© NIAB 2008
ISSN 1479-2621

Plant Genetic Resources: Characterization and Utilization 6(1); 41–51
DOI: 10.1017/S1479262108913903

Assessment of genetic diversity of *Guizotia abyssinica* (L.f.) Cass. (Asteraceae) from Ethiopia using amplified fragment length polymorphism

Mulatu Geleta[1,2]*, Tomas Bryngelsson[1], Endashaw Bekele[2] and Kifle Dagne[2]

[1] *Department of Plant Protection Biology, Swedish University of Agricultural Sciences, PO Box 44, SE 230 53 Alnarp, Sweden and* [2] *Department of Biology, Addis Ababa University, PO Box 1176, Addis Ababa, Ethiopia*

Received 25 June 2006; Accepted 28 March 2007

Abstract

Seventeen populations of niger (*Guizotia abyssinica*), representing all regions in Ethiopia where this crop is grown, were investigated using the amplified fragment length polymorphism (AFLP) technique, in order to determine the extent and distribution of its genetic diversity. A total of 539 AFLP loci were scored using seven primer combinations applied to 170 individual plants. Of these, 90% were polymorphic and all the individuals investigated were genetically unique. Despite the fact that most of the variation was within populations, populations were differentiated at a significant level (analysis of molecular variance; $P < 0.001$). There were no significant differences between populations in relation to the extent and altitude of cultivation. A significant positive correlation was revealed between Nei's standard genetic distance and geographic distance. Cluster analysis and principal coordinate analysis revealed that populations from the same regions were clustered together in most cases. Further collection of niger germplasm from areas underrepresented in gene bank collections is recommended.

Keywords: AFLP; AMOVA; genetic diversity; germplasm; *Guizotia abyssinica*; niger; population differentiation

Introduction

Guizotia abyssinica (L.f.) Cass. is an economically important edible oilseed crop that belongs to the tribe Heliantheae in the family Asteraceae. This crop, commonly known as 'noug' (in Amharic) or 'niger' (in English), is the only domesticated species of the small Afro-montane endemic and exclusively diploid ($2n = 30$) genus *Guizotia* (Baagøe, 1974; Hiremath and Murthy, 1992; Dagne, 1995). It is an annual crop widely cultivated in Ethiopia and India (Riley and Belayneh, 1989; Getinet and Sharma, 1996), and also on a small scale in several other African and Asian countries, as an edible oil crop (Murthy *et al.*, 1993; Getinet and Sharma, 1996), and in the USA, mainly as a component of birdseed (Kandel and Porter, 2002). Additionally, the oil is used for various industrial purposes such as soaps, paints, illuminants and lubricants (Baagøe, 1974; Riley and Belayneh, 1989; Kandel and Porter, 2002), and for cultural and medicinal purposes (Geleta *et al.*, 2002). Niger seed oil is rich in linoleic acid (Dutta *et al.*, 1994; Alemaw and Teklewold, 1995; Dagne and Jonsson, 1997; Ramadan and Mörsel, 2002). The species is strictly outcrossing via a self-incompatibility mechanism(s) (Riley and Belayneh, 1989; Nemomissa *et al.*, 1999), and is pollinated mainly by insects (particularly bees) (Fichtl and Adi, 1994; Geleta *et al.*, 2002). Some evidence has been presented which indicates that the domesticated species originated from *G. scabra* ssp. *schimperi* through selection and further

* Corresponding author. E-mail: gemulat@yahoo.com; E-mail: Mulatu.Geleta.Dida@ltj.slu.se

https://doi.org/10.1017/S1479262108913903 Published online by Cambridge University Press

cultivation (Baagøe, 1974; Hiremath and Murthy, 1988; Murthy *et al.*, 1993; Dagne, 1994, 1995, 2001).

Ethiopia is the centre of origin and diversity of niger (Harlan, 1969; Zeven and de Wet, 1982), where it has the longest history of cultivation (Baagøe, 1974; Hiremath and Murthy, 1988). It can be cultivated on waterlogged, marginal and poor soils where most other crops fail to grow (Getinet and Sharma, 1996), because of its ability to withstand salinity and anoxia (Abebe *et al.*, 1978). It is also known for its suitability for multiple cropping, especially as a border crop in fields of other crops (Geleta *et al.*, 2002). In Ethiopia, it is grown mainly from 1600 to 2200 m above sea level (asl), where the range in temperature is 15–23°C, and in annual rainfall is 500–1000 mm (Getinet and Sharma, 1996). Niger occupies about 50% of the total oil crop area and production volume in Ethiopia, where its production is mainly based on local landraces in need of genetic improvement in terms of various traits such as seed yield, seed oil content, oil fatty acid composition and pest/disease resistance. Such genetic improvement through breeding depends on the magnitude of genetic diversity and the extent to which this diversity is utilized. The few genetic diversity studies published to date are based on morphological characterization (Nayakar, 1976; Alemaw and Teklewold, 1995; Pradhan *et al.*, 1995; Genet and Belete, 2000), and the only published marker-based study used random amplified polymorphic DNA (RAPD) (Geleta *et al.*, 2007).

The present study was undertaken to investigate the extent of genetic variation within and among populations of Ethiopian niger using the amplified fragment length polymorphism (AFLP) marker technique. It also set out to identify any hotspots of diversity that may be important in optimizing both conservation strategies and the utilization of existing genetic diversity.

Materials and methods

Plant material and DNA extraction

Seventeen populations of niger were collected from farmers' fields in Ethiopia. A single farmer's field was considered as a population, and each population was represented by a single seed collected from each of ten individual plants. The populations were sampled from a wide range of altitudes (1590–2550 m asl) representing all regions where niger is currently cultivated. Each region was represented by at least one population (Fig. 1, see also Table 1). Seeds were grown in a greenhouse and fresh 15- to 30-day-old leaves were used for genomic DNA extraction, using the modified cetyl trimethyl ammonium bromide (CTAB) procedure described in Assefa *et al.* (2003).

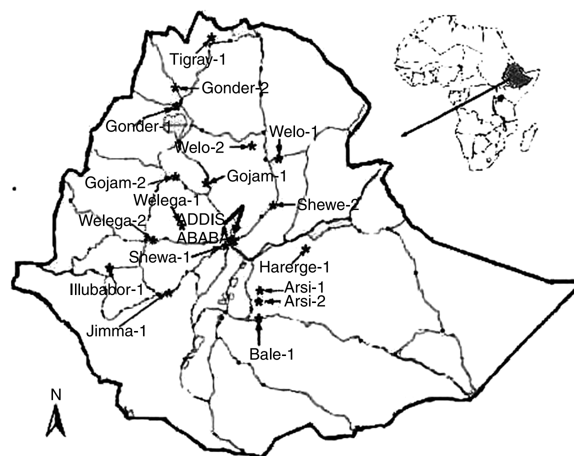

Fig. 1. The collection sites of the niger populations used in this study.

Amplified fragment length polymorphism

AFLP analysis was performed according to Vos *et al.* (1995) with modifications as follows. Genomic DNA (1 μg) was sequentially digested first with 5 U *Mse*I at 65°C for 1 h and then with 5 U *Eco*RI at 37°C for 90 min in a volume of 50 μl in 6.6 mM Tris-acetate, pH 7.9, 2 mM magnesium acetate, 13.2 mM potassium acetate and 20 ng/μl bovine serum albumin (BSA). Ligation was effected in a 10 μl mixture of 0.5 μmol *Eco*RI adapter, 6 μmol *Mse*I adapter, 50 mM Tris–HCl, pH 7.6, 10 mM $MgCl_2$, 1 mM adenosine triphosphate (ATP), 1 mM dithiothreitol (DTT), 5% (w/v) polyethylene glycol-8000 and 1 U T_4 DNA ligase, incubated for 3 h at 37°C. The product was diluted 1:2.3 with $T_{10}E_{0.1}$ (10 mM Tris–HCl, pH 8.0, 0.1 mM EDTA), of which 10 μl was used as a template in a 20 μl pre-amplification reaction containing 20 mM Tris–HCl, pH 8.55, 16 mM $(NH_4)SO_4$, 0.01% Tween20 and 2 mM $MgCl_2$, 30 ng each of *Eco*RI-A and *Mse*I-C primers, 0.2 mM dNTP, 1.5 mM $MgCl_2$ and 0.5 U Thermowhite *Taq* DNA polymerase (Saveen Werner AB). The reaction was subjected to 20 cycles of 92°C/1 min, 60°C/30 s and 72°C/1 min, diluted 1:25 with $T_{10}E_{0.1}$, and used as a template for selective amplification.

Seven selective primer combinations (PCs) were selected (see Table 2) from a set of 56, on the basis that they detected sufficient polymorphism and generated amplification profiles which were easy to score. The selective amplification primers carried three selective nucleotides (SNs). The 20 μl amplification reaction contained polymerase chain reaction (PCR) buffer (as above), 25 ng *Eco*RI primer + three SNs, 30 ng *Mse*I

https://doi.org/10.1017/S1479262108913903 Published online by Cambridge University Press

Table 1. Population codes, geographic description, number of polymorphic loci (NPL), percentage polymorphic loci (PPL), number of unique alleles (NUA), mean frequency of unique alleles (MFUA), H'_{loci} and H_j for the populations of *G. abyssinica*

Population code	Region of collection	Altitude of collection (m)	Place of collection	NPL	PPL	NUA	MFUA	H'_{loci}	H_j
A-1*	Arsi	2380	52 km from Bekoji to Chole	280	52.0	2	0.15	0.438	0.193
A-2*	Arsi	2470	13 km from Robe to Ticho	255	47.3	0	0.00	0.406	0.189
B-1*	Bale	2450	2.7 km from Adaba to Dodola	272	50.5	3	0.10	0.432	0.209
Gj-1**	Gojam	2550	75 km from Mota to Dejen	307	57.0	9	0.18	0.481	0.228
Gj-2**	Gojam	1890	35 km from Amanuel to Bure	270	50.1	0	0.00	0.425	0.203
Gr-1**	Gonder	2055	26 km from Gondar to K/Diba	273	50.7	1	0.10	0.432	0.200
Gr-2**	Gonder	1590	9 km from T/Dingay to Humera	262	48.6	0	0.00	0.400	0.190
H-1*	Harerge	1830	36 km from A/Teferi to Gelemso	271	50.3	2	0.15	0.426	0.204
I-1*	Illubabor	1865	3 km from Metu to Gore	297	55.1	0	0.00	0.477	0.226
J-1*	Jimma	1860	25 km from Jimma to A/Ababa	296	54.9	1	0.10	0.464	0.222
Sh-1**	Shewa	2155	18 km from A/Ababa to Weliso	289	53.6	1	0.10	0.449	0.202
Sh-2**	Shewa	1640	13 km from Sh/Robit to D/Sina	250	46.4	1	0.60	0.392	0.186
T-1*	Tigray	1972	9.5 km Shire to Shiraro	249	46.2	2	0.45	0.406	0.183
Wg-1**	Welega	2370	12 km from Fincha'a to Kombolcha	266	49.4	1	0.10	0.412	0.196
Wg-2**	Welega	1940	21 km from Nekemt to Ambo	279	51.8	1	0.20	0.439	0.208
Wl-1*	Welo	1650	31 km from Kombolcha to Bati	323	59.9	1	0.20	0.500	0.241
Wl-2*	Welo	2420	31 km from W/Tena to Gashena	261	48.4	0	0.00	0.395	0.188
Mean*				278.2	51.6	1.22	0.13	0.438	0.206
Mean**				274.5	51.0	1.75	0.16	0.429	0.202
Mean[a]				276.5	51.3	1.47	0.14	0.430	0.200
Total[b]				483	89.6			0.628	0.320

*Populations collected from minor niger-producing regions (MiNPR) and their mean values for different parameters.

**Populations collected from major niger-producing regions (MaNPR) and their mean values for different parameters.

[a] Mean of each parameter for the 17 populations (corresponding value of PPL is P_P).

[b] Corresponding values of each parameter when all individuals were considered together (corresponding value of PPL is P_S).

https://doi.org/10.1017/S1479262108913903 Published online by Cambridge University Press

Table 2. Summary of number of bands scored (NBS), number of polymorphic loci (NPL), percentage polymorphic loci (PPL), number of unique alleles (NUA), mean frequency of unique alleles (MFUA), various Shannon and gene diversity parameters estimates and AMOVA for seven AFLP primer combinations (PCs) and their corresponding means and total

PCs	NBS	NPL	PPL	NUA	MFUA	Shannon diversity estimates				Gene diversity estimates				AMOVA
						H'_{POP} [b]	H'_{SP} [b]	H'_{POP}/H'_{SP} [b]	G'_{ST} [a]	H_S [b]	H_T [b]	H_S/H_T [b]	G_{ST} [a]	F_{ST}
[c]E-AAG/[d]M-CTC	80	71	88.8	3	0.13	0.385	0.657	0.585	0.415	0.183	0.331	0.646	0.354	0.357
[c]E-ACA/[d]M-CTA	77	68	88.3	6	0.15	0.456	0.699	0.635	0.365	0.219	0.351	0.684	0.316	0.302
[c]E-ACA/[d]M-CTG	78	69	88.5	4	0.25	0.470	0.646	0.697	0.303	0.219	0.329	0.744	0.256	0.198
[c]E-ACG/[d]M-CTC	74	64	86.5	4	0.12	0.398	0.556	0.691	0.309	0.197	0.309	0.744	0.256	0.194
[c]E-ACT/[d]M-CAC	69	67	97.1	2	0.3	0.467	0.633	0.715	0.285	0.210	0.276	0.830	0.170	0.173
[c]E-AGG/[d]M-CAT	79	65	82.3	3	0.13	0.420	0.574	0.696	0.304	0.216	0.342	0.721	0.279	0.171
[c]E-AGG/[d]M-CTA	82	77	93.9	3	0.37	0.438	0.622	0.678	0.322	0.192	0.306	0.743	0.257	0.212
Overall mean	77	69	89.6	3.57	0.21	0.434	0.628	0.670*	0.330	0.205	0.320	0.731*	0.269	0.234
Total	539	483												

*Mean of values in column.
[a] Highly significant difference in G'_{ST} ($P = 0.000$) and G_{ST} ($P = 0.005$) between primer combinations.
[b] No significant difference between primer combinations as revealed by ANOVA.
[c] E = *Eco*RI primer (5′-GACTGCGTACCAATTC-3′).
[d] M = *Mse*I primer (5′-GATGAGTCCTGAGTAA-3′).

primer + three SNs, 0.2 M dNTP, 1.5 mM $MgCl_2$, 0.5 U *Taq* DNA polymerase and 10 μl diluted pre-amplification product. The amplification profile was 94°C/2 min, followed by 36 cycles as described in Ferdinandez and Coulman (2004), and a final step of 72°C/2 min. The amplified product was denatured by adding 15 μl of 98% formamide, 10 mM EDTA, 0.05% (w/v) each of bromophenol blue and xylene cyanol FF, and incubating at 96°C for 5 min. Seven microlitres of the amplification product was loaded on 5% (w/v) polyacrylamide gels and separated at 90 W constant power until the xylene cyanol FF dye had run two-thirds of the length of the plate. Before loading the samples, the gel was pre-run for 45 min. DNA bands were visualized using the silver staining technique of Caetano-Anollés and Gresshoff (1994) with the following modifications: (1) 10% acetic acid was used as fixer solution and stopping solution; and (2) the concentration of sodium thiosulphate in the developing solution was 2 mg/l.

Data scoring and analysis

Each AFLP fragment was considered as a single bi-allelic locus with one amplifiable and one null allele. Data were recorded as 1 for the presence and 0 for the absence of each amplified fragment in the size range 50–600 bp. Gels were routinely scored twice. Genetic diversity was calculated based on: (1) Shannon diversity index and (2) Nei's gene diversity with the modification provided by Lynch and Milligan (1994), as described in Geleta *et al.* (2007), using polymorphic loci only. Gene flow was estimated using Wright's (1951) equation, as modified by Crow and Aoki (1984). The NTSYSpc program (Rohlf, 2000) was used to calculate genetic distances, matrix correlation coefficients, the Mantel test, and to perform cluster and principal coordinate analyses. POPGENE version 1.31 (Yeh and Boyle, 1997) was used for analysis of number and percentage of polymorphic loci. Analysis of molecular variance (AMOVA) was conducted using Arlequin version 2 (Schneider *et al.*, 2000), and the FreeTree-Freeware program (Pavlicek *et al.*, 1999) was used for bootstrap analysis.

Results

Genetic polymorphism and AFLP primer combinations

A total of 539 fragments were detected among the 170 individual plants, of which 483 (over 89%) were polymorphic (Fig. 2, Table 1). The number of polymorphic loci averaged 69 per primer combination (PC). When each population was considered separately, the

https://doi.org/10.1017/S1479262108913903 Published online by Cambridge University Press

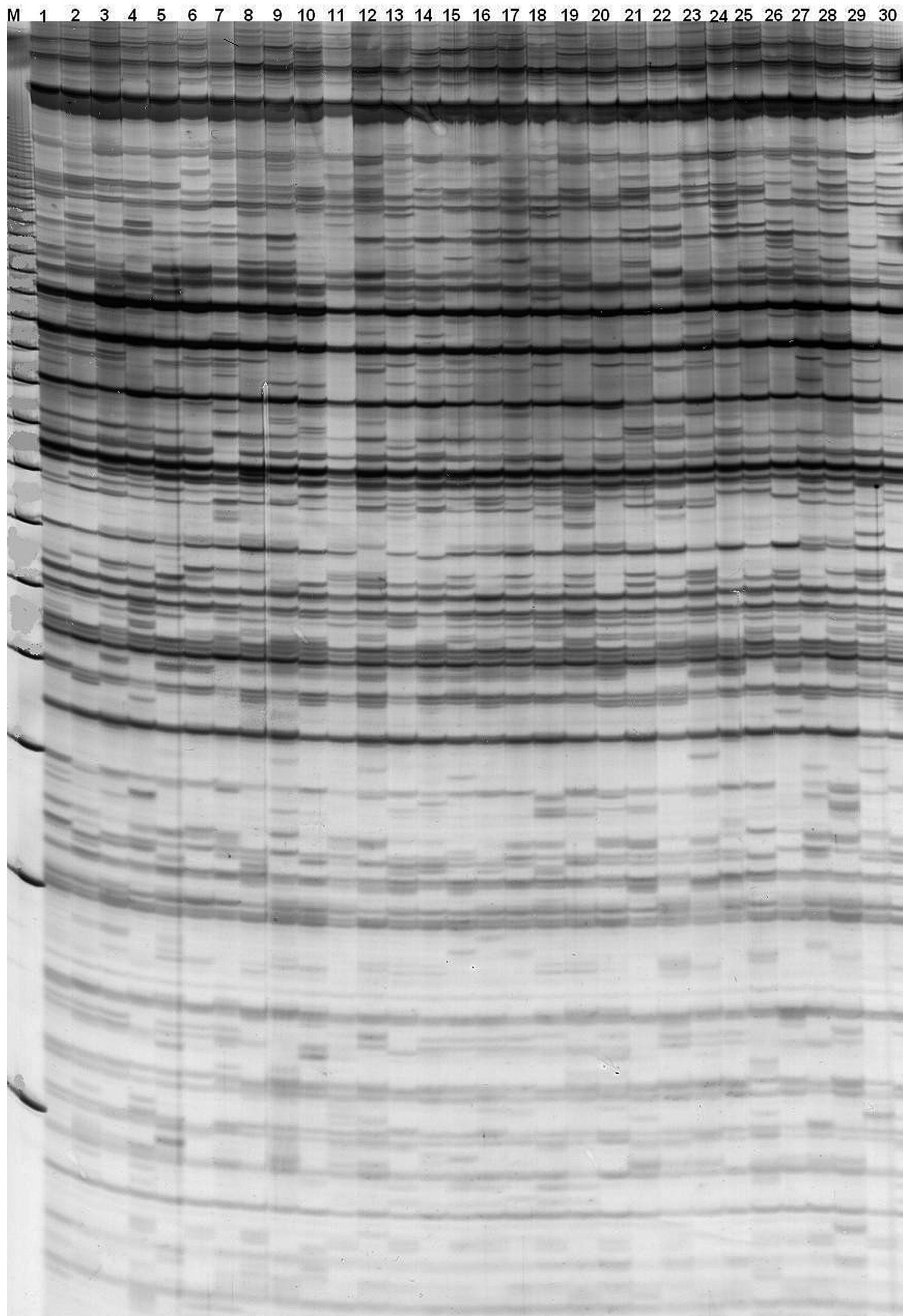

Fig. 2. Amplified fragment length polymorphism profiles generated by primer combination E-ACA/M-CTA for three populations. M, Molecular size marker; individuals of WI-1, WI-2 and I-1 populations are represented in lanes 1–10, 11–20 and 21–30, respectively.

https://doi.org/10.1017/S1479262108913903 Published online by Cambridge University Press

percentage of polymorphic loci ranged from 46% (T-1) to 60% (Wl-1) with the mean (P_P) of about 51% (Table 1). Comparisons of profiles of individuals across all loci revealed that each individual was genetically unique, implying the presence of a high level of genetic polymorphism.

AFLP PCs used in this study were significantly different in the number and percentage of polymorphic loci (NPL and PPL) they detected ($P < 0.01$). Of the seven PCs, the highest mean number of polymorphic loci (77) was revealed by E-AGG/M-CTA, while the highest number of unique alleles specific to a given population was revealed by E-ACA/M-CTA (Table 2). There was no significant difference between PCs in terms of total (H'_{SP} and H_T) and within-population (H'_{POP} and H_S) genetic variation. Contrary to this, there was a significant difference in G'_{ST} ($P < 0.001$) and G_{ST} ($P < 0.01$) between the PCs. Among the PCs, E-AAG/M-CTC revealed the highest estimate of population differentiation ($G'_{ST} = 0.415$; $G_{ST} = 0.354$; $F_{ST} = 0.357$; Table 2).

Total and within-population genetic variation

The overall genetic diversity estimated by Shannon diversity index as $\bar{H}'_{SP}$ and gene diversity estimate (Nei, 1978) as $\bar{H}_T$ was 0.628 and 0.320, respectively. Similarly, the overall within-population variation estimated by Shannon diversity index ($\bar{H}'_{POP}$) and Nei's gene diversity estimate ($\bar{H}_S$) were 0.434 and 0.205, respectively (Table 2). The extent of genetic diversity of each population was calculated using Shannon diversity and gene diversity estimates as H'_{loci} and H_j, respectively, which are the average values across the whole polymorphic loci. H'_{loci} ranged from 0.392 (Sh-2) to 0.500 (Wl-1), while H_j ranged from 0.183 (T-1) to 0.241 (Wl-1) (Table 1). Taking the two parameters into consideration, Sh-2, T-1 and Wl-2 showed lower genetic diversity as compared to other populations, while Wl-1 showed the highest genetic diversity, followed by Gj-1.

The evaluation of the AFLP fingerprints revealed unique alleles in 12 of the populations (Table 1). We grouped the 17 populations according to the major (MaNPR) and minor (MiNPR) niger-producing regions to determine whether there is any significant difference in the level of genetic variation between them. The mean H'_{loci} and H_j for populations from MiNPR were 0.438 and 0.206, respectively, while these parameters were 0.430 and 0.200, respectively, for populations from MaNPR, indicating a similar level of genetic variation in both groups.

Genetic variation between populations and groups

The population differentiation was calculated as G'_{ST} from Shannon diversity index, as G_{ST} from gene diversity estimates (Nei, 1973) and as F_{ST} from AMOVA, which resulted in the overall corresponding means of 0.330 ($\bar{G}'_{ST}$), 0.269 ($\bar{G}_{ST}$) and 0.234 ($\bar{F}_{ST}$) (Table 2). AMOVA revealed that the observed genetic variation among populations is highly significant ($P < 0.001$; Table 3A). On the other hand, AMOVA showed that the genetic variation between MaNPR and MiNPR populations was less than 1% of the total variation ($P > 0.100$) (Table 3B). Similarly, AMOVA conducted by grouping the populations into a higher altitude group (>2000 m asl) and a lower altitude group (<2000 m asl) revealed no significant difference between them (Table 3C). We also grouped populations into five groups based on their geographic proximity and better access to germplasm exchange (Table 3D), where AMOVA revealed significant variation between the groups (7.5%; $P < 0.001$). The presence of unique alleles in each group contributed to the significant variation obtained. For example, 12 unique alleles were recorded in group II (Gj-1 and Gj-2) with frequencies ranging from 0.05 to 0.35 (data not shown). The estimate of gene flow (Nm), calculated based on AMOVA-derived F_{ST}, was 0.924 (Table 3).

Genetic distance, cluster analysis and principal coordinate analysis (PCoA)

The significant population differentiation was further analysed using genetic distance coefficient and multivariate analyses to identify populations that are more differentiated from the majority and to reveal their clustering pattern. Nei's standard genetic distance coefficient (Nei, 1972) was used to evaluate the extent of genetic similarity between each pair of populations. A more than fourfold variation in genetic distance between pairs of populations that ranged from 0.040 (Gr-1 versus Gr-2; Gr-2 versus Wg-1) to 0.175 (H-1 versus T-1) was obtained, with the overall mean genetic distance of 0.118 (Supplementary Table 1, available online only at http://journals.cambridge.org). The comparison of matrices of Nei's standard genetic distances and geographic distances through normalized Mantel statistics (Mantel, 1967) with 1000 permutations revealed a significant positive correlation ($r = 0.258$; $P < 0.01$).

The cophenetic correlation coefficients between genetic distance and its cophenetic value matrix, and between genetic distance and distance matrix calculated from eigen vector matrix were 0.91 and 0.90, respectively (see Fig. 3a, b). Three clusters (I, II and III) were revealed in the unweighted pair group method with arithmetic average (UPGMA) cluster analysis (Fig. 3a), which was also clearly depicted in the PCoA (Fig. 3b). Clusters I, II and III consist of nine, three and five populations, respectively. The bootstrap value for branching of cluster

https://doi.org/10.1017/S1479262108913903 Published online by Cambridge University Press

Table 3. AMOVA for the 17 populations of *G. abyssinica* based on AFLP data: (A) without grouping the populations; (B) by grouping the populations into those from (i) minor niger-producing regions (MiNPR) and (ii) major niger-producing regions (MaNPR); (C) by grouping the populations into (i) higher and (ii) lower altitude groups; and (D) by grouping the populations into five groups based on geographic proximity and access to gene flow

Groups	Source of variation	Degrees of freedom	Variance components	% of variation	Fixation indices	Significance tests (*P* value)	Gene flow (Nm)
A. Without grouping the populations	AP	16	Va = 16.14	23.44	$F_{ST} = 0.234$	Va and $F_{ST} = 0.000$	0.924
	WP	153	Vb = 52.73	76.56			
	Total	169	68.87				
B. Grouping the populations into	AG	1	Va = 0.56	0.82	$F_{ST} = 0.237$	Va and $F_{CT} = 0.195$	
(i) A-1, A-2, B-1, H-1, I-1, J-1,	APWG	15	Vb = 15.84	22.91	$F_{SC} = 0.231$	Vb and $F_{SC} = 0.000$	
T-1, W-1, W-2 and	WP	153	Vc = 52.73	76.27	$F_{CT} = 0.008$	Vc and $F_{ST} = 0.000$	
(ii) Gj-1, Gj-2, Gr-1, Gr-2, Sh-1,	Total	169	69.13				
Sh-2, Wg-1, Wg-2							
C. Grouping the populations into	AG	1	Va = −0.96	−1.40	$F_{ST} = 0.229$	Va and $F_{CT} = 0.923$	
(i) A-1, A-2, B-1, Gj-1, Gr-1, Sh-1,	APWG	15	Vb = 16.65	24.33	$F_{SC} = 0.240$	Vb and $F_{SC} = 0.000$	
Wg-1, Wl-2 and	WP	153	Vc = 52.73	77.07	$F_{CT} = -0.014$	Vc and $F_{ST} = 0.000$	
(ii) Gj-2, Gr-2, H-1, I-1, J-1, Sh-2,	Total	169	68.42				
T-1, Wg-2, Wl-1							
D. Grouping the populations into	AG	4	Va = 5.22	7.48	$F_{ST} = 0.244$	Va and $F_{CT} = 0.000$	
(i) A-1, A-2, B-1, H-1;	APWG	12	Vb = 11.80	16.92	$F_{SC} = 0.183$	Vb and $F_{SC} = 0.000$	
(ii) I-1, J-1, Sh-1, Wg-1, Wg-2;	WP	153	Vc = 52.73	75.59	$F_{CT} = 0.075$	Vc and $F_{ST} = 0.000$	
(iii) Sh-2, W-1, W-2;	Total	169	69.75				
(iv) Gj-1, Gj-2; and							
(v) Gr-1, Gr-2, T							

AP, among populations; WP, within populations; AG, among groups; APWG, among populations within groups.

https://doi.org/10.1017/S1479262108913903 Published online by Cambridge University Press

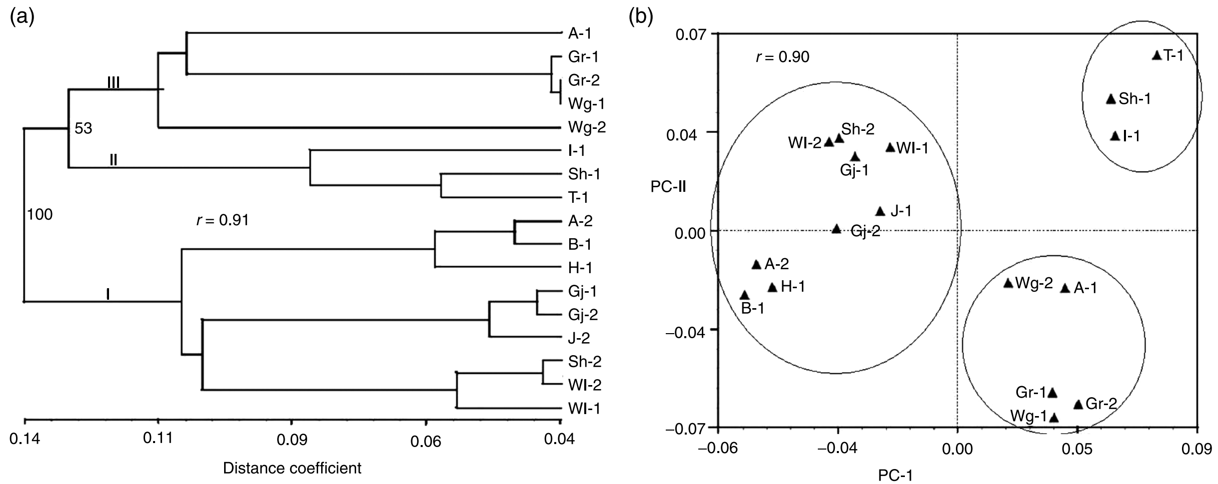

Fig. 3. (a) UPGMA phenogram of the 17 populations of niger based on Nei's (1972) standard genetic distance, estimated from amplified fragment length polymorphism (AFLP) profiles. (b) Two-dimensional plot generated by principal coordinate analysis.

I from the other two clusters was the maximum (100), while the other two clusters were separated from one another with a lower bootstrap value (53; Fig. 3a). In the case of PCoA, the first three principal coordinate axes explained 67% of the total variation in the AFLP data. The first principal coordinate axis (PC-I) explained 34% of the total variation, and the three clusters were more clearly discriminated on this axis (Fig. 3b). The second axis (PC-II) explained 20% of the total variation and was better than PC-I in discriminating populations within clusters.

Discussion

Genetic polymorphism and within-population genetic variation

The overall genetic polymorphism within the species was high, with about half of the loci in each population being polymorphic. This allows for an easy means to distinguish between niger populations and even between individuals within populations. Several different approaches have been used to estimate within-population genetic variation. We used both H'_{POP} and H_{S} in order to broaden the comparison with previous studies. The overall means of these parameters revealed in this study were slightly higher than that obtained from a RAPD-based study (Geleta *et al.*, 2007), indicating the superiority of AFLP over RAPD in detecting genetic variation in niger populations. The proportion of within-population variation reported by Geleta *et al.* (2007) for AMOVA and Shannon diversity index was 65% and 57%, respectively, which is lower than that revealed in the present study (77% and 67%, in that order). Despite the relatively higher proportion of within-population variation revealed by AFLP as compared to RAPD, both marker systems demonstrated that a higher proportion of the total variation is to be found within, rather than between, populations.

The extent of genetic variation in niger populations was wide ranging and the number of unique alleles per population also varied. Despite its highest genetic diversity, only one unique allele with a frequency of 0.2 was recorded in Wl-1 (Table 1). Thus, the highest genetic variation revealed in this population is mainly due to the fact that both alleles were maintained in a relatively higher frequency per locus. Here, it is interesting to note that Wl-1 (designated as Wl-2 in Geleta *et al.*, 2007) showed the highest diversity of the 70 populations studied using RAPD, indicating a reasonable degree of agreement between the two marker systems in detecting the extent of genetic diversity. The second most genetically diverse population revealed in the present study was Gj-1 (Table 1). This population was different from Wl-1 in that it contained several unique alleles revealed by five of the seven PCs. The maximum possible values of H'_{loci} and H_{j} that could be obtained for dominant markers of two alleles at each locus for a population represented by ten individuals are 1.000 and 0.538, respectively (Geleta *et al.*, 2007). Comparing the mean H'_{loci} (0.430) and H_{j} (0.200) obtained in the present study to these maximum possible diversity estimates leads to the conclusion that Ethiopian niger has sufficient genetic variability to be able to breed varieties with desirable traits.

The significantly higher genetic variation in MiNPR populations over MaNPR populations reported for a RAPD-based study (Geleta *et al.*, 2007) was not

https://doi.org/10.1017/S1479262108913903 Published online by Cambridge University Press

supported by the present study. The fact that the extent of genetic diversity was not associated with either the altitude or the extent of cultivation, both in this and in the RAPD-based study, leads to the conclusion that the existing genetic diversity of Ethiopian niger is distributed within all growing regions regardless of the altitude and the extent of cultivation.

Genetic variation and genetic distance between populations and groups

The main evolutionary forces responsible for population differentiation are selection, gene flow and genetic drift, which operate within the historical and biological context of each plant species (Loveless and Hamrick, 1984). Thus, the extent of population differentiation depends on the relative strength of these individual forces in interaction with the type of mating system and other life history traits of the species. Genetic differentiation of populations may occur for any genetically variable trait that is favoured under the existing selection conditions (Bossdorf *et al.*, 2005) and the estimate of such population differentiation can be calculated using various parameters such as G'_{ST}, G_{ST} and F_{ST}. Nybom (2004) analysed eight AFLP-based studies of outcrossing species and obtained a mean F_{ST} and G_{ST} value of 0.23 and 0.24 respectively, which is comparable with that of the present study. Therefore, this study demonstrated an average level of population differentiation with significant variation among populations.

Some loci were polymorphic in only one population but monomorphic in all others. Such unique alleles may serve as population-specific markers in future generations, provided that they are favoured under both natural and artificial selection conditions and that gene flow between populations is limited. The significant genetic variation between groups of populations based on geographical proximity and access of gene flow is strong evidence to suggest a considerable degree of 'regional' differentiation of niger populations. Such population differentiation into ecotypes is important for the selection of parental materials to maximize heterozygosity in the progeny. Furthermore, genetically differentiated populations are often suggested as candidates for genetic conservation to prevent the loss of unique genetic variants. Thus, conserving a large number of populations from all its growing regions *ex situ* in gene banks as a complement to on-farm conservation is the best policy for conserving a high level of unique genetic variants in the gene pool.

Although the populations were differentiated to a significant degree, the among-population variation was less than the within-population variation. Moreover, population-specific monomorphic markers were not detected for all the 483 polymorphic loci. The absence of such population-specific markers is an indication of strong gene flow between niger populations ($Nm = 0.924$; Table 3), most likely through germplasm exchange. The lower proportion of among-population variation as compared to the proportion of within-population variation is likely a result of the high level of genetic variability maintained by the outcrossing nature of the plant, 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; Nybom, 2004).

The overall mean Nei's genetic distance (0.118) revealed in this study is lower than that of the RAPD-based study (0.176) and significant positive correlations between geographic distances and Nei's genetic distance were obtained. Positive correlations between geographic distance and genetic distance in outcrossing species have been reported by several authors (e.g. Ayres and Ryan, 1997; Shim and Jørgensen, 2000).

UPGMA cluster analysis and PCoA

It has been suggested that the use of cluster analysis in combination with PCoA helps to extract maximum information from molecular data (Messmer *et al.*, 1992) as PCoA facilitates the detection of intermediate populations (Lessa, 1990). Our cluster analysis and PCoA fit well with the genetic distance data, as shown by high cophenetic correlation coefficients (Rohlf, 2000). PCoA is used to allow for a visualization of differences among the populations and the identification of possible groups, as long as the first two or three axes (PCs) explain most of the variation (Mohammadi and Prasanna, 2003). Three clusters were obtained, by applying the principle of an 'acceptable number of clusters' (i.e. where the within-cluster genetic distance is less than the overall mean genetic distance and where the between-cluster distances are greater than the within-cluster distance of the two clusters involved; Brown-Guedira *et al.*, 2000). All three clusters (Fig. 3a, b) contain populations from geographically distinct regions, which may indicate long-distance gene flow along with human movement. On the other hand, populations from the same region were clustered together, except for Arsi and Shewa, where the two populations from each region were placed into different clusters. This further supports the, previously reported, considerable degree of population clustering according to region of origin (Geleta *et al.*, 2007).

This study generated comprehensive information regarding the genetic diversity of niger and demonstrated that AFLP is an appropriate technique for its evaluation. With more than 20% of the total genetic variation found

https://doi.org/10.1017/S1479262108913903 Published online by Cambridge University Press

between populations, we conclude that all the populations have unique genetic properties that make each niger population a significant unit for conservation and breeding purposes. Thus, our recommendation is that as many populations as possible should be conserved, as this reduces the risk of losing unique genetic variants due to shifting of cultivation practices and other factors. Furthermore, conserving a large number of genetically differentiated populations would also preserve a larger evolutionary potential of the crop that exists due to co-adaptation of gene complexes and local adaptation of populations. The extent and distribution of genetic variation in niger accessions conserved *ex situ* could be evaluated reasonably by as few as two AFLP PCs. If populations are to be ranked, emphasis should be given to those with high genetic variation and genetic distance, to capture unique genetic variation. The study also strengthens our previous recommendation of further germplasm collection by giving special emphasis to regions and areas under-represented in the gene bank collections.

Acknowledgements

The authors wish to thank The Swedish International Development Agency (SIDA/SAREC) for financing this work and the 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 assistance with the laboratory work.

References

Abebe M, Yermanos DM and Bingham FT (1978) The ecophysiology of noog (*Guizotia abyssinica* Cass.). *African Journal of Agricultural Science* 5: 55–66.

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

Assefa K, Merker A and Hailu T (2003) Inter simple sequence repeat (ISSR) analysis of genetic diversity in tef [*Eragrostis tef* (Zucc.) Trotter]. *Hereditas* 139: 174–183.

Ayres DR and Ryan FJ (1997) The clonal and population structure of a rare endemic plant, *Wyethia reticulate* (Asteraceae): allozyme and RAPD analysis. *Molecular Ecology* 6: 761–772.

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

Bossdorf O, Auge H, Lafuma L, Rogers WE, Sicmann E and Prati D (2005) Phenotypic and genetic differentiation between native and introduced plant populations. *Oecologia* 144: 1–11.

Brown-Guedira GL, Thompson JA, Nelson RL and Warburton ML (2000) Evaluation of genetic diversity of soybean introductions and North American ancestors using RAPD and SSR markers. *Crop Science* 40: 815–823.

Caetano-Anollés G and Gresshoff PM (1994) Staining nucleic acid with silver: an alternative to radioisotopic and fluorescent labeling. *Promega Notes Magazine* 45: 13–20.

Crow JF and Aoki K (1984) Group selection for a polygenic behavioral trait: estimating the degree of population subdivision. *Proceedings of the National Academy of Sciences, USA* 81: 6073–6077.

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 Systematics and Evolution* 195: 121–135.

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

Dagne K and Jonsson A (1997) Oil content and fatty acid composition of seeds of *Guizotia* Cass. (Compositae). *Journal of the Science of Food and Agriculture* 73: 274–278.

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

Ferdinandez YSN and Coulman BE (2004) Genetic relationships among smooth bromegrass cultivars of different ecotype detected by AFLP markers. *Crop Science* 44: 241–247.

Fichtl R and Adi A (1994) *Honeybee Flora of Ethiopia*. Germany: Margraf Verlag.

Geleta M, Asfaw Z, Bekele E and 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.

Geleta M, Bryngelsson T, Bekele E and Dagne K (2007) Genetic diversity of *Guizotia abyssinica* (L.f.) Cass. (Asteraceae) from Ethiopia as revealed by random amplified polymorphic DNA (RAPD). *Genetic Resources and Crop Evolution* 54: 601–614.

Genet T and Belete K (2000) Phenotypic diversity in the Ethiopian noug germplasm. *African Crop Science Journal* 8: 137–143.

Getinet A and Sharma SM (1996) *Niger (*Guizotia abyssinica *(L.f.) Cass.): Promoting the Conservation and Use of Underutilized and Neglected Crops. 5*. Gatersleben/Rome: Institute of Plant Genetics and Crop Plant Research/ International Plant Genetic Resources Institute.

Hamrick JL and Godt MJW (1996) Effects of life history traits on genetic diversity in plant species. *Philosophical Transactions of the Royal Society of London Series B* 351: 1291–1298.

Harlan JR (1969) Ethiopia: a centre of diversity. *Economic Botany* 23: 309–314.

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

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

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

Lessa EP (1990) Multidimensional analysis of geographic genetic structure. *Systematic Zoology* 39: 242–252.

https://doi.org/10.1017/S1479262108913903 Published online by Cambridge University Press

Loveless MD and Hamrick JL (1984) Ecological determinants of genetic structure in plant populations. *Annual Review of Ecology and Systematics* 15: 65–95.

Lynch M and Milligan BG (1994) Analysis of population genetic structure with RAPD markers. *Molecular Ecology* 3: 91–99.

Mantel NA (1967) The detection of disease clustering and a generalized regression approach. *Nucleic Acids Research* 21: 1111–1115.

Messmer MM, Melchinger AE, Boppenmaier J, Herrmann RG and Brunklaus-Jung E (1992) RFLP analyses of early-maturing European maize germplasm: I. Genetic diversity among flint and dent inbreds. *Theoretical and Applied Genetics* 83: 1003–1012.

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

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

Nayakar NY (1976) Genetic variability and heritability for six quantitative characters in niger (*Guizotia abyssinica* Cass.). *Mysore Journal of Agricultural Science* 10: 553–558.

Nei M (1972) Genetic distance between populations. *American Naturalist* 106: 283–292.

Nei M (1973) Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Sciences, USA* 70: 3321–3323.

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

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

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

Pavlicek A, Hrda S and 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 and Paikary RK (1995) Genetic variability and character association in niger. *Indian Journal of Genetics and Plant Breeding* 55: 457–459.

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

Riley KW and Belayneh H (1989) Niger. In: Röbbelen G Downey RK and Shri A (eds) *Oil Crops of the World*. New York: McGraw Hill, pp. 394–403.

Rohlf FJ (2000) NTSYSpc, Numerical Taxonomy and Multivariate Analysis System, Version 2.1. Exeter Software, Setauket, New York, USA.

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

Shim SI and Jørgensen RB (2000) Genetic structure in cultivated and wild carrots (*Daucus carota* L.) revealed by AFLP analysis. *Theoretical and Applied Genetics* 101: 227–233.

Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M and Zabeau M (1995) AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research* 23: 4407–4414.

Wright S (1951) The genetical structure of populations. *Annals of Eugenics* 15: 323–354.

Yeh FC and Boyle TJB (1997) Population genetic analysis of codominant and dominant markers and quantitative traits. *Belgium Journal of Botany* 129: 157.

Zeven AC and de Wet JMJ (1982) *Dictionary of Cultivated Plants and their Regions of Diversity*. Dehra Dun, India: International Book Distributors.

https://doi.org/10.1017/S1479262108913903 Published online by Cambridge University Press