Genetica (2017) 145:189–200
DOI 10.1007/s10709-017-9957-y

ORIGINAL PAPER

Genetic diversity and structure of the zombi pea (*Vigna vexillata* (L.) A. Rich) gene pool based on SSR marker analysis

Sujinna Dachapak[1] · Prakit Somta[1] · Supalak Poonchaivilaisak[2] · Tarika Yimram[1] · Peerasak Srinives[1]

Received: 9 May 2016 / Accepted: 6 February 2017 / Published online: 23 February 2017

Abstract Zombi pea (*Vigna vexillata* (L.) A. Rich) is an underutilized legume species and a useful gene source for resistance to biotic and abiotic stresses, although there is little understanding on its genetic diversity and structure. In this study, 422 (408 wild and 14 cultivated) accessions of zombi pea from diverse origins (201 from Africa, 126 from America, 85 from Australia, 5 from Asia and 5 from unknown origin) were analyzed with 20 simple sequence repeat (SSR) markers to determine its genetic diversity and genetic structure. The SSR markers detected 273 alleles in total with a mean of 13.6 alleles per locus. Polymorphism information content values of the markers varied from 0.58 to 0.90 with an average of 0.76. Overall gene diversity was 0.715. Gene diversity and average allelic richness was highest in Africa (0.749 and 8.08, respectively) and lowest in America (0.435 and 4.10, respectively). Nei's genetic distance analysis revealed that the highest distance was between wild Australia and cultivated Africa (0.559), followed by wild West Africa and wild Australia (0.415). STRUCTURE, neighbor-joining (NJ), and principal coordinate analyses consistently showed that these zombi pea accessions were clustered into three major groups, viz. America, Africa and Asia, and Australia. NJ tree also suggested that American and Australian accessions are originated from East African zombi peas, and that the cultivated accessions from Africa and Asia were genetically distinct, while those from America were clustered with some cultivated accessions from Africa. These results suggest that Africa is the center of origin and diversity of zombi pea, and that domestication of this pea took place more than once in different regions.

Keywords Zombi pea · Diversity · Domestication · Microsatellite marker · Underutilized legume

Introduction

The genus *Vigna* is an agro-sociological important plant taxon. It comprises more than 100 leguminous plant species distributed in wide and diverse areas of Africa, America, Australia and Asia (Lewis et al. 2004). Nine *Vigna* species are considered as domesticated crops, including cowpea (*V. unguiculata* (L.) Walps.), zombi pea (*V. vexillata* (L.) A. Rich), Bambara groundnut (*V. subterranean* (L.) Verdc.), mungbean (*V. radiata* (L.) Wilczek), azuki bean (*V. angularis* (Ohwi) Ohwi and Ohashi), rice bean (*V. umbellata* (Thunb.) Ohwi and Ohashi), black gram (*V. mungo* (L.) Hepper), moth bean (*V. aconitifolia* (Jaqc.) Maréchal), and créole bean (*V. reflexo-pilosa* Hayata) (Tomooka et al. 2002, 2011). These species are grown mainly for dry seeds by small farmers in several cropping systems of tropical and sub-tropical regions. In addition, some other *Vigna* species are cultivated and/or harvested for seeds and pods as food, and for leaves and stems as feedsby local people in Africa, Asia, Australia and America (Tomooka et al. 2011).

Electronic supplementary material The online version of this article (doi:10.1007/s10709-017-9957-y) contains supplementary material, which is available to authorized users.

✉ Prakit Somta
pksomta@gmail.com; agrpks@ku.ac.th

[1] Department of Agronomy, Faculty of Agriculture at Kamphaeng Saen, Kasetsart University, Nakhon Pathom 73140, Thailand

[2] Program in Plant Biotechnology, Faculty of Agriculture at Kamphaeng Saen, Kasetsart University, Nakhon Pathom 73140, Thailand

Zombi pea is an underutilized herbaceous legume belonging to the subgenus *Plectrotropis* of the genus *Vigna*. Wild forms of zombi pea occur widely in Africa, America and Australia and Asia, while the cultivated form occurs in a few localities of Africa and Asia (Tomooka et al. 2011). Wild zompi pea often develops storage roots which are edible and consumed by people in Africa, India, Australia and Southeast Asia (Duke 1981; Lawn and Cottrell 1988; Sasikumar and Sardana 1988; Karuniawan et al. 2006), although seeds are sometimes consumed. Edible tubers of *V. vexillata* contain a protein content of about 15% which is roughly three times higher than that of potato and yam and six times than that of cassava (Chandel et al. 1972). Moreover, yield of *V. vexillata* tuberous roots has been reported to be as high as 7000 kg/hectare which is comparable to sweet potato (Bhattacharyya et al. 1984). There are two types of cultivated zombi pea. One form is grown for both edible seeds and tuberous roots, while the other is grown solely for tuberous roots. The former is reported in Sudan (Ferguson 1954), while the latter is reported in Bali and Timor islands of Indonesia (Karuniawan et al. 2006) and India (Bhattacharyya et al. 1984; Asati and Yadav 2004).

Zombi pea is a very heterogeneous legume species. Eight taxonomic varieties including *macrosperma, vexillata, angustifolia, dolichomena, yunnanensis, plurifora, lobatifloria* and *ovata* have been described/recognized (Maréchal et al. 1978; Pasquet 2001; Pienaar and Kok 1991; Maxted et al. 2004). Variety *macrosperma* is the cultivated type, while the others are wild type. Vanderborght (1989) examined germination habits of zombi pea germplasm and found that accessions from America were epigeal whereas those from Africa (except Nigeria) and Australia were hypogeal. Due to its wide distribution, zombi pea has been reported to adapt well to environmental stress conditions such as infertile soil (Karuniawan et al. 2006), alkaline soil (Lawn and Watkinson 2002), acid soil (Vanderborght 1989), saline soil (Roecklein and Leung 1987), drought (Roecklein and Leung 1987), and waterlogging (Miller and Williams 1980). Zombi pea is also reported resistant to several insect pests including *Callosobruchus maculatus, Zabrotes subfasciatus, Maruca testulalis* and *Clavigralla tomentosicollis* (Birch et al. 1986; Jackai and Oghiakhe 1989) which are major pests of legume crops. In addition, this legume is resistant to viral diseases such as cowpea mottle carmovirus (CPMoV), cowpea aphid borne mosaic virus (CAbMV) and cowpea yellow mosaic virus (CYMV) (Gomathinayagam et al. 1998; Thottappillyet al. 1994). Thus, zombi pea is potentially a gene source in breeding for resistance to biotic and abiotic stresses, as well as a new future crop.

So far, there are a limited number of reports on genetic diversity of zombi pea. The previous genetic diversity analyses in zombi pea were conducted using molecular markers, viz. seed storage proteins (Piergiovanni 1998), isozymes (Garba and Pasquet 1998; Jaaska 2001) and RAPD markers (Spinosa et al. 1998). The results from these studies suggested that Africa is the center of diversity of zombi pea. Nonetheless, these studies have provided little information on the extent of genetic diversity and genetic structure of zombi pea because these studies used a small number (<150) of accessions and the molecular markers used were low informative.

Simple sequence repeat (SSR) or microsatellite is the marker of choice for molecular genetics study in crops because of its advantages as being co-dominant, multiallelic, reliable, PCR-based, and easy to score. SSR markers from one species can also be used in other related species. In *Vigna* species, SSR markers developed from azuki bean, cowpea, and mungbean were shown to have moderate to high rates of amplifcation in other *Vigna* species (Chaitieng et al. 2006; Gupta and Gopalakrishna 2010; Somta et al. 2009; Tangphatsornruang et al. 2009).

In this study, we report a genetic diversity analysis of 422 zombi pea accessions from various origins using SSR markers from azuki bean, cowpea and mungbean. The objectives of this study were to assess genetic diversity and genetic structure of zombi pea. The results will be useful for plant geneticists/breeders to better understand the diversity and domestication of this legume species.

Materials and methods

V. vexillata germplasm and DNA extraction

A total of 422 (408 wild and 14 cultivated) accessions of *V. vexillata* originating from various origins, including Africa (201 accessions), America (126 accessions), Australia (85 accessions), Asia (5 accessions) and unknown (5 accessions) were used in this study (Table 1 and supplementary Table S1). They were sown in an experimental field of Kasetsart University, Kamphaeng Saen campus, Nakhon Pathom, Thailand. In addition, two seeds of each accession were germinated on a germination tray using peat moss as the culture media, and their germination habit was recorded as hypogeal or epigeal.

Young leaves from one plant of each accession were collected and extracted for total genomic DNA. The DNA extraction was carried out using a CTAB method described by Lodhi et al. (1994). Quality and quantity of the DNA were measured on 0.8% agarose gel electrophoresis by comparing with a known concentration of lambda DNA.

SSR marker analysis

One thousand and twenty-four SSR markers from azuki bean (Wang et al. 2004; Chankaew et al. 2014), mungbean

Table 1 Source and number of *V. vexillata* accessions used in this study with number of alleles, gene diversity (H_E) and observed heterozygosity (H_O) and average allelic richness (*Ar*) calculated from allelic data of 20 SSR markers

Status	Origin	No. of accessions	No. of alleles	H_O	H_E	*Ar*
Wild	All	408	273	0.169	0.713	7.51
	Africa	193	258	0.207	0.749	8.08
	East	100	237	0.208	0.741	7.98
	West	21	121	0.229	0.672	5.62
	South	49	170	0.198	0.692	6.64
	Central	23	119	0.204	0.673	5.45
	America	123	136	0.083	0.435	4.10
	South	93	119	0.081	0.423	3.79
	Central	30	84	0.088	0.413	3.72
	Australia	85	150	0.102	0.563	5.07
	East	75	138	0.103	0.547	4.83
	West	7	81	nd	nd	nd
	North	3	39	nd	nd	nd
	Asia	2	37	nd	nd	nd
	Unknown	5	79	nd	nd	nd
Cultivated	All	14	97	0.143	0.641	4.85
	Africa	8	61	nd	nd	nd
	America	3	28	nd	nd	nd
	Asia	3	40	nd	nd	nd
Average		–	–	–	–	13.65
Overall		422	273	0.148	0.715	

Ar was calculated based on 14 individuals

nd not determined

(Tangphasornruang et al. 2009; Somta et al. 2009), and cowpea (Li et al. 2001; Kongjaimun et al. 2012) were screened for amplification and polymorphism in six accessions (CIAT4295, CIAT4525, CIAT4338, Aus-TRCF320518, JP235863, and TVNu1583) of *V. vexillata*. These accessions were randomly selected from different geographical regions. A PCR mixture was prepared following Somta et al. (2008) with minor modification. In brief, the PCR mixture of 10 µl contained 10 ng genomic DNA, 5 pmol of each forward and reverse primers, 1 × *Taq* buffer, 2 mM dNTPs, 1.5 mM $MgCl_2$, and 1 U *Taq* DNA polymerase (Thermo Scientific). The PCR amplification was carried out in a GeneAmp PCR System 9700 Thermal Cycler (Applied Biosystems, Foster City, Calif., USA) or PTC-200 Thermal Cycler (MJ Research, Waltham, MA, USA) using cycling profile as followed: 94 °C for 2 min, followed by 35 cycles of 94 °C for 30 s, 55 °C for 30 s, 72 °C for 1 min and the final extension at 72 °C for 10 min. Amplified products were separated on 5% denaturing polyacrylamide gel (w/v; 19:1 acrylamide-bisacrylamide) with 7 mol/L urea and 1× Tris-borate-EDTA buffer. Electrophoresis was run at 70 W constant power for 1.5–2 h (depending on allele size) using a pBR322 DNA/*Bsu*RI (*Hae*III) (Thermo Scientific) as the size standard. The amplified products were visualized by silver staining. Twenty markers (supplementary Table S2) showing polymorphic DNA bands were used to analyze DNA of all the 422 *V. vexillata* accessions.

Genetic data analysis

Allele data were used to calculate number of alleles per locus, observed heterozygosity, gene diversity (expected heterozygosity), and average allelic richness using FSTAT 2.9.3.2 software. The data were also used to calculate the polymorphism information content (PIC) (Anderson et al. 1993).

Genetic distance (D_A) (Nei et al.1983) for all pairs of accessions and for all pairs of geographical groups was calculated using POPULATIONS 1.2.32 software. The D_A for all pairs of accessions was then subjected to a neighbor-joining analysis using MEGA 6 software (Tamura et al. 2013) and to a principle coordinate analysis using package Factoextra in R-program 3.0.1 software (R Development Core Team 2013) to reveal genetic relationships among accessions. Also, D_A for all pairs of geographical groups comprising more than 7 accessions was calculated using POPULATIONS 1.2.32 software.

The genetic structure of the 422 *V. vexillata* accessions was determined by STRUCTURE 2.3.4 software that employs a model-based Bayesian Markov Chain Monte

Carlo (MCMC) algorithm. Initially, 20 simulation runs were carried out using a number of assumed populations (K) of 1 to 10 with a burn-in period of 10,000 and 50,000 replicates of MCMC. The true K was determined by the *ad-hoc* ΔK method (Evanno et al. 2005). Subsequently, a run with optimum K, burn-in period of 100,000 and 500,000 replicates of the MCMC algorithm was performed to assign each individual to a cluster.

Results

Variation in germinations of *V. vexillata* germplasm

Germination was epigeal in 186 and hypogeal in 236 *V. vexillata* accessions (supplementary Table S1). Wild accessions showed both germination types, while cultivated accessions showed only hypogeal germination. Wild accessions from Africa were very heterogeneous for the germination habit. In contrast, the accessions from America and Australia were each highly homogeneous for the germination habit. All except one of the American accessions were of epigeal type, while all except two of the Australian accessions were of hypogeal type. Only two Asian accessions were available in this study and showed one germination type each.

SSR variation in *V. vexillata*

Of the 1,024 SSR markers used to screen for polymorphism in 6 accessions of zombi pea, 595 were able to amplify these DNA, 217 showed polymorphism. Twenty SSR markers (supplementary Table S2) were chosen to analyze the 422 accessions of *V. vexillata* and detected 273 alleles in total (Tables 1, 2). All the alleles, except one, were found in the wild germplasm. The number of alleles detected per marker was between 6 (CEDG091, CEDG174) and 20 (CEDG107), with an average of 13.65 (Table 2). Among the wild accessions with known origins, those from Africa possessed the highest number of unique alleles (alleles found in one origin but not in the others; 73), followed by the germplasm from Australia (7) and America (6).

The polymorphism information content (PIC) values of the markers varied from 0.132 (CEDG174) to 0.915 (CEDG214) with an average of 0.704 (Table 2). Among the wild accessions, average PIC value was highest in those from Africa but not different from those from Australia (Table 2), while in the cultivated germplasm, average PIC value was 0.601 (Table 2).

Observed heterozygosity (H_O) over all accessions was 0.148. H_O values between wild and cultivated accessions were similar, being 0.169 and 0.143, respectively. In the wild germplasm, H_O was highest in African accessions,

Table 2 Linkage groups (LG), number of detected alleles per locus, allele size range and polymorphic information content (*PIC*) of the 20 polymorphic SSR markers used in diversity analysis

Primer	LG	No. of alleles	Allele size range (bp)	Wild					Cultivated
				Overall	Africa	America	Australia	Unknown	Overall
CEDG011	4	14	140–170	0.704	0.775	0.549	0.576	0.320	0.603
CEDG015	6	18	200–230	0.860	0.844	0.419	0.778	0.780	0.640
CEDG043	3	9	170–200	0.296	0.306	0.246	0.504	0.480	0.457
CEDG091	4	6	200–230	0.641	0.625	0.653	0.589	0.640	0.612
CEDG098	11	19	150–170	0.796	0.879	0.271	0.706	0.780	0.768
CEDG102	1	8	150–170	0.600	0.780	0.255	0.018	0.660	0.674
CEDG107	4	20	130–150	0.831	0.917	0.146	0.686	0.320	0.676
CEDG147	10	18	240–260	0.863	0.865	0.752	0.821	0.700	0.712
CEDG174	7	6	180–200	0.132	0.171	0.000	0.298	0.780	0.357
CEDG181	4	19	170–200	0.900	0.920	0.546	0.875	0.740	0.706
CEDG214	1	17	200–240	0.915	0.932	0.736	0.869	0.740	0.636
CEDG244	2	15	160–180	0.628	0.668	0.729	0.439	0.560	0.681
CEDG248	6	12	130–150	0.625	0.749	0.295	0.589	0.480	0.489
CEDG264	5	17	180–200	0.907	0.919	0.866	0.829	0.720	0.574
CEDG304	9	12	120–150	0.711	0.728	0.103	0.736	0.820	0.460
VES0675	–	16	260–320	0.685	0.882	0.738	0.813	0.700	0.595
VES0476	1	9	200–300	0.786	0.683	0.074	0.488	0.640	0.520
VM24	8	11	220–260	0.801	0.788	0.335	0.719	0.660	0.564
VM27	9	15	170–200	0.699	0.787	0.362	0.821	0.700	0.584
VR304	12	12	180–240	0.704	0.724	0.577	0.655	0.700	0.710
Average		13.65		0.704	0.747	0.433	0.640	0.646	0.601
Overall		273							

followed by Australian and American accessions, respectively (Table 1).

Gene diversity (H_E) among all accessions was 0.72. H_E values in the wild and cultivated accessions were comparable. Among the wild germplasm, H_E was highest in the accession from Africa (0.75) and lowest from America (0.44) (Table 1). Wild germplasm from Africa showed highest average allelic richness (*Ar*), followed by Australia and America (Table 1). Of the African wild germplasm, East African germplasm had highest H_E and *Ar*, although not much different from the other African regions.

Neighbor-joining analysis

A phylogenetic tree reconstructed from neighbor-joining (NJ) analysis using D_A showed that, in general, accessions from the same geographical region were clustered together showing three major clusters (I, II and III) (Fig. 1). Cluster I was the largest cluster comprising mainly African accessions, all Asian accessions, 15 Australian accessions and 8 American accessions. All cultivated accessions were included in this cluster. Cluster II comprised principally accessions from America together with 5, 1, 1, and 2 accessions from East Africa, Central Africa, Australia, and unknown origin, respectively. Cluster III comprised most accessions from Australia and six accessions from East Africa.

The 14 cultivated accessions were separated into three sub-clusters within the cluster I (Fig. 1). Sub-cluster I comprised only accessions from Bali island of Indonesia, subcluster II included only accessions from Sudan, and subcluster III contained all accessions from America and one accession each from Sudan, Botswana and Central African Republic. Interestingly, the accessions from Bali were clustered with wild accessions from various origins.

In the wild African germplasm which is highly heterogeneous for germination habits, germplasms with the same germination type generally clustered together (supplementary Fig. S1).

Principal coordinate analysis

PCoA analysis based on D_A revealed that the first three PCs together accounted for 37.3% of the total variation, with PC1, PC2 and PC3 explained 18.2, 11.2 and 7.9, respectively. A scatter plot of the zombi pea germplasm based on PC1 and PC2 showed that accessions from Africa, America and Australia were largely differentiated, except for some accessions from America and Australia which were associated with those from Africa. Accessions from America and Australia were clearly distinct, while accessions from Asia mingled with those from Africa.

Wild accessions from East, West, Central and South Africa were not clearly differentiated. Among them, wild accessions from East Africa showed the widest distribution, whereas wild accessions from Central and South Africa showed a narrow distribution and were not differentiated. Wild accessions from Australia were clustered into two subgroups; one comprised accessions from East, North and West Australia, and the other included only accessions from East and West Australia. The latter group showed close genetic relationships with accessions from Africa. Wild accessions from Asia showed clear genetic differentiation, although there were only two accessions from this region, they showed close genetic relationships with accessions from Africa. All 14 cultivated accessions in this study showed a wide distribution and a clear distinction from each other. Even the wild African accessions showing the same germination type by NJ analysis were not grouped together by PCoA analysis.

Population structure analysis

Bayesian clustering of the zombi pea germplasm was performed using STRUCTURE software. Based on Evanno's *ad hoc* ΔK method (Evanno et al. 2005), there were three subpopulations among the 422 zombi pea accessions (Fig. 3). Subpopulation I was the largest subpopulation with 231 accessions. This subpopulation comprised almost all accessions from wild Africa (193 accessions), some wild accessions from America and Australia, all accessions from Asia and all cultivated accessions. Subpopulation II comprised 122 accessions; all except four (three from Africa and one from Australia) accessions are from America. Subpopulation III comprised 70 accessions, all from Australia. In most cases, the results of germplasm clustering using NJ, PCoA and STRUCTURE analyses were congruent.

Genetic differentiation among geographical groups

D_A analysis among groups of the wild zombie pea germplasm (Table 3) revealed that their genetic distance was low (0.151 for South Africa vs. East Africa) to relatively high (0.415 for wild West Africa vs. wild Australia). Among the African germplasm, East African germplasm showed the closest genetic relationship with South African germplasm, while South African germplasm showed the widest genetic relationship with Central African germplasm. Wild American germplasm showed the closest genetic relationship with East and West African germplasm, suggesting that the American germplasm was genetically originated from East or West Africa or both regions. Wild Australian germplasm showed the closest genetic relationship with East African germplasm, and is likely to be genetically derived from East Africa.The cultivated germplasm from Africa showed

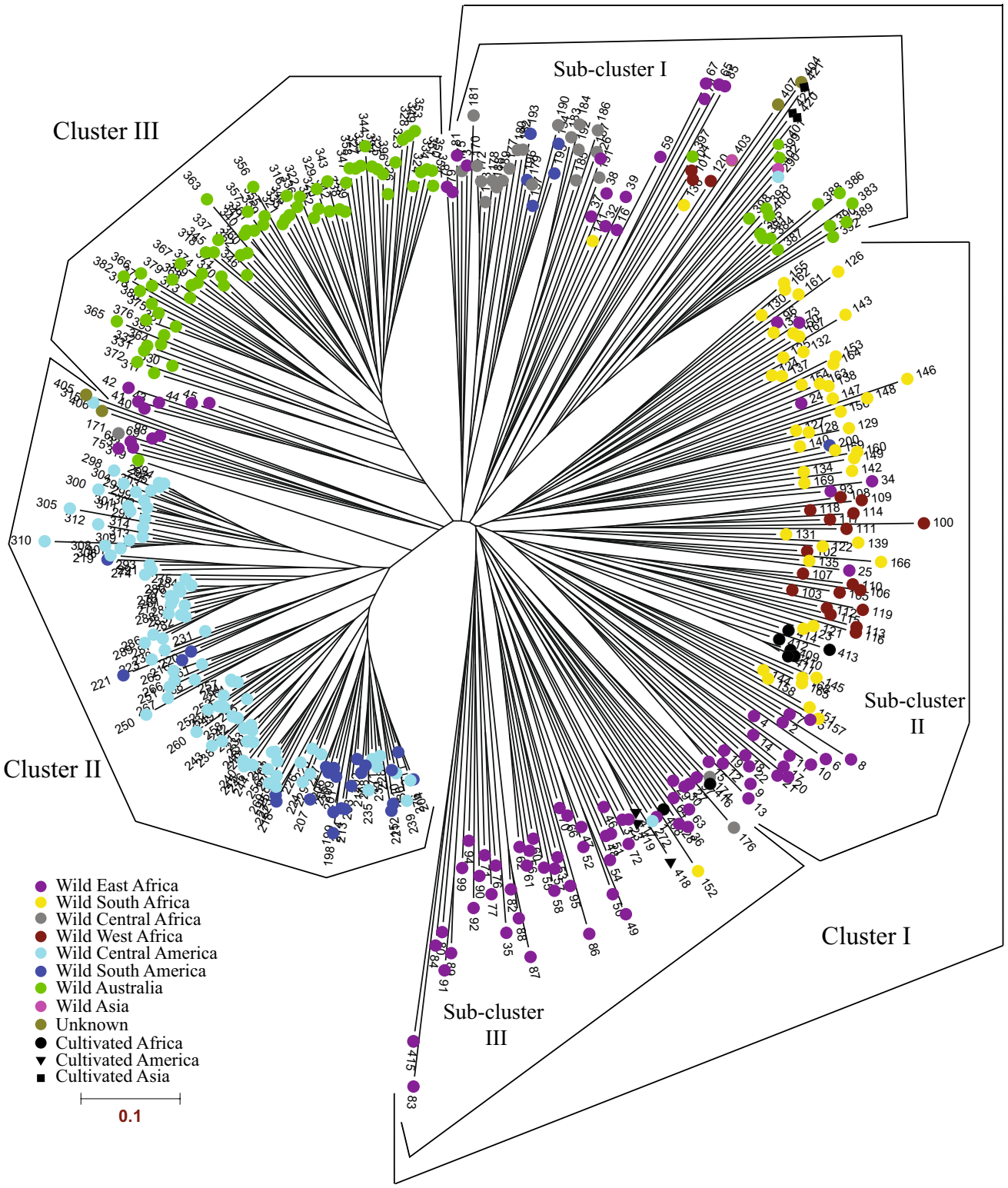

Fig. 1 A Neighbor-joining tree constructed from Nei's genetic distance (D_A) showing genetic relationship among the 422 accessions of *Vigna vexillata*. D_A was computed from allelic data of 20 simple sequence repeat markers

Table 3 Pairwise Nei's genetic distance (D_A) among the six geographical groups of *V. vexillata* based on 20 SSR markers

	Wild East Africa	Wild West Africa	Wild South Africa	Wild Central Africa	Wild America	Wild Australia
Wild West Africa	0.244					
Wild South Africa	0.151	0.252				
Wild Central Africa	0.211	0.265	0.308			
Wild America	0.271	0.305	0.325	0.306		
Wild Australia	0.272	0.415	0.346	0.378	0.410	
Cultivated Africa	0.351	0.366	0.373	0.399	0.372	0.559

closest genetic relationship with wild germplasm from East Africa ($D_A = 0.351$) and highest differentiation with the wild germplasm from Australia.

Discussion

Association between germination habits and geographical origins in zombi pea

Both epigeal and hypogeal germinations have been reported in zombi pea (Vanderborght 1989) and these germination habits are associated with geographical origins (Spinosa et al. (1998). Similar results were found in this study where zombi pea germplasm from Africa showed both germination types, while most Australian accessions had epigeal germination and American accessions had hypogeal germination (supplementary Table S1). This implies that the zombi pea germplasm from Africa possesses higher genetic diversity than the germplasm from America and Australia, and the genetic structure of *V. vexillata* from America and Australia is associated with germination habits.

Center of origin, diversity and domestication of *V. vexillata*

Maréchal et al. (1978) studied the morphological diversity of *V. vexillata* and proposed that this legume originated in Africa. In this study, SSR allelic data clearly revealed that wild *V. vexillata* germplasm from Africa had the greatest gene diversity (Table 1). Our results support the proposal of Maréchal et al. (1978). *V. vexillata* accessions from East Africa showed higher gene diversity and broader distribution on the PCoA scatter plot than the other germplasm of African origins (Fig. 2), suggesting that East Africa is more likely the center of origin and center of diversity of *V. vexillata*.

There are two types of cultivated *V. vexillata*; African and Asian (Bali and East Timor in Indonesia) types. The African type shows erect and determinate growth habits, photoperiod-insensitive, early flowering, and hairy pods, while the Bali type shows viny and indeterminate growth habits, photoperiod-sensitive, late flowering, and glabrous pods. As compared to the African type, the Asian type has larger seed size and longer seed development duration. Ferguson (1954) noted that *V. vexillata* is a unique crop of Sudan where it is called "Babun" and grown only on a small scale in one locality for edible root and seeds. Karuniawan et al. (2006) reported that *V. vexillata* is grown in some localities of Bali and Timor Islands of Indonesia for edible root. Damayanti et al. (2010) studied variation of agronomic and morphological traits in some *V. vexillata* germplasm, including wild form and African and Bali cultivated forms, and found that the Bali form was distinct from the other forms, although small number of accessions were studied. Recently, Takahashi et al. (2016) determined phylogenetic relationship of many *Vigna* accessions including 9 accessions of *V. vexillata* using DNA sequences of nuclear DNA-ITS and chloroplast *atpB-rbcL* spacer regions and found that the Bali cultivated accessionsare genetically different from the other accessions. These suggested that Bali cultivated *V. vexillata* has been domesticated independently. In this study, cultivated *V. vexillata* accessions from Africa, America and Asia (Bali) were used and similar results were obtained. Neighbor-joining tree (Fig. 2) and PCoA scatter plot (Fig. 3) separated cultivated accessions into three different groups, although they were all clustered with wild African accessions. The Bali accessions were in a clear separated group. Several accessions from Sudan were grouped together, while one accession from this country was grouped with all accessions from America and accessions from Botswana and Central African Republic. Based on the information gathered by Damayanti et al. (2010) the group of the Sudanese cultivated accessions was traced back to the accession named 'Babun'. Nonetheless, these results suggested that cultivated *V. vexillata* were domesticated twice and independently; once in Africa and once in Asia. The results supported the proposal of Garba and Pasquet (1998) based on herbarium specimens, morphological and isozyme variation that domestication of *V. vexillata* occurred independently in Africa and Asia. The clear grouping of cultivated accessions from Sudan (Fig. 1)

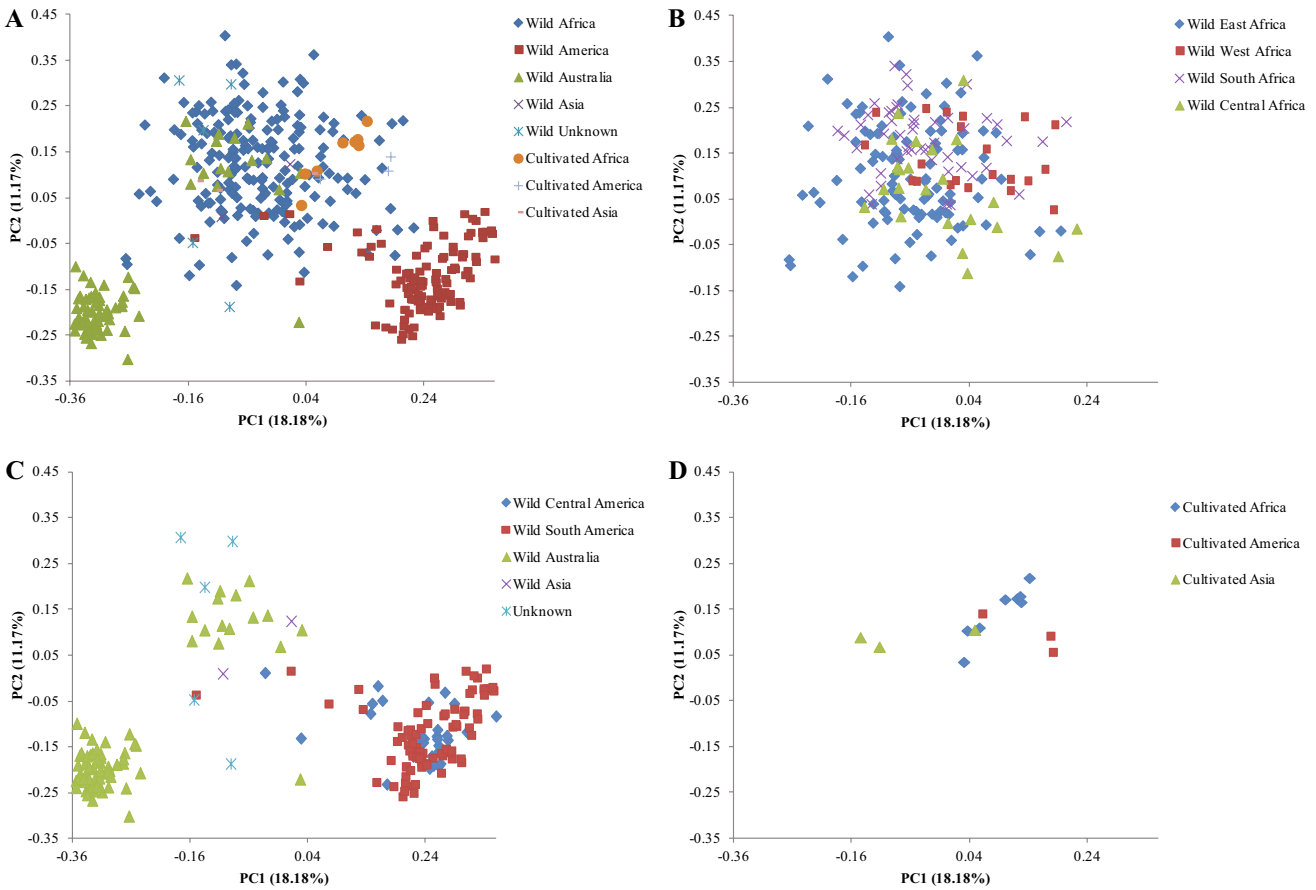

Fig. 2 A scatter-plot of PC1 and PC2 from principal coordinate analysis based on Nei's genetic distance (D_A) showing genetic relationship among the 422 accessions of *Vigna vexillata*. D_A was calculated from allelic data of 20 simple sequence repeat markers

supported the note of Ferguson (1954) that *V. vexillata* was first cultivated in Sudan. These together indicated that Sudan is one of the places of *V. vexillata* domestication in Africa and that the American cultivated accessions originated from the African cultivated accession, possibly from Sudan. This supported the report of Damayanti (2010) who traced the origins of American cultivated *V. vexillata* (also used in our study) and found that they were originally from Sudan.

In the case of the cultivated Bali accessions, they were genetically close to wild accessions from India (Aus-TRCF66514), Papua New Guinea (CIAT4944) and the Philippines (TVNu1632) (Fig. 1) which all have hypogeal germination. We found that hybridizing the accession Aus-TRCF66514 as the female parent by Bali cultivated accession (JP235863) as the male parent, resulted in partial self-fertile F_1 plants, despite a low number of hand-pollinated flowers (Dachapak and Somta, unpublished data). This suggested that the Bali cultivated accessions constitute the primary or secondary gene pool of *V. vexillata* and they are domesticated from wild *V. vexillata* of India. This is possibly due to the fact that (1) *V. vexillata* is cultivated in India (Bhattacharyya et al. 1984; Asati and Yadav 2004), (2) plant material similar to Bali cultivated *V. vexillata* exists in Sri Lanka (Damayanti 2000; personal communication with Dr. Rémy Pasquet by Dr. Robert Lawn), and (3) there was a past cultural and social (Hinduism) relationship between South Asia and Indonesia, especially Bali. Additional diversity study using more wild and cultivated germplasm of *V. vexillata* from Asia in particular from India and Indonesia is necessary to elucidate the domestication site of Bali cultivated *V. vexillata*.

Genetic structure of zombi pea

As many as 8 taxonomic varieties have been described for *V. vexillata* due to the high morphological variation of this species (Maxted et al. 2004). Accessions characterized as variety *angustifloria, macrosperma, vexillata*, and *ovata* were included in this study. NJ tree and PCoA plot showed that accessions of the same varieties were not always clustered together (Figs. 2, 3; see also supplementary Table S1). In stead, the genetic structure of this species is geographical dependent. Similar finding was reported

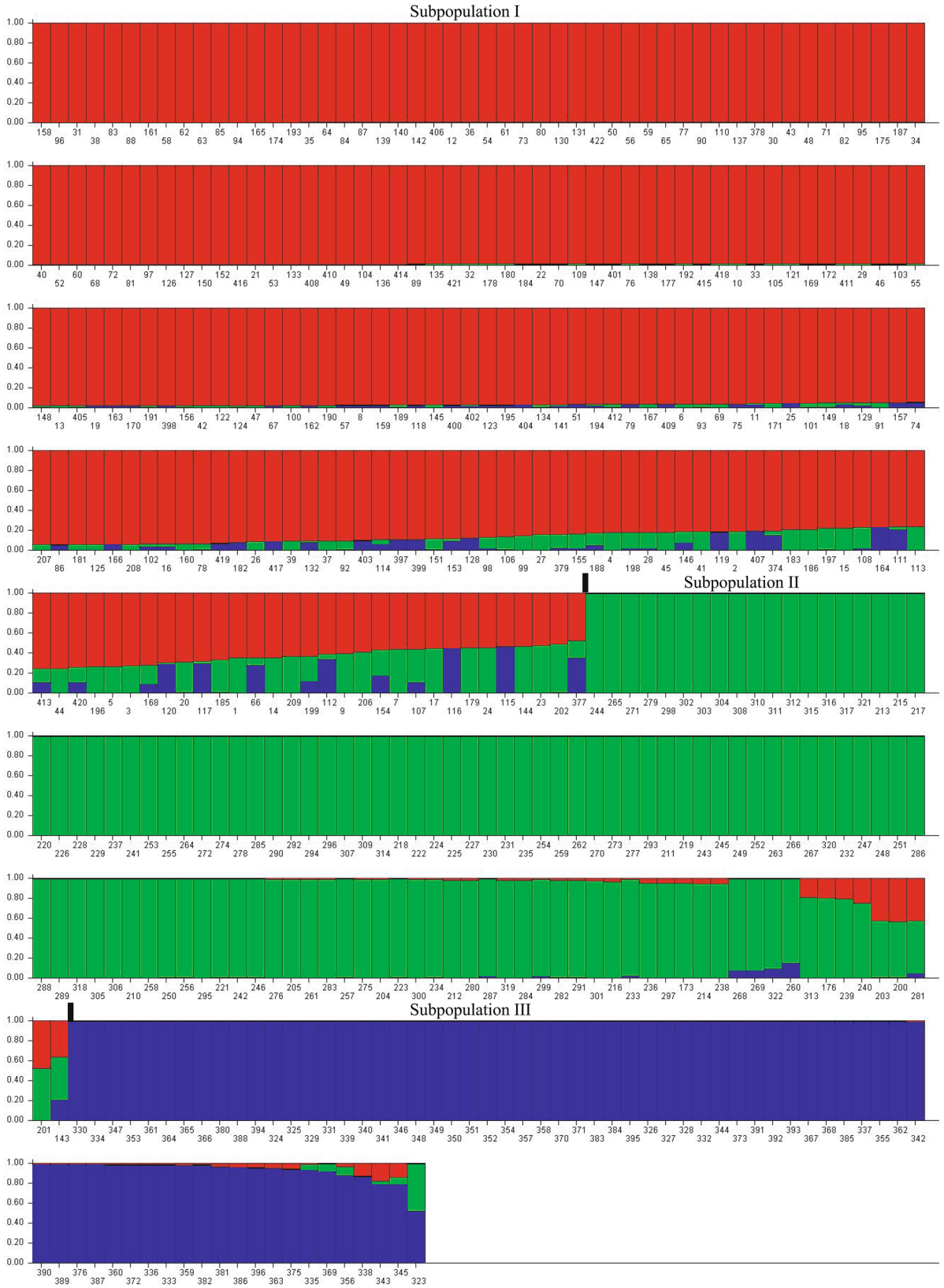

Fig. 3 Genetic structure of the 422 accessions of *Vigna vexillata* determined by STRUCTURE analysis based on simple sequence repeat allelic data at 20 loci. *Each bar* represents one accession

Table 4 Comparison of number of alleles per locus, gene diversity (H_E), outcrossing rate (t), PIC value and average allelic richness (Ar) between *Vigna vexillata* and other *Vigna* species

Species	No. of accessions	SSR loci used	No. of alleles (mean per locus)	H_E	*PIC*	*Ar*	References
V. vexillata	422	20	273 (13.65)	0.72	0.71	13.65	This study
V. angularis	616	13	311 (23.92)	0.76	0.74	16.09	Xu et al. (2008)
V. radiata	615	19	309 (16.26)	0.62	0.59	4.82	Sangiri et al. (2007)
V. mungo	534	22	199 (9.05)	0.60	0.60	9.01	Kaewwongwal et al. (2015)
V. umbellata	472	13	168 (12.92)	0.61	0.57	13.10	Tian et al. (2013)

earlier by Spinosa et al. (1998) who used RAPD markers to analyze *V. vexillata* accessions of var. *angustifloria, macrosperma* and *vexillata* and found no relationship between botanical varieties and geographical origins.

Damayanti et al. (2010) showed that the Bali cultivated *V. vexillata* is not cross-compatible with the wild *V. vexillata* or African cultivated *V. vexillata*. Based on this result the authors concluded that the Bali cultivated accessions are not in the primary gene pool of *V. vexillata*. As mentioned above, hybridization between a Bali cultivated accession (JP235863) and an Indian wild accession (Aus-TRCF66514) resulted in partially self-fertile F_1 plants. However, it is worth noting that when we conducted this hybridization (November 2014 to February 2015) the Bali accession showed very low (<3%) pod setting, despite setting a lot of flowers. However when the same accession was grown in the following year (November 2015 to February 2016) it set more of pods (~15%) (Sujinna and Somta, unpublished data). Environmental conditions appear to influence fertility of the Bali cultivated *V. vexillata* and possibly its offspring. Additional hybridization study is necessary to determine genepool of the Bali cultivated *V. vexillata*.

The previous diversity studies in *V. vexillata* using isozyme markers (Garba and Pasquet 1998; Jaaska 2001) and RAPD markers (Spinosa et al. 1998) showed that the African and American accessions are genetically different. The same result is obtained in this study using SSR markers (Figs. 1, 2, 3). By incorporating morphological variation with isozyme marker data, Pasquet (2000) proposed that *V. vexillata* from America and Australia were introduced from West and East Africa, respectively. Similarly, Jaaska (2001) proposed that American germplasm was derived from the Western Gondwanaland *V. vexillata* population by vicariance through a continental drift, and after that evolutionary change in the American germplasm was mostly at intraspecific level. These proposals are supported by our results which demonstrated that *V. vexillata* from America is genetically similar to that from East Africa (Figs. 1, 2; Table 3), thus suggesting that the American germplasm was derived from East African *V. vexillata*. Nonetheless, our results (Figs. 1, 2; Table 3) supported the proposal that the Australian accessions originated from East Africa *V. vexillata*. It is worth noting that the number of West African *V. vexillata* used in our study was only 21 accessions and different result may be obtained when the sample size from this region is increased.

By using isozyme marker analysis, Garba and Pasquet (1997, 1998) illustrated that *V. vexillata* accessions from Asia are genetically different from those from Africa and suggested that Asian *V. vexillata* can be considered a distinct subspecies. These are in contrast to our findings where accessions from Asia were clustered with those from Africa and showed closer genetic relationship than those from Australia and America (Figs. 2, 3). Thus our findings suggested that Asian *V. vexillata* should not be raised as a new subspecies, although not many accessions of Asian germplasm was used in this study. Again, further study using a wider range of germplasm from Asia is required to determine its genepool and taxonomic rank.

Comparison of genetic diversity in zombi pea with other *Vigna* species

Previous genetic studies in *Vigna* species using SSR markers on a large number of germplasm accessions included azuki bean (Xu et al. 2008), black gram (Kaewwongwal et al. 2015), mungbean (Sangiri et al. 2007), and rice bean (Tian et al. 2013). We compared our results in zombi pea with these *Vigna* species (Table 4). Average number of alleles per locus in zombi pea (13.65) was lower than azuki bean (23.92) and mungbean (16.26), but higher than black gram (9.05) and rice bean (12.92). Gene diversity of zombi pea (0.72) was close to azuki bean (0.76), and greater than those of mungbean, blackgram and rice bean. Average allelic richness of zombi pea (13.65) was higher than mungbean (4.82) and blackgram (9.01), but almost the same as rice bean (13.10) and lower than azuki bean (16.09). The high diversity found in zombi pea is likely due to wide geographical distribution of the germplasm used in this study.

Acknowledgements This research was supported by the Royal Golden Jubilee Ph.D. Scholarship and the TRF Research Career

Development Grant (RSA5880051) both of which co-funded by the Thailand Research Fund (TRF) and Kasetsart University to P. Somta and S. Dachapak.

References

Anderson JA, Churchill GA, Autrique JE, Tanksley SD, Sorrells ME (1993) Optimizing parental selection for genetic linkage maps. Genome 36(1):181–186

Asati BS, Yadav DS (2004) Diversity of horticultural crops in north eastern region. ENVIS Bull Himal Eco 12:1–11

Bhattacharyya PK, Ghosh AK, Sanyal B, Deb Ray G (1984). Grow *Vigna vexillata* for protein-rich tuber-cum-pulse crop in North-eastern hill region Seeds Farms 10:33–36

Birch ANE, Fellow LE, Evans SV, Doherty K (1986) *Para*-Aminophenylalanine in *Vigna* : possible taxonomic and ecological significance as a seed defence against bruchids. Phytochemistry 25(12):2745–2749

Chaitieng B, Kaga A, Tomooka N, Isemura T, Kuroda Y, Vaughan DA (2006) Development of a black gram [*Vigna mungo* (L.) Hepper] linkage map and its comparison with an azuki bean [*Vigna angularis* (Willd.) Ohwi and Ohashi] linkage map. Theor App Genet 113:1261–1269

Chandel KPS, Arora RK, Joshi BS (1972) *Vigna capensis* Walp. (*V. vexillata*) an edible root legume. Curr Sci 41:537

Chankaew S, Isemura T, Isobe S, Kaga A, Tomooka N, Somta P, Hirakawa H, Shirasawa K, Vaughan DA, Srinives P (2014) Detection of genome donor species of neglected tetraploid crop *Vigna reflexo-pilosa* (créole bean) and genetic structure of diploid species based on newly developed EST-SSR markers from azuki bean (*Vigna angularis*). PLoS One. doi:10.1371/journal.pone.0104990

Damayanti F (2010) Comparativestudies on cultivated and wild accessions of *Vigna vexillata* (L.) A. Rich. Masters (Research) Thesis, James Cook University

Damayanti F, Lawn RJ, Bielig LM (2010) Genotypic variation in domesticated and wild accessions of the tropical tuberous legume *Vigna vexillata* (L.) A. Rich. Crop Pasture Sci 61(10):771–784

Duke JA (1981) *Vigna unguiculata* (L.) Walp. spp. *unguiculata*. In: Okeson ON (ed) Legumes of world economic importance. Plenum Press, New York, pp 303–305

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

Ferguson H (1954) The food crops of the Sudan and their relation to environment. In: A paper presented at the 1953 Conference of the Philosophical Society of the Sudan on "Foodand Society in the Sudan", Philosophical Society of the Sudan Khartoum, McCorquodaleand Co. (Sudan), Ltd

Garba M, Pasquet RS (1997) Isozyme diversity in *Vigna vexillata* (L.) A. Rich (Fabaceae) complex. S Afr J Bot 64(3):163–175

Garba M, Pasquet RS (1998) The *Vigna vexillata* (L.) A. Rich. genepool. In: Sorensen M, Estrella JE, Hammann OJ, Rios Ruiz SA (eds) 2nd International symposium on tuberous legume. 5–8 August 1996. Celaya, Mexico, pp 61–71

Gomathinayagam P, Ganesh Ram S, Rathnaswamy R, Ramaswamy NM (1998) Interspecific hybridization between *Vigna unguiculata* (L.) Walp. and *V. vexillata* (L.) A. Rich. through *in vitro* embryo culture. Euphytica 102:203–209

Gupta SK, Gopalakrishna T (2010) Development of unigene-derived SSR markers in cowpea (*Vigna unguiculata*) and their transferability to other *Vigna* species. Genome 53(7):508–523

Jaaska V (2001) Isoenzyme diversity and phylogenetic relationships among the American beans of the genus *Vigna* Savi (Fabaceae). Biochem Syst Ecol 29:1153–1173

Jackai LEN, Oghiakhe S (1989) Pod wall trichomes and resistance of two wild cowpea, *Vigna vexillata*, accessions to *Maruca testualis* (Geyer) (Lepidoptera: Pyralidae) and *Clavigralla tomentosicollis* Stal (Hemiptera: Coreidae). B Entomol Res 79:595–605

Kaewwongwal A, Kongjaimun A, Somta P, Chankaew S, Yimram T, Srinives P (2015) Genetic diversity of the black gram [*Vigna mungo* (L.)] gene pool as revealed by SSR markers. Breed Sci 65:127–137

Karuniawan A, Iswandi A, Kale PR, Heinzemann J, Grüneberg WJ (2006) *Vigna vexillata* (L.) A. Rich. cultivated as a root crop in Bali and Timor. Genet Resour Crop Evol 53(1):213–217

Kongjaimun A, Kaga A, Tomooka N, Somta P, Shimizu T, Shu Y, Isemura T, Vaughan DA, Srinives P (2012) An SSR-based linkage map of yardlongbean (*Vigna unguiculata* (L.) Walp. subsp. *unguiculata* Sesquipedalis group) and QTL analysis of pod length. Genome 55(2):81–92

Lawn RJ, Cottrell A (1988) Wild mungbean and its relatives in Australia. Biologist (London) 35:267–273

Lawn RJ, Watkinson AR (2002) Habitats, morphological diversity and distribution of the genus *Vigna* Savi in Australia. Aust J Agri Res 53:1305–1316

Lewis GP, Shrine B, Mackinder B, Lock JM (2004) Legumes of the world. Royal Botanical Garden, Kew, London

Li CD, Fatokun CA, Ubi B, Singh BB, Scoles GJ (2001) Determining genetic similarities and relationships among cowpea breeding lines and cultivars by microsatellite markers. Crop Sci 41(1):189–197

Lodhi MA, Ye GN, Weeden NF, Reisch BI (1994) A simple and efficient method for DNA extraction from grapevine cultivars and *Vitis* species. Plant MolBiol Rep 12:6–13

Maréchal R, Mascherpa JM, Stainier F (1978) Etude taxonomique d'un grouped'especes des genres *Phaseolus* et *Vigna* (Papilonaceae) sur la base des donneesmorphologiques et polliques, traitees pour l'analyseinformatique. Boissiera 28:1–273

Maxted N, Mabuza-Diamini P, Moss H, Padulosi S, Jarvis A, Guarino L (2004) An ecogeographic study: African *Vigna*, systematic and ecogeographic studies of crop genepool 10. International Plant Genetic Resources Institutes, Rome, Italy

Miller IL, Williams WT (1980) Tolerance of some tropical legume to six months of simulated waterlogging. Trop Grasslands 15:39–43

Nei M, Tajima F, Tateno Y (1983) Accuracy of estimated phylogenetic trees from molecular data. J MolE 19:153–170

Pasquet RS (2000) Allozyme diversity of cultivated cowpea *Vigna unguiculata* (L.) Walp. Theor Appl Genet 101:211–219

Pasquet RS (2001) *Vigna* Savi. In: Mackinder B, Pasquet R, Polhill R, Verdcourt B (eds) Flora Zambesiaca, volume part Phaseoleae. Royal Botanic Gardens Kew, London, pp 121–156

Pienaar BJ, Kok PDF (1991) The *Vigna vexillata* complex (Fabaceae) in Southern Africa. S Afri J Bot 57:236–245

Piergiovanni AR (1998) *Vigna vexillata* (L.) A. Rich. seed protein: heterogeneity in subunits of globulin fraction. Genet Resour Crop Evol 45:97–103

R Development Core Team (2013) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna

Roecklein JC, Leung P (1987) A profile of economic plants. Transaction Publishers, USA

Sangiri C, Kaga A, Tomooka N, Vaughan DA, Srinives P (2007) Genetic diversity of the mungbean (*Vigna radiata*, Leguminosae) genepool on the basis of microsatellite analysis. Aus J Bot 55:837–847

Sasikumar B, Sardana S (1988) *Vigna vexillata* (Fabaceae), A pulse cum tuber crop of northeastern hill region of India. Econ Bot 42:292

Somta P, Mucsh W, Kongsamai B, Chanprame S, Nakasathien S, Toojinda T, Sorajjapinu W, Seehalak W, Tragoonrung S, Srinives P (2008) New microsatellite markers isolated from mungbean (*Vigna radiata* (L.) Wilczek). Mol Ecol Res 8(5):1155–1157

Somta P, Seehalak W, Srinives P (2009) Development, characterization and cross-species amplification of mungbean (*Vigna radiata*) genic microsatellite markers. Conserv Genet 10:1939–1943

Spinosa A, Pignone D, Sonnante G (1998) Assessment of genetic variation in a working collection of *Vigna vexillata* (L.) A. Rich. by isozyme and RAPD analyses. Genet Resour Crop Evol 45:347–354

Takahashi Y, Somta P, Muto C, Iseki K, Naito K, Pandiyan M, Natesan S, Tomooka N (2016) Novel genetic resources in the genus *Vigna* unveiled from gene bank accessions. PLoS One 11(1):e0147568. doi:10.1371/journal.pone.0147568

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: molecular evolutionary genetics analysis version 6.0. MolBiolE 30:2725–2729

Tangphatsornruang S, Somta P, Uthaipaisanwong P, Chanprasert J, Sangsrakru D, Seehalak W, Sommanas W, Tragoonrung S, Srinives P (2009) Characterization of microsatellites and gene contents from genome shotgun sequences of mungbean (*Vigna radiata* (L.) Wilczek). BMC Plant Biol 9:137. doi:10.1186/1471-2229-9-137

Thottappilly G, Ng NQ, Russel HW (1994) Screening germplasm of *Vigna vexillata* for resistance to cowpea mottle virus. Int J Trop Plant Dis 12:75–80

Tian J, Isemura T, Kaga A, Vaughan DA, Tomooka N (2013) Genetic diversity of the rice bean (*Vigna umbellata*) genepool as assessed by SSR markers. Genome 56:717–727

Tomooka N, Vaughan DA, Moss H, Maxted N (2002) The Asian *Vigna*: genus *Vigna* subgenus *Ceratotropis* genetic resources. Kluwer Academic Publishers, Dordrecht

Tomooka N, Kaga A, Isemura T, Vaughan DA, Srinives P, Somta P, Thadavong S, Bounphanousay C, Kanyavong K, Inthapanya P, Pandiyan M, Senthil N, Ramamoorthi N, Jaiwal PK, Tian J, Umezawa K, Yokoyama T (2011) *Vigna* genetic resources. In The 14th NIAS International Workshop on Genetic Resources: Genetic Resources and Comparative Genomics of Legumes (*Glycine* and *Vigna*). 14 September 2009, National Institute of Agrobiological Sciences, Japan, pp 11–21

Vanderborght T (1989) Some observations on seedlings of *Vigna vexillata* (L.) A. Rich. (Fabaceae). Bull Jard Bot Nat Belg 59:179–187

Wang XW, Kaga A, Tomooka N, Vaughan DA (2004) The development of SSR markers by a new method in plants and their application to gene flow studies in azuki bean [*Vigna angularis* (Willd.) Ohwi & Ohashi]. Theor Appl Genet 109(2):352–360

Xu HX, Jing T, Tomooka N, Kaga A, Isemura T, Vaughan DA (2008) Genetic diversity of the azuki bean (*Vigna angularis* (Willd.) Ohwi & Ohashi) gene pool as assessed by SSR markers. Genome 51(9):728–738