Genet Resour Crop Evol (2019) 66:1813–1824
https://doi.org/10.1007/s10722-019-00825-2

RESEARCH ARTICLE

Genetic diversity and population structure of enset (*Ensete ventricosum* Welw Cheesman) landraces of Gurage zone, Ethiopia

Fetta Negash Gerura · Beira Hailu Meressa · Kyallo Martina · Abush Tesfaye · Temesgen Magule Olango · Yao Nasser

Received: 25 January 2019 / Accepted: 23 August 2019 / Published online: 14 September 2019

Abstract Enset (*Ensete ventricosum* (Welw.) Cheesman), which feeds around 20 million Ethiopian people, is a unique crop; with all parts of the plant are utilizable. It is, arguably, less researched crop and the mode of production remained conventional. Understanding the extent of genetic diversity in the crop, especially making use of genotyping data, is a very important first step in the genetic improvement of the crop. Twelve polymorphic enset SSR markers were used to assess the genetic diversity and population structure of 79 cultivated landraces and four wild enset individuals collected from different enset growing locations of Ethiopia. The polymorphic information content of markers ranged from 0.62 to 0.77 with a mean value of 0.69. A total of 77 alleles were identified, and the average observed heterozygosity varied from 0.51 to 0.67. A mean gene diversity of 0.59 was recorded ranging from 0.55 to 0.62. The AMOVA revealed that within population allelic variations contributed more to the genetic diversity than among population variations. Discriminant Analysis of Principal Components and population structure analysis grouped the 83 enset germplasms into three major clusters, where the wild individuals clustered distinctly. Outcomes of this research provide valuable information for enset conservation and breeding strategies especially for development of resistance for bacterial wilt and nematode attacks.

Keywords Polymorphic information · SSR markers · Genetic distance · Gene diversity · Discriminant analysis

F. N. Gerura (✉) · B. H. Meressa
Plant Disease Diagnostics Lab, Jimma University, P. O. Box 307, Jimma, Ethiopia
e-mail: fettanegash@yahoo.com

K. Martina · Y. Nasser
Biosciences Eastern and Central Africa-International Livestock Research Institute (BecA-ILRI) Hub, P.O. Box 30709-00100, Nairobi, Kenya

A. Tesfaye
International Institute of Tropical Agriculture, PMB 5320, Oyo road, Ibadan 200001, Nigeria

T. M. Olango
School of Plant and Horticulture, Hawassa University, Hawassa, Ethiopia

Introduction

Enset (*Ensete ventricosum* (Welw.) Cheesman) is perennial herb plant having a close resemblance with banana, but is distinguished by its giant size, dilated bases and the edible parts are the starchy corm and pseudo-stem (Brandt et al. 1997). It is a diploid species with chromosome number of 2n = 18, and belongs to the order Zingiberales, family *Musaceae* and genus *Ensete* (Tomlinson 1969). About 20 species are recognized within the genus Ensete, but *Ensete*

ventricosum is the only economically important species (Baker and Simmonds 1953; Taye 1993). Entirely, enset is vegetatively propagated through production of suckers (Yemataw et al. 2018) although it produce seeds.

Enset is a food security crop in which every part of the plant is utilizable (Bizuayehu 2002; Brandt et al. 1997). It is drought tolerant and can be harvested throughout the year and stored for long without spoilage (Brandt et al. 1997). More than 20% (nearly 18 million) of Ethiopian population depends on enset for food, fiber, animal forage, construction materials and medicines (Brandt et al. 1997; Zippel 2002; Central Statistical Authority 2010). According to Central Statistics Authority (CSA) 2010 report of Ethiopia, a total of 395,632 hectare of land is covered by the crop and yielded 8015,531 quintals.

However, enset production is declining in terms of productivity and area coverage and facing genetic erosion because of severe drought, diseases and population pressure (Tsegaye and Struik 2002; Brandt et al. 1997; Tenaye and Geta 2009; Almaz et al. 2002). Bacterial wilt disease caused by *Xanthomonas campestris* Pv. musacearum is the most threatening problem to enset production system (Bobosha 2003). Root lesion nematode (*Pratylenchus goodeyi*) and root knot nematodes (*Meloidogyne* sp.) are also commonly attacking the crop (Peregrine and Bridge 1992). Due to low research attention, the crop is the least studied and no effective technology is delivered to the farming community, and hence, the mode of production remained conventional (Bobosha 2003).

Genetic diversity study of the available enset germplasm would help to improve the crop in subsequent breeding programs through exploiting the genetic variation in the landraces, and effective genetic resources conservation. Most of the enset genetic diversity studies, so far, mainly focused on morphological markers (Bizuayehu 2002; Olango et al. 2014; Fetta 2007; Worku 1996). There are also few molecular studies to assess genetic diversity of enset landraces in Ethiopia using AFLP (Negash et al. 2002), RAPD (Birmeta et al. 2002), ISSR (Tobiaw and Bekele 2011) and SSR (Olango et al. 2015; Getachew et al. 2014) markers. However, these molecular diversity studies did not include enset landraces which are found in Gurage zone. Gurage district, with diverse agro-ecology and agricultural production systems. Gurage is among the high potential enset producing areas in Ethiopia, where it is rich in landrace diversity (Fetta 2007; Worku 1996). Therefore, the study was carried out with the objectives of describing the genetic diversity of Enset landraces of Gurage zone using simple sequence repeat (SSR) markers that will allow the search for bacterial wilt and nematode resistance and other enset improvement researches.

Materials and methods

Plant materials and total genomic DNA extraction

A total of 83 leaf samples (79 cultivated and 4 wild enset germplasms) were collected from Wolkite University (WkU) and Areka Agricultural Research Center (AARC) enset germplasm conservation sites (Table 1). The cultivated germplasms included: cultivated landraces originally collected from Gurage (exhaustively gathered), Yem Area (included for comparison) and released varieties from AARC (also included for comparison). The wild enset individuals were originally collected from forests of Dawro and Gurage zones and are not edible (Table 1). The leaf tissues were kept in 38% NaCl-CTAB solution with a ratio of 3.0 g CTAB and 35 g NaCl dissolved in 100 ml of distilled water (Storchova et al. 2000). After the NaCl-CTAB solution was thoroughly washed off with double distilled water at BecA-ILRI hub laboratory, isolation of total genomic DNA was done using Zymo Quick-DNA^{TM} Plant/Seed Minprep Kit (The Epigenetics COMPANY^{TM}, USA). The quality and concentration of extracted DNA were estimated using the NanoDrop 2000c Spectrophotometer (Thermo Fisher Scientific, Walthum, MA) and visualized and photographed under gel documentation system (Azure Biosystems, c280 CA, USA) after fragmented at 5 V/cm in 0.8% agarose gel. DNA was stored at − 20 °C until further use.

SSR analysis

A total of nineteen SSR markers, thirteen of which developed from enset (Olango et al. 2015), and six cross-transferred from banana (Lagoda et al. 1998) and previously used for enset genetic diversity study (Getachew et al. 2014), were used in the study (Table 2). The SSR markers were tested for polymorphism level. The 5′ end of the forward primers of all

Table 1 Enset plant materials used for genetic diversity analysis

S. no.	Population and no. of germplasm	Landraces	S. no.	Population and no. of germplasm	Landraces	S. no.	Population and no. of germplasm	Landraces
1	East *Gurage* (16)	*Ageremariam*	29		*Ankefuye*	57		*Kibnar*
2		*Andawutiye*	30		*Astara*	58		*Lemat*
3		*Arka*	31		*Ayidas*	59		*Nechiwe*
4		*Ashakit*	32		*Aytquaqof*	60		*Oret*
5		*Aywegne*	33		*Badedet*	61		*Sapara*
6		*Chama*	34		*Bazeriye*	62		*Shebe-shertiye*
7		*Gefetano*	35		*Bekuret*	63		*Shebirat*
8		*Kekere*	36		*Benet*	64		*Shertiye*
9		*Keribote*	37		*Benezhe*	65		*Tedrader*
10		*Kibo*	38		*Bosere*	66	West *Gurage*	*Tegaded*
11		*Kose*	39		*Buangye-agade*	67		*Teriye*
12		*Sebara*	40		*Chehuyet*	68		*Tetteret*
13		*Semay-leged*	41		*Deree*	69		*Tikurina*
14		*Siraro*	42		*Derewetiye*	70		*Yedemert*
15		*Tikur-badedet*	43	West *Gurage* (51)	*Ehire*	71		*Yeqyeswe*
16		*Zigbot*	44		*Emirye*	72		*Yeshraqinqe*
17	Released varieties (8)	*Endale 02*	45		*Eniba*	73		*Yezire-badedet*
18		*Gewada 01*	46		*Fereziye*	74		*Yiregiye*
19		*Gewada 02*	47		*Gaznar*	75		*Zober*
20		*Kelisa 04*	48		*Genbenye*	76	Wild (4)	*Erpha 88*
21		*Messena 01*	49		*Gimbwe*	77		*Erpha 92*
22		*Yambule 01*	50		*Guadye*	78		*Erpha 93*
23		*Yambule 02*	51		*Guarye*	79		*Furinzir*
24		*Zereta 01*	52		*Gubira*	80		*Acharqa*
25	West Gurage (51)	*Adiya-ehire*	53		*Kanchiwe*	81	*Yem* (4)	*Argemo*
26		*Agade*	54		*Kekari*	82		*Gesero*
27		*Amerat*	55		*Kembatiye*	83		Lobo
28		*Anichara*	56		*Kese-kembat*			

*West and East Gurage and Yem enset landraces and one wild individual were from Wolkite University preservation site

*Released varieties and three of wild accessions were obtained from Areka Agricultural Research Center preservation site

*Three of the wild individuals were collected from Dawro while one from Gurage zones of Southern Ethiopia

the SSR loci were labeled with fluorescent dyes (6-FAM = blue, PET = red, VIC = green, and NED = yellow).

SSR genotyping

The amplification of each SSR loci was done using a standard PCR (Eppendorf Mastercycler nexus GSX1) to determine the appropriate annealing temperature. A four-primer multiplex PCR was designed based on their expected amplicon size and dye colour. The PCR program was composed of a volume of 15 µl containing 3 µl of genomic DNA (15 ng/µl), an average of 0.5 µl of each primer (2 pm/µl) (Sigma-Aldrich), 7.5 µM AccuPower PCR PreMix without dye (Bioneer, Daejeon, Republic of Korea) and additional of

Table 2 SSR markers used in the study with their repeat motif and size

ID	Marker name	Forward primer sequence (5′–3′)	Reverse primer sequence (5′–3′)	Repeat motif	Size range (bp)	References
1	Evg-01	AGTCATTGTGCGCAGTTTCC	GGAGGACTCCATGTGGATGAG	(CTT)8	100–120	Olango et al. (2015)
2	Evg-02	GGAGAAGCATTTGAAGGTTCTTG	TTCGCATTTATCCCTGGCAC	(AG)12	118–153	
3	Evg-03	ACAGCATAAGCGAAATAGCAG	ACAGCATAAGCGAAATAGCAG	(AG)12	107–123	
4	Evg-04	GCCATCGAGAGCTAAGGGG	GGCAAGGCCGTAAGATCAAC	(AG)21	113–147	
5	Evg-05	AGTTGTCACCAATTGCACCG	CCATCCTCCACACATGCC	(GA)22	103–141	
6	Evg-06	CCGAAGTGCAACACCAGAG	TCGCTTTGCTCAACATCACC	(CA)17	138–148	
7	Evg-07	GGTTGTCCTCAAGAACGTGG	TGATGCCTAATGCCTCTCCC	(GTG)9	73–94	
8	Evg-08	CCATCGACGCCTTAACAGAG	TGAACCTCGGGAGTGACATAAG	(GA)21	164–190	
9	Evg-09	GCCTTTCGTATGCTTGGTGG	ACGTTGTTGCCGACATTCTG	(GA)13	141–175	
10	Evg-10	CAGCCTGTGCAGCTAATCAC	CAGCAGTTGCAGATCGTGTC	(AG)21	191–210	
11	Evg-11	GGCCTAGTGACATGATGGTG	TGATGCTAGATTCAAAGTCAAGG	(AC)13	135–160	
12	Evg-12	TGCAACCCTTTGCTGCATTC	AGCATCATTCGCCATGGTTG	(TG)14	135–154	
13	Evg-17	GCGTCTGGTATGCTCAACTG	TCGGGAATGATACAGAGGCG	(TCA)8	111–154	
14	pMaCIR 431	AATTGAATTTTAGGTGCCCCTACTG	AACATGGAGTGAAGATGAGACGATT	(GA)10	100–150	Getachew et al. (2014)
15	mMa-1-6	CAAGAACCCAACCGGTCAC	TTGTCATCACCATCCGTCATT	(TC)10	250–258	
16	mMa-3-90	GCACGAAGAGGCATCAC	GGCCAAATTTGATGGACT	(CA)6	168–170	
17	mMaCIR 24	ATCTTTTCTTATCCTTCTAACG	ATTAGATCACCGAAGAACT	(TC)7	242–262	
18	pMaCIR 232	GTTTACTGGCACTCTGTCCTC	TGCGATCAGAAGACCAAACAT	(TC)11	123	
19	pMaCIR 332b	CAGTTTGGCCGCTTGATCTT	GGGGTCAACATGTTAAGTTCT	(TG)4TTT(CT)9	333	

0.3 mM $MgCl_2$ (Promega, Venlo, The Netherlands). The PCR amplification program consisted of an initial denaturing step of 94 °C for 3 min; followed by 30 cycles of 94 °C for 30 s, 57 °C annealing temperature for 1 min, 72 °C for 2 min; and a final extension step of 72 °C for 20 min, using a GeneAmp® PCR System 9700 thermocycler (Applied Biosystems, Foster City, CA). The PCR products were fragmented at 7.1 V/cm in 1.5% agarose gel (in 0.5% TBE stained with 2.5 µl GelRed/100 ml of buffer) and finally the gel was visualized and photographed under gel documentation system (Azure Biosystems, c280). Samples were prepared for capillary electrophoresis in LIZ-HiDi cocktail solution in a proportion of 1 ml of Hi-Di™ Formamide (Applied Biosystems®, Thermo Fisher Scientific, Waltham, MA, USA) and a 15 µl of GeneScan_500 LIZ® Size standard (Applied Biosystems®, Thermo Fisher Scientific, Waltham, MA, USA). Multiplex PCR products with a volume of 1.5 µl were diluted to 9 µl of LIZ-HiDi solution. The mixed products were denatured at 95 °C for 2 min in a GeneAmp® PCR System 9700 (Applied Biosystems) thermocycler and snap-chilled on ice for 5 min and sent to Segolip laboratory of ILRI-BecA/Kenya for capillary electrophoresis. The generated data were then analyzed manually using the GeneMapper® Software version 4.1 (Applied Biosystems®, Thermo Fisher Scientific, Waltham, MA, USA) and the allele size was scored in base pairs (bp) based on the relative migration of the internal size standard.

SSR polymorphism and genetic diversity analysis

Polymorphic SSR markers were used to analyze the genetic diversity of the 83 studied enset germplasms. The basic statistics, such as observed number of alleles (Na), number of effective alleles (Ne), number of private alleles (Npa), unbiased expected heterozygosity (uHe), observed heterozygosity (Ho), expected heterozygosity (He) and polymorphic information content (PIC) were computed using GenAlEx 6.503 (Peakall and Smouse 2012). GenAlEx 6.503 was also used to estimate genetic differentiation using Analysis of Molecular Variance (AMOVA) to partition total genetic variation into within and among population subgroups. Gene flow (Nm) that describes the average number of individuals per generation migrating between populations will be calculated by Nm = {(1/FST) − 1}/4 using GenAlEx and it will be taken as the major determinant of population structure when Nm > 0.5 where FST is the standardized genetic variance among (sub)populations in the total sample (Wright 1951). Because the sample size taken from the studied populations was different, rarefied allelic richness (Ra) and rarefied private allelic richness (Rpa) per population were estimated using rarefaction procedure implemented in the program HP-Rare 1.1 (Kalinowski 2005) to clearly describe allelic diversity. To investigate population differentiation, fixation index (F_{ST}) of the total populations and pairwise F_{ST} among all pairs of populations were computed, and significance was tested based on 1000 bootstraps. To estimate the discriminatory power of the SSR markers, the PIC for each marker was computed by PIC = $1-\sum Pi^2$, where Pi^2 referred to the sum of the ith allelic frequency of each microsatellite locus for the genotypes (Nei 1987). Discriminant Analysis of Principal Components (DAPC) were implemented in the adegenet package version 1.4.1 in R (Jombart 2008). Clustering was performed using neighbor joining method with arithmetic average (Nei 1972) using Darwin V6.0 (Perrier et al. 2003) to generate a phylogenetic tree with branch robustness tested using 1000 bootstraps. The pattern of population structure and detection of admixture were inferred using a Bayesian model-based clustering algorithm implemented in STRUCTURE version 2.3.4 (Pritchard et al. 2000), Structure Harvester (Earl and vonHoldt 2012), CLUster Matching and Permutation Program (CLUMPP) version 1.1.2 (Jakobsson and Rosenberg 2007) and online application of Structure Plot v2 (Ramasamy et al. 2014).

Results

Genetic diversity

Out of the total 19 SSR tested markers, seven-two from enset (Evg-03 and Evg-05) and five from banana (pMaCIR 431, mMa-3-90, mMaCIR 24, and pMaCIR 232) were found to be monomorphic. Hence, the remaining twelve (Table 3) were selected to evaluate allelic diversity of the selected 83 enset germplasms. All primers used yielded an average Polymorphic Information Content (PIC) value of 0.69 ranging from 0.62 to 0.77 (Table 3). The selected twelve polymorphic SSR markers detected a total of 77 alleles (58

Table 3 The 12 SSR markers used in the study with number of alleles, size range and PIC values

Locus	No. of alleles	Size range (bp)	PIC
Evg-01	6	102–119	0.63
Evg-02	8	118–146	0.72
Evg-04	11	121–166	0.73
Evg-06	4	199–209	0.67
Evg-07	5	072–094	0.67
Evg-08	12	154–184	0.77
Evg-09	6	150–172	0.76
Evg-10	6	191–209	0.70
Evg-11	3	153–163	0.62
Evg-12	6	131–147	0.72
Evg-17	5	136–153	0.68
mMa-1-6	5	124–150	0.65

being different) in the studied 83 enset germplasms. The mean allelic richness per locus was 6.4, ranging from 3 (Evg-11) to 12 (Evg-08) alleles (Table 4).

The genetic diversity indices for the five studied enset populations are summarized in Table 4. The West Gurage (WG) population had the highest number of different allele (Na), number of effective allele (Ne) and number of private allele (NPa) values. Similarly, expected heterozygosity was highest in WG and Wild (W) populations. Relatively higher Shannon's Index (I) was observed in both East and West Gurage (I = 1.17) populations while it was least in Yem (I = 0.93) population. The percentage polymorphic loci (PL) ranged from 91.7% (RV, W and Y) to 100% (EG and WG), with a mean value of 95%. The highest fixation index (F) value of 0.17 was recorded for Wild Accession population; while negative values for the same parameter (− 0.02 to − 0.14) was identified for the cultivated populations. However in situations with populations under investigation possess unequal size, rarefied allelic richness and rarefied private allelic richness are better to clearly describe and compare the genetic diversity between the studied populations. Accordingly, Wild enset accessions exhibited a higher rarefied allelic richness (3.44) and rarefied private allelic richness (0.74) than all the rest studied enset populations.

Population genetic structure and gene flow

Analysis of molecular variance (AMOVA) showed 93% total variation due to observed variation within

Table 4 Diversity parameters estimated for enset populations using 12 SSR markers

Populations	Diversity parameters									
	Na	Ra	Ne	I	Ho	He	uHe	NPa	Rpa	F	PL
East Gurage (N = 16)	4.58 ± 0.34	2.89 ± 0.32	2.89 ± 0.32	1.17 ± 0.09	0.64	0.61 ± 0.04	0.63	0.17 ± 0.11	0.21 ± 0.13	− 0.02	100
Released variety (N = 8)	3.58 ± 0.36	2.85 ± 0.34	2.72 ± 0.29	1.03 ± 0.12	0.64	0.57 ± 0.06	0.61	0.25 ± 0.13	0.41 ± 0.11	− 0.14	91.7
West Gurage (N = 49)	5.17 ± 0.53	2.82 ± 0.35	2.88 ± 0.27	1.17 ± 0.10	0.67	0.62 ± 0.04	0.62	0.58 ± 0.19	0.17 ± 0.15	− 0.06	100
Wild (N = 4)	3.58 ± 0.26	3.44 ± 0.28	3.04 ± 0.24	1.13 ± 0.11	0.51	0.62 ± 0.06	0.72	0.58 ± 0.15	0.74 ± 0.12	0.17	91.7
Yem (N = 4)	2.92 ± 0.26	2.83 ± 0.27	2.50 ± 0.22	0.93 ± 0.11	0.59	0.55 ± 0.06	0.64	0.08 ± 0.08	0.25 ± 0.09	− 0.09	91.7

Na No. Alleles, *Ra* rarefied allelic richness, *Ne* no. effective alleles, *I* information index, *Ho* observed heterozygosity, *He* expected heterozygosity, *uHe* unbiased expected heterozygosity, *Npa* no. private alleles, *Rpa* rarefied private alleles and *F* fixation index, *PL* percentage of polymorphic loci

Table 5 AMOVA among and within populations of enset populations based on the 12 SSR markers

Source	Df	SS	MS	Est. var.	Variation (%)	*P* value
Among populations	4	61.185	15.296	0.605	7	< 0.001
Within populations	78	633.3.9	8.119	8.119	93	< 0.001
Total	82	694.494		8.725	100	

Table 6 Nei's genetic distance (below diagonal) and Nei's genetic identity (above diagonal), among five population groups

	EG	W	RV	WG	Y
EG		0.826	0.896	0.997	0.910
W	0.191		0.868	0.858	0.989
RV	0.109	0.142		0.868	0.868
WG	0.003	0.153	0.142		0.949
Y	0.094	0.012	0.142	0.052	

EG East *Gurage*, *W* wild, *RV* released varieties, *WG* West *Gurage*, *Y Yem*

enset individual landraces. Only 4% of the total molecular variance was ascribed to the five population groups and 3% due to genetic variations among accessions in each population group (Table 5). The genetic differentiation among populations (F_{ST} = 0.037 at $P < 0.001$) was highly significant. The average estimate of the gene flow among populations (Nm) was 3.35, indicating high gene flow between populations.

The average genetic distances and similarities between the studied five enset populations is presented in Table 6. The greatest genetic distance (0.191) was observed between East Gurage and Wild Accession populations, and between West Gurage and Wild Accession populations (0.153). On the other hand, the least genetic distance (0.003) was detected between genotypes from West and East Gurage populations. The mean genetic distance among the studied germplasm was 0.17, ranging between 0.00 and 0.33. The highest genetic distance of 0.33 and 0.32 was observed between *Erpha* 88 (wild) and *Anichara* (cultivated) and *Furinzir* (wild) and *Yezire-badedet* (cultivated) germplasms, respectively. The least genetic distance (0.00) was observed between different cultivated enset genotypes. Cultivated enset landraces that did not differ in their SSR profile for the tested primers (with 0.00 genetic distance) include: (1) *Ashakit* and *Shertiye*; (2) *Agade* and *Buangye-agade*; (3) *Adiya-ehire*, *Bekuret* and *Chehuyet*; (4) *Benet*, *Gaznar* and *Sapara*; (5) *Genbenye* and *Shebe-shertiye*; (6) *Kekere* and *Yedemert*; (7) *Yeshiraqinqe* and *Gubira* and (8) *Arka* and *Chama* (Fig. 1). Genetic identity and distance between the studied enset germplasms are also clearly displayed in minimum spanning network (Fig. 2).

Unweighted neighbor-joining dendrogram (Fig. 1) clustered the studied 83 enset germplasms into three definite populations. Discriminant Analysis of Principal Components (DAPC) showed that the studied five enset populations were also grouped into three inferred clusters (K = 3) (Fig. 3a). The first cluster placed all the wild enset accessions distinctly away from the four cultivated enset landraces; while the second cluster included most of the released varieties, and the third cluster contained all the cultivated enset landraces and few of the released varieties (Fig. 3b). No distinct cluster pattern was observed within the cultivated landraces taken from different geographical locations.

Discussion

A polymorphic information content (PIC) value ≥ 0.5 is considered highly informative and useful in distinguishing degree of polymorphism (Prabakaran et al. 2010). The PIC values detected by the 12 SSR markers were in the range of 0.62–0.77; making them highly informative. Genetic distance is described by the observed total number of independent alleles (Kalinowski 2002). The 12 SSR markers detected a total of 77 alleles in the studied enset germplasm. This shows the recorded number of alleles is high compared to 61 alleles identified from 220 enset germplasm, using 11 Musa markers (Getachew et al. 2014). This number of alleles might be considered low relative to the 202 alleles detected from 60 cultivated enset landraces and six wild germplasm, using 34 enset SSR markers

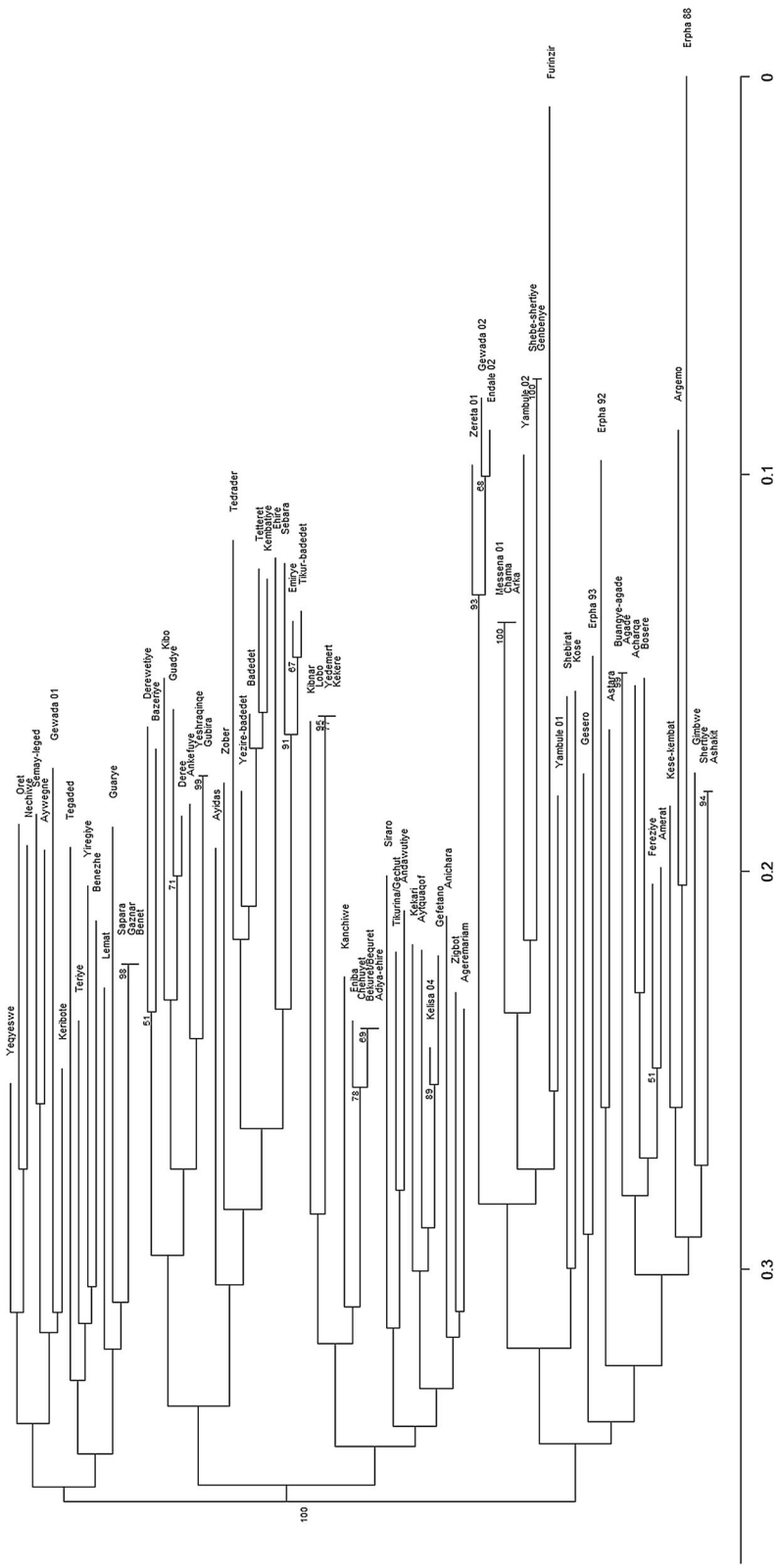

Fig. 1 Unweighted neighbor-joining phylogeny Bruvo distance tree based on 12 SSR markers for the 83 studied landraces of enset

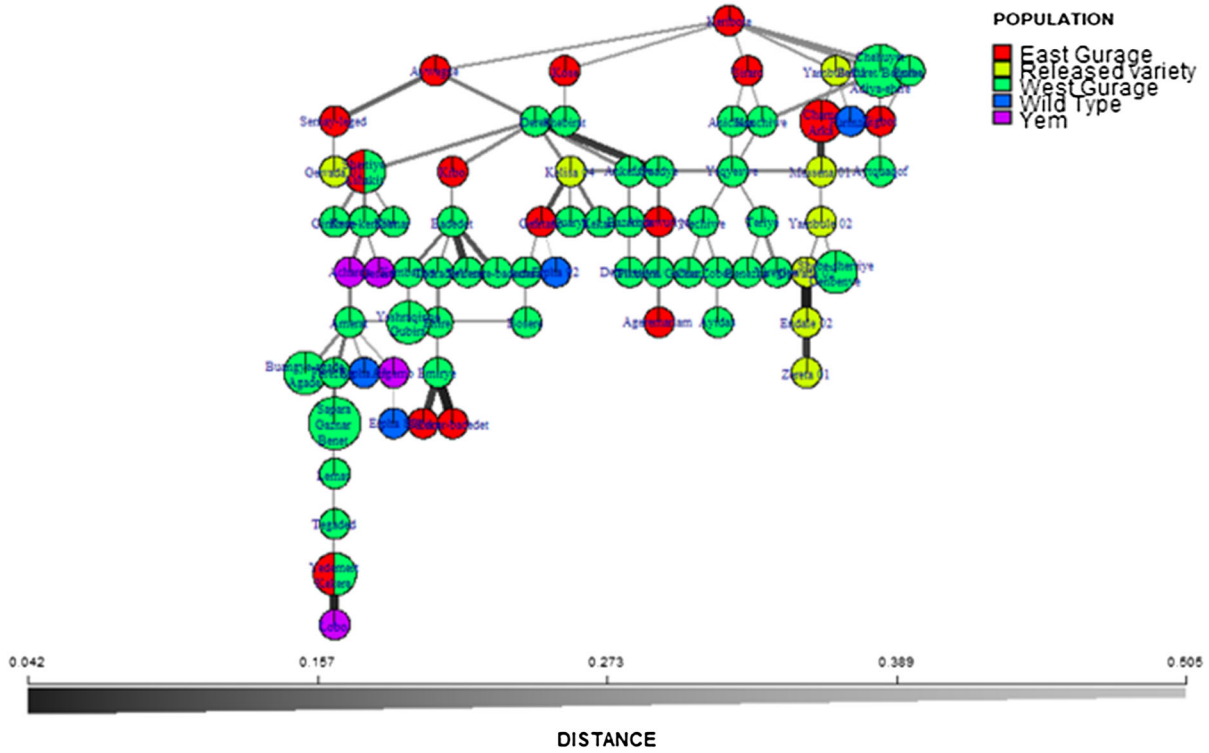

Fig. 2 Minimum spanning network based on the twelve SSR markers for the 83 studied landraces of enset

(Olango et al. 2015). The observed allele variation can be attributed to the use of different types and number of genotypes and primers applied in each study.

Among the studied enset populations, the Wild enset population was found to be relatively more diverse than others as it retained the highest value for rarefied allelic richness and rarefied private allelic richness values. Whereas, the rest four studied populations were to possess equivalent values indicating that they are almost at similar level of genetic diversity.

Shannon's diversity index of studied enset populations was > 1, implying the existence of high genetic diversity. The genetic diversity of the whole studied populations, as expressed by the mean expected heterozygosity (He = 0.59), is exactly similar with the genetic diversity (He = 0.59) recorded on enset (Olango et al. 2015) using enset SSR markers. Moreover, it is slightly higher than the genetic diversity (He = 0.55) recorded on enset diversity study, using cross-transferred Musa SSR markers (Getachew et al. 2014). The level of genetic diversity estimated in this study is also higher than previous reports of enset genetic diversity estimated using other DNA markers of ISSR (Tobiaw and Bekele 2011), AFLP (Negash et al. 2002) and RAPD (Birmeta et al. 2002). However, it is hard to make direct comparison between these studies, and draw general conclusions, since the number and type of the studied germplasm and the DNA markers used were different. The higher fixation index (F) value of 0.17 was recorded for Wild population, which indicates population differentiation, as the Wild population is sexually multiplied by seeds. Negative values (− 0.02 to − 0.14) of fixation index (Table 4) for the same parameter was identified for the cultivated populations, which is taken as F_{ST} = zero indicating high amount of gene flow between the studied populations equalizing the allele frequencies.

The AMOVA revealed that the variations within landraces of enset contributed to the majority (93%) of the total molecular variance; while the five population groups and the variations among landraces in each population group, contributed 4 and 3% of the total molecular variance, respectively. This result is consistent with the previous enset diversity studies (Olango et al. 2015; Getachew et al. 2014; Birmeta et al. 2004; Negash et al. 2002; Tobiaw and Bekele 2011). The within population variation was very high,

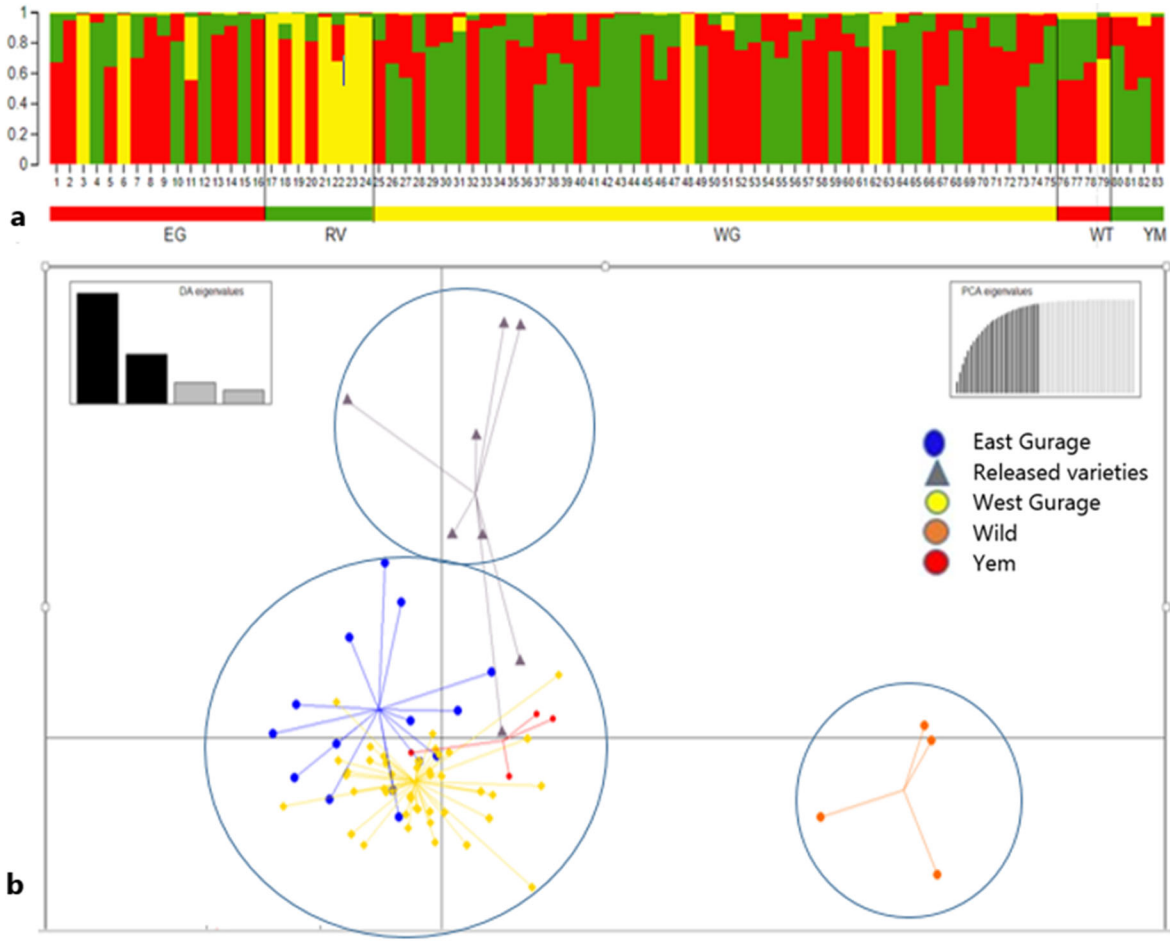

Fig. 3 Population structure based on 12 polymorphic SSR markers indicating estimated group structure with individual group membership values (1-83 following arrangement of landraces in Table 1) (**a**) and DAPC scatter plot for the studied 83 enset germplasms (**b**)

which indicates the need to stress within-population diversity in the conservation and exploitation of the existing enset diversity in the study area. Enset landraces of the five populations were sub-divided into three inferred cluster populations. Structure analysis, DAPC, and the Unweighted Neighbor-joining algorithm shown admixtures among the five populations. The observed low F_{ST} value (0.037) between the enset populations in this study indicated low differentiation among the groups that might be attributed to gene flow among regions. A low degree of differentiation in populations of enset shown in this study might be attributed to the exchange of planting material irrespective of geographical distances (Tesfaye and Ludders 2003). These may cause migration and gene flow between populations that resulted in admixture among enset populations.

The DAPC scatter plot (Fig. 3a, b) revealed a high level of differentiation between wild and cultivated enset germplasms. Other enset diversity studies have also reported a high level of genetic dissimilarity between cultivated and wild enset germplasm (Birmeta et al. 2004; Olango et al. 2015). The lowest (0.003) genetic distance was identified between West and East Gurage enset populations that showed a closer resemblance to one another that might be explained by frequent germplasm exchange, which might be due to their high cultural relations, despite the Zebidar mountain barrier between the two locations. In the study, duplication was found, as some of

the cultivated enset landraces did not differ in their SSR profile for the tested primers (with 0.00 genetic distance), including *Ashakit/Shertiye*, *Adiya-ehire/Bekuret/Chehuyet*, *Kekere/Yedemert*, *Genbenye/Shebe-shertiye*, *Agade/Buangye-agade*, *Yeshi-raqinqe/Gubira*, *Benet/Gaznar/Sapara* and *Arka/Chama*.

Conclusion

The mean diversity indices (He = 0.59, Ho = 0.61, I > 1) confirmed the existence of high genetic diversity in the studied *Gurage* and surrounding enset landraces. The studied wild and cultivated enset genotypes were genetically diverse. Similarly, most of the released varieties also genetically varied with the cultivated landraces. The West and East *Gurage* enset population showed a closer resemblance to one another, which might have resulted from germplasm exchange, due to their high cultural ties, though geographically separated by *Zebidar* mountain. In the present study 18 enset landraces found to be grouped into eight genotypes as they possessed identical genetic makeup across the studied loci indicating that they are simply duplications of one another. The result contributes important information for the implementation of appropriate conservation strategies and breeding programs of enset genetic resources for the basic enset production constraints such as nematode and bacterial wilt diseases and overall productivity improvement of the crop.

Acknowledgements This work was supported by BecA-ILRI Hub through the Africa Biosciences Challenge Fund (ABCF) program. The ABCF program is funded by the Australian Department for Foreign Affairs and Trade (DFAT) through the BecA-CSIRO partnership; the Syngenta Foundation for Sustainable Agriculture (SFSA); the Bill and Melinda Gates Foundation (BMGF); the UK Department for International Development (DFID) and the Swedish International Development Cooperation Agency (Sida). The authors are grateful to Jean-Baka Domelevo Entfellner for his valuable support on data analysis. Areka Agricultural Research Center and Wolkite University are kindly acknowledged for the provision of wild and cultivated enset germplasms.

Funding This study was funded by BecA-ILRI Hub through the Africa Biosciences Challenge Fund (ABCF) program (02/RF/18/4857).

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

Almaz N, Admasu T, Treuren RV, Visser B (2002) AFLP analysis of enset clonal diversity in south and southwestern Ethiopia for conservation. Crop Sci 42:1105–1111

Baker RED, Simmonds NW (1953) The genus ensete in Africa. Kew Bull 8:405–416

Birmeta G, Nybom H, Bekele E (2002) RAPD analysis of genetic diversity among clones of the Ethiopian crop plant *Ensete ventricosum*. Euphytica 124(3):315–325

Birmeta G, Nybom H, Bekele E (2004) Distinction between wild and cultivated Enset (*Ensete ventricosum*) gene pools in Ethiopia using RAPD markers. Hereditas 140:139–148

Bizuayehu T (2002) Studies on landrace diversity in-vivo and in-vitro regeneration of enset (*Ensete ventricosum* Welw.). Dissertation, Humbolkt University, Berlin

Bobosha K (2003) Characterization of *Xanthomonas campestis* Pv. musacearum isolates: causal agent of enset bacterial wilt disease. MSc Thesis, Addis Ababa University, Ethiopia

Brandt SA, Spring A, Hiebisch C, McCabe JT, Tabogie E, Diro M, Wolde-Michael G, Ynttso G, Shigeta M, Tesfaye S (1997) The tree against hunger. Enset based agricultural systems in Ethiopia. American Association for the Advancement of Science, Washington DC

Central Statistical Authority (CSA) (2010) Area, production and yield of crops for private peasant holdings for meher season 2009/2010. Addis Ababa, Ethiopia

Earl DA, von Holdt BM (2012) Structure harvester: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv Genet Resour 4(2):359–361. https://doi.org/10.1007/s12686-011-9548-7

Fetta N (2007) Diversity and indigenous management of enset (*Ensete ventricosum* Welw. Cheesman) in Gurage Zone, Ethiopia. MSc Thesis, Hawassa University, Ethiopia

Getachew S, Firew M, Belayneh A, Segenet K, Sisay K, Kefyalew N, Djikeng A, Nzuki I (2014) A look into genetic diversity of enset (*Ensete ventricosum* (Welw.) Cheesman) using transferable microsatellite sequences of banana in Ethiopia. Crop Improv 28(2):159–183. https://doi.org/10.1080/15427528.2013.861889

Jakobsson M, Rosenberg NA (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics 23:1801–1806

Jombart T (2008) Adegenet: R package for the multivariate analysis of genetic markers. Bioinformatics 24:1403–1405

Kalinowski ST (2002) How many alleles per locus should be used to estimate genetic distances? Hered 88:62–65

Kalinowski ST (2005) HP-Rare 1.0: a computer program for performing rarefaction on measures of allelic diversity. Mol Ecol Notes 5:187–189

Lagoda PJL, Noyer JL, Dambier D, Baurens FC, Grapin A, Lanaud C (1998) Sequence tagged microsatellite site (STMS) markers in the Musaceae. Mol Ecol 7:657–666

Negash A, Tsegaye A, van Treuren R, Visser B (2002) AFLP analysis of enset clonal diversity in south and southwestern Ethiopia for conservation. Crop Sci 42:1105–1111

Nei M (1972) Genetic distance between populations. Am Nat 106(949):283–292

Nei M (1987) Molecular evolutionary genetics. Columbia University Press, New York

Olango TM, Tesfaye B, Catellani M, Pè ME (2014) Indigenous knowledge, use and on-farm management of enset (*Ensete ventricosum* (Welw.) Cheesman) diversity in Wolaita, Southern Ethiopia. J Ethnobiol Ethnomed 10:41

Olango TM, Tesfaye B, Pagnotta MA, Pè ME, Catellani M (2015) Development of SSR markers and genetic diversity analysis in enset (*Ensete ventricosum* (Welw.) Cheesman), an orphan food security crop from Southern Ethiopia. BMC Genet 16:98. https://doi.org/10.1186/s12863-015-0250-8

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

Peregrine WTH, Bridge J (1992) The lesion nematode, *Pratylenchus goodeyi,* an important pest of enset in Ethiopia. Trop Pest Manag 38:325–326

Perrier X, Flori A, Bonnot F (2003) Data analysis methods. In: Hamon P, Seguin M, Perrier X, Glaszmann JC (eds) Genetic diversity of cultivated tropical plants. Enfield Science Publishers, Montpellier, pp 43–76

Prabakaran A, Paramasivam K, Rajesh T, Rajarajan D (2010) Molecular characterization of rice landraces using SSR markers. J Plant Breed 1(4):512–516

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multi locus genotype data. Genetics 155:945–959 **PMID: 10835412**

Ramasamy RK, Ramasamy S, Bindroo BB, Naik VG (2014) Structure plot: a program for drawing elegant STRUCTURE bar plots in user friendly interface. Springer Plus 3:431

Storchova H, Hrdlickorva R, Chrtek J, Tetera M, Fitze D, Fehrer J (2000) An improved method of DNA isolation from plants collected in the field and conserved in saturated NaCL/CTAB solution. Taxon 48:79–84

Taye B (1993) An overview on enset research and future technological needs for enhancing its production and utilization. In: Tsedeke A, Hibisch K, Brandt S (eds) Proceeding of the first international workshop on enset. IAR, Addis Ababa, pp 1–12

Tenaye A, Geta E (2009) Analysis of vulnerability and determinants of enset production in Wolaita, Southern Ethiopia. Acta Hortic 806:663–668. https://doi.org/10.17660/ActaHortic.2009.806.82

Tesfaye B, Ludders P (2003) Diversity and distribution patterns of enset landraces in Sidama, Southern Ethiopia. Genet Res Crop Evol 50:359–371

Tobiaw DC, Bekele E (2011) Analysis of genetic diversity among cultivated enset (*Ensete ventricosum*) populations from Essera and Kefficho, southwestern part of Ethiopia using inter simple sequence repeats (ISSRs) marker. Afr J Biotechnol 70:15697–15709

Tomlinson PB (1969) Anatomy of monocotyledons. III. Commelinales-Zingiberales. Clarendon Press, Oxford

Tsegaye A, Struik P (2002) Analysis of enset (*Ensete ventricosum*) indigenous production methods and farm-based biodiversity in major enset growing regions of southern Ethiopia. Exp Agric 38(3):291–315

Worku N (1996) The Gurage perception of enset. In: Tsedeke et al. Enset based sustainable agriculture in Ethiopia. Institute of Agricultural research, Addis Ababa, Ethiopia, pp 121–131

Wright S (1951) The genetical structure of populations. Ann Eugen 15:323–354

Yemataw Z, Tawle K, Blomme G, Jacobsen K (2018) Traditional enset (*Ensete ventricosum* (Welw.) Cheesman) sucker propagation methods and opportunities for crop improvement. Int J Trop Subtrop Hortic 73(6):342–348

Zippel K (2002) Enset (*Ensete ventricosum* (Welw.) Cheesm.) In subsistence farming systems in Ethiopia. In: Conference on international agricultural resources for development, Deutscher Tropentag, Witzenhausen, Germany