Genet Resour Crop Evol (2007) 54:601–614
DOI 10.1007/s10722-006-0018-0

RESEARCH ARTICLE

Genetic diversity of *Guizotia abyssinica* (L. f.) Cass. (Asteraceae) from Ethiopia as revealed by random amplified polymorphic DNA (RAPD)

Mulatu Geleta · Tomas Bryngelsson · Endashaw Bekele · Kifle Dagne

Received: 13 July 2005 / Accepted: 24 February 2006 / Published online: 19 July 2006

Abstract Genetic diversity of 70 populations of niger (*Guizotia abyssinica*) representing all its growing regions in Ethiopia was investigated using random amplified polymorphic DNA (RAPD) to reveal the extent of its populations genetic diversity. Ninety-seven percent of the loci studied was revealed to be polymorphic for the whole data set. The within population diversity estimated by Shannon diversity index and Nei gene diversity estimates was revealed to be 0.395 and 0.158, respectively. The extent of genetic variation of populations from major niger producing regions was significantly lower than that of populations from other regions; however, it is distributed regardless of altitude of growth. Genetic differentiation between populations was estimated with Shannon index as G'_{ST} (0.432), Nei's G_{ST} (0.242) and AMOVA based F_{ST} (0.350) and appears to be equivalent to the average values calculated from various RAPD based studies on outcrossing species. Higher proportion of the variation detected by AMOVA resided within populations (64.58%) relative to the amount of variation among populations (35.42%). UPGMA cluster analysis showed that most of the populations were clustered according to their region of origin. However, some populations were genetically distant from the majority and seem to have unique genetic properties. It is concluded that the crop has a wide genetic basis that may be used for the improvement of the species through conventional breeding and/or marker assisted selection. Collection of germplasm from areas not yet covered and/or underrepresented is the opportunity to broaden the genetic basis of genebank collection.

Keywords AMOVA · Genetic diversity · Germplasm · *Guizotia abyssinica* · Niger · Population · RAPD

M. Geleta · T. Bryngelsson
Department of Crop Science, Swedish University of Agricultural Sciences, P.O. Box 44, Alnarp SE 230 53, Sweden

E. Bekele · K. Dagne · M. Geleta (✉)
Department of Biology, Addis Ababa University, P.O. Box 1176, Addis Ababa, Ethiopia
e-mail: gemulat@yahoo.com

Introduction

Guizotia abyssinica (L. f.) Cass., commonly known as 'niger', belongs to the family Asteraceae, tribe Heliantheae and subtribe Coreopsidinae. Niger is the only domesticated species of the small exclusively diploid genus, *Guizotia* Cass. It is an economically important oilseed crop species with 2n=30 chromosomes (Hiremath and Murthy 1992; Dagne 1995). This crop is widely cultivated particularly in Ethiopia but also in India (Riley and Belayneh 1989; Hiremath and Murthy 1992). It is also cultivated in small scale in several other African and Asian countries as an edible oil crop (Murthy et al. 1993; Getinet and

Sharma 1996) and currently in the United States of America, mainly as bird seed (Kandel and Porter 2002). Niger is an outcrossing species with self-incompatibility mechanisms (Riley and Belayneh 1989; Nemomissa et al. 1999). *Guizotia abyssinica* is suggested to be evolved from *Guizotia scabra* (Vis.) Chiov. ssp. *schimperi* (Sch. Bip. in Walp.) Baag. based on evidences from phytogeography, cytogenetics and morphology (Baagøe 1974; Hiremath and Murthy 1988; Murthy et al. 1993; Dagne 1994; Dagne 1995; Dagne 2001).

Niger seed is used for various purposes. Its yellow, edible, semi-drying oil is used mainly for cooking purpose. It is also used for soaps, paints, illuminants, and lubricants (Baagøe 1974; Riley and Belayneh 1989; Kandel and Porter 2002). The Press cake left after the oil has been extracted is found to be an excellent animal feed as it contains about 33% protein (Kandel and Porter 2002). The oil content of the niger seed vary considerably and the most abundant fatty acid is linoleic acid as reported by different authors (Dutta et al. 1994; Alemaw and Teklewold 1995; Dagne and Jonsson 1997; Ramadan and Mörsel 2002).

Ethiopia is the center of origin and diversity of *G. abyssinica* (Harlan 1969; Zeven and DeWet 1982), where it has been under cultivation for much longer than any other places (Baagøe 1974; Hiremath and Murthy 1988) and stands first with coverage of about 60% of the total area and production volume of oil crops. Niger is suitable for multiple cropping, especially as border crop (Geleta et al. 2002), which is a strategy to stabilize production, especially under unfavorable environmental conditions. In Ethiopia, it is grown mainly in an altitudinal range of 1600–2200 m asl with temperature range of 15–23 °C, and annual rainfall range of 500–1000 mm (Getinet and Sharma 1996).

Crop improvement through breeding depends on the magnitude of the genetic diversity and the extent to which this diversity is utilized. Although there are few morphological studies on genetic diversity of niger (Nayakar 1976; Alemaw and Teklewold 1995; Pradhan et al. 1995, Genet and Belete 2000) no molecular marker based genetic diversity study has been reported at least for Ethiopian niger. Random amplified polymorphic DNA (RAPD) technique is a method of choice for studying genetic diversity for crop species where there is little or no molecular genetics research (Nybom 2004), as it does not require sequence information for the target species. Furthermore, it is specially suited for studying large number of samples as it is relatively simple, fast and cheap.

Hence, the present study was under taken to investigate the extent of genetic variation within populations and genetic differentiation among populations of Ethiopian niger using RAPD markers with the goal of identifying hotspots for conservation and utilization of its genetic diversity.

Material and methods

Plant material

Seventy populations of niger, with 10 individuals each, collected from 11 regions of Ethiopia were used for this study (Fig. 1; Table 1). All populations were collected directly from farmers' fields from early November to the end of December 2003. The crop on a single farmer's field was considered as a population. Individuals were sampled at equidistant along the longest line found across the field. Different number of populations was sampled per region depending on the extent of cultivation. The sampled populations represent the altitudinal ranges and geographic regions where niger is currently grown within the country.

DNA extraction

Young and fresh leaves from 15 days to 1 month old plants grown in the greenhouse were used for genomic DNA extraction. DNA was extracted by modified CTAB procedures as described by Aga et al. (2003).

PCR amplification and electrophoresis

Several protocols that have been used for plant DNA amplification were tested and the one with best amplification profiles was chosen. PCR components, DNA concentration and PCR amplification temperature profiles were then optimized. After protocol optimization, 150 RAPD primers from QIAGEN (QIAGEN Operon GmbH, Germany) were tested with the objective of screening primers that can de-

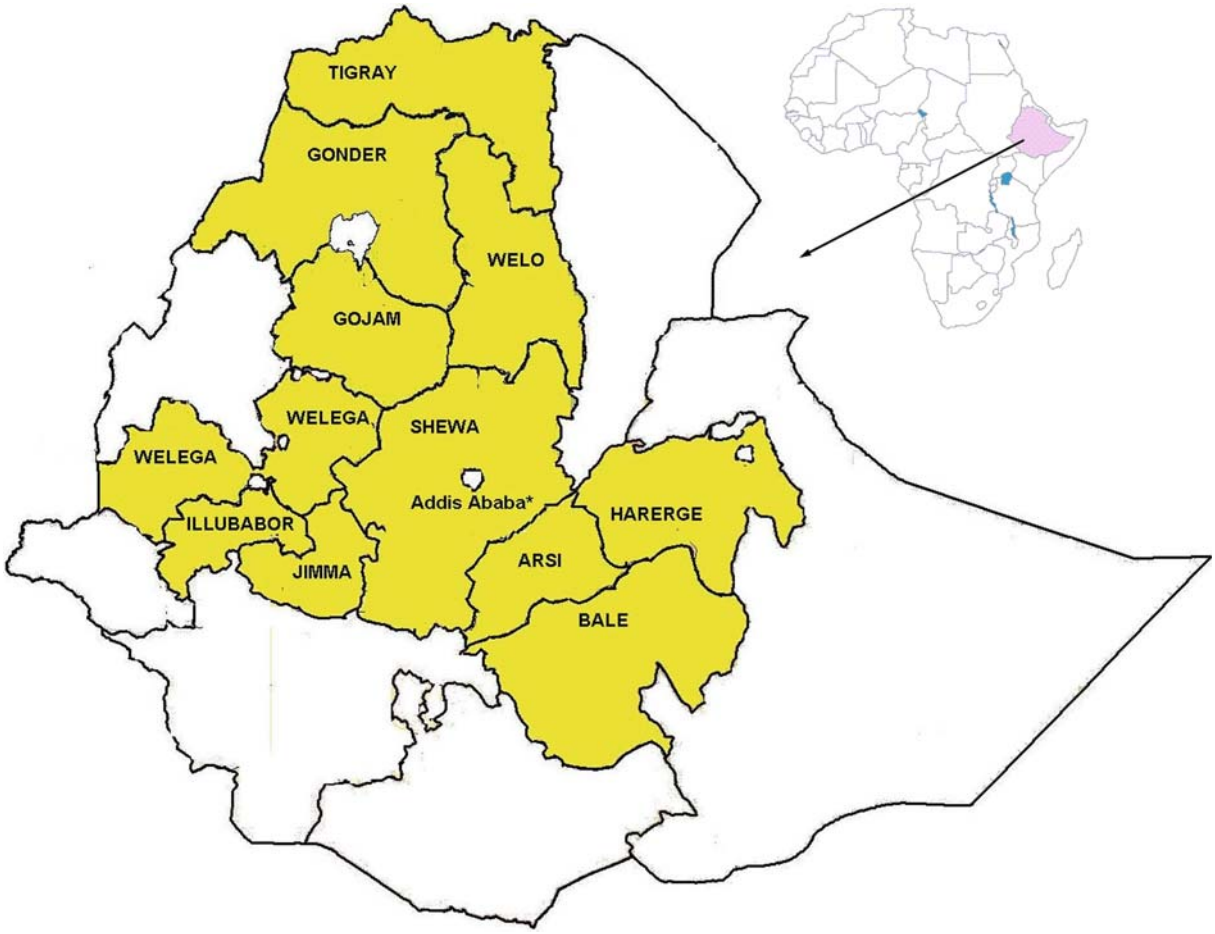

Fig. 1 Map of Ethiopia showing the 11 regions from where the *Guizotia abyssinica* populations were sampled (Shaded regions). *Capital City

tect polymorphism, show clearly resolvable banding patterns and amplify larger number of loci per sample. Twenty five primers that fulfilled the above criteria were initially screened. This was followed by repeated checking for the reproducibility of the banding patterns using replicated samples within a gel and across different gels, as suggested by Skroch and Neinhuis (1995). This led us to consider 10 primers that were highly reproducible (see Table 2).

The DNA amplification reaction was performed in a total volume of 20 µL containing 1× reaction buffer (75 mM Tris–HCl pH 8.8, 20 mM $(NH_4)_2SO_4$, 0.01% (v/v) Tween 20), 3.5 mM $MgCl_2$, 20 ng primer, 0.1 mM of each dNTPs, 0.6 units of *Taq* DNA polymerase from *ABgene* (ABgene house, UK), and 25 ng of sample DNA. RAPD amplification was carried out using GeneAMP PCR system 9700 thermocycler with the following temperature profiles: initial 3 min denaturing at 94 °C and final 10 min extension at 72 °C with the intervening 45 cycles of 1 min denaturing at 94 °C, 1 min of primer annealing at 37 °C, and 2 min of primer extension at 72 °C. The amplified product was stored at 4 °C until electrophoresis.

The amplified product was loaded on 1.5% (w/v) agarose gels after adding 5 µL loading buffer (0.024% (w/v) bromophenol blue, 0.024% (w/v) xylene cyanol and 5% glycerol) and electrophoresed in 1× TAE (0.04 M Tris–Acetate and 0.002 M EDTA) buffer for 3 h at constant voltage of 90 V. A 100-basepair ladder was loaded on marginal and central lanes to estimate the size of the fragments. After electrophoresis, the gel was stained in ethidium bromide (5 µg/mL) on a shaker for 30 min and then washed in water for

Table 1 Population codes, altitudes and collection sites of the *Guizotia abyssinica* populations studied

Popn[a]* code	Altitude (m asl)	Place of collection	Popn[a]* code	Altitude (m asl)	Place of collection
A-1	2380	52 km from Bekoji to Chole	J-2	1860	25 km from Jimma to A/Ababa
A-2	2570	9 km from Diksis to Kula	S-1	2660	53 km from G/guracha to D/Tsige
A-3	2565	13 km from Robe to Ticho	S-2	1590	5 km NW of Ataye town
A-4	2510	7 km from Robe to Diksis	S-3	2540	2 km from G/Tsion to G/Guracha
A-5	2470	21 km South of Adele	S-4	2372	78 km from A/Ababa to Weliso
B-1	1460	2 km D/Mena to Goba	S-5	2500	7 km from Gedo to Ambo
B-2	2450	2.7 km from Adaba to Dodola	S-6	2030	19 km from Welkite to Hos'ana
B-3	2425	63 km from Ginir to Gasera	S-7	2460	21 km from Ambo to A/Ababa
B-4	1835	5.5 km NE of Ginir	S-8	2470	6 km from Holeta to Muger
B-5	2410	35 km SE of Bale-Robe	S-9	2340	56 km from M/Turi to A/Ketema
Gj-1	2550	29 km from Dejen to D/Markos	S-10	2155	18 km from A/Ababa to Weliso
Gj-2	2550	75 km from Mota to Dejen	S-11	1640	13 km from Sh/Robit to D/Sina
Gj-3	2055	19 km from Dangla to B/Dar	S-12	2570	1 km from Sendafa to Ch/Donsa
Gj-4	1890	35 km from Amanuel to Bure	T-1	1972	9.5 km Shire to Shiraro
Gj-5	2495	10 km from Bichena to Dejen	T-2	2180	10 km from Adwa to Axum
Gj-6	2205	34 km from Injibara to Dangla	T-3	2070	11 km from Slakleka to Shire
Gj-7	2250	8 km from D/Markos to Dangla	T-4	2150	18 km from Adwa to Axum
Gj-8	2070	33 km from F/Selam to Injibara	T-5	1400	86 km from Shire to A/Arkay
Gj-9	2610	10 km from Injibara to Dangla	Wg-1	2450	26 km from Shambu to Finch'a
Gj-10	2300	36 km from B/Dar to Adet	Wg-2	2190	18 km from Nekemt to Gimbi
Gr-1	2100	17 km from Azezo to Aykel	Wg-3	1486	77 km from Nekemt to Gimbi
Gr-2	2055	26 km from Gondar to K/Diba	Wg-4	2310	50 km from Bako to Shambu
Gr-3	2630	41 km from Werota to D/Tabor	Wg-5	2370	12 km from Fincha'a to Kombolcha
Gr-4	2130	27 km from Werota to B/Dar	Wg-6	2325	46 km from Fincha'a to Gedo
Gr-5	1880	14 km from A/Zemen to Werota	Wg-7	2685	54 km m Bedele to Argo
Gr-6	1400	42 km from A/Arkay to Debark	Wg-8	2210	18 km from Bako to Shambu
Gr-7	1590	9 km from T/Dingay to Humera	Wg-9	1940	21 km from Nekemt to Ambo
Gr-8	1975	48 km from Werota to B/Dar	Wl-1	1620	18 km from Wergesa to Mersa
Gr-9	2000	32 km from Gondar to B/Dar	Wl-2	1650	31 km from Kombolcha to Bati
H-1	1860	49 km from A/Teferi to Gelemso	Wl-3	1740	25 km from Woldia to Kobo
H-2	1830	36 km from A/Teferi to Gelemso	Wl-4	1820	58 km from Dessie to W/Tena
I-1	1865	3 km from Metu to Gore	Wl-5	2135	10 km Haik to Bistima
I-2	2095	82 km from Agaro to Bedele	Wl-6	1800	4 km from Karakore to Kemisie
I-3	1900	53 km from Gore to Bure	Wl-7	1882	18 km from Woldia to Lalibela
J-1	1900	3 km from Asendabo to O/Nadda	Wl-8	2420	31 km from W/Tena to Gashena

[a]Population code. *Letter(s) in population code indicate(s) the region from where the corresponding population was collected. A=Arsi, B=Bale, Gj=Gojam, Gr=Gonder, H= Harerge, I=Illubabor, J=Jimma, S=Shewa, T=Tigray, Wg=Welega, Wl=Welo

20 min. The stained gel was photographed using Saveen Werner AB UV camera equipped with Sony Black and white Monitor SSM930CE and Sony Video graphic printer UP-895CE. The photograph was printed and also saved on floppy diskettes and transferred into computer for band scoring.

Data scoring and analysis

Each RAPD band was considered as a single bi-allelic locus with one amplifiable and one null allele. Data were scored as 1 for the presence and 0 for the absence of a DNA band for each locus across the 700 genotypes. The locus was considered polymorphic when the frequency of present allele or null allele is less than 95% across the whole genotypes investigated. Genetic diversity was calculated based on (1) Shannon diversity index using both monomorphic and polymorphic loci and (2) Nei's unbiased gene diversity (Nei 1978) with the modification provided by Lynch and Milligan (1994) using polymorphic loci only.

Shannon index was calculated for each locus for each population as $H'_j = -\sum p_i \log_2 p_i$, where p_i is the frequency of the presence or absence of RAPD band in that population. The average diversity per population for each locus was calculated as $H'_{\text{pop}} = \sum H'_j/n$, where n is the number of populations, while the mean observed Shannon

Table 2 RAPD primers used for the study and size of fragment amplified by each primer. Primer-wise Shannon diversity indices, Nei gene diversity estimates and F_{ST} obtained with AMOVA

Primer	Base sequence	SRAF[a]	NLA[b]	NPL[c]	Shannon diversity estimate					Nei gene diversity estimate			AMOVA
					H'_{pop}	H'_{sp}	H'_{pop}/H'_{sp}	$1-H'_{pop}/H'_{sp}$	G'_{ST}	H_S	H_T	G_{ST}	F_{ST}
OPA-11	5′-CAATCGCCGT-3′	510–1900	18	18	0.413	0.699	0.590	0.410	0.424	0.162	0.235	0.244	0.33
OPA-14	5′-TCTGTGCTGG-3′	500–1300	6	6	0.465	0.789	0.590	0.410	0.438	0.215	0.298	0.299	0.37
OPB-18	5′-CCACAGCAGT-3′	430–1880	16	15	0.523	0.720	0.715	0.285	0.288	0.215	0.250	0.123	0.22
OPB-20	5′-GGACCCTTAC-3′	400–2000	21	21	0.445	0.698	0.637	0.363	0.374	0.193	0.257	0.192	0.27
OPD-20	5′-ACCCGGTCAC-3′	710–2300	15	14	0.386	0.733	0.526	0.474	0.482	0.193	0.319	0.322	0.38
OPF-5	5′-CCGAATTCCC-3′	220–1930	18	18	0.330	0.670	0.493	0.507	0.504	0.139	0.193	0.252	0.41
OPF-10	5′-GGAAGCTTGG-3′	380–1900	32	31	0.352	0.691	0.510	0.490	0.491	0.151	0.209	0.242	0.39
OPG-2	5′-GGCACTGAGG-3′	220–2050	24	24	0.430	0.769	0.559	0.441	0.440	0.194	0.287	0.263	0.36
OPG-16	5′-AGCGTCCTCC-3′	510–1850	24	22	0.420	0.699	0.597	0.403	0.408	0.176	0.242	0.219	0.34
OPG-17	5′-ACGACCGACA-3′	340–1850	20	19	0.375	0.711	0.519	0.481	0.473	0.164	0.258	0.302	0.41
Total			194	188	$\overline{H'}_{pop}$	$\overline{H'}_{sp}$			$\overline{G}'_{ST}$	$\overline{H}_S$	$\overline{H}_T$	$\overline{G}_{ST}$	$\overline{F}_{ST}$
Mean					0.400[d]	0.712[d]	0.574[e]	0.426[e]	0.436[d]	0.176[d]	0.248[d]	0.242[d]	0.350[d]

[a]Size range of amplified fragments

[b]Number of loci amplified

[c]Number of polymorphic loci

[d]Mean obtained by averaging per locus values across all loci

[e]Mean of values in column

diversity of each population was calculated as $H'_{\text{loci}} = \sum H'_j/L$, where L is the number of loci studied. Similarly, the overall Shannon diversity within the species for each locus was calculated as $H'_{sp} = -\sum p_s \log_2 p_s$ where p_s is the frequency of the presence or the absence of RAPD band across the 700 genotypes. The overall diversity within the species (H'_{sp}) was then partitioned into proportion of within and between population genetic diversity for each locus as $H'_{\text{pop}}/H'_{\text{sp}}$ and $(1 - H'_{\text{pop}}/H'_{\text{sp}})$, respectively. The extent of population differentiation was calculated as G'_{ST} for each polymorphic loci and the overall mean was given as $\overline{G'}_{ST}$.

Gene diversity was calculated for each population for every locus according to Lynch and Milligan (1994) as $H_j(i) = 2q_j(i)[1 - q_j(i)] + 2\text{Var}[q_j(i)]$ where q is the frequency of null allele at a locus for a given population. q was calculated from x, which is the frequency of individuals within a population that lack the RAPD band, as $q = x^{1/2}\left[1 - \frac{\text{Var}(x)}{8x^2}\right]^{-1}$. $\text{Var}(q)$ and $\text{Var}(x)$ were calculated as $(1 - x)/4N$ and $x(1 - x)/N$, respectively, where N is number of individuals per population. The mean observed gene diversity within each population was then calculated as $H_j = 1/L\sum_{i=1}^{L} H_j(i)$, where L is the number of polymorphic loci. Estimates of genetic differentiation between populations (G_{ST}) (Nei 1973) was calculated as $G_{\text{ST}} = (H_{\text{T}} - H_{\text{S}})/H_{\text{T}}$, where H_{S} is the mean gene diversity per population averaged across all populations for each polymorphic locus while H_{T} is the total gene diversity calculated from the overall frequency of the amplifiable and null allele for each polymorphic locus.

NTSYSpc program (Rohlf 2000) was used to calculate Nei (1972) standard genetic distances and unweighted pair group method with arithmetic average (UPGMA) cluster analysis using the sequential agglomerative hierarchical nested cluster analysis (SAHN). POPGENE version 1.31 (Yeh and Boyle 1997) was used for analysis of percentage of polymorphic loci for each population. Analysis of molecular variance (AMOVA) was done using Arlequin version 2 (Schneider et al. 2000). FreeTree-Freeware program (Pavlicek et al. 1999) was used to generate Nei (1972) standard genetic distances between regions for bootstrap analysis and the generated trees were viewed using TreeView (Win32) 1.6.6 program (Page 1996). Minitab statistical software (MINITAB release 14) was used for further analysis of some of the outputs from other software.

Results and discussion

RAPD primers and percentage of polymorphic loci

Table 2 summarizes the characteristics of the 10 primers used in this study, genetic variation and extent of population differentiation revealed by each primer used. The primers applied to the 700 individuals generated 194 consistently amplified loci, of which 188 (97%) were polymorphic when the whole material was considered. When each population was taken into account, percent polymorphic loci ranged from 37% for a population from Gonder (Gr-2) to 66% for a population from Welo (Wl-2), with a mean of 52% (Table 3).The high percentage of polymorphic loci and the wide range of percentage of polymorphic loci for populations investigated suggest the existence of high genetic polymorphism in niger. The polymorphism detected as percent of polymorphic loci in the present study was higher than the 64% of polymorphic loci reported by Sivolap et al. (1998) using RAPD marker for analysis of molecular genetic polymorphism in sunflower (*Helianthus annuus* L.).

When the diversity recorded for each locus was averaged across the loci within a primer, it became clear that primer OPB-18 detected the highest within population variation, which means that this primer might be better suited for niger to detect the within population variation when one wishes to select genotypes with desirable traits from the heterogeneous landrace populations for breeding purposes. On the other hand, primer OPD-20, primer-OPF-5 and primer-OPG-17 seem to be better suited to identify populations that have lost amplifiable RAPD alleles, as they revealed a relatively higher proportion of between population variations.

Genetic diversity

The evaluation of the RAPD fingerprints by 10 primers resulted in the overall species diversity of 0.712 and 0.248 respectively when estimated by Shannon diversity index as $\overline{H'}_{\text{sp}}$ and Nei's unbiased (Nei 1978) gene diversity with modification of Lynch

Table 3 Mean Shannon diversity estimates (H_{loci}), mean gene diversity estimates (H_W) and mean percent polymorphic loci (%PL) for each population (Popn) and region

Popn/region code*	Mean H'_{loci}	Mean H^a_W	%PL	Popn/region code*	Mean H'_{loci}	Mean H^a_W	%PL	Popn/region code*	Mean H'_{loci}	Mean H^a_W	%PL
A-1	0.425	0.202	54.12	Gr-5	0.372	0.150	49.48	S-12	0.425	0.203	53.61
A-2	0.429	0.195	55.67	Gr-6	0.302	0.112	39.69	**S**	**0.402**	**0.183**	**51.81**
A-3	0.392	0.169	51.03	Gr-7	0.338	0.139	45.88	T-1	0.386	0.171	48.97
A-4	0.375	0.175	48.45	Gr-8	0.371	0.156	52.06	T-2	0.409	0.177	53.09
A-5	0.496	0.222	63.92	Gr-9	0.418	0.176	55.15	T-3	0.462	0.196	58.25
A	**0.4234**	**0.193**	**54.64**	**Gr**	**0.349**	**0.142**	**46.79**	T-4	0.372	0.168	51.03
B-1	0.401	0.178	51.03	H-1	0.377	0.175	48.45	T-5	0.445	0.189	59.28
B-2	0.443	0.195	57.73	H-2	0.372	0.156	48.45	**T**	**0.415**	**0.180**	**54.12**
B-3	0.441	0.193	54.64	**H**	**0.375**	**0.165**	**48.45**	Wg-1	0.388	0.176	48.45
B-4	0.410	0.182	54.12	I-1	0.404	0.185	51.55	Wg-2	0.413	0.183	52.58
B-5	0.388	0.186	50.52	I-2	0.431	0.192	57.73	Wg-3	0.330	0.152	43.81
B	**0.417**	**0.187**	**53.61**	I-3	0.416	0.195	55.67	Wg-4	0.383	0.161	52.58
Gj-1	0.347	0.156	47.94	**I**	**0.417**	**0.191**	**54.98**	Wg-5	0.387	0.164	50.52
Gj-2	0.344	0.149	46.91	J-1	0.384	0.176	50.52	Wg-6	0.345	0.160	45.36
Gj-3	0.328	0.152	43.3	J-2	0.449	0.205	61.34	Wg-7	0.359	0.160	49.48
Gj-4	0.360	0.151	47.42	**J**	**0.417**	**0.190**	**55.93**	Wg-8	0.351	0.158	45.36
Gj-5	0.357	0.165	48.97	S-1	0.409	0.191	52.58	Wg-9	0.438	0.191	54.64
Gj-6	0.399	0.189	52.06	S-2	0.395	0.185	52.06	**Wg**	**0.377**	**0.167**	**49.2**
Gj-7	0.373	0.161	48.97	S-3	0.382	0.167	47.94	Wl-1	0.451	0.199	58.76
Gj-8	0.345	0.147	42.78	S-4	0.368	0.170	48.45	Wl-2	0.518	0.245	65.98
Gj-9	0.348	0.146	44.85	S-5	0.394	0.192	51.55	Wl-3	0.482	0.209	61.34
Gj-10	0.361	0.150	46.91	S-6	0.396	0.175	51.55	Wl-4	0.397	0.196	52.06
Gj	**0.356**	**0.157**	**47.01**	S-7	0.377	0.179	47.94	Wl-5	0.410	0.183	51.55
Gr-1	0.344	0.133	47.42	S-8	0.371	0.167	48.97	Wl-6	0.468	0.221	57.73
Gr-2	0.282	0.119	36.6	S-9	0.383	0.172	47.94	Wl-7	0.400	0.193	52.58
Gr-3	0.317	0.134	43.3	S-10	0.512	0.214	65.46	Wl-8	0.468	0.218	61.34
Gr-4	0.397	0.163	51.55	S-11	0.408	0.175	53.61	**W1**	**0.449**	**0.208**	**57.67**
Overall mean									**0.395**	**0.158**	**52.2**

[a]No correlation with altitudes and no significant difference between groups classified based on altitudes of collection

*Population/region code

and Milligan (1994) as ($\overline{H'}_T$) (Table 2). The overall within population genetic diversity estimated with Shannon diversity index as $\overline{H'}_{pop}$ as was revealed to be 0.400 (Table 2), which is higher when compared with the mean (0.260) for the 24 outcrossing species reported in a review by Nybom and Bartish (2000). Similarly, the overall within population genetic diversity estimated by Nei's gene diversity as $\overline{H'}_S$ was found to be 0.176 (Table 2).

Nei's gene diversity is one of the most commonly used approaches to estimate within population diversity (Nybom and Bartish 2000; Nybom 2004). A modification of the Nei's formula proposed by Lynch and Milligan (1994) was applied to our data set as it is more unbiased estimate for RAPD markers. However, The pruning of loci suggested by them based on the 3/*N* criterion was not used due to various reasons including the small sample size used that allow the pruning of large number of loci with null allele frequency of below 0.3.

Use of the AMOVA to analyze RAPD marker variation has been shown to be effective for population analysis of highly heterogeneous, outcrossing plant species (Huff et al. 1993). AMOVA calculated using the present/absent (phenotypic) data for the populations of the 11 regions revealed percentage of variations among regions, among populations within regions and within populations to be 13.68%, 22.64% and 63.68%, respectively, which was found to be highly significant at all the three hierarchical levels ($P<0.00001$) (Table 5). Additionally, partitioning of the total Shannon genetic diversity into within and between population diversity components revealed that 57% of the variation resided within populations (Table 2).

The extent of diversity of each population was estimated using both Shannon and gene diversity as H'_{loci} and H_W, respectively, which is the average value across the whole loci. H'_{loci} ranged from 0.282 (Gr-2) to 0.518 (Wl-2). The H_W for these two populations was 0.119 (Gr-2) and 0.245 (Wl-2), which is still the two extremes with exception of another population from Gonder (Gr-6) with H_W of 0.112 (Table 3). Thus, The extent of genetic diversity for niger populations was found to be more than twofold, which is in agreement with Buckler IV and Thornsberry (2002) that described the substantial variation in extent of polymorphism between populations/species and sampled loci.

Genetic diversity of this crop at regional level was also estimated by averaging H'_{loci} and H_W over the populations within a region. Welo stands first with diversity estimates of 0.449 (mean of H'_{loci}) and 0.208 (mean of H_W), while Gonder stands last with diversity estimates of 0.349 (mean of H'_{loci}) and 0.142 (mean of H_W) (Table 3). As indicated above, AMOVA calculated by grouping populations into their respective regions revealed highly significant variation between regions ($P<0.0001$). Pair-wise comparison of regions also showed that 45% of pairs were significantly different from one another in the extent of diversity as revealed by analysis of variance (ANOVA) calculated based on Nei's gene diversity estimates (Table 4). When the extreme cases were considered it is interesting to note that the mean within population genetic diversity for

Table 4 Nei's standard genetic distances between regions

	Arsi	Bale	Gojam	Gonder	Harerge	Illubabor	Jimma	Shewa	Tigray	Welega	Welo
Arsi	0										
Bale	0.1	0									
Gojam	0.1**	0.1**	0								
Gonder	0.1**	0.1**	0.028	0							
Harerge	0.1	0.1*	0.095	0.095	0						
Illubabor	0.1	0.1	0.145**	0.146**	0.157	0					
Jimma	0.1	0.1	0.125**	0.129*	0.101	0.142	0				
Shewa	0.1	0.1	0.079**	0.073**	0.108	0.112	0.078	0			
Tigray	0.1	0.1	0.058**	0.063**	0.073	0.112	0.079	0.057	0		
Welega	0.1*	0.1**	0.078	0.073**	0.146	0.14*	0.12	0.055*	0.086	0	
Welo	0.1	0.1*	0.088**	0.095**	0.082*	0.1	0.058	0.065**	0.044*	0.098**	0

Superscripts are codes for ANOVA of Nei's gene diversity estimates for a pair of regions

*Variation in Nei's gene diversity was significant ($0.01<P<0.05$)

**Variation in Nei's gene diversity estimates was highly significant ($P<0.01$)

Gojam, a major niger producing region, was significantly lower than the means for the seven of the ten regions ($P<0.01$) (Table 4), suggesting that the extent of cultivation of crop's landraces does not always result in higher genetic diversity. Contrary to Gojam, the mean within population genetic diversity for Welo was found to be the highest and significantly higher than most of the means of other regions.

Thus, this study provides further illustration of the reported agromorphological variation (Alemaw and Teklewold 1995; Genet and Belete 2000) and variation in oil content (Dutta et al. 1994; Dagne 1994; Alemaw and Teklewold 1995) for Ethiopian niger. Genet and Belete (2000) estimated the diversity of Ethiopian niger using phenological and morphological characters and indicated that regionally the highest Shannon diversity index was recorded for Gojam. They also described that the major niger producing regions (Gojam, Shewa and Welega), which they considered to be the center of diversity for the species, have greater diversity. Unlike their report, our RAPD based study revealed that the diversity estimated using both Shannon and gene diversity estimates was not higher for these regions in comparison with other regions, rather the reverse was the result. This was also supported by lower percentage of polymorphic loci for these regions (Table 3).

The 70 Populations studied were grouped into populations from major niger producing regions and populations from minor niger producing regions and analyzed using AMOVA. About 6% of the total variation was resided between these groups (Table 5). This analysis revealed significantly lower level of genetic variation in populations from major niger producing regions as compared with the other group, which might be due to stronger selection for desirable trait such as yield and quality as this crop is produced mainly as cash crop in major niger producing regions. Our results in this study suggest that minor niger producing regions such as Arsi, Bale and Illubabor need to get due attention as these populations might have unique genetic properties. This helps to further enrich the gene pool at genebank, increases the genetic diversity of populations in major niger producing regions and broaden the genetic bases of breeding material.

In contrary, AMOVA computed by grouping populations into higher altitude group (>2000 m asl) and lower altitude group (<2000 m asl) revealed no significant differences between the groups. Furthermore, correlation analysis revealed the absence of correlation between extent of genetic diversity and

Table 5 AMOVA calculated for 70 populations (a) without grouping (b) by grouping the populations into 11regions, (c) by grouping into major and minor niger producing regions, and (d) by grouping the populations into two lower and higher altitude groups

Source of variation	DF*	Sum of squares	Variance components	% of variation	Fixation indices	*P* value
(a) Without grouping						
Among populations	69	8789.40	Va: 10.77	35.42	FST: 0.35	0.000
Within populations	630	12,373.50	Vb: 19.64	64.58		
Total	699	21,162.90	30.41			
(b) By grouping the populations into 11regions						
Among regions	10	3510.57	Va: 4.22	13.68	FSC: 0.26	0.000
Among populations within regions	59	5278.83	Vb: 6.98	22.64	FST: 0.36	0.000
Within populations	630	12,373.50	Vc: 19.64	63.68	FCT: 0.14	0.000
Total	699	21,162.90	30.84			
(c) By grouping into major and minor niger producing regions						
Among groups	1	798.27	Va: 1.99	6.32	FSC: 0.33	0.000
Among populations within groups	68	7991.13	Vb: 9.79	31.16	FST: 0.37	0.000
Within populations	630	12,373.50	Vc: 19.64	62.52	FCT: 0.06	0.000
Total	699	21,162.90	31.41			
(d) By grouping the populations into two lower and higher altitude groups						
Among groups	1	152.82	Va: 0.08	0.26	FSC: 0.35	0.138
Among populations within groups	68	8636.58	Vb: 10.74	35.25	FST: 0.35	0.000
Within populations	630	12,373.50	Vc: 19.64	64.48	FCT: 0.001	0.000
Total	699	21,162.90	30.46			

altitude, indicating that the existing genetic variation within the species was distributed in all growing regions regardless of the wide range of altitudes of collection (1400–2685 m asl) (Table 1).

The extent of diversity per population generated from our data could be explained in terms of the maximum possible level of diversity to be attained for biallelic loci from 10 individuals per population. Under this condition the maximum possible value for Shannon diversity is 1.00, and it is attained when the frequency of band presence and band absence is 0.5 each. However, the highest per population value obtained in our study was 0.518, and the overall mean was 0.400 (40% of the maximum possible value) (Table 2). Using the same formula Birmeta et al. (2004) reported an overall mean of 0.630 for wild enset (*Ensete ventricosum* (Welw.) Cheesman) while Bussell (1999), reported an overall mean of 0.043 for *Isotoma petraea* F. Muell. Similarly, the maximum Nei's gene diversity to be attained under our condition is 0.538, which could be achieved when the frequency of the null allele and present allele is 0.471 and 0.529, respectively. However, the highest gene diversity recorded in our study was 0.245, and the overall mean was 0.176 (32.7% of the maximum value) (Table 2). Thus, comparing the extent of genetic diversity revealed in our study with these maximum possible values and other RAPD based studies led us to the conclusion that the extent of genetic diversity in Ethiopian niger might be sufficient enough for producing varieties of great interest through breeding.

Population differentiation

The overall mean of population differentiation calculated from Shannon diversity ($\overline{G'}_{ST}$) and from gene diversity ($\overline{G}_{ST}$) was 0.436 and 0.242 respectively (Table 2). Additionally, the mean F_{ST} value obtained from AMOVA was 0.350 which is highly significant ($P<0.00001$) (Table 2). ($\overline{G'}_{ST}$) was revealed to be higher than $\overline{G}_{ST}$ and mean F_{ST} obtained with AMOVA, which is in agreement with Nybom et al. (2001). On the other hand, F_{ST} was higher than G_{ST} unlike the report by Nybom (2004) who indicated that G_{ST}, and F_{ST} obtained by AMOVA usually produce very similar estimates when applied to the same plant material using the same set of marker data. The G'_{ST} estimates revealed in this study was about average when compared with other RAPD based G'_{ST} estimates for several outcrossing species (see Bussell 1999). Similarly, the G_{ST} (0.242) revealed by our study was found to be equivalent to the mean G_{ST} (0.23) obtained for 18 outcrossing species as reported in a review of RAPD based studies (Nybom and Bartish 2000).

The relatively low level of population differentiation observed between niger populations would therefore seem to result from a high level of genetic variability maintained by outcrossing nature of the plant, which is in agreement with the general understanding that outcrossing species tend to be more diverse within, with less genetic differentiation between populations (Hamrick and Godt 1996), which was also concluded specifically from RAPD based studies (Nybom 2004).

Genetic distance and cluster analysis

Nei's standard genetic distance between pair of populations ranged from 0.05 (Gr-6 vs. Gr-7 and Gr-3 vs. Gr-2) to 0.30 (S-9 vs. A-1) with the mean of 0.176 (see Fig. 2). The cluster analysis was tested for its goodness of fit to genetic distance estimates. The cophenetic correlation between the genetic distance matrix and its cophenetic distance matrix was found to be 0.751, which indicates that the goodness of fit of the cluster analysis to genetic distance estimates is not good, as described in Rohlf (2000). However, it does not mean that clustering is not possible, but only indicates that some distortion might have occurred (Mohammadi and Prasanna 2003). This low cophenetic correlation coefficient might be due to larger number of populations used for clustering as this value decreases as the number of populations increases up to 50 (Rohlf and Fisher 1968).

The UPGMA clustering, based on Nei's standard genetic distance, for the 70 populations revealed three major clusters (III, IV, V), three minor clusters (I, II,VI) and one solitary (VII) at mean genetic distance of 0.176 (Fig. 2). This clustering pattern is interesting in that the majority of the populations from the same region were clustered together. Furthermore, populations from adjacent regions were clustered together to a considerable degree (Fig. 3). For example, all the 19 populations from the two neighboring regions (Gojam and Gonder) clustered

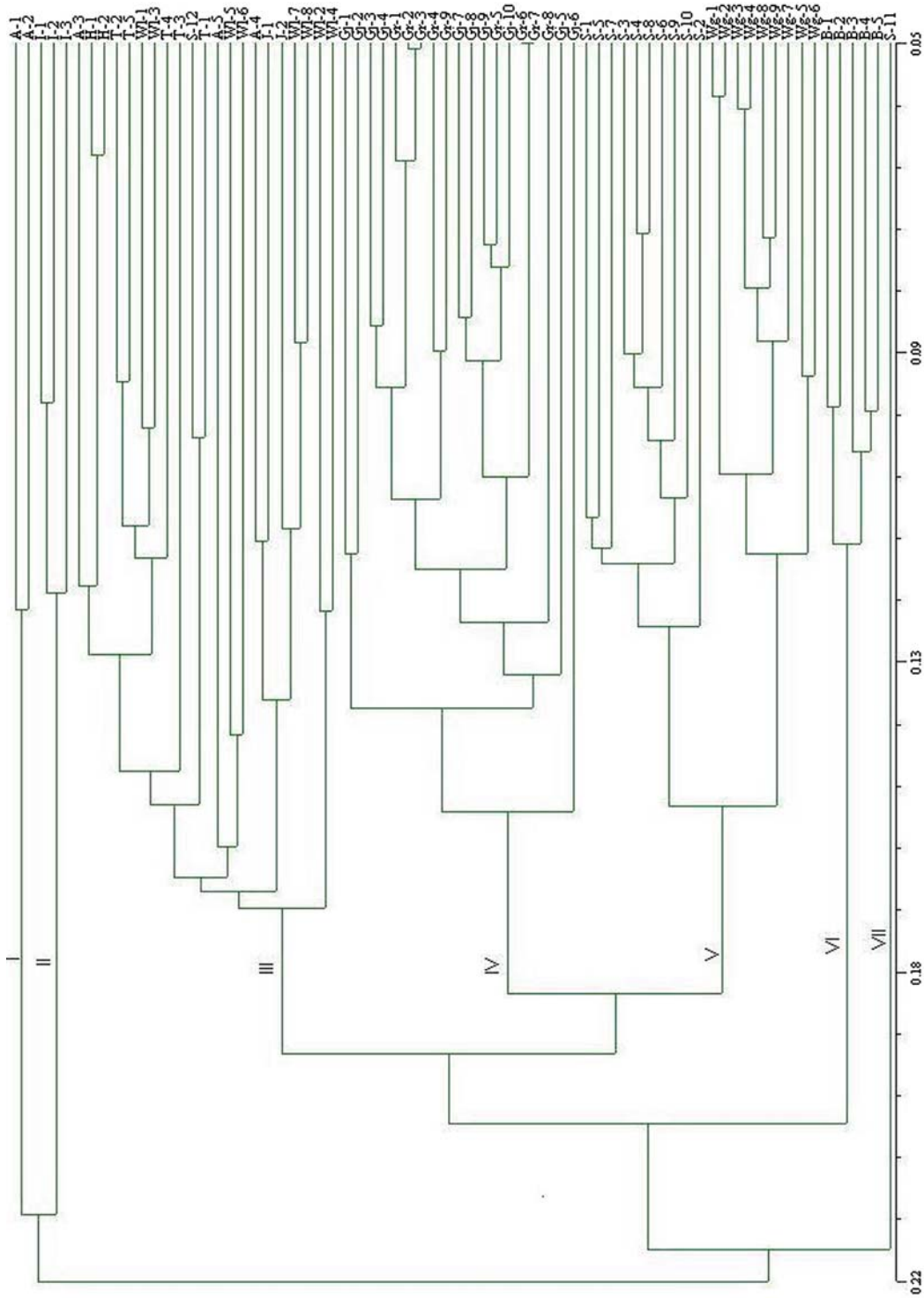

Fig. 2 UPGMA Phenogram of the 70 populations of *G. abyssinica* based on Nei's standard genetic distance estimated from RAPD band profiles

together (Cluster IV). In such cases it seems reasonable to suggest that geographic proximity contribute a lot to genetic similarity between populations of adjacent regions, due to high chance of gene flow. However, this was not always the case as reflected in cluster III that contains populations from geographically distant regions. The clustering together of populations from geographically distant regions might be due to the movement of niger populations together with movement of human populations to new settlement areas in the past. Cluster V contained all populations from Welega and 10 of the 12 populations from Shewa, which were completely separated into different clusters at a genetic distance of about 0.150.

Nei's standard genetic distance was also calculated by grouping populations into their respective regions. It ranged from 0.028 (Gojam vs. Gonder, which are neighbor regions) to 0.157 (Harerge vs. Illubabor, located east and south-west of the country respectively) with the mean of 0.092. Thus, genetic distances between regions was found to be lower than that of between populations, which is also supported by AMOVA due to the fact that the contribution of among regions variation to the total variation was 13.68% while the contribution of among populations within region variation was 22.64% (Table 5). Bootstrap analysis with 500 repetitions generated 285 different trees, with the most frequent tree and the original tree generated 19 and 13 times, respectively. Bale and Illubabor regions were separated from the rest regions at higher genetic distances, with Illubabor region separating first, which further strengthens the importance of conserving populations from minor niger producing regions as described above.

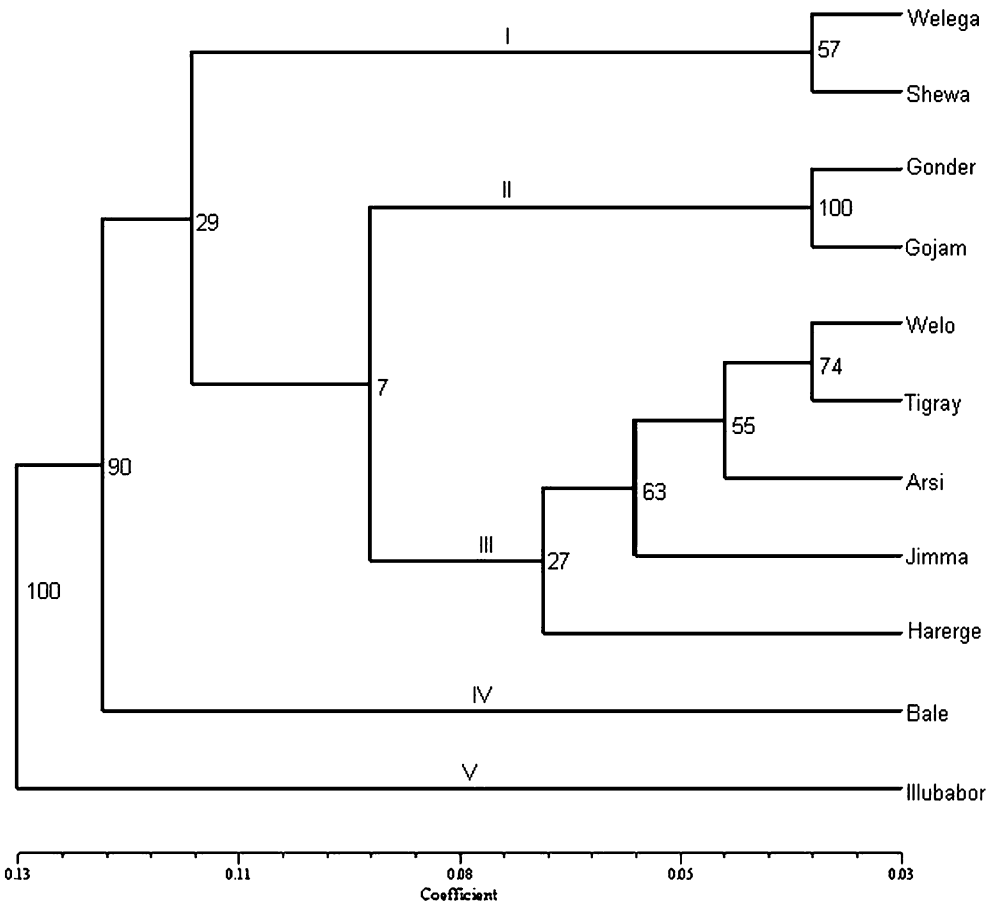

Fig. 3 UPGMA Phenogram of *G. abyssinica* of the 11 regions based on Nei's standard genetic distance. Numbers near branches are bootstrap values

Generally, Ethiopian niger seems to have a wide genetic basis, regardless of altitude and extent of cultivation that makes it suitable for the adaptation to diverse environmental conditions as the chance of finding adaptive genotypes is high due to its high percentage of polymorphic loci. Some of these polymorphic loci might be linked to economically important traits, and can thus be used for marker assisted selection (MAS), which is much more powerful than conventional breeding methods (Snowdon and Friedt 2004). Populations from some regions such as Bale and Illubabor, where niger cultivation seems to be declining, were found to be at a relatively higher genetic distance from other populations and have higher genetic variation within populations. However, these regions are underrepresented in the genebank collection. Thus, future germplasm collection mission should give special attention for such regions to further broaden the genetic basis of genebank collection and to increase the chance of conserving important genotypes that can be used for breeding programs.

Acknowledgements The authors wish to thank The Swedish International Development Agency (SIDA/SAREC) for financing this work and International Science Program (ISP) for coordinating the program and facilitating the working atmosphere for the first author. We are grateful to Mrs. Britt Green for her great assistance in the laboratory work.

References

Aga E, Bryngelsson T, Bekele E, Salomon B (2003) Genetic diversity of forest arabica coffee (*Coffea arabica* L.) in Ethiopia as revealed by random amplified polymorphic DNA (RAPD) analysis. Hereditas 138:36–46

Alemaw G, Teklewold A (1995) An agronomic and seed-quality evaluation of noog (*Guizotia abyssinica* Cass.) germplasm from Ethiopia. Plant Breed 114:375–376

Baagøe J (1974) The genus *Guizotia* (Compositae). A taxonomic revision. Bot Tidsskr 69:1–39

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

Buckler ES IV, Thornsberry JM (2002) Plant molecular diversity and applications to genomics. Curr Opin Plant Biol 5:107–111

Bussell JD (1999) The distribution of random amplified polymorphic DNA (RAPD) diversity amongst populations of *Isotoma petraea* (Lobeliaceae). Mol Ecol 8:775–789

Dagne K (1994) Meiosis in interspecific hybrids and genomic interrelationships in *Guizotia* Cass. (Compositae). Hereditas 121:119–129

Dagne K (1995) Karyotypes, C-banding and nucleolar numbers in *Guizotia* (Compositae). Plant Syst Evol 195:121–135

Dagne K (2001) Cytogenetics of new *Guizotia* Cass. (Compositae), interspecific hybrids pertaining to genomic and phylogenetic affinities. Plant Syst Evol 230:1–11

Dagne K, Jonsson A (1997) Oil content and fatty acid composition of seeds of *Guizotia* Cass. (Compositae). J Sci Food Agric 73:274–278

Dutta PC, Helmersson S, Kebedu E, Alemaw G (1994) Variation in lipid composition of niger seed (*Guizotia abyssinica* Cass.) samples collected from different regions in Ethiopia. JAOCS 71:839–843

Geleta M, Asfaw Z, Bekele E, Teshome A (2002) Edible oil crops and their integration with the major cereals in North Shewa and South Welo, central Highlands of Ethiopia: an ethnobotanical perspective. Hereditas 137:29–40

Genet T, Belete K (2000) Phenotypic diversity in the Ethiopian noug germplasm. Afr Crop Sci J 8:137–143

Getinet A, Sharma SM (1996) Niger (*Guizotia abyssinica* (L. f.) Cass. Promoting the conservation and use of underutilized and neglected crops. 5. Institute of Plant Genetics and Crop Plant Research, Gatersleben/International Plant Genetic Resources Institute, Rome

Hamrick JL, Godt MJW (1996) Effects of life history traits on genetic diversity in plant species. Philos T Roy Soc B 351:1291–1298

Harlan JR (1969) Ethiopia: a centre of diversity. Econ Bot 23:309–314

Hiremath SC, Murthy HN (1988) Domestication of niger (*Guizotia abyssinica*). Euphytica 37:225–228

Hiremath SC, Murthy HN (1992) Cytological studies in *Guizotia* (Asteraceae). Caryologia 45:69–82

Huff DR, Peakall R, Smouse PE (1993) RAPD variation within and among natural populations of outcrossing buffalograss [*Buchloe dactyloides* (Nutt.) Engelm.]. Theor Appl Genet 7:927–934

Kandel H, Porter P (eds) (2002) Niger (*Guizotia abyssinica)* (L. f.) Cass. Production in northwest Minnesota. University of Minnesota Extension Service

Lynch M, Milligan BG (1994) Analysis of population genetic structure with RAPD markers. Mol Ecol 3:91–99

Mohammadi SA, Prasanna BM (2003) Analysis of genetic diversity in crop plants-salient statistical tools and considerations. Crop Sci 43:1235–1248

Murthy HN, Hiremath SC, Salimath SS (1993) Origin, evolution and genome differentiation in *Guizotia abyssinica* and its wild species. Theor Appl Genet 87:587–592

Nayakar NY (1976) Genetic variability and heritability for six quantitative characters in niger (*Guizotia abyssinica* Cass). Mysore J Agric Sci 10:553–558

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

Nei M (1973) Analysis of gene diversity in subdivided populations. Proc Natl Acad Sci USA 70:3321–3323

Nei M (1978) Estimation of average hetrozygosity and genetic distance from a small number of individuals. Genetics 89:583–590

Nemomissa S, Bekele E, Dagne K (1999) Self-incompatibility system in Ethiopian populations of *Guizotia abyssinica* (L. f.) Cass. (niger). Sinet: Ethiop J Sci 22:67–88

Nybom H (2004) Comparison of different nuclear DNA markers for estimating intraspecific genetic diversity in plants. Mol Ecol 13:1143–1155

Nybom H, Bartish IV (2000) Effects of life history traits and sampling strategies on genetic diversity estimates obtained from RAPD markers in plants. Perspect Plant Ecol 3:93–114

Nybom H, Werlemark G, Olsson ÅME (2001) Between and within population diversity in dogrose species. Acta Hortic 546:139–144

Page RDM (1996) TREEVIEW. An application to display phylogenetic trees on personal computers. Comput Appl Biosci 12:357–358

Pavlicek A, Hrda S, Flegr J (1999) FreeTree – Freeware program for construction of phylogenetic trees on the basis of distance data and bootstrap/jackknife analysis of the tree robustness. Application in the RAPD analysis of the genus *Frenkelia*. Folia Biol, Prague 45:97–99

Pradhan K, Mishra RC, Paikary RK (1995) Genetic variability and character association in niger. Indian J Genet Pl Br 55:457–459

Ramadan MF, Mörsel J-T (2002) Proximate neutral lipid composition of niger (*Guizotia abyssinica* Cass.) seed. Czech J Food Sci 20:98–104

Riley KW, Belayneh H (1989) Niger. In: Röbbelen G, Downey RK, Shri A (eds) Oil crops of the world. McGraw Hill Publishing Company, New York, pp 394–403

Rohlf FJ (2000) NTSYSpc, Numerical taxonomy and multivariate analysis system. Version 2.1. Exeter Software, Setauket, New York, USA

Rohlf FJ, Fisher DR (1968) Tests for hierarchical structure in random data sets. Syst Zool 17:407–412

Schneider S, Roessli D, Excoffier L (2000) Arlequin: a software for population genetics data analysis. Ver 2.000. Genetics and Biometry Lab, Dept. of Anthropology, University of Geneva

Sivolap YM, Solodenko AE, Burlov VV (1998) RAPD analysis of molecular-genetic polymorphism in sunflower *Helianthus annuus*. Genetika 34:266–271

Skroch P, Neinhuis J (1995) Impact of scoring error and reproducibility of RAPD data on RAPD based estimates of genetic distance. Theor Appl Genet 91:1086–1091

Snowdon RJ, Friedt W (2004) Molecular markers in *Brassica* oilseed breeding: current status and future possibilities. Plant Breed 123:1–8

Yeh FC, Boyle TJB (1997) Population genetic analysis of codominant and dominant markers and quantitative traits. Belg J Bot 129:157

Zeven AC, DeWet JMJ (1982) Dictionary of cultivated plants and their regions of diversity. International Book Distributors, Dehra Dun, India