectively), a position of the broadest leaf that was moderately toward the leaf base (62.5% and 66.1%, respectively), the distribution of secondary color on the upper side of the leaves near the margin and veins (58.3% and 81.4%, respectively), no blotch on the leaf blade (58.3% and 86.4%, respectively), glomerulate-type inflorescences (89.6% and 81.4%, respectively), a medium density of glomerules (62.5% and 62.7%, respectively), a longer bract than utricle (91.7% and 93.2%, respectively), betalain coloration on the stem base (79.2% and 81.4%, respectively), and a floury seed type (100.0% and 98.3%, respectively). Most of the common-type accessions had green upper sides of the leaves (68.8%), green stems (56.3%), light green stripes on the stems (54.2%), green leaves at the stage when 50% of plants were in full blossom (66.7%), green inflorescences (66.7%), no betalain coloration on the petioles (47.9%), and determinate inflorescences (79.2%), whereas the majority of

unique-type accessions had purple upper sides of the leaves (55.9%), purple stems (72.9%), purple stripes on the stems (71.2%), purple leaves at the stage when 50% of plants were in full blossom (49.2%), purple inflorescences (66.1%), strong and very strong betalain coloration on the petioles (62.7%), and indeterminate inflorescences (67.8%).

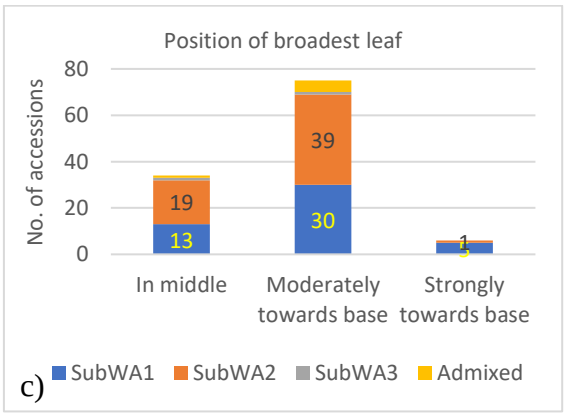
Most of the VA accessions had medium or strong levels of disease resistance (75.7% and 16.5%, respectively), while half had medium or strong levels of resistance to insects (40.0% and 14.8%, respectively; figure 4.1u and v). Eight of the accessions were found to be strongly resistant to both disease and insects (VA216, VA240, VA221, VA185, VA188, VA193, VA194, and VA253). The number of accessions that had a medium resistance to disease was higher for the unique type (n = 49) than for the common type (n = 31). However, slightly fewer of the unique type accessions had strong resistance to disease than the common type (n = 8 and 11, respectively). Similarly, more of the unique-type accessions had medium resistance to insects than common type accessions (n = 26 and 16, respectively), while the reverse was true for strong resistance to insects (n = 7 and 10, respectively). In total, 45.8% of the common type accessions and 44.1% of the unique-type accessions showed low tolerance to insects.


Betalains coloration of cotyledon
No. of accessions
80
60
40
20
0
6
2
30
29
29
19
Absent
Present
SubWA1
SubWA2
SubWA3
Admixed

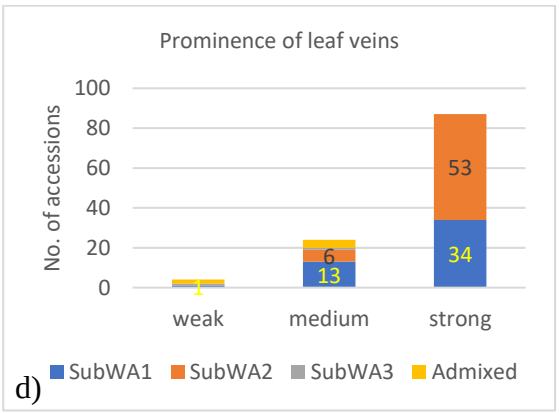
a)


Betalains intensity of hypocotyl coloration
No. of accessions
40
30
20
10
0
2
9
9
2
1
15
8
2
1
19
16
16
15
absent
weak
medium
strong
SubWA1
SubWA2
SubWA3
Admixed

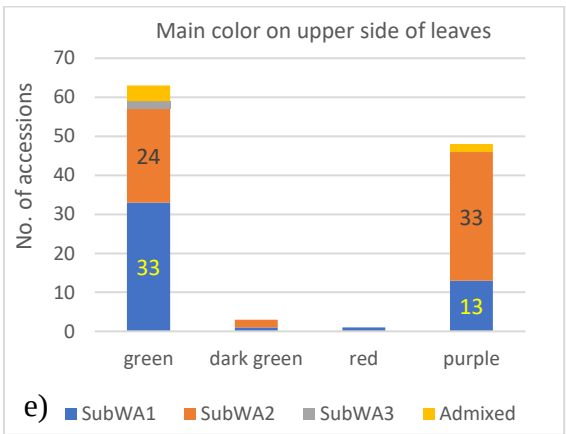
b)


Position of broadest leaf
No. of accessions
80
60
40
20
0
19
13
39
30
1
In middle
Moderately towards base
Strongly towards base
SubWA1
SubWA2
SubWA3
Admixed

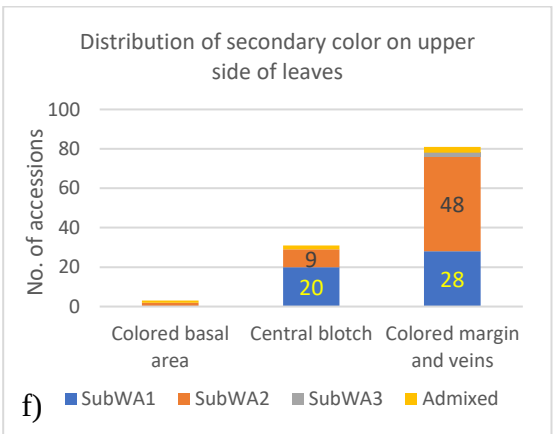
c)


Prominence of leaf veins
No. of accessions
100
80
60
40
20
0
1
6
13
53
34
weak
medium
strong
SubWA1
SubWA2
SubWA3
Admixed

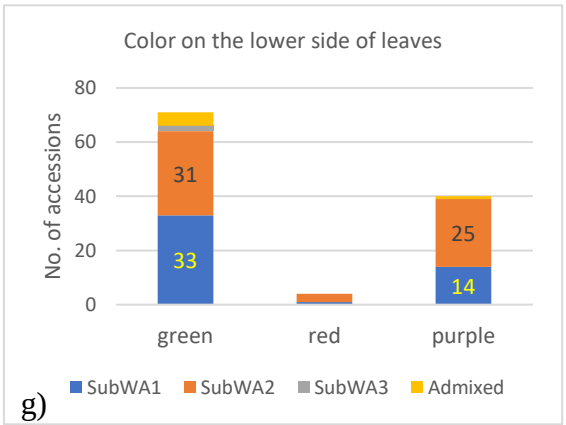
d)


Main color on upper side of leaves
No. of accessions
70
60
50
40
30
20
10
0
24
33
33
13
green
dark green
red
purple
SubWA1
SubWA2
SubWA3
Admixed

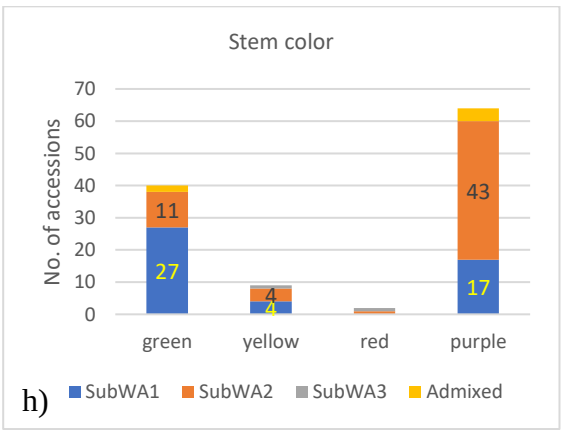
e)


Distribution of secondary color on upper side of leaves
No. of accessions
100
80
60
40
20
0
9
20
48
28
Colored basal area
Central blotch
Colored margin and veins
SubWA1
SubWA2
SubWA3
Admixed

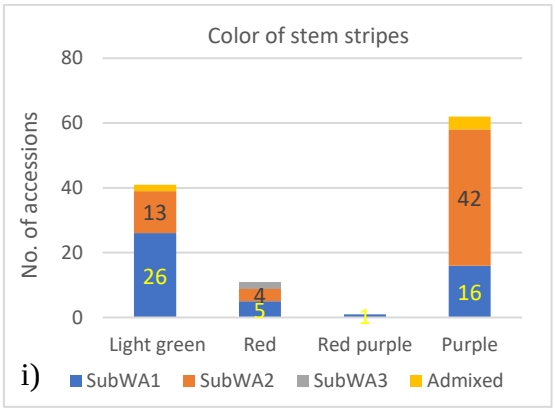
f)


Color on the lower side of leaves
No. of accessions
80
60
40
20
0
31
33
25
14
green
red
purple
SubWA1
SubWA2
SubWA3
Admixed

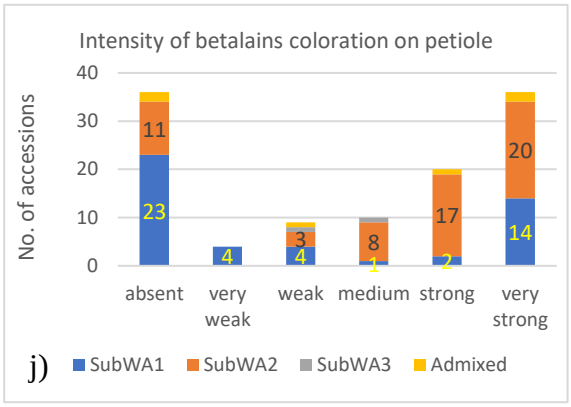
g)


Stem color
No. of accessions
70
60
50
40
30
20
10
0
11
27
4
4
43
17
green
yellow
red
purple
SubWA1
SubWA2
SubWA3
Admixed

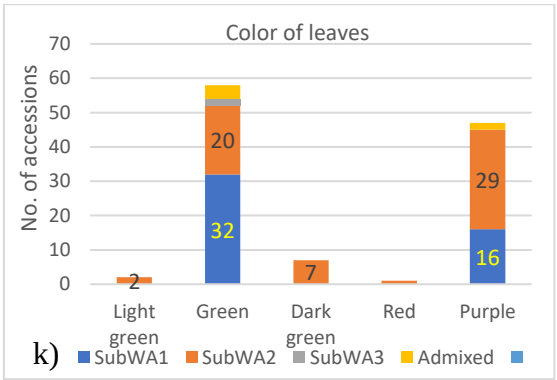
h)


Color of stem stripes
No. of accessions
80
60
40
20
0
13
26
4
5
1
42
16
Light green
Red
Red purple
Purple
SubWA1
SubWA2
SubWA3
Admixed

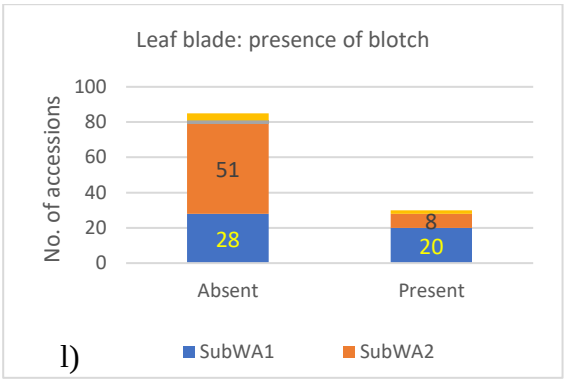
i)


Intensity of betalains coloration on petiole
No. of accessions
40
30
20
10
0
11
23
4
3
4
8
1
17
2
20
14
absent
very weak
weak
medium
strong
very strong
SubWA1
SubWA2
SubWA3
Admixed

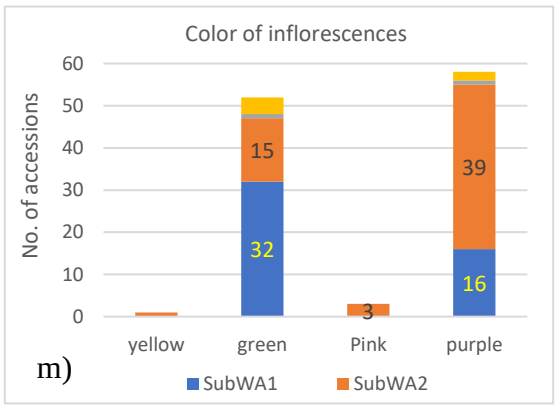
j)


Color of leaves
No. of accessions
70
60
50
40
30
20
10
0
2
20
32
7
29
16
Light green
Green
Dark green
Red
Purple
SubWA1
SubWA2
SubWA3
Admixed

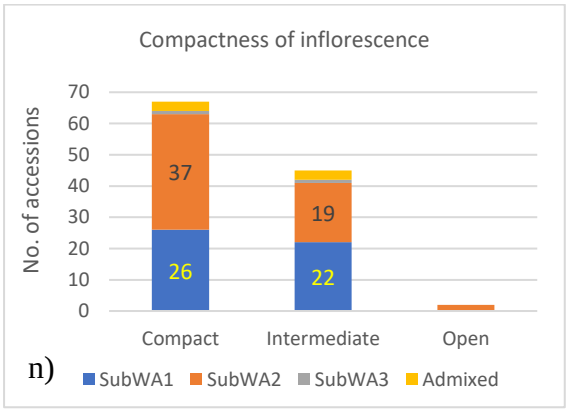
k)


Leaf blade: presence of blotch
No. of accessions
100
80
60
40
20
0
51
28
8
20
Absent
Present
SubWA1
SubWA2

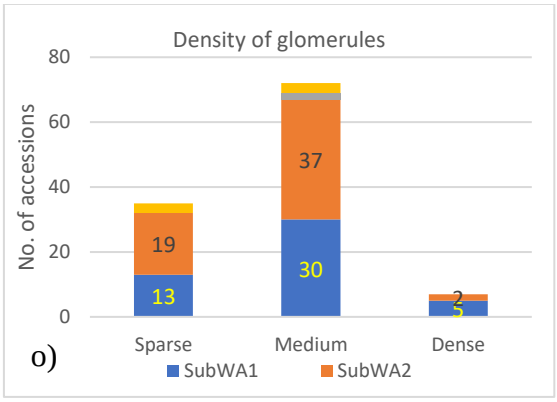
l)


Color of inflorescences
No. of accessions
60
50
40
30
20
10
0
15
32
3
39
16
yellow
green
Pink
purple
SubWA1
SubWA2

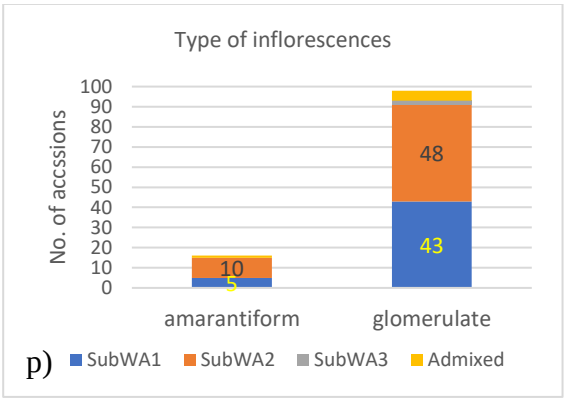
m)


Compactness of inflorescence
No. of accessions
70
60
50
40
30
20
10
0
37
26
19
22
Compact
Intermediate
Open
SubWA1
SubWA2
SubWA3
Admixed

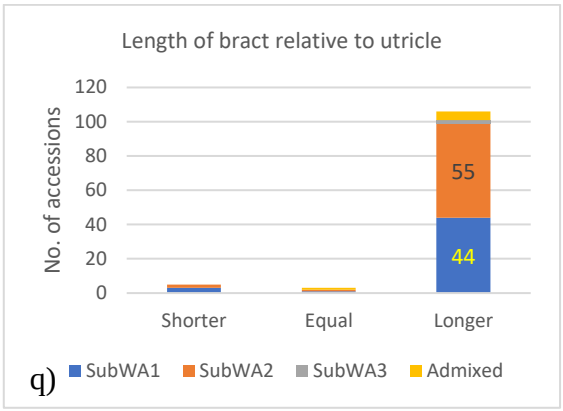
n)


Density of glomerules
No. of accessions
80
60
40
20
0
19
13
37
30
2
5
Sparse
Medium
Dense
SubWA1
SubWA2

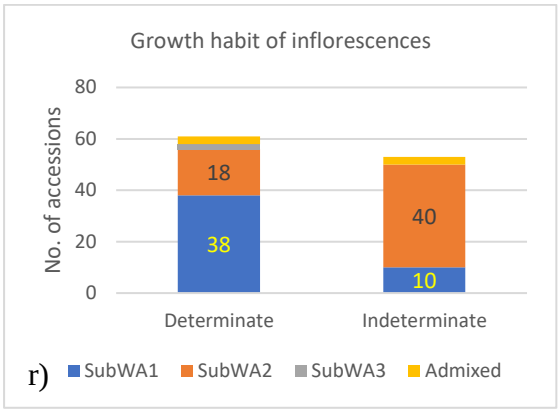
o)


Type of inflorescences
No. of accssions
100
90
80
70
60
50
40
30
20
10
0
10
5
48
43
amarantiform
glomerulate
SubWA1
SubWA2
SubWA3
Admixed

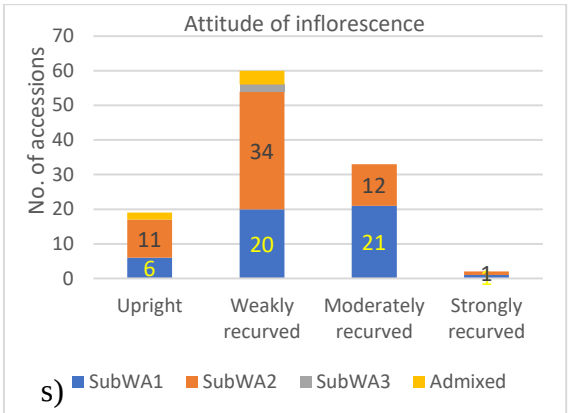
p)


Length of bract relative to utricle
No. of accessions
120
100
80
60
40
20
0
55
44
Shorter
Equal
Longer
SubWA1
SubWA2
SubWA3
Admixed

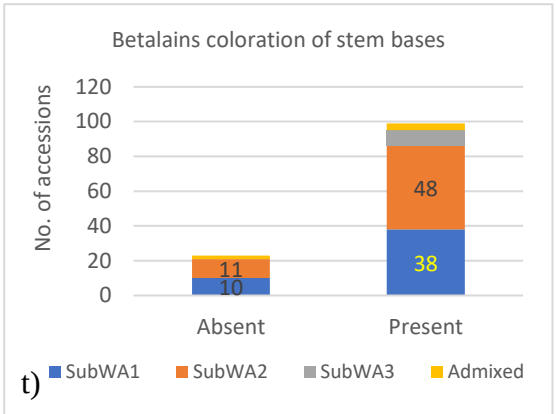
q)


Growth habit of inflorescences
No. of accessions
80
60
40
20
0
18
38
40
10
Determinate
Indeterminate
SubWA1
SubWA2
SubWA3
Admixed

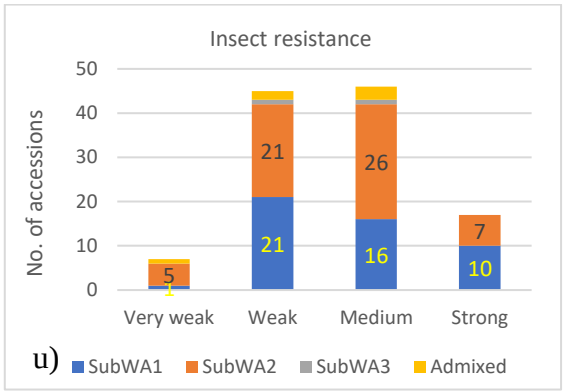
r)


Attitude of inflorescence
No. of accessions
70
60
50
40
30
20
10
0
11
6
34
20
12
21
1
Upright
Weakly recurved
Moderately recurved
Strongly recurved
SubWA1
SubWA2
SubWA3
Admixed

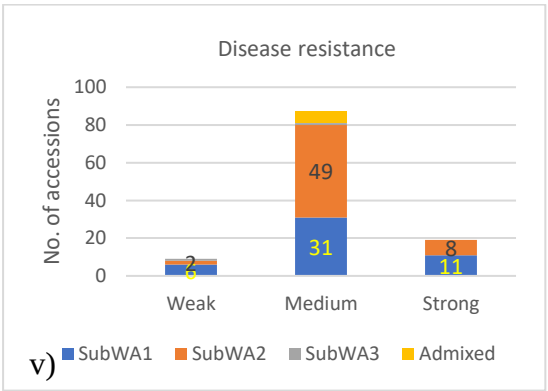
s)


Betalains coloration of stem bases
No. of accessions
120
100
80
60
40
20
0
11
10
48
38
Absent
Present
SubWA1
SubWA2
SubWA3
Admixed

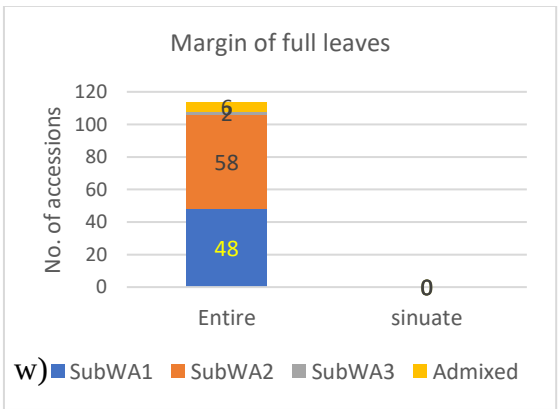
t)


Insect resistance
No. of accessions
50
40
30
20
10
0
5
1
21
21
26
16
7
10
Very weak
Weak
Medium
Strong
SubWA1
SubWA2
SubWA3
Admixed

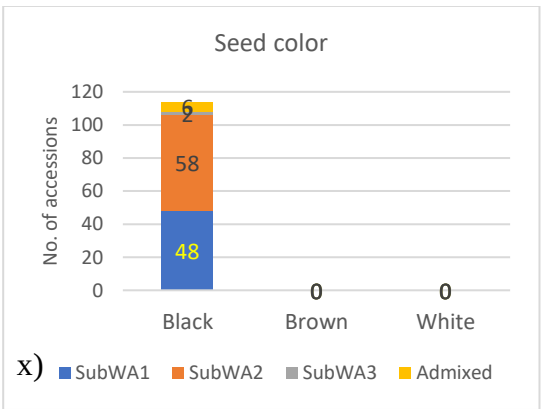
u)


Disease resistance
No. of accessions
100
80
60
40
20
0
2
0
49
31
8
11
Weak
Medium
Strong
SubWA1
SubWA2
SubWA3
Admixed

v)


Margin of full leaves
No. of accessions
120
100
80
60
40
20
0
6
2
58
48
0
Entire
sinuate
SubWA1
SubWA2
SubWA3
Admixed

w)


Seed color
No. of accessions
120
100
80
60
40
20
0
6
2
58
48
0
0
Black
Brown
White
SubWA1
SubWA2
SubWA3
Admixed

x)

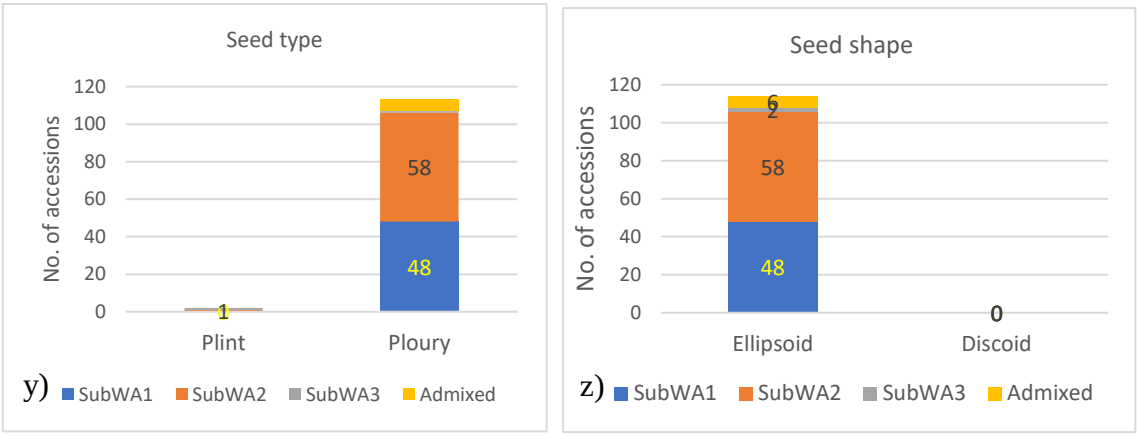


Figure 4.1. Stacked column charts for the morphological traits of *Amaranthus tricolor* accessions in the Vietnamese *Amaranthus* (VA) collection (n = 115). SubWA1, SubWA2, SubWA3, and Admixed represent subgroup 1, subgroup 2, subgroup 3, and the admixed group based on the STRUCTURE analysis at K = 4 that was presented in Chapter 3.

4.4. Discussion

Most of the traits that were evaluated exhibited variation, with the exception of the margins of the leaves and the color and shape of the seeds (Figure 4.1), indicating that *A. tricolor* accessions in the VA collection are diverse in their phenotypic traits. This agrees with the results of Shukla *et al.* (2006a), who found considerable variability in 29 strains of *A. tricolor* that were grown for two seasons based on an evaluation of nine morphological and quality traits. This also partly supports the results of several previous morphological assessments of the genus *Amaranthus*, which indicated that there is a wide diversity in phenotypic traits not only among *Amaranthus* species but also among genotypes within the same species (Wu *et al.* 2000, Thapa and Blair 2018).

The two types of *A. tricolor* in the VA collection possessed many similar characteristics, such as an entire margin of the full leaves, a strong prominence of leaf veins, the broadest position toward the leaf base, the distribution of secondary color on the upper side of the leaves near the margin and veins, an absence of blotches on the leaf blade, glomerulate inflorescences, a medium density of glomerules, a longer bract than utricle, the presence of betalain coloration on the stem base, and black, discoid, floury seeds. These characteristics are mostly similar to those listed by Pham (1999) in a description of *A. tricolor* species in Vietnam, which were as follows: leaves arranged spirally, simple, broadly ovate to lanceolate, 4–12 cm in length, broadest position toward the base, acute at apex; greenor purple leaves; inflorescences bear glomerules at axillary and a terminal spike, with male and female flowers intermixed; flowers unisexual, with male flowers having three stamens and female flowers haing a superior single-celled ovary with three stigmas; pyxis fruit with a single seed; and seeds black, smooth, discoid shaped, and 1 mm in length. It was also found that the majority of common type

accessions had green leaves, stems or stripes on the stems, and inflorescences; an absence of betalain coloration on the petioles; and determinate inflorescences. By contrast, the majority of unique type accessions had purple leaves and petioles, stems or stripes on the stems, and inflorescences; and indeterminate.

The selection of disease-resistant cultivars is one of the most important objectives of breeding programs, particularly if they are to respond to climate change and reduce the amounts of pesticide residues on agricultural products in Vietnam. The findings of the present study showed that most of the VA accessions had a medium or strong resistance to disease, while half were resistant to insects. Furthermore, eight of the VA accessions were strongly resistant to both disease and insects, seven of which (VA185, VA188, VA193, VA194, VA216, VA240, and VA221) belonged to the unique type and one of which (VA253) belonged to the common type. Thus, the unique type appears to have superior performance in terms of pest resistance. Together, these findings indicate that the unique-type accessions may be useful for enhancing the genetic resistance of modern varieties to pests and that these accessions may also be landraces that can adapt well to the local environment, possessing genes with a tolerance to both biotic and abiotic stresses.

In conclusion, the morphological analysis of 115 *A. tricolor* accessions that the VA collection of *A. tricolor* is phenotypically diverse. Although both types of *A. tricolor* in the VA collection possessed many similar characteristics, there were differences in the coloration of the leaves, stems or stripes on the stems, and inflorescences, amount of betalain coloration on the petioles, and the type of inflorescence between the common-type and unique-type accessions. It was also found that most of the VA accessions had medium or strong levels of disease resistance and half of the VA accessions were resistant

to insects, with eight accessions being strongly resistant to both, most of which were of the unique type.

CHAPTER 5: General Discussion

Climate change is considered one of the greatest environmental threats to sustainable agricultural production and food security in many parts of the world but is particularly an issue in developing countries (FAO 2008). As in other developing countries, crop productivity in Vietnam has decreased as a result of the increase in temperature and variation in rainfall, which has increased the frequencies of heatwaves and droughts during the summer. In addition, the high amounts of pesticide residues on agricultural products is a major issue in Vietnam, costing the economy approximately US$700 million per year (Hoi *et al.* 2016).

The development of new crop cultivars with a high resistance and tolerance to abiotic and biotic stresses is an effective strategy to respond to climate change and reduce pesticide use in agricultural production, and amaranths are widely considered promising crops for the 21st century owing to their excellent nutritional value, high adaptability to severe conditions, and lack of major diseases (Shukla *et al.* 2006a, Rastogi and Shukla 2013, Achigan-Dako *et al*. 2014, Assad *et al*. 2017). These crops are extremely adaptable to challenging growing conditions, such as heat stress, drought, and grows well in agriculturally marginal lands (Shukla *et al.* 2006a, Rastogi and Shukla 2013). Furthermore, they are generally considered to be tolerant of nematodes, even being recommended as a rotation crop to reduce nematode populations for subsequent crops, and their leaves can produce a sustainable yield even under the pressure of numerous chewing insects (O'Brien and Price 1983). However, amaranths are currently underexploited as vegetable crops and limited effort has been placed in their genetic improvement (Shukla *et al.* 2006a). Thus, in Vietnam, amaranths have not been subjected

to systematic breeding and most of the amaranth germplasms that are grown by farmers are landraces with that exhibit some desirable phenotypes.

Evaluation of the genetic diversity is a key prerequisite for any crop breeding program and is also useful for the management and exploitation of genetic resources from gene banks (Shukla *et al.* 2010). Plant genetic diversity can be estimated from the morphological, biochemical, cytological, and molecular traits (Sammour *et al.* 2012). In the present study, the amaranth genetic resources in Vietnam were classified and evaluated based on both molecular markers (barcoding and SSR markers) and morphological traits. All of the information obtained in this study will provide a basis for the effective utilization and conservation of these resources and for the improvement of leafy amaranths in the future.

The correct classification of a crop at the species level is the starting point for any breeding program. *Amaranthus* has traditionally been considered a "difficult" genus to classify because of the overall similarity among many of the species, the small sizes of the diagnostic parts, and the presence of intermediate forms (Costea and DeMason 2001, Achigan-Dako *et al*. 2014, Assad *et al*. 2017). Therefore, in the present study, the *matK* marker, which is considered a universal barcode for flowering plants (Lahaye *et al.* 2008), was used to identify the species of *Amaranthus* in unknown accessions from Vietnam. As expected, the phylogenetic tree classified 68% of the VA accessions to the species level, among which 120 were assigned to *A. tricolor* with high bootstrap support (85%), and a further 41 were assigned to *A. dubius*, 18 to *A. blitum*, and 6 to *A. viridis*. Thus, *A. tricolor* accounted for the largest proportion of accessions (44%) in this collection, supporting the widely held view that this is one of the most popular vegetable amaranths in Vietnam.

Use of only the *matK* marker was able to identify 7 of the 13 *Amaranthus* species analyzed in this study, including 3 leafy amaranths (*A. tricolor*, *A. blitum*, and *A. viridis*), 3 weed amaranths (*A. spinosus*, *A. retroflexus*, and *A. powellii*), and *A. dubius* (a relative of the grain amaranths). However, it could not clarify the relationship among the three grain amaranths (*A. hypochondriacus*, *A. caudatus*, and *A. cruentus*) and their putative ancestors (*A. hybridus* and *A. quitensis*), which make up the so-called "hybridus complex." This limitation has also been observed in numerous previous studies in which the phylogeny of the genus *Amaranthus* has been analyzed using a certain number of molecular markers (Lanoue *et al.* 1996, Chan and Sun 1997, Yudina *et al.* 2005, Lee *et al.* 2008, Mallory *et al.* 2008, Ogundipe and Chase 2009, Chandi *et al.* 2013, Singh *et al.* 2013, Wang and Park 2013, Kietlinski *et al.* 2014, Suresh *et al.* 2014), although this has been overcome through the use of genotyping by sequencing in some studies (Stetter and Schmid 2017, Wu and Blair 2017). Therefore, it appears that the *matK* marker may be useful for the phylogenetic analysis of leafy and weed amaranths but that high-throughput sequencing technology, such as genotyping by sequencing and restriction site-associated DNA sequencing (RAD-Seq), needs to be used to elucidate the phylogenetic relationships within the hybridus complex.

SSRs are currently considered the most powerful molecular markers for plant genetic analyses (Varshney *et al.* 2005, Kumar *et al.* 2009, Guichoux *et al.* 2011, Vieira *et al.* 2016, Taheri *et al.* 2018). Most previous studies on the genetic diversity of amaranths have used only 11–14 SSR markers (Lee *et al.* 2008, Mallory *et al.* 2008, Khaing *et al.* 2013, Singh *et al.* 2013, Wang and Park 2013, Kietlinski *et al.* 2014, Suresh *et al.* 2014), which were developed by Mallory *et al.* (2008) and Lee *et al.* (2008) based on the genomic DNA of *A. hypochondriacus*. In the present study, SSR markers were

developed using the PacBio sequence data of *A. tricolor* ‘Biam’ to analyze the genetic diversity of this leafy amaranth, which resulted in the successful development of 21 SSR markers that amplified single loci with relatively high levels of polymorphism. The recent advent of next-generation sequencing (NGS) technologies has enabled millions of genome-wide SNPs to be detected for genetic diversity studies, bypassing the time-consuming process of marker development (Chung *et al.* 2017). However, the high costs of this technology hinders its wider use, particularly for minor crops such as amaranths. Thus, the SSR markers that were developed in this study may prove useful for future studies on the genetic diversity of leafy amaranths, particularly in countries with limited technological resources and research funds, such as Vietnam.

The VA collection of *A. tricolor* was found to be more closely genetically related to the East Asian collection than the other collections (Southeast Asian, South Asian, and American). This may be due to the long-distance dispersal of *A. tricolor* from East Asia (China) to Vietnam, partly through the migrations of the ancestors of most of today’s Vietnamese ethnic groups from south of the Yangtze River in China (Hall and Patrinos 2012) and the introduction of this crop over the long period of Chinese domination in Vietnam from 111 BC to 938 AD (Le *et al.* 1991). The VA collection had a lower level of genetic diversity than the WA collection but a higher level of genetic diversity than the East Asian collection. Furthermore, it could be divided into two major types: a common type in East Asia and a unique type in Vietnam.

Based on these results, it can be hypothesized that the genetic diversity in Vietnamese *A. tricolor* was mainly established through long-distance dispersal events from East Asia and adaptation to the local environments through on-site cross pollination, as well as the retention of cross-pollinated seeds for the next season without any artificial selection.

Alongside *de novo* mutations and the standing variation, which will widen the pool of genetic variation that is available for adaptation, adaptive introgression could be an important source of variation for the evolution of Vietnamese leafy amaranths. According to Suarez-Gonzalez *et al.* (2018), the environmental context in which hybridization occurs, together with the dispersion distance and natural selection are key factors underlying the geographical patterns and spatial scales of introgression. Although amaranths are predominantly self-pollinated plants, outcrossing rates vary from 3.5% to 34% and interspecific hybridization and high fertility of the interspecific progeny has been widely reported within the genus (Brenner *et al.* 2000). In Vietnam, *A. tricolor* is grown as a vegetable in gardens or small plots in open fields, but five other amaranth species (*A. viridis*, *A. spinosus*, *A. hybridus*, *A. blitum*, and *A. retroflexus*) also grow nearby as wild species (Pham 1999). In addition, Vietnamese farmers often keep amaranth seeds for the next season without any artificial selection. Therefore, the environment has been suitable for adaptative introgression between *A. tricolor* and other amaranth species.

In general, breeding for resistance to diseases and extreme conditions, such as droughts, floods, and high and low temperatures, is an important objective of nearly all breeding programs, and this is particularly important for responding to climate change and reducing the amounts of pesticide residues in agricultural products in Vietnam. The present study showed that half of the VA *A. tricolor* accessions are unique in Vietnam. These accessions could be landraces that are well adapted to the local environment, possessing genes with a tolerance to biotic and abiotic stresses. The morphological data analysis also found that most of the VA *A. tricolor* accessions had a medium or strong resistance to disease and half were resistant to insects, with eight accessions exhibiting

resistance to both of these. Thus, these will be useful for enhancing the genetic resistance of modern varieties to pests and environmental stresses. However, a lower diversity was observed in the VA germplasm resources than in the WA collection. In modern cultivar development, a narrow genetic base for a given crop being more vulnerable to disease and having limited adaptability, so accessions exhibiting the greatest amount of genetic differentiation are used in breeding programs to maximize the introgression of diversity into new varieties (Roorkiwal *et al.* 2014). Therefore, improvement of the gene pool of leafy amaranths in Vietnam will require germplasm to be imported from South Asia and America, both of which were relatively genetically different from the VA accessions.

There was relatively little variation in *A. tricolor* among the geographical groups in Vietnam and no relationship between geographical distance and genetic diversity. Thus, it appears that a high level of genetic exchange occurred among the *A. tricolor* accessions. This may have resulted from the historical demographic event known as "southern expansion" from the north to the south of Vietnam along the coast, which occurred from the 10^{th} to 11^{th} centuries until the mid-18^{th} century (Pischedda.S *et al.* 2017).

In the present study, the genetic diversity of VA accessions was identified and evaluated in comparison with amaranth accessions from other parts of the world based on genotypic (*matK* and SSR) and phenotypic data. This research has provided a basic understanding of the genetic diversity of VA germplasm resources, which will facilitate the management and exploitation of gene banks and the genetic improvement of amaranths. However, the identification of amaranth species in the VA collection was only based on molecular data. Therefore, this should be checked using morphological data in the future. In addition, 87 of the VA accessions could not be classified using *matK*, indicating that high-throughput sequencing technology (genotyping by sequencing or

RAD-Seq) will need to be used to analyze their phylogenetic relationships. Finally, the morphological data of the 120 *A. tricolor* accessions were obtained and evaluated separately over a 3-year period. Therefore, all of these accessions should be evaluated under the same environmental conditions in the future to confirm these results.

Summary

Climate change and high levels of pesticide residues on agricultural products are currently considered serious issues for agricultural production in Vietnam. The development of new crop cultivars with a high resistance and tolerance to abiotic stress and disease is one method that can be used to effectively address these issues. Species in the genus *Amaranthus* are considered promising crops for the 21st century owing to their excellent nutritional value, high adaptability to severe conditions, and lack of major diseases. Leafy amaranths, particularly *A. tricolor*, are a rich and inexpensive source of protein, carotenoids, vitamins, dietary fiber, and a wide range of minerals. However, *A. tricolor* has not been subjected to systematic breeding in Vietnam. Therefore, in this study, amaranth genetic resources in Vietnam were classified and assessed using molecular markers (barcoding and SSR markers) and morphological traits. All of the information obtained in this study will provide a basis for effectively utilizing and conserving these resources, as well as improving leafy amaranths in the future.

Alignment of the *matK* gene sequences from 272 VA, 27 WA, and 45 GRIN amaranth accessions gave a total length of 434 bp per accession, which included 96.8% constant, 3.2% variable, and 2.5% parsimoniously informative sites. The phylogenetic dendrogram was able to classify 7 of the 13 *Amaranthus* species analyzed in this study, including 3 leafy amaranths (*A. tricolor*, *A. blitum*, and *A. viridis*), 3 weed amaranths (*A. spinosus*, *A. retroflexus*, and *A. powellii*), and *A. dubius* (a relative of the grain amaranths). The phylogenetic analysis determined the species for 68% of the VA accessions, 120 of which were assigned to *A. tricolor* with high bootstrap support (85%), while a further 41 were assigned to *A. dubius*, 18 to *A. blitum*, and 6 to *A. viridis*. Thus, *A. tricolor* accounted for the highest proportion of accessions (44%) in the VA collection. However, *matK* alone

could not clarify the relationship among the three grain amaranths (*A. hypochondriacus*, *A. caudatus*, and *A. cruentus*) and their putative ancestors (*A. hybridus* and *A. quitensis*).

The 21 SSR markers that were developed in the present study allowed the overall genetic diversity of the VA and WA collections of *A. tricolor* and the genetic relationships among the geographic groups to be evaluated. These markers successfully amplified 153 alleles, with a mean allelic richness of 7.29 per marker and mean values of 0.14 for H_o, 0.38 for H_e, and 0.35 for PIC.

There were significant differences between the VA collection and all of the overseas groups except the African collection. However, while the VA collection was found to be divergent from the South Asian and American collections, it was relatively closely genetically related to the East Asian collection and commercial cultivars. Furthermore, it was genetically closer to the East Asian accessions than the Southeast Asian accessions. This may be related to the dispersal events of amaranths from East Asia (China) to Vietnam during the migrations of the ancestors of most of today's Vietnamese ethnic groups from south of the Yangtze River of China and the introduction of this crop during the long period of Chinese domination in Vietnam from 111 BC to 938 AD.

The STRUCTURE analysis indicated that the accessions could be grouped into two or four subpopulations. At $K = 4$, subgroup 1 mostly included accessions from East Asia and Vietnam, and contained 90.7% of all of the East Asian accessions and 88.9% of all of the commercial cultivars; subgroup 2 mostly included accessions from Vietnam; subgroup 3 mostly included accessions from South Asia; and subgroup 4 included 60% of the accessions from America. These results suggest that there is a relationship between the amaranth accessions and their origins among the VA collection and most of the overseas groups owing to the low level of genetic exchange of *A. tricolor* among these

geographical areas. There was also a significant but small amount of variation among the geographical groups in Vietnam. However, there was no relationship between geographical distance and genetic diversity, indicating a high level of genetic exchange among the VA *A. tricolor* accessions.

The STRUCTURE analysis indicated that the VA collection of *A. tricolor* could be divided into two major types: a common type in East Asia and a unique type in Vietnam. Thus, it appears that the genetic diversity of Vietnamese *A. tricolor* was established by dispersal events mainly from East Asia and adaptation to the local environments through on-site cross pollination, as well as the retention of cross-pollinated seeds for the next season without any artificial selection.

The morphological analysis indicated that *A. tricolor* accessions in the VA collection exhibited phenotypic traits. The common-type accessions had green leaves, stems or stripes on the stems, and inflorescences; an absence of betalain coloration on the petioles; and determinate inflorescences. By contrast, the majority of the unique-type accessions had purple leaves and petioles, stems or stripes on the stems, and inflorescences; and indeterminate inflorescences. In addition, while most of the VA accessions exhibited medium or strong resistance to disease and half were resistant to insects, eight accessions were strongly resistant to both of these, most of which belonged to the unique type.

In conclusion, the genetic diversity of amaranth germplasm resources in Vietnam were successfully identified and evaluated based on their genotypic and phenotypic traits, which may be useful for effectively managing and exploiting the genetic resources in gene banks and for genetically improving amaranths in the future.

Acknowledgments

Firstly, I am deeply indebted to my supervisor Professor Dr. Ryo Ohsawa, who provided a precious opportunity to study in Plant Breeding Laboratory of the University of Tsukuba, and who provided rigorous guidance during my study. My genuine gratitude is also extended to Associate Professor Dr. Yosuke Yoshioka of Plant Breeding Laboratory of the University of Tsukuba for his unlimited support, kindness and patience to guide me during my PhD program. I also would like to express my thanks to Associate Professor Dr. Naoya Fukuda of the University of Tsukuba and Professor Dr. Junichi Tanaka of National Agriculture and Food Research Organization for their precious comments and suggestions to help me complete my research.

I want to express my gratitude to Dr. Tran Danh Suu and Dr. Tran Thi Thu Hoai of the Plant Resources Center of Vietnamese Academy of Agricultural Sciences for their valuable comments and suggestions. I also want to say many thanks to Vuong Thi Anh Tuyet, the staff of the PRC who help me to a lot in collecting morphological data of Vietnamese amaranth collection.

I would like to express my profound thanks to my tutor Ruikun Chen, lab staff Noriko Nishi and other lab members during my study in Tsukuba, they helped me a lot to complete my study as well as to overcome social and cultural challenges during the studying period in Japan.

I also would like to express my thanks to Agriculture and Forestry Research Center of University of Tsukuba provided the field for cultivation experiments. We are grateful to the PRC, Vietnam the USDA National Plant Germplasm System for providing the seed materials.

This work was equally supported by a grant (PGRAsia Project 2014-2019) from the Ministry of Agriculture, Forestry and Fisheries of the Government of Japan, and Monbukagakusho (MEXT) scholarship provided by the Japanese Ministry of Science, Education and Culture.

Finally, my deepest and genuine gratitude is extended to my family, specially my wife and my son who provide me affection, support, motivation during my study.

References

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

Aellen, P. L., 1959. *Amaranthus L.* In: Hegi, G.E. (Ed.), Illustrierte Flora von Mitteleuropa 3(2), Carl Hanser Verlag, München, pp. 465-516.

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

Andini, R., Yoshida, S., Yoshida, Y. and Ohsawa, R. (2013) *Amaranthus* genetic resources in Indonesia: morphological and protein content assessment in comparison with worldwide amaranths. Genetic Resources and Crop Evolution 60: 2115-2128.

Assad, R., Reshi, Z. A., Jan, S. and Rashid, I. (2017) Biology of Amaranths. The Botanical Review 83: 382-436.

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

Bressani, R., Gonzhles, J. M., Zuiiga, J., Breuner, M. and Elias, L. G. (1987) Yield, Selected Chemical Composition and Nutritive Value of 14 Selections of Amaranth Grain Representing Four Species. Journal of the Science of Food and Agriculture 38: 347-356.

Bui, V. L., Nguyen, T. D. M., Sadik, S., Tu, S., Vidal, J. and Tran, T. V. (1998) Rapid plant regeneration of a C4 dicot species: *Amaranthus edulis*. Plant Science 132: 45-54.

Cai, Y., Sun, M. and Corke, H. (2005) Characterization and application of betalain pigments from plants of the *Amaranthaceae*. Trends in Food Science and Technology 16: 370-376.

Chan, K. F. and Sun, M. (1997) Genetic diversity and relationships detected by isozyme and RAPD analysis of crop and wild species *Amaranthus*. Theoretical and Applied Genetics 95: 865-873.

Chandi, A., Milla-Lewis, S. R., Jordan, D. L., York, A. C., Burton, J. D., Zuleta, M. C., Whitaker, J. R. and Culpepper, A. S. (2013) Use of AFLP Markers to Assess Genetic Diversity in Palmer Amaranth (*Amaranthus palmeri*) Populations from North Carolina and Georgia. Weed Science 61: 136-145.

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

Chung, Y. S., Choi, S. C., Jun, T. H. and Kim, C. (2017) Genotyping-by-Sequencing: a Promising Tool for Plant Genetics Research and Breeding. Horticulture Environment and Biotechnology 58: 425-431.

Costea, M. and DeMason, D. A. (2001) Stem morphology and anatomy in *Amaranthus* L. (*Amaranthaceae*) - Taxonomic significance. Journal of the Torrey Botanical Society 128: 254-281.

Cumhur, A. and Malcolm, S. C. (2008) The Effects of Global Climate Change on Agriculture. American-Eurasian Journal of Agricultural and Environmental Sciences 3: 672-676.

Das, S. (2012) Systematics and taxonomic delimitation of vegetable, grain and weed amaranths: a morphological and biochemical approach. Genetic Resources and Crop Evolution 59: 289-303.

Dasgupta, S., Laplante, B., Meisner, C., Wheeler, D. and Yan, J., 2007. The Impact of Sea Level Rise on Developing Countries. A Comparative Analysis. Policy Research Working Paper. World Bank, Washington, DC. pp 51.

Ebert, A. W., 2013. Ex situ Conservation of Plant Genetic Resources of Major Vegetables. Comservation of Tropical Plant Species. Springer, pp. 373-417.

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

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

Faircloth, B. C. (2008) MSATCOMMANDER: detection of microsatellite repeat arrays and automated, locus-specific primer design. Molecular Ecology Resources 8: 92-94.

FAO, 2008. Climate Change and Food Security: A Framework Doccument. Rome, Italy. pp. 93.

FAO, 2011. Climate Change Impacts on Agriculture in Vietnam. Food and Agriculture Organization, Hanoi, Vietnam. pp. 10.

Fasuyi, A. O. and Akindahunsi, A. O. (2009) Nutritional Evaluation of *Amaranthus cruentus* Leaf Meal Based Broiler Diets Supplemented with Cellulase/Glucanase/Xylanase Enzyme. American Journal of Food Technology 4: 108-118.

Govindaraj, M., Vetriventhan, M. and Srinivasan, M. (2015) Importance of Genetic Diversity Assessment in Crop Plants and Its Recent Advances: An Overview of Its Analytical Perspectives. Genetics Research International.

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

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

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

Hilu, K. W. and Liang, H. P. (1997) The matK gene: Sequence variation and application in plant systematics. American Journal of Botany 84: 830-839.

Ho, C., 2018. The Climate Change in Vietnam and Its Impact on Agricultural sector in Vietnam. https://www.researchgate.net/profile/Cuong_Ho5/publication/329024730_pp. 30

Hoi, P. V., Mol, A. P. J., Oosterveer, P., Van Den Brink, P. J. and Huong, P. T. M. (2016) Pesticide use in Vietnamese vegetable production: a 10-year study. International Journal of Agricultural Sustainability 14: 325-338.

Hoi, V. P., Mol, A. and Oosterveer, P. (2013) State governance of pesticide use and trade in Vietnam. NJAS - Wageningen Journal of Life Sciences 67: 19-26.

Iamonico, D. (2015) Taxonomic revision of the genus *Amaranthus* (*Amaranthaceae*) in Italy. Phytotaxa 199: 1-84.

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

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

Khandaker, L., Akond, A., Ali, M. B. and Oba, S. (2010) Biomass yield and accumulations of bioactive compounds in red amaranth (*Amaranthus tricolor* L.) grown under different colored shade polyethylene in spring season. Scientia Horticulturae 123: 289-294.

Kietlinski, K. D., Jimenez, F., Jellen, E. N., Maughan, P. J., Smith, S. M. and Pratt, D. B. (2014) Relationships between the Weedy *Amaranthus hybridus* (*Amaranthaceae*) and the Grain Amaranths. Crop Science 54: 220-228.

Kumar, P., Gupta, V. K., Misra, A. K., Modi, D. R. and Pandey, B. K. (2009) Potential of Molecular Markers in Plant Biotechnology. Plant Omics 2: 141-162.

Kumar, S., Stecher, G. and Tamura, K. (2016) MEGA7: Molecular evolutionary genetics analysis version 7.0 for Bigger Datasets. Molecular Biology and Evolution 33: 1870-1874.

Lahaye, R., Van Der Bank, M., Bogarin, D., Warner, J., Pupulin, F., Gigot, G., Maurin, O., Duthoit, S., Barraclough, T. G. and Savolainen, V. (2008) DNA barcoding the

floras of biodiversity hotspots. Proceedings of the National Academy of Sciences of the United States of America 105: 2923-2928.

Lan, T. T. N., Le, N. P. and Phong, N. T. (2014) State pesticide management in Thai Binh province. Journal of Science and Development 14: 836-843.

Lanoue, K. Z., Wolf, P. G., Browning, S. and Hood, E. E. (1996) Phylogenetic analysis of restriction-site variation in wild and cultivated *Amaranthus* species (*Amaranthaceae*). Theoretical and Applied Genetics 93: 722-732.

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

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

Liu, F. and Stutzel, H. (2004) Biomass partitioning, specific leaf area, and water use efficiency of vegetable amaranth (*Amaranthus* spp.) in response to drought stress. Scientia Horticulturae 102: 15-27.

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

Luo, P. P., Mu, D. R., Xue, H., Ngo-Duc, T., Dang-Dinh, K., Takara, K., Nover, D. and Schladow, G. (2018) Flood inundation assessment for the Hanoi Central Area, Vietnam under historical and extreme rainfall conditions. Scientific Reports 8. *12623* | DOI:10.1038/s41598-018-30024-5

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

Maughan, P. J., Smith, S. M., Fairbanks, D. J. and Jellen, E. N. (2011) Development, characterization, and linkage mapping of single nucleotide polymorphisms in the grain Amaranths (*Amaranthus* sp.). Plant Genome 4: 92-101.

Mertz, O., Halsnaes, K., Olesen, J. E. and Rasmussen, K. (2009) Adaptation to climate change in developing countries. Environmental Management 43: 743-752.

Mosyakin, S. L. and Robertson, K. R. (1996) New infrageneric taxa and combinations in *Amaranthus* (*Amaranthaceae*). Annales Botanici Fennici 33: 275-281.

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

Myers, R. L., 1996. Amaranth: New crop opportunity. J. Janick (ed.), Progress in new crops. ASHS Press, Alexandria,VA, pp. 207-220.

Nguyen, T. M., N.T.T., L., J., H. and D. B., H. (2018) Pesticide use in vegetable production: A survey of Vietnamese farmers' knowledge. Plant Protection Science 54: 203-214.

Nguyen, V. T. and Tran, K. T. (1996) The guide for vegetable growers, Agricultural Publishing House, p. 114.

Ogundipe, O. T. and Chase, M. (2009) Phylogenetic analyses of *Amaranthaceae* based on matK DNA sequence data with emphasis on West African species. Turkish Journal of Botany 33: 153-161.

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

O'Brien, G. K. and Price., M. L., 1983. Amaranth: Grain and vegetable type. Educational Concerns for Hunger Organization (ECHO) Technical Note #2, PP. 9.

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

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

Pischedda.S, Barral-Arca, R., Gómez-Carballa, A., Pardo-Seco, J., Catelli, M. L., Álvarez-Iglesias, V., Cárdenas, J. M., Nguyen, N. D., Ha, H. H., *et al.* (2017) Phylogeographic and genome-wide investigations of Vietnam ethnic groups reveal signatures of complex historical demographic movements. Scientific Reports 7. *12630* | DOI:10.1038/s41598-017-12813-6

Prc, 2018. Database of plant genetic resources Plant Resources Center, Vietnam.

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

Rastogi, A. and Shukla, S. (2013) Amaranth: A new millennium crop of nutraceutical values. Critical Reviews in Food Science and Nutrition 53: 109-125.

Reddy, V. V. P. and Varalakshmi, B. (1998) Heterosis and combining ability for leaf yield and its components in vegetable amaranth (*Amaranthus tricolor*). Indian Journal of Agricultural Sciences 68: 773-775.

Robertson, K. R. (1981) The genera of *Amaranthaceae* in the Southeastern United States. Journal of the Arnold Arboretum 62: 267-313.

Roorkiwal, M., Von Wettberg, E. J., Upadhyaya, H. D., Warschefsky, E., Rathore, A. and Varshney, R. K. (2014) Exploring germplasm diversity to understand the domestication process in Cicer spp. using SNP and DArT Markers. PLoS ONE 9: e102016.

Sammour, R. H., Radwan, S. A. and Mira, M. (2012) Genetic diversity in genus *Amaranthus*: From morphology to genomic DNA. Research & Reviews in BioSciences 6: 351-360.

Sauer, J. D. (1950a) Amaranths as dye plants among the pueblo peoples. Southwestern Journal of Anthropology 6: 412-415.

Sauer, J. D. (1950b) The grain amaranths: A survey of their history and classification. Annals of the Missouri Botanical Garden 37: 561-632.

Sauer, J. D. (1955) Revision of the dioecious amaranths. Madroño 13: 5–46.

Sauer, J. D. (1967) The grain amaranths and their relatives: A revised taxonomic and geographic survey. Annals of the Missouri Botanical Garden 54: 103-137.

Sauer, J. D. (1969) Identity of archaeologic grain amaranths from the Valley of Tehuacan, Puebla, Mexico. American Antiquity 34: 80-81.

Sauer, J. D. (1993) Historical geography of crop plants: A Select Roster CRC Press, Boca Raton. pp. 320.

Schonfeldt, H. C. and Pretorius, B. (2011) The nutrient content of five traditional South African dark green leafy vegetables-A preliminary study. Journal of Food Composition and Analysis 24: 1141-1146.

Schweig, S. R. and Brown, R. N. (2018) Vegetable amaranths for summer greens production in the Northeastern United States. Horttechnology 28: 399-406.

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

Shukla, S., Bhargava, A., Chatterjee, A., Pandey, A. C. and Mishra, B. K. (2009) Diversity in phenotypic and nutritional traits in vegetable amaranth (*Amaranthus tricolor*), a nutritionally underutilised crop. Journal of the Science of Food and Agriculture 90: 139-144.

Shukla, S., Bhargava, A., Chatterjee, A. and Singh, S. P. (2004) Estimates of genetic parameters to determine variability for foliage yield and its different quantitative and qualitative traits in vegetable amaranth (*A. tricolor*). Journal of Genetics and Breeding 58: 169-176.

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

Shukla, S., Bhargava, A., Chatterjee, A., Srivastava, J., Singh, N. and Singh, S. P. (2006b) Mineral Profile and Variability in Vegetable Amaranth (*Amaranthus tricolor*). Plant Foods for Human Nutrition 61: 21-26.

Shukla, S. and Singh, S. P. (2000) Studies on genetic parameters in vegetable amaranth. Journal of Genetics and Breeding 54: 133-135.

Silva, P. I., Martins, A. M., Gouvea, E. G., Pessoa-Filho, M. and Ferreira, M. E. (2013) Development and validation of microsatellite markers for Brachiaria

ruziziensisobtained by partial genome assembly of Illumina single-end reads. BMC Genomics 14: 17.

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

Srivastava, R. (2017) An updated review on phyto-pharmacological and pharmacognostical profile of Amaranthus tricolor: A herb of nutraceutical potentials. Pharma Innov 6: 127-129.

Stetter, M. G. and Schmid, K. J. (2017) Analysis of phylogenetic relationships and genome size evolution of the Amaranthus genus using GBS indicates the ancestors of an ancient crop. Molecular Phylogenetics and Evolution 109: 80-92.

Suarez-Gonzalez, A., Lexer, C. and Cronk, Q. C. B. (2018) Adaptive introgression: a plant perspective. Biology Letters 14: 8.

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

Taheri, S., Abdullah, T. L., Yusop, M. R., Hanafi, M. M., Sahebi, M., Azizi, P. and Shamshiri, R. R. (2018) Mining and Development of Novel SSR Markers Using Next Generation Sequencing (NGS) Data in Plants. Molecules 23: 20.

Thapa, R. and Blair, M. W. (2018) Morphological Assessment of Cultivated and Wild Amaranth Species Diversity. Agronomy-Basel 8. pp. 13.

Tran, T., T., N. V., Huynh, T. L. H., Mai, V. K., Nguyen, X. H. and P., D. H., 2016. Climate Change and Sea Level Rise Scenarios for Viet Nam - Summary for Policymakers. Ministry of Natural Resources and Environment, Hanoi, Vietnam.

UPOV, 2008. Grain Amaranth. In: International Union for the protection of new varieties of plants (Ed.), TG/247/1. pp. 35.

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

Venskutonis, P. R. and Kraujalis, P. (2013) Nutritional Components of Amaranth Seeds and Vegetables: A Review on Composition, Properties, and Uses. Comprehensive Reviews in Food Science and Food Safety 12: 381-412.

Vieira, M. L. C., Santini, L., Diniz, A. L. and Munhoz, C. D. (2016) Microsatellite markers: what they mean and why they are so useful. Genetics and Molecular Biology 39: 312-328.

Viljoen, E., Odeny, D. A., Coetzee, M. P. A., Berger, D. K. and Rees, D. J. G. (2018) Application of Chloroplast Phylogenomics to Resolve Species Relationships Within the Plant Genus *Amaranthus*. Journal of Molecular Evolution 86: 216-239.

Wang, H., Moustier, P., Loc, N. T. T. and Tho, P. T. H. (2012) Quality Control of Safe Vegetables by Collective Action in Hanoi, Vietnam. Procedia Economics and Finance 2: 344-352.

Wang, X. Q. and Park, J. Y. (2013) Comparison of Genetic Diversity among Amaranth Accessions from South and Southeast Asia using SSR Markers Korean J. Medicinal Crop Science 21: 220-228.

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

Wiens, J. J. (2004) The role of morphological data in phylogeny reconstruction. Systematic Biology 53: 653-661.

Wu, H. X., Sun, M., Yue, S. X., Sun, H. L., Cai, Y., Huang, R. H., Brenner, D. and Corke, H. (2000) Field evaluation of an *Amaranthus* genetic resource collection in China. Genetic Resources and Crop Evolution 47: 43-53.

Wu, X. B. and Blair, M. W. (2017) Diversity in grain amaranths and relatives distinguished by genotyping by sequencing (GBS). Frontiers in Plant Science 8: 12.

Xu, F. X. and Sun, M. (2001) Comparative analysis of phylogenetic relationships of grain amaranths and their wild relatives (*Amaranthus; amaranthaceae*) using internal transcribed spacer, amplified fragment length polymorphism, and double-primer fluorescent intersimple sequence repeat markers. Molecular Phylogenetics and Evolution 21: 372-387.

Yu, J., Xue, J. H. and Zhou, S. L. (2011) New universal matK primers for DNA barcoding angiosperms. Journal of Systematics and Evolution 49: 176-181.

Yudina, R. S., Zheleznova, N. B., Zakharova, O. V., Zheleznov, A. V. and Shumny, V. K. (2005) Isozyme analysis in a genetic collection of amaranths (*Amaranthus* L.). Russian Journal of Genetics 41: 1395-1400.

Zalapa, J. E., Cuevas, H., Zhu, H. Y., Steffan, S., Senalik, D., Zeldin, E., Mccown, B., Harbut, R. and Simon, P. (2012) Using Next generation sequencing approaches to isolate simple sequence repeat (SSR) loci in the plant sciences. American Journal of Botany 99: 193-208.

Supplemental materials

Supplemental Table 1. List of the 272 Vietnamese *Amaranthus* (VA) accessions that were used in this study

Code	Accession number	Local name	Seed distributor	Collection area	Scientific name based on *matK* barcode phylogeny
VA001	3700	Rau den do	PRC	Northeast	*Amaranthus tricolor* L.
VA002	3701	Rau den tim	PRC	Northeast	Hybrida complex
VA003	3702	Rau den tam	PRC	Northeast	*Amaranthus tricolor* L.
VA004	3764	Rau den do	PRC	Northeast	*Amaranthus tricolor* L.
VA005	3848	Rau den xanh	PRC	Northeast	Hybrida complex
VA006	3879	Rau den	PRC	North	*Amaranthus tricolor* L.
VA007	5484	Rau den xanh	PRC	Northeast	*Amaranthus dubius* Mart. ex Thell.
VA008	5485	Peec ngan sai	PRC	Northeast	*Amaranthus dubius* Mart. ex Thell.
VA009	5486	Rau den	PRC	Northeast	*Amaranthus dubius* Mart. ex Thell.
VA010	5487	Rau den do	PRC	Northeast	*Amaranthus tricolor* L.
VA011	5489	Den đo	PRC	Northwest	*Amaranthus tricolor* L.
VA012	5490	Rau den ca	PRC	Northwest	*Amaranthus tricolor* L.
VA013	5491	Rau den ta	PRC	Northwest	*Amaranthus dubius* Mart. ex Thell.
VA014	5492	Rau den trang	PRC	Northwest	*Amaranthus tricolor* L.
VA015	5493	Rau den mao	PRC	Northwest	*Amaranthus tricolor* L.
VA016	5496	Rau den do	PRC	Northeast	*Amaranthus tricolor* L.
VA017	5497	Rau den	PRC	Southeast	*Amaranthus tricolor* L.
VA018	5499	Rau den dai	PRC	Central	Hybrida complex
VA019	5500	Rau den dại	PRC	Central	*Amaranthus tricolor* L.
VA020	5501	Rau den	PRC	Southeast	*Amaranthus tricolor* L.
VA021	5502	Rau den	PRC	Southwest	*Amaranthus tricolor* L.
VA022	5503	Rau den	PRC	Southwest	*Amaranthus tricolor* L.
VA023	5504	Rau den dai	PRC	Southwest	*Amaranthus tricolor* L.
VA024	5505	Rau den do	PRC	Southwest	*Amaranthus tricolor* L.
VA025	5506	Rau den	PRC	Southwest	*Amaranthus tricolor* L.
VA026	5507	Rau den	PRC	Southwest	*Amaranthus tricolor* L.
VA027	5508	Den do	PRC	Southeast	*Amaranthus tricolor* L.
VA028	5509	Den trang	PRC	Southeast	*Amaranthus tricolor* L.
VA029	5510	Den trang	PRC	Southeast	*Amaranthus tricolor* L.
VA030	5511	Den tia	PRC	Southwest	*Amaranthus tricolor* L.
VA031	5512	Rau den	PRC	Red River	*Amaranthus tricolor* L.
VA032	6543	Rau den Trung	PRC	Northwest	*Amaranthus dubius* Mart. ex Thell.
VA033	6544	Rau den	PRC	Southwest	*Amaranthus tricolor* L.
VA034	6615	Rau den	PRC	Central	*Amaranthus tricolor* L.
VA035	6616	Rau den trang	PRC	Central	*Amaranthus tricolor* L.
VA036	6617	Den do	PRC	Central	*Amaranthus tricolor* L.
VA037	6618	Den tia	PRC	Central	Hybrida complex
VA038	6619	Den trang	PRC	Northeast	Hybrida complex
VA039	6620	Rau den	PRC	Northeast	*Amaranthus tricolor* L.
VA040	6753	Rau den trang	PRC	Northeast	*Amaranthus viridis* L.
VA041	6882	Rau den do	PRC	North	*Amaranthus tricolor* L.

Code	Accession number	Local name	Seed distributor	Collection area	Scientific name based on *matK* barcode phylogeny
VA042	6883	Rau den do	PRC	North	*Amaranthus tricolor* L.
VA043	6884	Rau den trang	PRC	North	*Amaranthus tricolor* L.
VA044	6885	Rau den tia	PRC	North	*Amaranthus tricolor* L.
VA045	7961	Rau gien tia	PRC	North	Hybrida complex
VA046	7962	Pac hom	PRC	North	*Amaranthus dubius* Mart. ex Thell.
VA047	7963	Phac hom	PRC	Northwest	*Amaranthus dubius* Mart. ex Thell.
VA048	8999	Rau den	PRC	Northeast	Hybrida complex
VA049	9000	Rau den	PRC	Red River	*Amaranthus tricolor* L.
VA050	9001	Rau den	PRC	Red River	*Amaranthus tricolor* L.
VA051	9002	Rau den	PRC	Red River	*Amaranthus tricolor* L.
VA052	9003	Rau den	PRC	Red River	*Amaranthus tricolor* L.
VA053	9004	Rau den ta	PRC	Northwest	*Amaranthus tricolor* L.
VA054	9005	Rau den trang	PRC	Northwest	*Amaranthus tricolor* L.
VA055	9006	Rau den mao	PRC	Northwest	*Amaranthus tricolor* L.
VA056	9007	Rau den mao	PRC	Northwest	*Amaranthus tricolor* L.
VA057	9008	Rau den mao	PRC	Northwest	*Amaranthus tricolor* L.
VA058	9009	Rau den mao	PRC	Northwest	*Amaranthus tricolor* L.
VA059	9010	Rau den mao	PRC	Northwest	*Amaranthus tricolor* L.
VA060	9011	Rau den do	PRC	Northeast	*Amaranthus tricolor* L.
VA061	9012	Rau den	PRC	Central	*Amaranthus dubius* Mart. ex Thell.
VA062	9013	Phac hom nang	PRC	Northwest	*Amaranthus tricolor* L.
VA063	9014	Den tia	PRC	North	*Amaranthus tricolor* L.
VA064	9015	Gien đo	PRC	Red River	*Amaranthus tricolor* L.
VA065	9806	Rau gien trang	PRC	North	*Amaranthus dubius* Mart. ex Thell.
VA066	12512	Rau den	PRC	Central	Hybrida complex
VA067	12518	Rau gien	PRC	Southwest	*Amaranthus tricolor* L.
VA068	12861	Phac hom	PRC	Northwest	*Amaranthus dubius* Mart. ex Thell.
VA069	13712	Den tia	PRC	Central	Hybrida complex
VA070	13713	Rau gien do	PRC	Northeast	*Amaranthus tricolor* L.
VA071	13714	Tac trenh	PRC	Northwest	*Amaranthus dubius* Mart. ex Thell.
VA072	15333	Tac trenh cang	PRC	Northwest	*Amaranthus tricolor* L.
VA073	15334	Tac trenh thia	PRC	Northwest	*Amaranthus tricolor* L.
VA074	15338	Rau den trang	PRC	Northeast	*Amaranthus dubius* Mart. ex Thell.
VA075	15339	Pac hom	PRC	Northwest	*Amaranthus blitum* L.
VA076	15341	Phac hom	PRC	Northwest	*Amaranthus tricolor* L.
VA077	15342	Pac hom	PRC	Northwest	*Amaranthus blitum* L.
VA078	15343	Phac hom	PRC	Northwest	Hybrida complex
VA079	15344	Phac hum khao	PRC	Northwest	*Amaranthus blitum* L.
VA080	15345	Phac hom	PRC	Northwest	*Amaranthus blitum* L.
VA081	15346	Xu ntu	PRC	Northwest	Hybrida complex
VA082	15347	Xu ntu	PRC	Northwest	Hybrida complex
VA083	15348	Ru gien	PRC	Northwest	*Amaranthus tricolor* L.
VA084	15349	Cha ni go chu	PRC	Northwest	Hybrida complex
VA085	15350	Song lu	PRC	Northwest	*Amaranthus blitum* L.
VA086	15351	Phac hom	PRC	Northwest	Hybrida complex
VA087	15352	Xu tu	PRC	Northwest	Hybrida complex
VA088	15353	Chun tu	PRC	Northwest	*Amaranthus dubius* Mart. ex Thell.
VA089	15354	Phắc hôm đanh	PRC	Northwest	*Amaranthus tricolor* L.
VA090	15355	Xung tu	PRC	Northwest	*Amaranthus tricolor* L.

Code	Accession number	Local name	Seed distributor	Collection area	Scientific name based on *matK* barcode phylogeny
VA091	15356	Phac hom	PRC	Northwest	*Amaranthus blitum* L.
VA092	15357	Phac hom	PRC	Northwest	*Amaranthus blitum* L.
VA093	15358	Phac hom	PRC	Northwest	*Amaranthus blitum* L.
VA094	15359	Phac hom	PRC	Northwest	*Amaranthus blitum* L.
VA095	15361	Su ntu	PRC	Northeast	Hybrida complex
VA096	15363	Thanh my la	PRC	Northeast	Hybrida complex
VA097	15364	Pac hom	PRC	Northeast	*Amaranthus blitum* L.
VA098	15365	Pac hom	PRC	Northeast	*Amaranthus blitum* L.
VA099	15366	Phac hom	PRC	Northwest	*Amaranthus blitum* L.
VA100	16697	Den trang	PRC	Central	*Amaranthus tricolor* L.
VA101	16698	Den do	PRC	Central	Hybrida complex
VA102	16700	Phac khum	PRC	Northwest	Hybrida complex
VA103	16701	Rau den	PRC	Northwest	*Amaranthus dubius* Mart. ex Thell.
VA104	16702	Phac hum	PRC	Northwest	Hybrida complex
VA105	16703	Xu tu	PRC	Northeast	Hybrida complex
VA106	16704	Su du	PRC	Northeast	Hybrida complex
VA107	16705	Xuong tu	PRC	Northeast	Hybrida complex
VA108	16706	Thanh li	PRC	Northeast	Hybrida complex
VA109	16707	Su tu	PRC	Northeast	Hybrida complex
VA110	16708	Rau gien do	PRC	Northeast	*Amaranthus tricolor* L.
VA111	16709	Den do	PRC	South	*Amaranthus tricolor* L.
VA112	16710	Den trang	PRC	South	Hybrida complex
VA113	16711	Larom	PRC	South	Hybrida complex
VA114	16712	Ro hum dom	PRC	Central	Hybrida complex
VA115	16713	Ro hum bo	PRC	Central	Hybrida complex
VA116	16714	Rơhum dom po	PRC	Central	Hybrida complex
VA117	16715	Rohum dom	PRC	Central	Hybrida complex
VA118	16716	Rohum bo	PRC	Central	Hybrida complex
VA119	16717	Rohum pro he	PRC	Central	Hybrida complex
VA120	16718	Rau den	PRC	Central	*Amaranthus tricolor* L.
VA121	16719	Ro hum brah	PRC	Central	Hybrida complex
VA122	20460	Soongs tu	PRC	Northeast	Hybrida complex
VA123	20462	Xe mi say	PRC	Northwest	Hybrida complex
VA124	20464	Sung tu	PRC	Northwest	Hybrida complex
VA125	20465	Rau huan	PRC	Northwest	Hybrida complex
VA126	20467	Thung tu	PRC	Northeast	Hybrida complex
VA127	20468	Phec gien	PRC	Northeast	Hybrida complex
VA128	20472	Phec hom	PRC	Northeast	*Amaranthus tricolor* L.
VA129	20473	Pe dup	PRC	South	*Amaranthus viridis* L.
VA130	20475	Rau gien	PRC	Northeast	*Amaranthus tricolor* L.
VA131	20476	Rau gien	PRC	South	*Amaranthus tricolor* L.
VA132	20478	Rau gien	PRC	Northeast	*Amaranthus tricolor* L.
VA133	20479	Rau gien	PRC	Northeast	*Amaranthus tricolor* L.
VA134	20488	Rau den la do	PRC	South	*Amaranthus tricolor* L.
VA135	20489	Den do	PRC	South	*Amaranthus tricolor* L.
VA137	20492	Song tu	PRC	Northeast	Hybrida complex
VA138	22078	Nong lu	PRC	Northwest	Hybrida complex
VA139	22079	Phac hom	PRC	Northwest	*Amaranthus blitum* L.
VA140	22080	Phac hom	PRC	Northwest	*Amaranthus blitum* L.

Code	Accession number	Local name	Seed distributor	Collection area	Scientific name based on *matK* barcode phylogeny
VA141	22081	Hang sai	PRC	Northeast	Hybrida complex
VA142	22082	Rhum	PRC	Central	Hybrida complex
VA143	22083	Rau den trang	PRC	Southwest	*Amaranthus dubius* Mart. ex Thell.
VA144	22084	Pe rup	PRC	South	*Amaranthus tricolor* L.
VA145	22094	Phac hom	PRC	Northwest	*Amaranthus tricolor* L.
VA146	22095	Rau gien	PRC	Northwest	*Amaranthus tricolor* L.
VA147	22096	Giay hen thi	PRC	Northeast	Hybrida complex
VA148	T00813	Rau den	PRC	Northeast	*Amaranthus dubius* Mart. ex Thell.
VA149	T08318	Rhum	PRC	Central	Hybrida complex
VA150	T08323	Rau den	PRC	Central	*Amaranthus dubius* Mart. ex Thell.
VA151	T08324	Den trang	PRC	Central	Hybrida complex
VA152	T09381	Gay hen	PRC	Northeast	Hybrida complex
VA153	T09677	Phyac hom	PRC	Northeast	Hybrida complex
VA154	T09725	Rau gien	PRC	Northeast	Hybrida complex
VA155	T09726	Rau thun tum	PRC	Northeast	Hybrida complex
VA156	T09779	Song du	PRC	Northeast	Hybrida complex
VA158	T09797	Rau den	PRC	Red River	*Amaranthus tricolor* L.
VA159	T09826	Rau den	PRC	Red River	*Amaranthus tricolor* L.
VA160	T09827	Rau den	PRC	Red River	*Amaranthus tricolor* L.
VA161	T12249	Rau khong dau	PRC	Northeast	Hybrida complex
VA162	T12251	Kla dun	PRC	Northwest	*Amaranthus tricolor* L.
VA163	T12252	Phac cat	PRC	Northwest	*Amaranthus viridis* L.
VA164	T12257	Xe mi say	PRC	Northwest	Hybrida complex
VA165	T12259	Thung tu	PRC	Northeast	Hybrida complex
VA166	T12262	Phac hom	PRC	Northeast	Hybrida complex
VA167	T12266	Gien com	PRC	Northeast	*Amaranthus viridis* L.
VA168	T12270	Gien mat	PRC	Northeast	*Amaranthus tricolor* L.
VA170	T15696	Rau gien	PRC	Northeast	*Amaranthus tricolor* L.
VA171	T15706	Rau gien	PRC	Northeast	*Amaranthus tricolor* L.
VA172	T17582	Rau gien com	PRC	Northeast	*Amaranthus blitum* L.
VA173	T21421	Rau gien	PRC	Southwest	*Amaranthus tricolor* L.
VA174	T21422	Rau gien	PRC	Southwest	*Amaranthus dubius* Mart. ex Thell.
VA175	T21424	Rau den	PRC	Southwest	*Amaranthus tricolor* L.
VA176	T21425	Hlahnhem	PRC	Central	*Amaranthus tricolor* L.
VA177	T21426	Rau den	PRC	Central	*Amaranthus tricolor* L.
VA178	5498	Rau den	PRC	Central	Hybrida complex
VA179	16699	Rau gien	PRC	WorldVeg	Hybrida complex
VA180	20469	Gien 10	PRC	Northeast	*Amaranthus tricolor* L.
VA181	20470	Gien tia	PRC	Northeast	*Amaranthus tricolor* L.
VA182	20481	Den com	PRC	Northwest	*Amaranthus blitum* L.
VA183	20482	Rau den do	PRC	Central	*Amaranthus tricolor* L.
VA184	20483	Rau den tieu	PRC	South	*Amaranthus tricolor* L.
VA185	20484	Rau den do	PRC	South	*Amaranthus tricolor* L.
VA186	20485	Rau den do	PRC	South	*Amaranthus tricolor* L.
VA187	20486	Rau den la	PRC	South	*Amaranthus tricolor* L.
VA188	20487	Rau den la do	PRC	South	*Amaranthus tricolor* L.
VA189	20493	Sien mi Sai	PRC	Northeast	Hybrida complex
VA190	22085	Pac hom lanh	PRC	Northwest	*Amaranthus tricolor* L.
VA191	22086	Den trang	PRC	North	*Amaranthus tricolor* L.

Code	Accession number	Local name	Seed distributor	Collection area	Scientific name based on *matK* barcode phylogeny
VA192	22087	Den trang	PRC	North	*Amaranthus tricolor* L.
VA193	22088	Den do	PRC	North	*Amaranthus tricolor* L.
VA194	22089	Den tím	PRC	North	*Amaranthus tricolor* L.
VA195	22090	Den tim	PRC	North	*Amaranthus tricolor* L.
VA196	22091	Den tía	PRC	North	*Amaranthus tricolor* L.
VA197	22093	Den kim	PRC	North	*Amaranthus tricolor* L.
VA198	22094	Rau gien	PRC	Northeast	*Amaranthus tricolor* L.
VA199	22096	Giay hen thi	PRC	Northeast	*Amaranthus tricolor* L.
VA200	T05901	Phac hom	PRC	Northwest	*Amaranthus blitum* L.
VA201	T08318	Rhum	PRC	Central	*Amaranthus dubius* Mart. ex Thell.
VA202	T08952	Phac lam	PRC	Northeast	Hybrida complex
VA203	T09675	Quy go	PRC	Northeast	Hybrida complex
VA204	T09676	Quy go pu	PRC	Northeast	*Amaranthus dubius* Mart. ex Thell.
VA205	T12239	Gien xanh	PRC	Northeast	*Amaranthus dubius* Mart. ex Thell.
VA206	T12261	Phec hom	PRC	Northeast	*Amaranthus tricolor* L.
VA207	T12263	Xi xanh	PRC	Northeast	Hybrida complex
VA208	T12264	Xi xanh	PRC	Northeast	Hybrida complex
VA209	T12265	Sac ham	PRC	Northeast	Hybrida complex
VA210	T12268	Gien gai trang	PRC	Northeast	Hybrida complex
VA211	T15687	Phec hom	PRC	Northeast	*Amaranthus dubius* Mart. ex Thell.
VA212	T15689	Phac hom	PRC	Northeast	Hybrida complex
VA213	T15690	Vi ky	PRC	Northeast	*Amaranthus dubius* Mart. ex Thell.
VA214	T15691	Hla tre	PRC	Central	*Amaranthus tricolor* L.
VA215	T15692	Lah Hrum	PRC	Central	Hybrida complex
VA216	T15699	Rau gien	PRC	Central	*Amaranthus tricolor* L.
VA217	T15702	Gien com	PRC	South	Hybrida complex
VA218	T15703	Rau gien cơm	PRC	South	Hybrida complex
VA219	T15705	Gien cơm	PRC	South	*Amaranthus tricolor* L.
VA220	T17583	Rau den do	PRC	Northeast	*Amaranthus tricolor* L.
VA221	T17584	Den do	PRC	Northwest	*Amaranthus tricolor* L.
VA222	T17589	Rau den do	PRC	South	*Amaranthus dubius* Mart. ex Thell.
VA223	T17596	Rau den la	PRC	South	*Amaranthus tricolor* L.
VA224	T17597	Den 3 mau	PRC	South	*Amaranthus tricolor* L.
VA225	T17598	Den trang	PRC	South	Hybrida complex
VA226	T17858	Den com	PRC	WorldVeg	*Amaranthus viridis* L.
VA227	T17859	Rau den trang	PRC	WorldVeg	Hybrida complex
VA228	T17860	Rau den	PRC	WorldVeg	Hybrida complex
VA229	T20471	Den do la	PRC	North	*Amaranthus tricolor* L.
VA230	T20475	Den do	PRC	North	*Amaranthus tricolor* L.
VA231	T20481	Phac hom	PRC	Northwest	*Amaranthus tricolor* L.
VA232	T20485	Pac ham	PRC	Northeast	*Amaranthus dubius* Mart. ex Thell.
VA233	T20486	Pac ham	PRC	Northeast	*Amaranthus tricolor* L.
VA234	T20487	Rau gien	PRC	Northwest	*Amaranthus tricolor* L.
VA235	T20490	Ro ham	PRC	Central	*Amaranthus tricolor* L.
VA236	T20494	Rau gien	PRC	Northeast	Hybrida complex
VA237	T20495	Tau le	PRC	Northeast	Hybrida complex
VA238	T20496	Phec hom dinh	PRC	Northeast	*Amaranthus tricolor* L.
VA239	T20536	Ray hieen	PRC	Northeast	*Amaranthus dubius* Mart. ex Thell.
VA240	T21423	Rau gien	PRC	Southwest	*Amaranthus tricolor* L.

Code	Accession number	Local name	Seed distributor	Collection area	Scientific name based on *matK* barcode phylogeny
VA241	T17601	Den Mui	PRC	South	*Amaranthus blitum* L.
VA242	T17605	Ro sun tu	PRC	Northwest	Hybrida complex
VA243	T23028	Pac ham	PRC	Northeast	*Amaranthus tricolor* L.
VA244	2016.01	Rau den ta	PRC	North	*Amaranthus dubius* Mart. ex Thell.
VA245	2016.02	Rau gien gai	PRC	North	Hybrida complex
VA246	2016.03	Rau gien gai	PRC	North	Hybrida complex
VA247	2016.04	Rau gien gai	PRC	North	Hybrida complex
VA248	2016.05	Pha hom	PRC	North	Hybrida complex
VA249	2016.13	Lai Le xi	PRC	North	*Amaranthus tricolor* L.
VA251	2016.16	Lai Le pua	PRC	North	*Amaranthus dubius* Mart. ex Thell.
VA252	2016.24	Phac hom le	PRC	North	*Amaranthus dubius* Mart. ex Thell.
VA253	2016.25	Phac hom danh	PRC	North	*Amaranthus tricolor* L.
VA254	2016.26	Phac hom	PRC	North	*Amaranthus dubius* Mart. ex Thell.
VA255	2016.34	Phac hom	PRC	North	*Amaranthus dubius* Mart. ex Thell.
VA256	2016.36	Phac hom	PRC	North	*Amaranthus dubius* Mart. ex Thell.
VA257	2016.37	Phac hom	PRC	North	*Amaranthus dubius* Mart. ex Thell.
VA258	2016.38	Phac hom	PRC	North	*Amaranthus dubius* Mart. ex Thell.
VA259	2016.39	Phac hom	PRC	North	Hybrida complex
VA260	2016.40	Phac hom	PRC	North	*Amaranthus dubius* Mart. ex Thell.
VA261	2016.41	Phac hom	PRC	North	*Amaranthus dubius* Mart. ex Thell.
VA263	2016.45	Sau trenh	PRC	North	*Amaranthus viridis* L.
VA264	2016.46	Rau gien canh	PRC	North	Hybrida complex
VA265	2016.47	Sau trenh	PRC	North	Hybrida complex
VA266	2016.49	Sau trenh kai	PRC	North	*Amaranthus dubius* Mart. ex Thell.
VA269	2016.52	Rau gien xanh	PRC	North	*Amaranthus dubius* Mart. ex Thell.
VA270	2016.53	Rau gien gai	PRC	North	Hybrida complex
VA271	2016.54	Rau gien do	PRC	North	*Amaranthus dubius* Mart. ex Thell.
VA272	2016.56	Phac hom	PRC	North	*Amaranthus dubius* Mart. ex Thell.
VA273	2016.66	Phac hom deng	PRC	North	*Amaranthus tricolor* L.
VA275	2016.68	Phac hom khau	PRC	North	*Amaranthus dubius* Mart. ex Thell.
VA276	2016.72	Phac hom nam	PRC	North	*Amaranthus tricolor* L.
VA277	2016.76	Clom	PRC	North	Hybrida complex
VA278	2016.77	Rau gien xanh	PRC	North	Hybrida complex
VA279	T20457	Phac hom deng	PRC	Unknown	Hybrida complex
VA280	T20463	Phac hom hao	PRC	Unknown	Hybrida complex

PRC, Plant Resources Center of Vietnam; WorldVeg, World Vegetable Center

Supplemental Table 2. List of the 182 world *Amaranthus* (WA) accessions that were used in this study

No.	Code	Name	Scientific name	Seed distributor	Country of collection
1	WA001	Irokoina	*Amaranthus tricolor* L.	University of Tsukuba	na
2	WA002	Biam	*Amaranthus tricolor* L.	University of Tsukuba	na
3	WA003	Green Pointed	*Amaranthus tricolor* L.	University of Tsukuba	na
4	WA004	Red Stripe	*Amaranthus tricolor* L.	University of Tsukuba	na
5	WA005	White Leaf	*Amaranthus tricolor* L.	University of Tsukuba	na
6	WA006	Tender Leaf	*Amaranthus tricolor* L.	University of Tsukuba	na
7	WA007	Green Round	*Amaranthus tricolor* L.	University of Tsukuba	na
8	WA008	Southern Red	*Amaranthus tricolor* L.	University of Tsukuba	na
9	WA009	Asia Red	*Amaranthus tricolor* L.	University of Tsukuba	na
10	WA010	New Aztec	*Amaranthus hypochondriacus* L.	University of Tsukuba	Japan
11	WA011	Ames-2051	*Amaranthus tricolor* L.	USDA	Indonesia
12	WA012	Ames-2091	*Amaranthus tricolor* L.	USDA	Nepal
13	WA013	Ames-5102	*Amaranthus tricolor* L.	USDA	Hong Kong
14	WA014	Ames-5368	*Amaranthus tricolor* L.	USDA	Bangladesh
15	WA015	Ames-29034	*Amaranthus tricolor* L.	USDA	Malaysia
16	WA016	PI-349553	*Amaranthus tricolor* L.	USDA	Papua New Guinea
17	WA017	PI-419057	*Amaranthus tricolor* L.	USDA	China
18	WA018	PI-599683	*Amaranthus tricolor* L.	USDA	India
19	WA019	PI-604669	*Amaranthus tricolor* L.	USDA	Taiwan
20	WA020	PI-607446	*Amaranthus tricolor* L.	USDA	Thailand
21	WA021	PI-608761	*Amaranthus tricolor* L.	USDA	India
22	WA022	I 942	*Amaranthus tricolor* L.	University of Tsukuba	Nepal
23	WA023	Y-3233(1)D	*Amaranthus tricolor* L.	University of Tsukuba	Nepal
24	WA024	Y-3233(2)D	*Amaranthus tricolor* L.	University of Tsukuba	Nepal
25	WA025	Y-3245(1)D	*Amaranthus tricolor* L.	University of Tsukuba	Nepal
26	WA026	Y-3245(2)D	*Amaranthus tricolor* L.	University of Tsukuba	Nepal
27	WA027	Amaranth	*Amaranthus tricolor* L.	University of Tsukuba	na
28	WA028	PI538322	*Amaranthus hypochondriacus* L.	USDA	US
29	WA031	PI419121	*Amaranthus tricolor* L.	USDA	China
30	WA036	PI603896	*Amaranthus tricolor* L.	USDA	US
31	WA039	PI604670	*Amaranthus tricolor* L.	USDA	Taiwan
32	WA040	Ames34159	*Amaranthus tricolor* L.	USDA	India
33	WA041	PI604668	*Amaranthus tricolor* L.	USDA	Taiwan
34	WA043	PI277269	*Amaranthus tricolor* L.	USDA	India
35	WA044	PI603897	*Amaranthus tricolor* L.	USDA	US
36	WA045	PI462128	*Amaranthus tricolor* L.	USDA	India
37	WA046	PI277267	*Amaranthus tricolor* L.	USDA	India
38	WA047	Ames2214	*Amaranthus tricolor* L.	USDA	US
39	WA048	Ames29505	*Amaranthus tricolor* L.	USDA	Brazil
40	WA049	Ames29504	*Amaranthus tricolor* L.	USDA	Brazil
41	WA050	PI462127	*Amaranthus tricolor* L.	USDA	India
42	WA055	PI636179	*Amaranthus tricolor* L.	USDA	Taiwan
43	WA056	Ames25153	*Amaranthus tricolor* L.	USDA	US
44	WA057	Ames25152	*Amaranthus tricolor* L.	USDA	US
45	WA059	PI667172	*Amaranthus tricolor* L.	USDA	India
46	WA060	Ames2103	*Amaranthus tricolor* L.	USDA	India

No.	Code	Name	Scientific name	Seed distributor	Country of collection
47	WA061	PI632237	*Amaranthus tricolor* L.	USDA	US
48	WA062	Ames15328	*Amaranthus tricolor* L.	USDA	US
49	WA063	Ames15330	*Amaranthus tricolor* L.	USDA	China
50	WA064	Ames29035	*Amaranthus tricolor* L.	USDA	India
51	WA068	Ames5311	*Amaranthus tricolor* L.	USDA	India
52	WA069	PI666331	*Amaranthus tricolor* L.	USDA	India
53	WA070	Ames26230	*Amaranthus tricolor* L.	USDA	China
54	WA071	Ames26229	*Amaranthus tricolor* L.	USDA	China
55	WA072	Ames26228	*Amaranthus tricolor* L.	USDA	China
56	WA073	Ames26226	*Amaranthus tricolor* L.	USDA	China
57	WA074	Ames26225	*Amaranthus tricolor* L.	USDA	China
58	WA075	Ames26223	*Amaranthus tricolor* L.	USDA	China
59	WA076	Ames26221	*Amaranthus tricolor* L.	USDA	China
60	WA077	Ames26220	*Amaranthus tricolor* L.	USDA	China
61	WA078	Ames26219	*Amaranthus tricolor* L.	USDA	China
62	WA079	Ames26218	*Amaranthus tricolor* L.	USDA	China
63	WA080	Ames26217	*Amaranthus tricolor* L.	USDA	China
64	WA081	Ames26216	*Amaranthus tricolor* L.	USDA	China
65	WA082	Ames26215	*Amaranthus tricolor* L.	USDA	China
66	WA083	Ames26214	*Amaranthus tricolor* L.	USDA	China
67	WA084	Ames26213	*Amaranthus tricolor* L.	USDA	China
68	WA085	Ames26212	*Amaranthus tricolor* L.	USDA	China
69	WA086	Ames26211	*Amaranthus tricolor* L.	USDA	China
70	WA087	Ames26210	*Amaranthus tricolor* L.	USDA	China
71	WA089	Ames26209	*Amaranthus tricolor* L.	USDA	China
72	WA090	Ames26232	*Amaranthus tricolor* L.	USDA	China
73	WA091	Ames26227	*Amaranthus tricolor* L.	USDA	China
74	WA092	Ames26222	*Amaranthus tricolor* L.	USDA	China
75	WA093	Ames26208	*Amaranthus tricolor* L.	USDA	China
76	WA094	PI619252	*Amaranthus tricolor* L.	USDA	India
77	WA099	PI603899	*Amaranthus tricolor* L.	USDA	US
78	WA101	NSL6100	*Amaranthus tricolor* L.	USDA	US
79	WA102	Ames5303	*Amaranthus tricolor* L.	USDA	US
80	WA112	PI669847	*Amaranthus tricolor* L.	USDA	India
81	WA116	Ames18049	*Amaranthus tricolor* L.	USDA	Nepal
82	WA117	PI566899	*Amaranthus tricolor* L.	USDA	India
83	WA118	PI527321	*Amaranthus tricolor* L.	USDA	China
84	WA119	PI478310	*Amaranthus tricolor* L.	USDA	China
85	WA120	PI462129	*Amaranthus tricolor* L.	USDA	India
86	WA121	PI462126	*Amaranthus tricolor* L.	USDA	India
87	WA123	PI277268	*Amaranthus tricolor* L.	USDA	India
88	WA124	PI214036	*Amaranthus tricolor* L.	USDA	India
89	WA125	Ames5354	*Amaranthus tricolor* L.	USDA	Madagascar
90	WA126	Ames5317	*Amaranthus tricolor* L.	USDA	Hong Kong
91	WA127	PI667171	*Amaranthus tricolor* L.	USDA	India
92	WA129	Ames5139	*Amaranthus tricolor* L.	USDA	US
93	WA130	PI674261	*Amaranthus tricolor* L.	USDA	India
94	WA131	Ames5134	*Amaranthus tricolor* L.	USDA	US
95	WA132	Ames5128	*Amaranthus tricolor* L.	USDA	US
96	WA133	Ames5126	*Amaranthus tricolor* L.	USDA	US

No.	Code	Name	Scientific name	Seed distributor	Country of collection
97	WA134	Ames5118	*Amaranthus tricolor* L.	USDA	Puerto Rico
98	WA135	Ames5117	*Amaranthus tricolor* L.	USDA	Puerto Rico
99	WA136	Ames5113	*Amaranthus tricolor* L.	USDA	Taiwan
100	WA137	Ames5111	*Amaranthus tricolor* L.	USDA	Papua New Guinea
101	WA138	Ames5110	*Amaranthus tricolor* L.	USDA	West Africa
102	WA139	Ames5101	*Amaranthus tricolor* L.	USDA	Hong Kong
103	WA140	Ames5100	*Amaranthus tricolor* L.	USDA	Hong Kong
104	WA141	Ames5099	*Amaranthus tricolor* L.	USDA	Hong Kong
105	WA142	Ames2230	*Amaranthus tricolor* L.	USDA	India
106	WA143	Ames2229	*Amaranthus tricolor* L.	USDA	India
107	WA144	Ames2228	*Amaranthus tricolor* L.	USDA	India
108	WA145	Ames2227	*Amaranthus tricolor* L.	USDA	India
109	WA146	Ames2226	*Amaranthus tricolor* L.	USDA	India
110	WA147	Ames2225	*Amaranthus tricolor* L.	USDA	India
111	WA148	Ames2224	*Amaranthus tricolor* L.	USDA	India
112	WA149	Ames2223	*Amaranthus tricolor* L.	USDA	India
113	WA150	Ames2222	*Amaranthus tricolor* L.	USDA	India
114	WA151	Ames2221	*Amaranthus tricolor* L.	USDA	India
115	WA152	Ames2209	*Amaranthus tricolor* L.	USDA	Hong Kong
116	WA153	PI674260	*Amaranthus tricolor* L.	USDA	Hong Kong
117	WA154	Ames2207	*Amaranthus tricolor* L.	USDA	Hong Kong
118	WA155	Ames2205	*Amaranthus tricolor* L.	USDA	Hong Kong
119	WA156	Ames2204	*Amaranthus tricolor* L.	USDA	Hong Kong
120	WA157	Ames2203	*Amaranthus tricolor* L.	USDA	Hong Kong
121	WA158	Ames2202	*Amaranthus tricolor* L.	USDA	Hong Kong
122	WA159	Ames2199	*Amaranthus tricolor* L.	USDA	Taiwan
123	WA160	Ames2198	*Amaranthus tricolor* L.	USDA	Hong Kong
124	WA161	Ames2197	*Amaranthus tricolor* L.	USDA	Hong Kong
125	WA162	Ames2196	*Amaranthus tricolor* L.	USDA	Hong Kong
126	WA163	Ames2149	*Amaranthus tricolor* L.	USDA	India
127	WA164	Ames2148	*Amaranthus tricolor* L.	USDA	India
128	WA165	Ames2147	*Amaranthus tricolor* L.	USDA	India
129	WA166	Ames2146	*Amaranthus tricolor* L.	USDA	India
130	WA167	Ames2145	*Amaranthus tricolor* L.	USDA	India
131	WA168	Ames2143	*Amaranthus tricolor* L.	USDA	India
132	WA169	Ames2142	*Amaranthus tricolor* L.	USDA	India
133	WA170	Ames2140	*Amaranthus tricolor* L.	USDA	India
134	WA171	Ames2139	*Amaranthus tricolor* L.	USDA	India
135	WA172	Ames2138	*Amaranthus tricolor* L.	USDA	India
136	WA173	Ames2135	*Amaranthus tricolor* L.	USDA	India
137	WA174	Ames2134	*Amaranthus tricolor* L.	USDA	India
138	WA175	Ames2132	*Amaranthus tricolor* L.	USDA	India
139	WA176	Ames2131	*Amaranthus tricolor* L.	USDA	India
140	WA177	Ames2130	*Amaranthus tricolor* L.	USDA	India
141	WA178	Ames2129	*Amaranthus tricolor* L.	USDA	India
142	WA179	Ames2128	*Amaranthus tricolor* L.	USDA	India
143	WA180	Ames2127	*Amaranthus tricolor* L.	USDA	India
144	WA181	Ames2126	*Amaranthus tricolor* L.	USDA	India
145	WA182	Ames2125	*Amaranthus tricolor* L.	USDA	India
146	WA183	Ames2124	*Amaranthus tricolor* L.	USDA	India

No.	Code	Name	Scientific name	Seed distributor	Country of collection
147	WA184	Ames2123	*Amaranthus tricolor* L.	USDA	India
148	WA185	Ames2122	*Amaranthus tricolor* L.	USDA	India
149	WA186	Ames2121	*Amaranthus tricolor* L.	USDA	India
150	WA187	Ames2120	*Amaranthus tricolor* L.	USDA	India
151	WA188	Ames2119	*Amaranthus tricolor* L.	USDA	India
152	WA189	Ames2118	*Amaranthus tricolor* L.	USDA	India
153	WA190	Ames2117	*Amaranthus tricolor* L.	USDA	India
154	WA191	Ames2116	*Amaranthus tricolor* L.	USDA	India
155	WA192	Ames2114	*Amaranthus tricolor* L.	USDA	India
156	WA193	Ames2113	*Amaranthus tricolor* L.	USDA	India
157	WA194	Ames2112	*Amaranthus tricolor* L.	USDA	India
158	WA195	Ames2110	*Amaranthus tricolor* L.	USDA	India
159	WA196	Ames2109	*Amaranthus tricolor* L.	USDA	India
160	WA197	Ames2108	*Amaranthus tricolor* L.	USDA	India
161	WA198	Ames2107	*Amaranthus tricolor* L.	USDA	India
162	WA199	Ames2106	*Amaranthus tricolor* L.	USDA	India
163	WA200	Ames2105	*Amaranthus tricolor* L.	USDA	India
164	WA201	Ames2104	*Amaranthus tricolor* L.	USDA	India
165	WA202	Ames2102	*Amaranthus tricolor* L.	USDA	India
166	WA203	Ames2101	*Amaranthus tricolor* L.	USDA	India
167	WA204	Ames2100	*Amaranthus tricolor* L.	USDA	India
168	WA206	Ames2040	*Amaranthus tricolor* L.	USDA	Malaysia
169	WA207	Ames2039	*Amaranthus tricolor* L.	USDA	Thailand
170	WA208	Ames2029	*Amaranthus tricolor* L.	USDA	China
171	WA209	Ames2024	*Amaranthus tricolor* L.	USDA	Taiwan
172	WA210	Ames2017	*Amaranthus tricolor* L.	USDA	China
173	WA211	Ames1998	*Amaranthus tricolor* L.	USDA	Taiwan
174	WA212	Ames1993	*Amaranthus tricolor* L.	USDA	Taiwan
175	WA213	Ames1983	*Amaranthus tricolor* L.	USDA	India
176	WA214	Ames1982	*Amaranthus tricolor* L.	USDA	India
177	WA215	Ames1980	*Amaranthus tricolor* L.	USDA	Zaire
178	WA217	PI337611	*Amaranthus hypochondriacus* L.	USDA	Uganda
179	WA221	Ames2141	*Amaranthus tricolor* L.	USDA	India
180	WA222	PI636187	*Amaranthus hypochondriacus* L.	USDA	India
181	WA223	Ames5689	*Amaranthus hypochondriacus* L.	USDA	Brazil
182	WA224	PI667174	*Amaranthus hypochondriacus* L.	USDA	Zimbabwe

USDA, United States Department of Agriculture National Plant Germplasm System; na, not available