001029

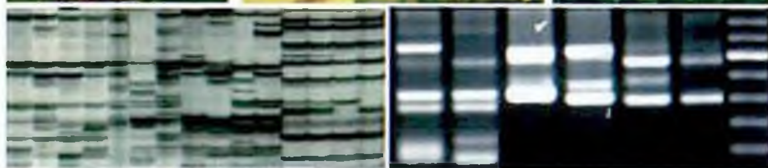
Acta Universitatis Agriculturae Sueciae

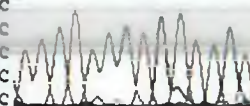
EIAR Library

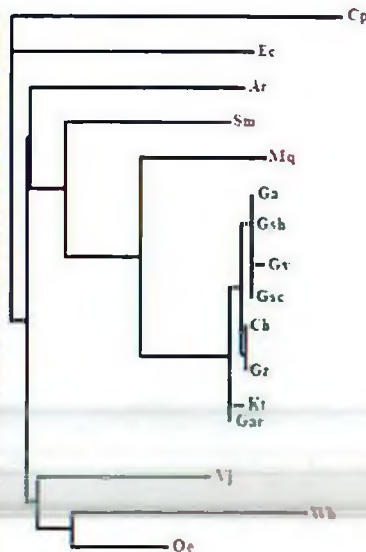
Doctoral Thesis No. 2007:27
Faculty of Landscape Planning, Horticulture
and Agricultural Science

Genetic Diversity, Phylogenetics and Molecular Systematics of *Guizotia* Cass. (Asteraceae)

Mulatu Geleta







Library - EIAR

Genetic Diversity, Phylogenetics and Molecular Systematics of *Guizotia* Cass. (Asteraceae)

Mulatu Geleta
Faculty of Landscape Planning, Horticulture and Agricultural Science
Department of Plant Protection Biology
Alnarp

To Dr. Kebebew Assefa

with best regards and wishes for the future!
Thank you very much for your kind help!

From Mulatu

Doctoral thesis
Swedish University of Agricultural Sciences
Alnarp 2007

Acta Universitatis Agriculturae Sueciae
2007:27

ISSN 1652-6880
ISBN 978-91-576-7326-8
© 2007 Mulatu Geleta, Alnarp
Tryck: SLU Service/Repro, Alnarp 2007

Genetic diversity, phylogenetics and molecular systematics of *Guizotia* Cass. (Asteraceae)

Mulatu Geleta

Akademisk avhandling som för vinnande av filosofie doktorsexamen kommer att offentligen försvaras i Crafoordsalen, SLU, Alnarp, fredagen den 11 maj 2007, kl 13.00

Abstract

The genus *Guizotia* belongs to the tribe Heliantheae in the family Asteraceae. It has been placed under different subtribes. The genus has its center of origin, distribution and genetic diversity in Ethiopia, where *G. abyssinica* (niger) has been domesticated. Amplified Fragment Length Polymorphism (AFLP), Random Amplified Polymorphic DNA (RAPD) and DNA sequencing were applied to study the genetic diversity, phylogenetics, and molecular systematics of this genus.

A large number of niger populations, representing all regions in Ethiopia where this crop is grown, was investigated using AFLP and RAPD molecular marker techniques. The extent of genetic variation in niger is distributed throughout its growing regions, regardless of the extent and altitude of cultivation. Despite the fact that most of the variation was within populations, significant population differentiation was obtained (AMOVA: $P < 0.001$) in all guizotias. It is concluded that both *G. abyssinica* and its wild and/or weedy relatives have wide genetic bases that need to be conserved and utilized for the improvement of *G. abyssinica*. Further collection of niger germplasm and exploration and conservation of highly localized guizotias are recommended.

Most of the diagnostic markers generated from AFLPs and RAPDs in this study were specific to *G. arborescens* and *G. zavattarii*. Phylogenetic analyses of the genus *Guizotia* were undertaken based on molecular sequence data from the internal transcribed spacers (ITS) and five chloroplast DNA regions. The analyses revealed a close phylogenetic relationship between *G. abyssinica* and *G. scabra* ssp. *schimperi* and support the previous suggestion that the latter is the progenitor of the former. According to this study, *G. scabra* ssp. *scabra*, *G. scabra* ssp. *schimperi*, and the Chelelu and Ketcha populations are best viewed at present as separate species within the genus *Guizotia*. Those perennial guizotias with highly localized geographic distribution appears to have evolved first during the evolutionary history of the genus. This study supports the placement of the genus *Guizotia* within the subtribe Milleriinae. It is suggested that the present species composition of *Guizotia* and the subtribal placement of the genus need to be redefined.

Key words: AFLP, cpDNA, DNA sequencing, genetic diversity, *Guizotia*, ITS, niger, phylogenetics, RAPD, systematics.

Distribution:
Swedish University of Agricultural Sciences
Department of Plant Protection Biology
SE-230 53 ALNARP, Sweden

Alnarp 2007
ISSN 1652-6880
ISBN 978-91-576-7326-8.

Genetic diversity, phylogenetics and molecular systematics of *Guizotia* Cass. (Asteraceae)

Mulatu Geleta

Akademisk avhandling som för vinnande av filosofie doktorsexamen kommer att offentligen försvaras i Crafoordsalen, SLU, Alnarp, fredagen den 11 maj 2007, kl 13.00

Abstract

The genus *Guizotia* belongs to the tribe Heliantheae in the family Asteraceae. It has been placed under different subtribes. The genus has its center of origin, distribution and genetic diversity in Ethiopia, where *G. abyssinica* (niger) has been domesticated. Amplified Fragment Length Polymorphism (AFLP), Random Amplified Polymorphic DNA (RAPD) and DNA sequencing were applied to study the genetic diversity, phylogenetics, and molecular systematics of this genus.

A large number of niger populations, representing all regions in Ethiopia where this crop is grown, was investigated using AFLP and RAPD molecular marker techniques. The extent of genetic variation in niger is distributed throughout its growing regions, regardless of the extent and altitude of cultivation. Despite the fact that most of the variation was within populations, significant population differentiation was obtained (AMOVA; $P < 0.001$) in all guizotias. It is concluded that both *G. abyssinica* and its wild and/or weedy relatives have wide genetic bases that need to be conserved and utilized for the improvement of *G. abyssinica*. Further collection of niger germplasm and exploration and conservation of highly localized guizotias are recommended.

Most of the diagnostic markers generated from AFLPs and RAPDs in this study were specific to *G. arborescens* and *G. zavattarii*. Phylogenetic analyses of the genus *Guizotia* were undertaken based on molecular sequence data from the internal transcribed spacers (ITS) and five chloroplast DNA regions. The analyses revealed a close phylogenetic relationship between *G. abyssinica* and *G. scabra* ssp. *schimperi* and support the previous suggestion that the latter is the progenitor of the former. According to this study, *G. scabra* ssp. *scabra*, *G. scabra* ssp. *schimperi*, and the Chelelu and Ketcha populations are best viewed at present as separate species within the genus *Guizotia*. Those perennial guizotias with highly localized geographic distribution appears to have evolved first during the evolutionary history of the genus. This study supports the placement of the genus *Guizotia* within the subtribe Milleriinae. It is suggested that the present species composition of *Guizotia* and the subtribal placement of the genus need to be redefined.

Key words: AFLP, cpDNA, DNA sequencing, genetic diversity, *Guizotia*, ITS, niger, phylogenetics, RAPD, systematics.

Distribution:
Swedish University of Agricultural Sciences
Department of Plant Protection Biology
SE-230 53 ALNARP, Sweden

Alnarp 2007
ISSN 1652-6880
ISBN 978-91-576-7326-8.

Abstract

Geleta, M. 2007. Genetic diversity, phylogenetics and molecular systematics of *Guizotia* Cass. (Asteraceae). Doctoral dissertation.
ISSN 1652-6880, ISBN 978-91-576-7326-8.

The genus *Guizotia* belongs to the tribe Heliantheae in the family Asteraceae. It has been placed under different subtribes. The genus has its center of origin, distribution and genetic diversity in Ethiopia, where *G. abyssinica* (niger) has been domesticated. Amplified Fragment Length Polymorphism (AFLP), Random Amplified Polymorphic DNA (RAPD) and DNA sequencing were applied to study the genetic diversity, phylogenetics, and molecular systematics of this genus.

A large number of niger populations, representing all regions in Ethiopia where this crop is grown, was investigated using AFLP and RAPD molecular marker techniques. The extent of genetic variation in niger is distributed throughout its growing regions, regardless of the extent and altitude of cultivation. Despite the fact that most of the variation was within populations, significant population differentiation was obtained (AMOVA; $P < 0.001$) in all guizotias. It is concluded that both *G. abyssinica* and its wild and/or weedy relatives have wide genetic bases that need to be conserved and utilized for the improvement of *G. abyssinica*. Further collection of niger germplasm and exploration and conservation of highly localized guizotias are recommended.

Most of the diagnostic markers generated from AFLPs and RAPDs in this study were specific to *G. arborescens* and *G. zavattarii*. Phylogenetic analyses of the genus *Guizotia* were undertaken based on molecular sequence data from the internal transcribed spacers (ITS) and five chloroplast DNA regions. The analyses revealed a close phylogenetic relationship between *G. abyssinica* and *G. scabra* ssp. *schimperi* and support the previous suggestion that the latter is the progenitor of the former. According to this study, *G. scabra* ssp. *scabra*, *G. scabra* ssp. *schimperi*, and the Chelelu and Ketcha populations are best viewed at present as separate species within the genus *Guizotia*. Those perennial guizotias with highly localized geographic distribution appears to have evolved first during the evolutionary history of the genus. This study supports the placement of the genus *Guizotia* within the subtribe Milleriinae. It is suggested that the present species composition of *Guizotia* and the subtribal placement of the genus need to be redefined.

Key words: AFLP, cpDNA, DNA sequencing, genetic diversity, *Guizotia*, ITS, niger, phylogenetics, RAPD, systematics.

Author's addresses: (1) Mulatu Geleta, Department of Plant Protection Biology, Swedish University of Agricultural Sciences (SLU), Box 44, SE-230 53 Alnarp, Sweden. E-mail: Mulatu.Geleta.Dida@ltj.slu.se; or (2) Department of Biology, Addis Ababa University (AAU), P. O. Box 1176, Addis Ababa, Ethiopia. E-mail: gemulat@yahoo.com.

Contents

Introduction 7

History of the genus *Guizotia* 7
Systematics of *Guizotia* Cass. 8
Some characteristics and geographical distribution of *Guizotia* species 9
Cytogenetics of *Guizotia* taxa (guizotias) 11
Reproductive mechanism and crossability among taxa 13
Domestication of *G. abyssinica* 14
Uses of niger and other guizotias 14
Distribution, ecology and agronomy of *G. abyssinica* 16
Pests and diseases of *G. abyssinica* 16
Breeding and biotechnology in niger 17
Oil content and fatty acid composition of guizotias 18
Genetic diversity and conservation of crops and their wild relatives 18
AFLP and RAPD 20
cpDNA and nDNA sequence data for phylogenetics and systematics 21

Objectives 22

Material and methods 22

Plant material and DNA extraction 22
AFLP and RAPD 23
Sequencing 23
Data analysis 24

Summary of results and discussions 24

Genetic diversity of *G. abyssinica* (papers I &II) 24
Genetic diversity of wild and/or weedy guizotias (paper III) 28
Phylogenetic relationship between guizotias (papers IV, V & VI) 30
The position of *Guizotia* Cass. within Heliantheae (papers V & VI) 35

Conclusions 36

Recommendations and future prospects 37

References 39

Acknowledgements 46

List of Appendices

Paper I-VI

The present thesis is based on the following papers, which will be referred to by their Roman numerals:

I. Geleta, M., Bryngelsson, T., Bekele, E. & Dagne, K. 2006. Genetic diversity of *Guizotia abyssinica* (L. f.) Cass. (Asteraceae) from Ethiopia as revealed by Random Amplified Polymorphic DNA (RAPD). *Genetic Resources and Crop Evolution* DOI 10.1007/s10722-006-0018-0.

II. Geleta, M., Bryngelsson, T., Bekele, E. & Dagne, K. Assessment of genetic diversity of *Guizotia abyssinica* (L.F.) Cass. (Asteraceae) from Ethiopia using Amplified Fragment Length Polymorphism (AFLP). *Plant Genetic Resources*. In press.

III. Geleta, M., Bryngelsson, T., Bekele, E. & Dagne, K. AFLP and RAPD analyses of genetic diversity of wild and/or weedy *Guizotia* (Asteraceae) from Ethiopia. *Hereditas* DOI:10.1111/j.2007.0018-0661.01983.x.

IV. Geleta, M., Bryngelsson, T., Bekele, E. & Dagne, K. 2007. Comparative analysis of genetic relationship and diagnostic markers of several taxa of *Guizotia* Cass. (Asteraceae) as revealed by AFLPs and RAPDs. *Plant Systematics and Evolution* DOI 10.1007/s00606-007-0521-6. In press.

V. Bekele, E., Geleta, M., Dagne, K., Jones, A. L., Barnes, I., Bradman, N. & Thomas, M.G. 2006. Molecular phylogeny of genus *Guizotia* (Asteraceae) using DNA sequences derived from ITS. *Genetic Resources and Crop Evolution* DOI 10.1007/s10722-006-9126-0

VI. Geleta, M., Bryngelsson, T., Bekele, E. & Dagne, K. Molecular Phylogeny of the genus *Guizotia* (Asteraceae) based on sequences derived from the *trn*T/*trn*L and *trn*L/*trn*F intergenic spacers, *trn*L and 3'*trn*K/*mat*K introns and *mat*K gene. Submitted.

Published papers are reproduced with the kind permissions of the corresponding publishers.

Introduction

History of the genus *Guizotia*

Guizotia Cass. is a small Afromontane endemic genus that belongs to the tribe Heliantheae in the family Asteraceae. The early history of the genus *Guizotia* is the history of its type species, *Guizotia abyssinica* (L. f.) Cass. (Baagøe, 1974), which is the only cultivated species of the genus. *G. abyssinica* (Fig. 1), commonly known as 'noug' (in Amharic), ramtil (in Hindi) and 'niger' (in English) is known in Ethiopia and India earlier than anywhere else. The Ethiopian and Indian niger had been described separately until 1877. On the Ethiopian side, the first valid name given to niger was *Polymnia abyssinica* L. f. in 1781 (Baagøe, 1974). According to Baagøe (1974), the name *Polymnia frondosa* was also given to this species in 1805. In 1821, a French botanist and naturalist, Count Alexandre Henri Gabriel de Cassini (1781-1832), described this species under the name of *Heliopsis platyglosa*; however, in 1929 he found that his *H. platyglosa* is the same as *P. abyssinica*. He then described a new genus, *Guizotia*, named after the French historian, orator and statesman Francois Pierre Guillaume Guizot (1787-1874; Seegeler, 1983). Cassini changed his *H. platyglosa* to *Guizotia abyssinica* L. f. and described it as a typical species of the genus. On the Indian side, the name *Verbesina sativa* was given to this species and published in 1807 (Baagøe, 1974). In 1834, the famous botanist Augustin Pyramus de Candolle (1778-1841) described a new genus *Ramtilla*, where he placed the Indian *V. sativa* under the new name of *Ramtilla olifera*. De Candolle then recognized that his *Ramtilla* and Cassini's *Guizotia* are the same and consequently changed the name *Ramtilla olifera* to *Guizotia olifera* (Baagøe, 1974). However, in 1877, the epithet "*abyssinica*" was maintained instead of "*olifera*" by Oliver and Hiern (Baagøe, 1974). The name *Guizotia* Cass. was conserved for the genus by the Vienna Congress in 1905, and *G. abyssinica* has been considered as the typical species of the genus *Guizotia* since 1930 (Briquet, 1935).

Figure 1. *Guizotia abyssinica* (A) plants and (B) flower.

Systematics of *Guizotia* Cass.

The genus *Guizotia* has been placed under different subtribes of tribe Heliantheae by different authors. Robinson (1981) stated that the restriction of the genus to the African region and its uncertain relationship to the western plant species make its position less certain in the tribe Heliantheae. Bentham (1873) placed the genus under the subtribe Coreopsidinae, where it had been treated until Baagøe (1974) suggested its transfer to the subtribe Verbesininae. Three years later, Stuessy (1977), after revising the tribe Heliantheae, maintained the genus within the Coreopsidinae, to which Baagøe later agreed (Baagøe, 1977). However, Robinson (1981) indicated that the terete, striate achenes, the ornamented seed coats, and the glanduliferous anther appendages are evidences against the placement of the genus within the Coreopsidinae. Consequently, he placed the genus under the subtribe Milleriinae based on close approximation of technical characters despite differences in habit and flower color. The placement of the genus under the subtribe Milleriinae was also asserted by Karis (1993). The transfer of an African *Sigesbeckia* species (*S. somalensis* S. Moore) that belong to the subtribe Milleriinae (Humbles, 1972) to the genus *Guizotia* by Schulz (1990) indirectly supports the placement of the genus *Guizotia* under the subtribe Milleriinae.

Figure 2. A young plant of *Guizotia arborescens*.

In her taxonomic revision of the genus *Guizotia*, Baagøe (1974) reduced the number of species considerably through merging two or more taxa, reducing the taxonomic status to subspecies level and excluding some species. Baagøe excluded *Guizotia bidentoides* Oliv. & Hiern and *Guizotia discoidea* Sch. Bip. in Schweinf. nom. nud., as they were found to be synonyms of *Bidens pinnatipartita* (O. Hoffm) Wild and *Sigesbeckia discoidea* (Vatke) Blake, respectively. *S. discoidea* was later moved to the genus *Micractis* by Schulz (1990). Baagøe circumscribed the genus to six species viz.: *G. abyssinica* (L. f) Cass., *G. arborescens* I. Friis, *G. jacksonii* (S. Moore) J. Baagøe comb. nov., *G. scabra* (Vis.) Chiov. ssp. *scabra*, *G. scabra* (Vis.) Chiov. ssp. *schimperi* (Sch. Bip. In Walp) J. Baagøe stat. nov., *G. villosa* Sch. Bip. in Walp, *G. zavattarii* Lanza in Chiov. & al. var. *zavattarii*, and

Dedicated to my parents: Geleta and Leakie

G. zavattarii Lanza in Chiov. & al. var. *angustata* Cuf. After Baagøe's (1974) taxonomic revision, two new populations of *Guizotia* were discovered in Ethiopia by K. Dagne who called them "Chelelu" and "Ketcha" (Dagne, 1995). These populations are distinct both from each other and from the recognized taxa of the genus (Dagne, 1995, 2001).

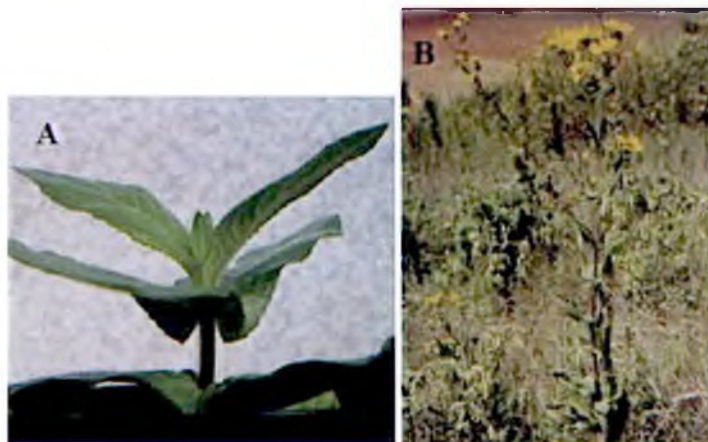


Figure 3. *Guizotia scabra* ssp. *scabra* (A) young plant and (B) mature plant in its natural habitat.

Some characteristics and geographical distribution of *Guizotia* species

The genus *Guizotia* is native to tropical Africa with most of the taxa restricted to East Africa, and with the highest concentration of species in Ethiopia (Baagøe, 1974). Ethiopia is a country in the horn of Africa located approximately between 3°N and 15°N latitude and 33°E and 48°E longitude. *G. abyssinica* is an annual herb with capitulum that consists of six to eight fertile female ray florets and 40-60 hermaphroditic disk florets (Getinet & Sharma, 1996). It is distinguished from other guizotias mainly by its corymbous cymes of heads, 5 broadly ovate-obovate outer involucral leaves, 5-nerved paleae and bigger achenes (Baagøe, 1974). Being a crop, niger has the widest geographic distribution among the guizotias. It has also been collected as a weed and as a wild plant in Ethiopia (Baagøe, 1974; Weiss, 1983). However, it is not clear whether this is a result of seeds dispersed from farmers' fields or a true wild/weedy form. Since we have not encountered any *G. abyssinica* in the wild during our extensive survey and collection of *Guizotia* species, it is possible that the reported *G. abyssinica* in the wild was just an escape from cultivation that died out.

Figure 4. A young plant of *Guizotia scabra* ssp. *schimperi*.

G. arborescens (Fig. 2) is a shrubby perennial endemic to southwest Ethiopia and mountain Imatongs in Sudan and Uganda within an altitudinal range of 1800-2600 m (Friis, 1971). Its arborescent habit, petiolate leaves, densely corymbous cymes and shrubby habit make it distinct from the rest of the taxa (Baagøe, 1974). *G. jacksonii* is a creeping perennial herb with sparsely branching habit grown within an altitudinal range of 2200-3700 m, endemic to Aberdares, Mt. Kenya and Mt. Elgon in Kenya and Uganda (Baagøe, 1974). Its creeping habit and solitary heads make it distinct from the other taxa. *G. scabra* ssp. *scabra* (Fig. 3) is a perennial, course, densely scabrous plant with stiff leaves, distributed, as part of natural vegetations, in East Africa, Cameron and Nigerian highlands, commonly within an altitudinal range of 1100-2700 m (Baagøe, 1974; Hiremath & Murthy, 1986). According to Baagøe (1974), it is the most variable taxon of the genus morphologically. *G. scabra* ssp. *schimperi* (Fig. 4) is an annual plant with foliaceous leaves native to Ethiopia and is a common weed of crops in mid and high altitude areas (1600-2300 m). *G. scabra* ssp. *scabra* and *G. scabra* ssp. *schimperi* are distinct from each other, for example, in having 50-120 and 35-50 disk florets and 8-15 and 6-8 ray florets, respectively (Baagøe, 1974).

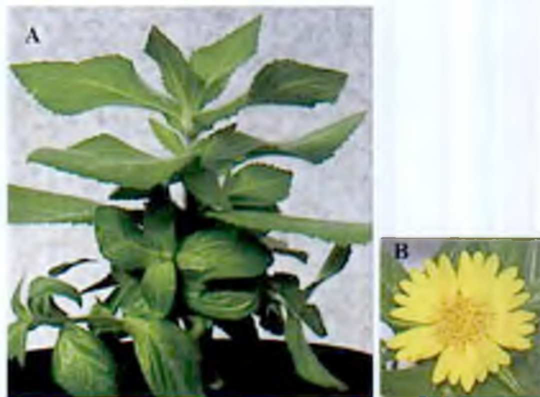


Figure 5. *Guizotia villosa* (A) young plant and (B) flower.

G. villosa (Fig. 5) is an annual herb restricted in distribution to the northern and north-western part of the Ethiopian highlands, and can be easily distinguished from other taxa by densely pilose indumentum, small heads and nearly square ray florets (Baagøe, 1974). *G. zavattarii* is a predominantly woody perennial plant restricted to southern Ethiopia (around mount Mega, Arero and Yabelo) and the Hurri Hills in northern Kenya, growing within the altitudinal range of 1500-2200 m. This species can be distinguished from the other taxa by its suffrutescent habit (Baagøe, 1974). *G. zavattarii* var. *angustata* was reported to be collected only in the vicinity of mount Mega (Baagøe, 1974) though we could not find it during our germplasm collections. The pandurate lacerate leaves are unique to *G. zavattarii* var. *zavattarii* (Fig. 6) within the genus. The geographic distribution of Chelelu and Ketcha is not fully explored; and they have not been botanically described. Chelelu (Fig. 7) is a riverine perennial plant that can be distinguished from other guizotias by its rhizomatous-like vegetative propagation and seed color (Dagne, 1995). It has been collected from the Chelelu river about 20 km north of Addis Ababa. Ketcha (Fig. 8) is a perennial plant that has been collected from Ketcha locality about 64 km from Bale-Goba to Delo-Mena in southeast Ethiopia.

Figure 6. A young plant of *Guizotia zavattarii*.

Cytogenetics of *Guizotia* taxa (guizotias)

All guizotias are diploids with 2n = 30 chromosomes and characterized by relatively small chromosomes with small differentiation between each complement (Dagne & Heneen, 1992; Hiremath & Murthy, 1992; Dagne, 1995). The number of satellite chromosomes in *Guizotia* species reported by Hiremath & Murthy (1992) was different from that of Dagne (1995). Hiremath & Murthy (1992) reported two satellite chromosomes in all *Guizotia* species they studied, while Dagne (1995) reported four to eight satellite chromosomes in these taxa. According to Dagne (1995), the discrepancy is likely due to chromosome preparation techniques and stages at which the satellite chromosomes were analyzed. A single B-chromosome was reported in some Ethiopian *G. scabra* ssp. *scabra* populations, a case that was not observed in populations of this taxon from other countries (Hiremath & Murthy, 1986) and in other guizotias. Later, Dagne

(1994b) reported three types of B-chromosomes in one population of this taxon in Ethiopia. About threefold variation in genome size was reported within the genus *Guizotia* (Hiremath, Murthy & Salimath, 1992) despite their similarity in number of chromosomes. The genome size of *G. abyssinica* was reported to be larger than that of its more closely related *Guizotia* taxa, including the suggested progenitor, *G. scabra* ssp. *schimperi*. The implication is that the genome size of *G. abyssinica* has increased over its evolutionary and domestication period without a change in number of chromosomes. Gene duplication (*e.g* Ohta, 1994; Meyerowitz, 1999) and retrotransposition (*e.g.* Vicient *et al.*, 1999; Bennetzen, Ma & Devos, 2005) are the likely mechanisms through which genome size variations have occurred within the genus *Guizotia*. The karyotypes of *G. abyssinica* and *G. scabra* ssp. *schimperi* are similar although the latter has relatively smaller chromosomes (Dagne & Heneen, 1992; Dagne, 1995). The difference in size of their chromosomes explains the variation in their genome size reported by Hiremath, Murthy & Salimath (1992). *G. abyssinica*, *G. scabra* ssp. *schimperi* and Chelelu are characterized by predominantly m type (symmetrical) chromosomes (Dagne, 1995). *G. arborescens* and *G. zavattarii* have relatively large chromosomes, which are predominantly asymmetrical (Hiremath & Murthy, 1992; Dagne, 1995). Although the genome size of *G. arborescens* has not been reported, the similarity with *G. zavattarii* in terms of chromosome size may indicate its relatively large genome size, as the genome size of the latter was reported to be the largest among the guizotias studied (Hiremath, Murthy & Salimath, 1992). The chromosomes of *G. scabra* ssp. *scabra*, *G. villosa* and Ketcha are more asymmetrical than those of *G. arborescens* and *G. zavattarii* (Dagne, 1995). The karyotype of *G. villosa* was first reported as symmetrical (Hiremath & Murthy, 1992); however, it was later reported that it is even more asymmetrical than that of the other guizotias (Dagne, 1995).

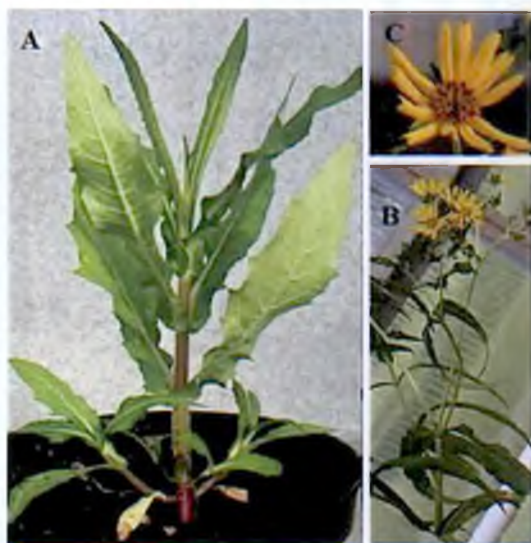


Figure 7. Plants of Chelelu at different growing stages (A &B) and its flower (C).

Figure 8. A young plant and flower of Ketcha.

Reproductive mechanism and crossability among taxa

Murthy, Hiremath & Salimath (1993) mentioned that *Guizotia* species in general are highly cross-pollinated and self-incompatible. The outcrossing nature of *Guizotia* species can be inferred from the cross-compatibility between various taxa that produces viable and fertile hybrids (Murthy, Hiremath & Salimath, 1993; Dagne, 1994a, 2001). However, direct evidence regarding self-incompatibility is only available for *G. abyssinica* (Ramachandran & Menon, 1979; Riley & Belayneh, 1989; Nemomissa, Bekele & Dagne, 1999). Self-incompatibility in niger is of the sporophytic type (Prasad, 1990; Nemomissa, Bekele & Dagne, 1999) that causes inhibition of pollen germination or twisting of pollen tube over the surface of the papillae (Prasad, 1990). Self-compatible niger genotypes were also reported in low frequencies within the Ethiopian gene pool, which can be as high as 5% in some populations (Getinet & Sharma, 1996; Nemomissa, Bekele & Dagne, 1999). Niger is mainly pollinated by insects (Fig. 9), particularly by bees (Ramachandran & Menon, 1979; Geleta *et al.*, 2002; Kandel & Porter, 2002), which are also the most likely agent of pollination in other guizotias.

Figure 9. Pollinators of *Guizotia abyssinica.*

Sexual hybrids can easily be obtained from crosses between *G. abyssinica, G. scabra* ssp. *schimperi, G. scabra* ssp. *scabra* and *G. villosa* (Dagne, 1994a). The crossing of these taxa with *G. zavattarii* was not successful, as the shriveled seeds

produced from these crosses failed to germinate (Murthy, Hiremath & Salimath, 1993; Dagne, 1994a). According to Dagne (1994a), the failure of the crosses involving *G. zavattarii* might be due to a postzygotic cross-incompatibility. Pollen viability of the F1 hybrid between *G. abyssinica* and *G. scabra* ssp. *schimperi* was about 80%, while that of the other F1 hybrids was less than 50% (Dagne, 1994a). Chelelu is cross compatible with *G. scabra* and *G. zavattarii* while *G. arborescens* is cross-compatible with *G. zavattarii* (Dagne, 2001). Therefore, almost all *Guizotia* taxa do not fulfill the definition of "biological/reproductive species", as viable and fertile hybrids can be formed between them at least in some combinations. Generally, hybrids are intermediate between their parents in terms of overall morphological appearance (Dagne, 1994a).

Domestication of *G. abyssinica*

Plant domestication is a gradual process that involves selective breeding through which plants are adapted to human needs. Plant domestication has been conducted for about 10,000 years in several regions of the world independently (Gepts, 2002). Ethiopia is one of the world centers where crop plant diversity is strikingly high and where some crop species were domesticated *de novo* (Vavilov, 1951; Harlan, 1969). The country has a unique position in the crop domestication history of a wide range of species, including oil crops (Harlan, 1969; Zeven & de Wet, 1982). Evidences regarding the origin and domestication of crop plants can be generated from various sources, such as history, linguistics, archeobotany, comparative morphology, phytogeography, cytogenetics and molecular biology. Although archeobotanical evidence regarding the origin and domestication of niger is lacking, based on morphological, phytogeographical, and cytological evidences, it was reported that *G. abyssinica* might have originated from *G. scabra* ssp. *schimperi* through selection and further cultivation (Baagøe, 1974; Hiremath & Murthy, 1988; Murthy, Hiremath & Salimath, 1993; Dagne, 1994a, 1995, 2001). Again based on these evidences, Ethiopia, where the crop has been under cultivation for much longer time than in any other place, has been considered as the center of origin, domestication and diversity of *G. abyssinica* (Baagøe, 1974; Weiss, 1983; Hiremath & Murthy, 1988). According to these authors, niger was domesticated in Ethiopia as early as 3,000 BC and taken to India by Ethiopian immigrants and/or through trade routes along with other crops. However, the magnitude of genetic diversity and use value of this crop imply a longer date of domestication than has been suggested.

Uses of niger and other guizotias

Niger seed is the major source of edible oil in Ethiopia. It also accounts for a considerable proportion of edible oil production in several African and Asian countries, including India. In India, it is frequently used as a substitute for sesame oil (Weiss, 1983). The seed contains 17-20% protein (Abebe, Yermanos & Bingham, 1978; Kandel & Porter, 2002) that offers an important source of protein and significantly contributes to the human dietary protein intake. It also contains 34-40% carbohydrate and 13.5% fiber and is an important source of thiamine,

riboflavin, and niacin (Kandel & Porter, 2002). Niger seed is consumed after being processed in various forms in Ethiopia and India (Seegeler, 1983; Getinet & Sharma, 1996). It is also used as a component of birdseed in USA and Europe (Kandel & Porter, 2002) and for cultural and medicinal purposes in Ethiopia (Geleta *et al.*, 2002). The press-cake left after oil extraction is an excellent poultry and livestock feed, as it contains 33-37% protein and is rich in inorganic constituents and crude fibers (Seegeler, 1983; Kandel & Porter, 2002). The whole plant is used as fodder and green manure (Weiss, 1983).

Refined niger seed oil is used for the preparation of soaps, paints, illuminants and lubricants and for cleaning machinery (Baagøe, 1974; Riley & Belayneh, 1989; Dutta *et al.*, 1994; Kandel & Porter, 2002). It is also used in perfumes as a carrier of the scents and fragrances (Kandel & Porter, 2002). The Ethiopian seed is superior to the Indian seed in its use for paints due to its relatively higher linoleic acid content. Niger seed oil can also be used as biodiesel through *trans* esterification of its long chain fatty acids with methanol that can partially substitute diesel oil and perform better with lower emission levels (Devi *et al.*, 2006).

The economic use of the wild and/or weedy guizotias is not well known. However, *G. scabra* ssp. *scabra* was reported to be used for human consumption (seeds) in Nigeria, as a medicine (seeds) and to make fishing nets (stem) in Congo (Baagøe, 1974). It has also been used as a treatment for hepatotoxicity and stomach disorders in Rwanda and neighboring countries (Fujimoto *et al.*, 1990). *G. scabra* ssp. *schimperi* was reported to be used as a substitute for *G. abyssinica* (Baagøe, 1974) and also as herbal medicine in some places in Ethiopia. The similarity in fatty acid composition between niger and the other guizotias (Dagne & Jonsson, 1997) may indicate the possibility of using wild and/or weedy guizotias for human consumption though further content analysis is required.

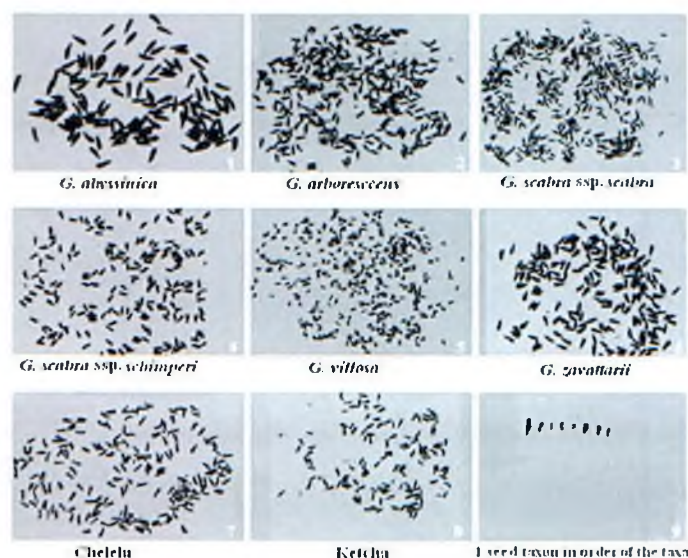


Figure 10. Seeds from eight *Guizotia* taxa.

Distribution, ecology and agronomy of *G. abyssinica*

Niger is mainly cultivated in Ethiopia and India. It is also cultivated in small scale in several other African (Sudan, Uganda, Zaire, Tanzania, Malawi and Zimbabwe) and Asian (Nepal, Bangladesh and Bhutan) countries, and in USA and West Indies (Weiss, 1983; Murthy, Hiremath & Salimath, 1993; Getinet & Sharma, 1996; Kandel & Porter, 2002). Niger has also been grown in Germany, Switzerland, France, Liechtenstein and the former USSR in the nineteenth and/or early twentieth century (Weiss, 1983).

In Ethiopia, niger is grown mainly from 1,600 masl to 2,200 masl, where the range in temperature is 15°C to 23°C, and where the annual rainfall is 500-1,000 mm (Getinet & Sharma, 1996). However, it can be cultivated at altitudes as low as 500 m and as high as 3000 m in the availability of enough rainfall (Weiss, 1983; Getinet & Sharma, 1996). The major niger producing regions in Ethiopia are Gojam, Gonder Shewa and Welega, which are followed by Tigray and Welo. The production of niger in Arsi, Bale, Harerge, Jimma and Illubabor is low and restricted to few districts. Ethiopian niger normally matures within 120-150 days after emergence, but varieties with a shorter maturation period of 90-120 days also exist (Weiss, 1983). Niger occupies about 50% of the total oil crop area and production volume in Ethiopia (Getinet and Sharma, 1996). It is grown either as a sole crop or inter/border crop in fields of other crops (Geleta *et al.*, 2002) usually without the application of fertilizer or herbicide. Fertilizer application promotes vegetative growth rather than increasing seed yield (Getinet & Sharma, 1996). Niger can grow on a wide variety of soils, but appears to thrive best on clayey loams or sandy clays (Weiss, 1983) within a pH range of 5.2-7.3 (Getinet & Sharma, 1996). It can grow on waterlogged, marginal and poor soils where most other crops fail to grow, as it is able to withstand salinity and low oxygen levels (Abebe, Yermanos & Bingham, 1978).

Despite its long cultivation history, niger has low seed yield due to several factors, including an indeterminate growth habit, lodging and shattering. Its low yield makes niger unsuccessful in Europe, the former USSR and Canada as a commercial crop (Weiss, 1983). On average, both Ethiopian and Indian niger yields 300-400 kg/ha when grown in pure stands although higher yield was reported for some varieties from Kenya (600 kg/ha; Kandel & Porter, 2002) and India (1,000 kg/ha; Weiss, 1983).

Pests and diseases of *G. abyssinica*

A large number of niger pests and diseases have been recorded though they are fewer than those recorded in most other oilseed crops (see Getinet & Sharma, 1996). Niger pests and diseases do not cause serious damage to the crop (Weiss, 1983; Kandel & Porter, 2002) due to its apparently considerable level of tolerance/resistance, which needs to be well investigated. Niger fly (*Dioxyna sororcula* and *Eutretosoma* spp.), black pollen beetles (*Meligethes* spp.) and dodder (*Cuscuta campestris*; a parasitic weed) are the most important pests of niger while niger blight (*Alternaria* sp.) and bacterial leaf spot (*Xanthomonas* sp.)

are the most serious diseases of niger (Gebre-Medhin & Mulatu, 1992; Getinet & Sharma, 1996).

Breeding and biotechnology in niger

Ethiopian and Indian niger are different in several characteristics, including plant height, days to maturity, and fatty acid composition (Weiss, 1983; Getinet & Sharma, 1996), which is the result of separate evolution since the separation of the two gene pools. Niger production in Ethiopia is mainly based on local landrace populations, which makes it more variable than Indian niger. According to Getinet & Sharma (1996), three landraces of niger are known in Ethiopia. These are 'abat' 'bungne' and 'mesno', which are medium to late maturing, early maturing and late maturing, respectively. Abat is the common landrace with higher yield than the other two varieties and with higher oil content than bungne (Getinet & Sharma, 1996). Lodging, shattering, indeterminate growth habit and self-incompatibility have been reported as major factors that contribute to low yield in *G. abyssinica*.

The low yield in niger makes it less competitive with other oil crops and hampers its improvement through breeding, as the interest of breeders has been low. In India, it was possible to improve seed yield by about three-fold (Weiss, 1983). Unlike the case in India, only few improved varieties have been released in Ethiopia. These varieties are not that much better than the landraces, in terms of seed yield, oil content and days to maturity (Getinet & Sharma, 1996). In this regard, niger breeding in Ethiopia is still in its relative infancy. In niger, seed yield is positively correlated with seed size, number of seeds/capitula, number of capitula/plant and number of primary branches in niger (*e.g.* Singh & Patra, 1989). Our unpublished agromorphological data showed the existence of several-fold variation in number of seeds/capitula, capitula/plant, and primary branches/plant both within and between Ethiopian niger populations. Some degree of variation in achene size also exists in the Ethiopian niger gene pool. Therefore, great effort is needed to use the existing genetic diversity for the improvement of the crop.

Biotechnology might be considered as an alternative approach for niger improvement. Several cell tissue culture studies have been conducted on niger. Regeneration has been reported from leaves (Sujatha, 1997; Jadimath *et al.*, 1998; Kumar *et al.*, 2000), cotyledons (Adda, Reddy & Kishor, 1993b, 1994b; Nikam & Shitole, 1997), and anthers (Adda, Reddy & Kishor, 1993a, 1994b). In India, dwarfs and large headed self-compatible doubled haploid niger genotypes were obtained from anther culture (Adda, Reddy & Kishor, 1994a). The protocol for *in vitro* plant regeneration for large-scale propagation of male sterile niger was published by Sujatha (1997). This protocol may facilitate the efficient use of pollen grains carrying desirable traits to pollinate the male sterile genotypes (Sujatha, 1997). The various regeneration protocols reported so far and the *Agrobacterium* mediated transformation protocol developed by Murthy *et al.* (2003) are useful for the future improvement of niger.

Oil content and fatty acid composition of guizotias

Seed lipids usually contain over 95% neutral storage lipids in the form of triacylglycerol (TAG) (Ohlrogge & Jaworski, 1997). The oil content of the niger seed is commonly within the range of 27-50%, as reported by various authors (*e.g.* Seegeler, 1983; Dutta *et al.*, 1994; Getinet & Teklewold, 1995; Dagne & Jonsson, 1997; Marini *et al.*, 2003; Ramadan & Mörsel, 2003; Asilbekova *et al.*, 2005), which is higher than that of wild and/or weedy guizotias (Dagne & Jonsson, 1997). However, up to 60% oil has been obtained from niger seed of improved varieties (Kandel & Porter, 2002).

The fatty acid composition of wild and/or weedy guizotias is similar to that of *G. abyssinica* apart from some degree of variation in their proportion (Dagne & Jonsson, 1997). The predominant fatty acid in niger seed oil is linoleic acid (LA) regardless of the differences among reports in terms of its proportion within the range of 54-85% (*e.g.* Weiss, 1983; Getinet & Teklewold, 1995; Dagne & Jonsson, 1997; Marini *et al.*, 2003; Ramadan & Mörsel, 2003). Niger seed oil has a higher proportion of LA and a lower proportion of oleic acid (OA) as compared to that of wild and/or weedy guizotias (Dagne & Jonsson, 1997), and sunflower and safflower (Dutta *et al.*, 1994). OA is the second major unsaturated fatty acid in niger seed oil, which commonly accounts for 5-13% of the fatty acids. This fatty acid is higher in the Indian than in the Ethiopian niger (Riley & Belayneh, 1989; Marini, 2003). Generally, LA, OA and the two major saturated fatty acids in niger seed oil (palmitic acid and stearic acid) make up more than 90% of the fatty acids in niger seed oil (*e.g.* Dutta *et al.*, 1994; Dagne & Jonsson, 1997; Ramadan & Morsel, 2003).

The high content of LA (an essential fatty acid) in niger seed oil is nutritionally highly valuable, as it is known to prevent cardiovascular disorders such as coronary heart diseases, arteriosclerosis and high blood pressure (Vles & Gottenbos, 1989). Niger seed oil is also a good source of vitamin E, as almost all of the tocopherols are α-tocopherol (94-96%; Dutta *et al.*, 1994; Ramadan & Mörsel, 2004). The oxidative stability of niger seed oil was reported to be lower than that of black cumin and coriander seed oils (Ramadan & Mörsel, 2004), which may limit its utilization in processed and fortified foods as well as nutritional supplement, as lipid oxidation negatively affects the flavor, odor, color and nutritional value of foods during storage. The lower oxidative stability of niger seed oil is partly explained by its high LA content that overcome the antioxidant activity expected from tocopherols and partly by its relatively lower level of phenolics and polar lipids that have anti-oxidant activity (Ramadan & Mörsel, 2004).

Genetic diversity and conservation of crops and their wild relatives

Crop genetic diversity is the genetic variation within and between individuals, populations and varieties of the cultivated species, which has occurred through mutation, introgression, recombination, adaptation to new environments and

continuous selection. In most cases, crop's genetic diversity is greatest in regions where it was originally domesticated and where its evolution has the longest record (Hawkes, 1983). The genetic diversity within crop species is an invaluable genetic resource for farmers, scientists and consumers elsewhere. The significance of such diversity can be measured, in part, by its role in increasing crop resistance to pests, diseases, and abiotic stresses, which in turn contribute to the production of higher, more stable and good quality yields (*e.g.* Bekele, 1985; Wood & Lenne, 1997; Jana, 1999). In other words, genetic diversity of crops is of paramount importance in maintaining and increasing agricultural productivity and product quality and helps to withstand newly emerging pests and pathogens, as it is a defense against the uncertain future. Genetic diversity in the crops' wild relatives is also very important to tackle various problems associated with crop failure. Crop plants have been improved through using wild relatives especially those within the primary gene pools of crops (Harlan & de Wet, 1971; Hawkes, 1977). Since the 1980's, crop improvement has been made by transferring desirable genes from distantly related and even non-related taxa through genetic engineering, thereby broadening the value of crops' wild relatives by expanding their usefulness into secondary and tertiary gene pools (Meilleur & Hodgkin, 2004).

There has been an increasing concern over the loss of genetic diversity in crops and their wild relatives in areas of crop domestication. Deforestation and habitat loss due to agricultural expansion, soil erosion and degradation are among the most serious environmental problems threatening the genetic diversity of crops' wild relatives (Jana, 1999), which is perceived as threats to the genetic base of world agriculture (Meilleur & Hodgkin, 2004). Unless genetic variation in crop plants and their wild relatives is properly conserved, they may be lost forever, to the great detriment of agriculture and human food security. Since several decades, there have been attempts to conserve crop genetic variation both *in situ* and *ex situ* (*e.g.* Brush, 1995; Bellon, 1996, Heal *et al.*, 2004). For example, 668 niger accessions and 510 accessions of wild and/or weedy guizotias are conserved *ex situ* at the Institute of Biodiversity Conservation (IBC) of Ethiopia (Adugna Abdi, personal communication). The *in situ* conservation of agricultural crops and their wild relatives requires conservation of their respective original habitats in their major centers of diversity (Jana, 1999). Therefore, identifying the genetic diversity center for a crop and its wild relatives is important for proper conservation.

Analyses of plant genetic diversity may lead to the identification of diverse parental combinations to create segregating progenies with maximum variability for further selection (*e.g.* Barrett & Kidwell, 1998; Brush & Meng, 1998), to introgression of desirable genes from diverse germplasms into the variety of interest (*e.g.* Thompson & Nelson, 1998) and to a reliable classification of accessions of germplasm collections (*e.g.* Ma *et al.*, 2006). Such an assessment of levels and patterns of plant genetic variation can be carried out based on various types of data, such as pedigree (*e.g.* Barata & Carena, 2006; Soleiman, Baum & Johnson, 2007), morphological (*e.g.* Ayana & Bekele, 1999; Talhinhas, Leitao & Neves-Martins, 2006), isozymes (*e.g.* Hamrick & Godt, 1997; Ayana, Bryngelsson & Bekele, 2001) storage proteins (*e.g.* Bekele *et al.*, 1995; Alvarez, Moral & Martin, 2006) and DNA markers (*e.g.* Assefa, Merker & Hailu, 2003; Cavagnaro *et al.*, 2006).

AFLP and RAPD

Amplified Fragment Length Polymorphism (AFLP) and Random Amplified Polymorphic DNA (RAPD) are among the most commonly used PCR based molecular marker techniques and they have various applications in plants, including genetic diversity (*e.g.* Nybom & Bartish, 2000; Raina *et al.*, 2001; Nybom, 2004), phenetic and phylogenetic analyses (*e.g.* Landry & Lapointe, 1996; Blattner *et al.*, 2001; Després *et al.*, 2003; Bänfer, Fiala & Weising, 2004; Kadereit & Kadereit, 2005) and species/cultivar identification (*e.g.* Johnson *et al.*, 2003; Kelly & Miklas, 1998; Boukar *et al.*, 2004).

The RAPD method makes use of single short oligomers (usually 10-mer) of arbitrary sequence which anneal to random homologous target sites within the genome. It allows the generation of a product only when two copies of a primer are annealed in an inverted fashion to each other within an amplifiable distance (Welsh & McClelland, 1990; Williams *et al.*, 1990). RAPD polymorphisms result from DNA sequence variation at primer binding sites and from DNA length differences between primer binding sites, which disrupt or displace homologous target sites and results in the loss of a product. This technique largely generates dominant markers, although co-dominant markers that are produced due to length polymorphisms caused by small insertions/deletions can occur at low frequencies (2–5%; Milbourne *et al.*, 1997). RAPD is specially suited to analyze large number of samples as it is relatively simple, fast and cheap. Problems associated with reproducibility (Jones *et al.*, 1997) and co-migration of some non-homologous fragments (Rieseberg, 1996) are drawbacks of the RAPD technique. However, reproducibility can be considerably improved by making the reaction conditions uniform across samples, and by thorough screening of primers for reproducibility (Skroch & Nienhuis, 1995). The problem associated with co-migration of non-homologous fragments is minimal when RAPD is applied to study closely related populations or species (Rieseberg, 1996).

AFLP is a molecular marker technique that involves digestion of target DNA by restriction endonucleases, ligation of adaptors to the restricted fragments at both ends and amplification of fragments (Vos *et al.*, 1995). In higher plants, fragment amplification is usually conducted in two steps: preamplification and selective amplification. AFLP primers possess complementary sequences to the adaptors and the adjacent restriction site that serve as primer binding sites for subsequent amplification of the restriction fragments. Additionally, these primers possess one (preamplification) or three (selective amplification) selective nucleotides at their 3'-end (Vos *et al.*, 1995). The presence or absence of these selective nucleotides in the genomic fragments being amplified and the restriction fragment size variation provide the basis for revealing polymorphism in AFLPs. Despite the fact that AFLPs are generally treated as dominant markers, codominance seems to be more common in AFLPs than in RAPDs, as 10% codominant AFLPs were reported in *Populus* (Yin *et al.*, 2002) and 14% in *Arabidopsis thaliana* (Alonso-Blanco *et al.*, 1998). The AFLP technique has the capacity to detect a higher number of polymorphic loci in a single assay than RFLPs or RAPDs (Powell *et al.*, 1996), has the highest discrimination efficiency in comparison to RAPD and ISSR

(Archak *et al.*, 2003), and produces highly reproducible results (Jones *et al.*, 1997).

The use of RAPD and AFLP techniques to assess genetic relationship between different groups of taxa, to assess genetic variability within and among populations and to screen for diagnostic markers is promising because many polymorphic loci can be obtained fairly easy, in a relatively short time and without any prior knowledge of the genome of the species under study (*e.g.* Vos *et al.*, 1995, Gupta *et al.*, 1999, Nybom & Bartish, 2000; Nybom, 2004). These marker techniques were reported to work well for phylogenetic studies in closely related species (Landry & Lapointe, 1996; Després *et al.*, 2003) and can be used especially when DNA sequencing data fails to resolve the phylogeny (Weising *et al.*, 2005).

cpDNA and nDNA sequence data for phylogenetics and systematics

DNA sequencing is the most direct method of detecting genetic variation at the DNA level. The conserved regions of the DNA are the bases for designing PCR and sequencing primers. The PCR-amplified target regions can be either directly sequenced or sequenced after cloning (Hillis, Moritz & Mable, 1996). Comparative DNA sequencing has become a widespread tool for systematic and phylogenetic studies as it is relatively fast and convenient, and provides highly informative, robust and reproducible data sets. It can be applied for comparing organisms at different taxonomic levels by choosing appropriate genomic target regions (Weising *et al.*, 2005). The majority of data used in plant molecular phylogenetic and systematic studies derive from chloroplast DNA (cpDNA) and nuclear ribosomal DNA (nrDNA) (Small, Cronn & Wendel, 2004). Slowly evolving DNA regions such as the chloroplast *rbc*L gene (Clegg, 1993; Soltis, Soltis & Chase, 1999) and the nrDNA genes (Kuzoff *et al.*, 1998) have been used at higher taxonomic levels.

For systematic and phylogenetic studies at infrafamilial taxonomic levels, the internal transcribed spacers (ITS1 and ITS2) of nrDNA (Baldwin, 1992; Baldwin *et al.*, 1995; Crawford & Mort, 2005), and various cpDNA regions such as non-coding regions of tRNA genes (*e.g.* *trn*T/*trn*L and *trn*L/*trn*F intergenic spacers and *trn*L intron) and *mat*K gene have been widely used (*e.g.* Taberlet *et al.*, 1991; Johnson & Soltis, 1994; Liang & Hilu, 1996; Bayer, Greber & Bagnall., 2002; Winkworth *et al.*, 2002; Crawford & Mort, 2005; Shaw & Small, 2005). DNA sequence data based phylogenetic analyses can be conducted by a variety of approaches and algorithms (*e.g.* Saitou & Nei, 1987; Archibald, Mort & Crawford, 2003; Huelsenbeck & Crandall, 1997; Hall, 2001).

The relatively simple genetics of cpDNA makes it a primary target for molecular phylogenetic analyses in plants (Small, Cronn & Wendel, 2004). A typical DNA sample contains a relatively high copy number of cpDNAs, as multiple copies of chloroplasts per leaf cell and multiple copies of cpDNAs per chloroplast exist. This facilitates PCR amplification of specific cpDNA regions, as high copy-number sequences are more readily accessible. Moreover, cpDNA is generally

stable, haploid, non-recombinant and uniparentally inherited, which facilitates its use in systematic studies (*e.g.* Birky, 1995; Small, Cronn & Wendel, 2004). However, the haploid nature and uniparental inheritance of cpDNA limit its use in phylogenetic studies of hybrids and polyploids, as it may incorrectly identify them as belonging to the clade of one of the two parents without revealing the hybrid history (Small, Cronn & Wendel, 2004). Ribosomal genes (including ITS) exist in tandem arrays of genes composed of hundreds to thousands of copies per array, which facilitates DNA sequencing directly from a PCR product (Baldwin *et al.*, 1995; Kuzoff *et al.*, 1998). At lower taxonomic level, ITS sequences generally provide greater levels of divergence and thus greater resolution and stronger support than an equivalent sample of non-coding cpDNA (Sang, Crawford & Stuessy, 1997; Whitten, Williams & Chase, 2000).

Objectives

The major objective of this doctoral thesis was to study the genetic diversity, phylogenetics and systematics of the genus *Guizotia* with the ultimate goal of conserving and utilizing its genetic diversity. The following are the specific objectives of the study:

1. to determine the genetic structure and the extent of genetic variation within and between populations of *G. abyssinica* and its wild and/or weedy relatives using AFLP and RAPD for their conservation and use value
2. to resolve the phenetic and phylogenetic relationship between *Guizotia* taxa and to identify the closest relatives to *G. abyssinica* using molecular markers and DNA sequence data
3. to determine the taxonomic status of two yet taxonomically untreated *Guizotia* populations (Chelelu and Ketcha)
4. to evaluate the taxonomic status of the two subspecies of *G. scabra*
5. to provide data and suggest the appropriate subtribal placement of the genus *Guizotia* within the tribe Heliantheae

Material and methods

Plant material and DNA extraction

The germplasm of all guizotias used in this study was collected in Ethiopia from early November to the end of December 2003. *G. abyssinica* populations (Table 1) were collected directly from farmers' fields and the crop on a single field was considered as a population. Individuals were sampled at an equidistant along the longest line found across the field. The sampled populations represent the altitudinal range and geographic regions where niger is currently grown within the country. In the case of wild and/or weedy guizotias, plants of the same taxon

found in their habitat without large gaps in between were considered as a population. Samples were collected randomly and systematically depending on the size of the populations. Different number of populations were studied per taxon based on the geographic range and abundance of the species. Seeds were grown in a greenhouse and fresh leaves from 15-30 days old plants were used for genomic DNA extraction. DNA was extracted by a modified CTAB procedure. The detailed procedure is given in paper V.

Table 1. *(i) Number of populations used for genetic diversity (GD) and genetic relationship (GR) study using AFLP and RAPD techniques, and (ii) number of individuals sequenced for nrDNA and cpDNA based phylogenetic analyses of eight Guizotia taxa*

Taxa	AFLP[a]		RAPD[a]		Sequencing[b]	
	GD	GR	GD	GR	nrDNA	cpDNA
G. abyssinica	17	11	70	11	2	4
G. arborescens	3	2	4	2	4	3
G. scabra ssp. *scabra*	4	3	9	3	1	3
G. scabra ssp. *schimperi*	4	3	9	3	1	3
G. villosa	3	2	6	2	2	3
G. zavattarii[c]	3	2	6	2	2	3
Chelelu	1	1	1	1	-	2
Ketcha	1	1	1	1	-	2

[a]Seven AFLP primer combinations (PCs) and ten RAPD primers were used for each taxon. [b]Five cpDNA regions (*trn*T/*trn*L intergenic spacer, *trn*L intron, *trn*L/*trn*F intergenic spacer, 3'*trn*K/*mat*K portion of the *trn*K intron and *mat*K gene) and two nrDNA regions (ITS1 and ITS2) were used for this study. [c]Only *Guizotia zavattarii* var. *zavattarii* was used.

AFLP and RAPD

DNA amplification, electrophoresis, staining and data scoring

After optimizing all procedures, 150 RAPD primers and 56 AFLP primer combinations (PCs) were screened, out of which ten RAPD primers and seven AFLP PCs were used for final analyses. RAPD based DNA amplification, electrophoresis and staining of the amplified product were described in paper I. The detailed descriptions of DNA enzyme restriction, adaptor ligation, preamplification, selective amplification and silver staining procedures of AFLP are given in papers II & IV. Each RAPD and AFLP band was considered as a single bi-allelic locus with one amplifiable and one null allele. Data was scored manually as 1 for the presence and 0 for the absence of a DNA band for each locus across the genotypes of each taxon. As a strategy to minimize scoring error, both within and among gels, data scoring was performed twice separately and discrepancy was reconsidered and corrected.

Sequencing

The entire region of internal transcribed spacers (ITS1 and ITS2) of nrDNA, and the *trn*T/*trn*L intergenic spacer, the *trn*L intron, the *trn*L/*trn*F intergenic spacer, the 3'*trn*K/*mat*K portion of the *trn*K intron and the *mat*K gene of cpDNA were amplified using specific primer-pairs and sequenced using corresponding sequencing primers. Details for cleaning amplified products, sequencing and sequence processing were described in papers V (ITS) and VI (cpDNA regions).

Data analysis

Data from the AFLP and RAPD studies was analyzed using various statistical programs. The details of Shannon diversity estimates and gene diversity parameters (Nei, 1973, 1978; Lynch & Milligan, 1994) were described in paper I. NTSYSpc (Rohlf, 2000) was used for genetic similarity and distance estimates (Jaccard, 1908; Nei, 1972, 1978), cluster analysis, principal coordinate analysis (PCoA) and comparison of matrices. POPGENE version 1.31 (Yeh & Boyle, 1997) was used for analysis of percentage of polymorphic loci while Arlequin version 2 (Schneider, Roessli & Excoffier, 2000) was used for analysis of molecular variance (AMOVA). FreeTree-Freeware (Pavlicek, Hrda & Flegr, 1999) and TreeView (Win32) 1.6.6 (Page, 1996) were used for bootstrap analysis and to view trees, respectively. MINITAB release 14 was used for further analyses of some of the outputs from other programs. DNA sequencing data was handled and manipulated using BIOEDIT version 7.0.5 (Hall, 2005) and SEQUENCE SCANNER version 1.0 (Applied Biosystems®), and aligned using CLUSTAL X version 1.81 (Thompson *et al.*, 1997) and SEQUENCHER (Gene Codes Corporation). Phylogenetic analysis of DNA sequence data was carried out using PAUP* 4.0 Beta 10 (Swofford, 2000).

Summary of results and discussions

Genetic diversity of *G. abyssinica* (papers I & II)

A total of 539 clear and unambiguous loci generated by seven AFLP PCs applied to 170 individual plants were scored, of which 483 (90%) loci were polymorphic. Similarly, a total of 194 loci that were consistently amplified by ten RAPD primers applied to 700 individuals were scored, of which 188 (97%) were polymorphic. Both studies revealed that, on average, about half of the loci in each population are polymorphic (Table 2). The high percentage of polymorphic loci revealed by both marker techniques suggests the existence of high genetic polymorphism in Ethiopian niger and was proven to be useful in distinguishing niger populations and even individuals within populations. The polymorphism detected in niger was higher than, for example, the RAPD based polymorphism (64%) reported by Sivolap, Solodenko & Burlov (1998) in sunflower (*Helianthus annuus* L.). The RAPD based study revealed an average of 19 polymorphic loci per primer while an average of 69 polymorphic loci per AFLP PC was obtained. Thus, the multiplex ratio (the number of loci analyzed simultaneously per experiment) of AFLPs is more than threefold when compared with that of RAPDs. The overall utility of AFLP was also higher than RAPD, as measured by the marker index (product of multiplex ratio and diversity index), which is in agreement with the report of Milbourne *et al.* (1997).

The overall gene diversity estimate (H_T; Nei, 1978) with the modification of Lynch & Milligan (1994) was 0.320 (AFLP) and 0.248 (RAPD). Similarly, the overall within population variation (H_S) was 0.205 (AFLP) and 0.176 (RAPD) and thus AFLP revealed relatively higher overall total and within population diversity

as compared to RAPD in *G. abyssinica*. The extent of genetic diversity of each population was calculated using gene diversity estimates as H_j, which ranged from 0.183 to 0.241 (AFLP) and from 0.112 to 0.245 (RAPD). It is interesting to note that the highest within population genetic variation obtained by using AFLP or RAPD refers to the same population (Wl-1), which shows an agreement between the two marker techniques. Niger populations from some regions such as Illubabor (Table 2) where niger cultivation seems to be declining (local informants), appears to have higher genetic variation within populations.

Table 2. *Mean percent polymorphic loci (PPL) and mean gene diversity estimates (H_j) generated from AFLP and RAPD data for eight Guizotia taxa*

Taxa	Regions (code)	*PPL*		H_j	
		AFLP	RAPD	AFLP	RAPD
G. abyssinica	Arsi (A)	49.7	54.6	0.19	0.19
G. abyssinica	Bale (B)	50.5	53.6	0.21	0.19
G. abyssinica	Gojam (Gj)	53.6	47.0	0.22	0.16
G. abyssinica	Gonder (Gr)	49.7	46.8	0.19	0.14
G. abyssinica	Harerge (H)	50.3	48.5	0.20	0.16
G. abyssinica	Illubabor (I)	55.1	55.0	0.27	0.19
G. abyssinica	Jimma (J)	54.9	55.9	0.22	0.19
G. abyssinica	Shewa (Sh)	50.0	51.8	0.19	0.18
G. abyssinica	Tigray (T)	46.2	54.1	0.18	0.18
G. abyssinica	Welega (Wg)	50.6	49.2	0.20	0.17
G. abyssinica	Welo (Wl)	54.2	57.7	0.21	0.21
G. abyssinica	Mean	51.3	52.2	0.20	0.16
G. arborescens		17.38	37.57	0.24	0.17
G. scabra ssp. *scabra*		61.09	51.20	0.24	0.17
G. scabra ssp. *schimperi*		64.57	51.13	0.24	0.17
G. villosa		62.63	56.85	0.24	0.19
G. zavattarii		23.69	34.84	0.17	0.16
Chelelu		40.82	74.68	0.37	0.31
Ketcha		80.64	91.67	0.33	0.30

Analysis of molecular variance of AFLP and RAPD data revealed a highly significant variation between populations ($P < 0.001$) with 23% and 35% of the total variation differentiating the populations, respectively. The RAPD based AMOVA calculated by grouping the populations into regions of origin revealed a highly significant variation between regions (13.9%; $P < 0.001$). Similarly, highly significant variation (7.5%; $P < 0.001$) was obtained when AFLP data was analyzed by grouping the populations based on geographic proximity and better access of gene flow. The result suggests considerable degree of regional differentiation of *G. abyssinica* populations. The RAPD based AMOVA calculated by grouping the populations into populations from major niger producing regions (MaNPRs) and populations from minor niger producing regions (MiNPRs) revealed significant variation ($P < 0.001$) between the groups, with the latter possessing higher diversity. Unlike RAPD, AFLP analysis did not result in significant variation between these groups. Thus, the results from both analyses do not agree with the report of Genet & Belete (2000) who concluded, based on phenological and morphological data, that niger populations from MaNPRs have higher genetic diversity. Rather, the result suggests that the extent of cultivation of landraces does not always result in higher genetic diversity. Both AFLPs and

RAPDs revealed that the level of genetic variation in populations from the higher altitude group (> 2000 masl) and the lower altitude group (< 2000 masl) are similar and that altitude and level of genetic diversity did not correlate. Based on these results, it is concluded that the extent of the existing genetic diversity in Ethiopian niger is distributed over all its growing regions regardless of altitude and extent of cultivation. These characteristics make niger suitable for adaptation to diverse environmental conditions as the chances of finding more adaptive genotypes are high.

Genetic differentiation of populations may occur for any genetically variable trait that is favored under the existing selection conditions (Bossdorf *et al.*, 2005) as well as due to random stochastic processes, mutation and migration and the degree of such population differentiation could be estimated using different parameters. The overall degree of population differentiation estimated as G_{ST} (Nei, 1973) was 0.27 and 0.24 for AFLPs and RAPDs, respectively (Table 3). The RAPD based G_{ST} (0.24) revealed in this study was almost similar to the RAPD based mean G_{ST} (0.23) obtained for 18 outcrossing species (Nybom & Bartish, 2000). Nybom (2004) analyzed eight AFLP-based studies of outcrossing species and obtained a mean G_{ST} of 0.24. Therefore, this study demonstrates an average level of population differentiation with significant variation among niger populations. Partitioning of the total genetic diversity into within and among population components showed that 65% (RAPD) and 77% (AFLP) of the total variation was due to genetic variation within populations (AMOVA). Despite the relatively higher value obtained from AFLPs as compared to RAPDs, both marker systems demonstrated that a higher proportion of the total genetic variation is within populations, in *G. abyssinica*. Generally, at least 20% of the total genetic variation in niger is found among populations, implying that each population has unique genetic properties which makes it a significant unit for conservation efforts and breeding purposes. Therefore, it is recommended that as many populations as possible should be conserved *ex situ*, as it reduces the risk of losing unique genetic variants due to shifting of cultivation practices and many other factors.

Genetic distance and cluster analyses

The AFLP and RAPD based Nei's standard genetic distance (Nei, 1972) between niger populations ranged from 0.040 to 0.175 (mean = 0.118) and from 0.050 to 0.300 (mean = 0.176), respectively. The lowest and the highest genetic distance were recorded between populations from the same region and different regions, respectively, both with AFLP and RAPD. The comparison of matrices of Nei's standard genetic distance (from combined data set of AFLP and RAPD) and geographic distance between populations using normalized Mantel statistics (Mantel, 1967) revealed a significant positive correlation ($r = 0.266$; $P < 0.01$). Such positive correlation between geographic distance and genetic distance in outcrossing species is not uncommon (*e.g.* Ayres & Ryan, 1997; Shim & Jørgensen, 1999). The RAPD based UPGMA cluster analysis for the 70 niger populations revealed that the majority of the populations from the same region were clustered together. Populations from adjacent regions were also clustered to a considerable degree. Almost similar results were obtained when AFLP based

cluster analysis and PCoA (Fig. 11) were calculated, despite the low correlation (r = 0.37) between the similarity matrices of AFLP and RAPD.

Table 3. *Summary of gene diversity and AMOVA for eight Guizotia taxa*

Taxon-marker technique	Gene diversity			AMOVA		
	H_T	H_S	G_{ST}	PVWP[a]	PVAP[b]	F_{ST}
G. abyssinica-AFLP	0.32	0.20	0.27	76.56	23.44	0.23
G. abyssinica-RAPD	0.25	0.18	0.24	64.58	35.42	0.35
G. arborescens-AFLP	0.37	0.24	0.30	70.69	29.31	0.29
G. arborescens-RAPD	0.28	0.15	0.34	54.56	45.44	0.45
G. scabra ssp. *scabra*-AFLP	0.32	0.24	0.18	77.51	22.49	0.22
G. scabra ssp. *scabra*-RAPD	0.23	0.17	0.19	63.70	36.30	0.36
G. scabra ssp. *schimperi*-AFLP	0.32	0.24	0.19	82.67	17.33	0.17
G. scabra ssp. *schimperi*-RAPD	0.22	0.17	0.20	62.20	37.80	0.38
G. villosa-AFLP	0.33	0.24	0.19	74.10	25.90	0.26
G. villosa-RAPD	0.23	0.19	0.15	68.83	31.17	0.31
G. zavattarii-AFLP	0.37	0.17	0.41	49.88	50.12	0.50
G. zavattarii-RAPD	0.28	0.14	0.40	42.62	57.38	0.57
Chelelu-AFLP	-	0.37	-	-	-	-
Chelelu-RAPD	-	0.31	-	-	-	-
Ketcha-AFLP	-	0.33	-	-	-	-
Ketcha-RAPD	-	0.30	-	-	-	-
Mean-AFLP	0.34	0.25	0.26	63.05	36.95	0.28
Mean-RAPD	0.25	0.20	0.25	61.88	38.13	0.40

[a]Percent variation within populations. [b]Percent variation among populations.

Figure 11. Two-dimensional plot for the 17 populations of *Guizotia abyssinica* generated by PCoA from (i) AFLP (green), RAPD (red) and combined data (black). Note: the codes of the populations refer to their regions of origin (see Table 2).

The general trend in Ethiopian niger is that genetic similarity between populations increases with geographic proximity due to the corresponding increase in rate of gene flow. Therefore, germplasm collections should consider geographic distance between populations as one major criterion of sampling. However, this is not always the case as populations from geographically distant regions were also clustered in some cases, which may be explained by long distance movement of niger populations along with the movement of humans into new settlement areas in the past. Out of the total of 733 loci generated by AFLP and RAPD, no population specific monomorphic marker was found which may indicate continuous gene flow between populations. Generally, this study generated comprehensive information regarding genetic diversity in *G. abyssinica*.

Genetic diversity of wild and/or weedy guizotias (paper III)

Various parameters were used to analyze the genetic diversity in seven wild and/or weedy guizotias. The mean percentage of polymorphic loci per population (P_p) for *G. arborescens*, *G. scabra* ssp. *scabra*, *G. scabra* ssp. *schimperi*, *G. villosa* and *G. zavattarii* was 17%, 61%, 65%, 62% and 24% in the case of AFLP and 38%, 51%, 51%, 57% and 35% in the case of RAPD, respectively. When all loci were considered for each taxon as a whole, the percentage of polymorphic loci (P_S) was 28.5%, 84.5%, 90.0%, 83.9% and 50.1% for AFLP and 86.6%, 99.6%, 99.6%, 98.5% and 92.7% for RAPD in that order for the above taxa. The percentage of polymorphic loci (P_P) for Chelelu and Ketcha was 41% and 81% in the case of AFLP and 75% and 92% in the case of RAPD, respectively. Both AFLPs and RAPDs revealed a relatively higher level of genetic polymorphism in those *Guizotia* taxa with relatively wider geographic distribution in Ethiopia (*G. scabra* ssp. *scabra* and *G. scabra* ssp. *schimperi*), which might be partly due to introgressive hybridization between them. The mean AFLP based estimate of the Jaccard's similarity coefficient between populations of each taxon ranged from 0.24 (*G. scabra* ssp. *schimperi)* to 0.51 (*G. arborescens*) while that of RAPD ranged from 0.51 (*G. villosa*) to 0.89 (*G. arborescens*). Despite higher similarity obtained using AFLPs, both marker systems revealed that the similarity between populations of *G. arborescens* and *G. zavattarii* is higher as compared to that of the other guizotias.

Considering AFLPs and RAPDs together, the extent of genetic similarity between populations of *G. scabra* ssp. *scabra*, *G. scabra* ssp. *schimperi* and *G. villosa* was almost similar. These guizotias have a similar level of total genetic diversity as estimated by using gene diversity estimate (H_T), which is also similar to that of *G. abyssinica* (Table 3). The G_{ST} and F_{ST} values generated from AFLPs and RAPDs indicated that all guizotias are more diverse within with less genetic differentiation between populations with the exception of *G. zavattarii* (Table 3), which is in agreement with the general understanding that outcrossing species tend to be more diverse within, with less genetic differentiation between populations (Hamrick & Godt, 1996; Nybom, 2004). Significant genetic variation between populations was obtained in all guizotias (AMOVA; $P < 0.001$). The estimate of the total genetic diversity was higher in *G. arborescens* and *G. zavattarii* as

compared to the other guizotias mainly due to their higher population differentiation (Table 3; Fig. 12).

A lower percentage of polymorphic loci and higher population differentiation in *G. arborescens* and *G. zavattarii* as compared to their relatively common congeners might be best explained by the fact that these species are rare with small population size and highly localized; as species with such ecological characteristics tend to have this type of population genetic structure (Loveless & Hamrick, 1984; Slatkin, 1987). The counter-example to this trend is the relatively high polymorphism and genetic variation revealed within the populations of Chelelu and Ketcha, regardless of their small population size and seemingly being isolated populations. The implication is that genetic drift could be counterbalanced by other factors that promote genetic variation (*e.g.* a high rate of mutation and interspecific hybridization) even in small populations. The relatively higher within population diversity in Chelelu and Ketcha (H_S; Table 3) might also be explained by the fact that they are perennials. The general trend is that perennials have higher within population diversity as compared to annuals, provided that they have similar mating systems (Hamrick & Godt, 1989; Nybom & Bartish, 2000). However, this does not explain the low within population variation in the other two perennial species (*G. arborescens* and *G. zavattarii*). Therefore, different factors such as life forms and geographic distribution with differential intensity of selection contributed differently to the population genetic structure of different taxa of the genus *Guizotia*.

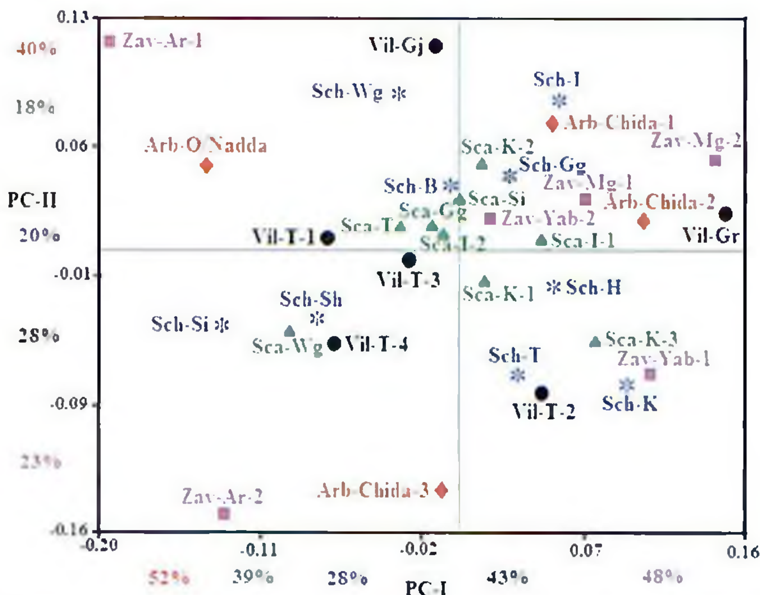


Figure 12. Two-dimensional plot for the populations of (i) *Guizotia arborescens* (red), *Guizotia scabra* ssp. *scabra* (green), *Guizotia scabra* ssp. *schimperi* (blue), *Guizotia villosa* (black) and *Guizotia zavattarii* (pink) generated by PCoA from RAPD band profiles. Note: The code for each population is derived and combined from taxon name-region/place-number. For example, Vil-T-4 refers to population number 4 of *G. villosa* collected from Tigray (See Table 1 of paper III).

Analysis of the extent of genetic diversity in each population and genetic similarity between populations is important to suggest genetic diversity centers for each taxon. The abundance of *G. arborescens* is relatively better in Chida than in Omo Nadda. Correspondingly, this study revealed that the genetic diversity of Chida populations is slightly higher than that of Omo Nadda (see Table 3 of paper III). This may suggest Chida as a micro genetic diversity center of this species and thus should be considered as a priority site for *in situ* conservation. On the other hand, the extent of genetic variability of the two subspecies of *G. scabra* seems to be distributed regardless of their geographic locations within the country. *G. scabra* ssp. *scabra* is relatively frequent in southwestern Ethiopia (*e.g.* in Jimma and Keficho). In this region, its distribution overlaps with that of *G. scabra* ssp. *schimperi* and *G. arborescens*. Therefore, this region might be considered for *in situ* conservation of the two subspecies of *G. scabra*. Although *G. villosa* is relatively common in Tigray, populations from Gonder and Gojam seem to possess higher genetic variation as compared to those of Tigray. Thus, future conservation activities, including germplasm collection, should also consider populations located at the periphery of the species distribution range. In the case of *G. zavattarii*, populations from Arero are genetically less similar to the other populations and seem to have slightly higher diversity. The fact that *G. zavattarii* populations are highly differentiated emphasizes the importance of conservation of populations in all the three sites (Mega, Yabelo and Arero). However, if priority needs to be given, Arero should be preferred, as it is a relatively stable community.

Analysis of population differentiation based on combined data of AFLPs and RAPDs revealed a significantly higher population differentiation ($P < 0.001$) in perennials ($G_{ST} = 0.32$; $F_{ST} = 0.40$) as compared to annuals ($G_{ST} = 0.18$; $F_{ST} = 0.33$). The mean G_{ST} for *Guizotia* taxa was higher than the average values reported for several outcrossing species (*e.g.* Bussell, 1999; Nybom, 2004), suggesting the relatively high population differentiation in wild and/or weedy guizotias, specifically in *G. arborescens* and in *G. zavattarii*. Since they are Afromontane species, they are restricted to Afromontane archipelago and some populations of each taxon are separated by lowlands, which contributed to the significant differentiation between populations. Generally, although the extent of within and between populations genetic variation varies among taxa, a substantial amount of overall genetic diversity was revealed in all guizotias. Acknowledging the existing genetic variation within and between populations is important for conservation and breeding purposes. The combination of a significant level of genetic variation and the cross-compatibility of *G. scabra* ssp. *schimperi*, *G. scabra* ssp. *scabra* and *G. villosa* with *G. abyssinica* is an opportunity for improvement of *G. abyssinica*. This study highlights the importance of molecular analysis in understanding the genetic diversity and population structure of various guizotias and contributes to the knowledge of conservation of genetic resources in the Ethiopian flora.

Phylogenetic relationship between guizotias (papers IV, V & VI)

Phenetic and phylogenetic relationships between *Guizotia* taxa were investigated based on data generated using AFLP, RAPD and DNA sequencing (ITS of nrDNA

and various cpDNA regions). A total of 658 AFLP and 353 RAPD loci generated by seven AFLP PCs and ten RAPD primers were used for this analysis. The extent of genetic similarity between the taxa was investigated based on molecular variance, genetic distance and F_{ST} from AFLPs and RAPDs (paper IV). Both AFLP and RAPD based AMOVA revealed a highly significant variation ($P < 0.001$) between the taxa. Cluster analyses revealed that intraspecific populations were clearly clustered together with higher genetic similarity between them as compared to the genetic similarity between any pair of interspecific populations and thus proved a clear differentiation among *Guizotia* taxa.

Complete sequences of the ITS were obtained for *G. abyssinica*, *G. arborescens*, *G. scabra* ssp. *scabra*, *G. scabra* ssp. *schimperi*, *G. villosa* and *G. zavattarii* (paper V). The length of the entire ITS region was 645 bp in these guizotias, except that one sample of *G. villosa* was shorter by one nucleotide. Forty parsimony informative characters were obtained within the genus *Guizotia* for the entire ITS. A total aligned length of the sequences of the five cpDNA regions for all guizotias and three outgroups was 2997 bp, of which 45 characters were parsimony informative (paper VI). A significant length variation between taxa of the genus *Guizotia* was obtained in the entire *trn*T/*trn*L intergenic spacer, ranging from 582 (*G. zavattarii*) to 634 (Ketcha) nucleotides. The complete sequence of the *trn*L intron is 433 bp for all guizotias. The *trnL*/F intergenic spacer is shorter than both the *trn*T/*trn*L spacer and the *trn*L intron (ranging from 345-347 nucleotides long). Sequences of 1255 bp long 5'-most portion of the *mat*K gene and 301 bp long 3'*trn*K/*mat*K portion of the *trn*K intron were also available for analysis. Four variable sites were revealed in the *mat*K coding region between *Guizotia* taxa (Table 4), two of which resulting in synonymous and the other two in non-synonymous amino acid substitutions.

Table 4. *Part of the first 1254 nucleotide sequences and 408 amino acid (AA) sequences of matK coding region of Guizotia taxa showing variable regions*

	1		255[a]		588[a]			1171[a]		1209[a]			
G. abyssinica-DNA	ATG	---	AAG	---	GGT	TCT	---	GTT	---	CGC	TTT	---	TCT
G. abyssinica-AA	M	---	K	---	G	S	---	V	---	R	F	---	S
G. arborescens-DNA	ATG	---	AAG	---	GGT	TCT	---	ATT	---	CGA	TTT	---	TCT
G. arborescens-AA	M	---	K	---	G	S	---	I	---	R	F	---	S
G. scabra ssp. *scabra*-DNA	ATG	---	AAG	---	GGT	TCT	---	GTT	---	CGC	TTT	---	TCT
G. scabra ssp. *scabra*-AA	M	---	K	---	G	S	---	V	---	R	F	---	S
G. scabra ssp. *schimperi*-DNA	ATG	---	AAG	---	GGT	TCT	---	GTT	---	CGC	TTT	---	TCT
G. scabra ssp. *schimperi*-AA	M	---	K	---	G	S	---	V	---	R	F	---	S
G. villosa-DNA	ATG	---	AAG	---	GGT	TCT	---	GTT	---	CGC	TTT	---	TCT
G. villosa-AA	M	---	K	---	G	S	---	V	---	R	F	---	S
G. zavattarii-DNA	ATG	---	AAG	---	GGC	TCT	---	ATT	---	CGC	TTT	---	TCT
G. zavattarii-AA	M	---	K	---	G	S	---	I	---	R	F	---	S
Chelelu-DNA	ATG	---	AAT	---	GGT	TCT	---	ATT	---	CGC	TTT	---	TCT
Chelelu-AA	M	---	N	---	G	S	---	I	---	R	F	---	S
Ketcha-DNA	ATG	---	AAT	---	GGT	TCT	---	ATT	---	CGC	TTT	---	TCT
Ketcha-AA	M	---	N	---	G	S	---	I	---	R	F	---	S
	1		85[b]					391b					

[a]Variable sites of DNA sequence. [b]Variable sites of amino acid sequence. Note: Variable nucleotides/amino acids are indicated in bold, and gaps are represented by '---'.

G. abyssinica versus other guizotias

The phylogenetic analysis of the cpDNA regions produced a clade that contains *G. abyssinica*, *G. scabra* ssp. *scabra*, *G. scabra* ssp. *schimperi* and *G. villosa*, which indicates their close phylogenetic relationship (Fig. 13C). These taxa are cross compatible and produce a viable and fertile hybrid (Dagne, 1994a). It is also interesting to note that these taxa possess identical sequences of the *mat*K coding region but differ at least at two of the four variable sites from the other taxa (Table 4). However, they appear to form two sub-clades when analyzed based on the combined data set of AFLP and RAPD (paper IV) and ITS data (paper V), in which *G. abyssinica* and *G. scabra* ssp. *schimperi* form one sub-cluster while *G. villosa* and *G. scabra* ssp. *scabra* form another (Fig. 13A, 13B). Although the bootstrap support at some branches of Figure 13A is low, a similar phenogram with high bootstrap support was generated when Unweighted Pair Group Method with Arithmetic mean (UPGMA) method was applied (see figure 2C of paper IV). This grouping is in agreement with the grouping based on karyotypes (Dagne, 1995) and meiotic behavior of their hybrids (Dagne, 1994a). The ITS data was different from the combined data of AFLP and RAPD in that it revealed two forms of *G. abyssinica* that fell into different sub-clades. Despite this difference, the two data sets agree in that at least one form of *G. abyssinica* is most closely related to *G. scabra* ssp. *schimperi* and that *G. scabra* ssp. *scabra* and *G. villosa* are closely related.

The lower genetic variation obtained in the cpDNA regions as compared to the ITS between *G. abyssinica*, *G. villosa* and the two subspecies of *G. scabra* is in line with the lower evolutionary rate of the chloroplast genome as compared to the nuclear genome (*e.g.* Small, Cronn & Wendel, 2004). Molecular marker techniques such as AFLP are useful tools to resolve phylogeny when DNA sequencing fails, as they sample a larger portion of the genome (Desprès *et al.*, 2003). The general similarity of results obtained by AFLP and DNA sequence data in this study substantiates the use of AFLPs for phylogenetic inference at lower taxonomic levels as previously suggested (Desprès *et al.*, 2003; Weising *et al.*, 2005). Based on various evidences, several authors suggest that *G. abyssinica* might have originated from *G. scabra* ssp. *schimperi* through selection and cultivation (Baagøe, 1974; Hiremath & Murthy, 1988; Murthy, Hiremath & Salimath, 1993; Dagne, 1994a, 1995, 2001). Similarly, this study showed that *G. scabra* ssp. *schimperi* is the most likely progenitor of *G. abyssinica* although *G. scabra* ssp. *scabra* and *G. villosa* might also have some contribution successively albeit at various stages of evolution to the origin of *G. abyssinica*.

G. scabra ssp. scabra, G. scabra ssp. schimperi and G. villosa

The combined data set of AFLPs and RAPDs have shown that *G. scabra* ssp. *scabra* and *G. scabra* ssp. *schimperi* are more closely related to each other as compared to most of the pairs of taxa compared in this study. However, AFLP based analysis showed that *G. scabra* ssp. *scabra* vs. *G. villosa* are slightly more closely related to one another than *G. scabra* ssp. *scabra* vs. *G. scabra* ssp. *schimperi* (Paper IV). Sequence divergence analysis based on the cpDNA regions revealed a similar level of divergence between any pair of *G. scabra* ssp. *scabra*,

G. scabra ssp. *schimperi*, *G. abyssinica* and *G. villosa*. On the other hand, the ITS data placed these subspecies under different sub-clades (Fig. 13B). As a result, all sets of data generated in this study agree that the genetic similarity between *G. scabra* ssp. *scabra* and *G. scabra* ssp. *schimperi* is not more than the similarity revealed, for example, between *G. scabra* ssp. *scabra* and *G. villosa* and between *G. scabra* ssp. *schimperi* and *G. abyssinica*. The combined result of this study and the previously reported morphological and karyotypic differences (Hiremath & Murthy, 1992; Dagne, 1995) lead to the conclusion that *G. scabra* ssp. *scabra* and *G. scabra* ssp. *schimperi* are best viewed as separate species rather than as the subspecies of the same species. Based on morphological similarities, Baagøe (1974) suggested that *G. villosa* might have been derived from *G. scabra* (without mentioning the subspecies). The close similarity between these taxa was also reported by Dagne (1994a, 1995). The AFLP, RAPD and ITS data revealing that *G. villosa* is more closely related to *G. scabra* ssp. *scabra* than to *G. scabra* ssp. *schimperi* thus disagree with the speculation of Hiremath & Murthy (1992) that *G. villosa* has evolved from *G. scabra* ssp. *schimperi*.

Chelelu and Ketcha

The Chelelu and Ketcha populations were first reported to belong to the genus *Guizotia* based on their morphology and karyotypes (Dagne, 1995), which was further strengthened by the study of cross-compatibility with other guizotias and chromosome pairing of the hybrids (Dagne, 2001). In this study, there was no complete agreement between different sets of data regarding the degree of relatedness of these populations to each other and to the taxonomically recognized *Guizotia* taxa. For example, when the overall cpDNA regions and the combined data set of AFLP and RAPD were considered, Chelelu and Ketcha appeared to be more closely related to each other than to the other taxa (Fig. 13A); but was not supported when only the two intergenic spacers of cpDNA were considered (Fig. 13C). The grouping of Chelelu together with *G. abyssinica* and *G. scabra* ssp. *schimperi* and the grouping of Ketcha together with *G. villosa* and *G. scabra* ssp. *scabra* by Dagne (1995, 2001) were not supported by this study, except that the cluster analysis based on RAPD data groups Chelelu and *G. abyssinica* together. In spite of these differences, both Chelelu and Ketcha were clearly separated from the other taxa and from each other. The AFLP and RAPD based genetic distance between these populations and the other guizotias was equivalent to the genetic distance between the taxonomically recognized species. Additionally, the cpDNA sequence divergence between Chelelu and Ketcha was higher than the sequence divergence between any pair of taxa within the group of *G. abyssinica*, *G. scabra* ssp. *scabra*, *G. scabra* ssp. *schimperi* and *G. villosa*. Generally, the phenetic and phylogenetic analyses of cpDNA data and the data from the two marker techniques confirm that these populations certainly belong to the genus *Guizotia* and that they are distinct enough to be treated as separate species (Fig. 13). It will be of interest to see the relationship between *Sigesbeckia* species moved to the genus *Guizotia* by Schulz (1990) and the species excluded from the genus *Guizotia* by Baagøe (1974) as well as Chelelu and Ketcha in the revision of the genus by considering as many characters as possible.

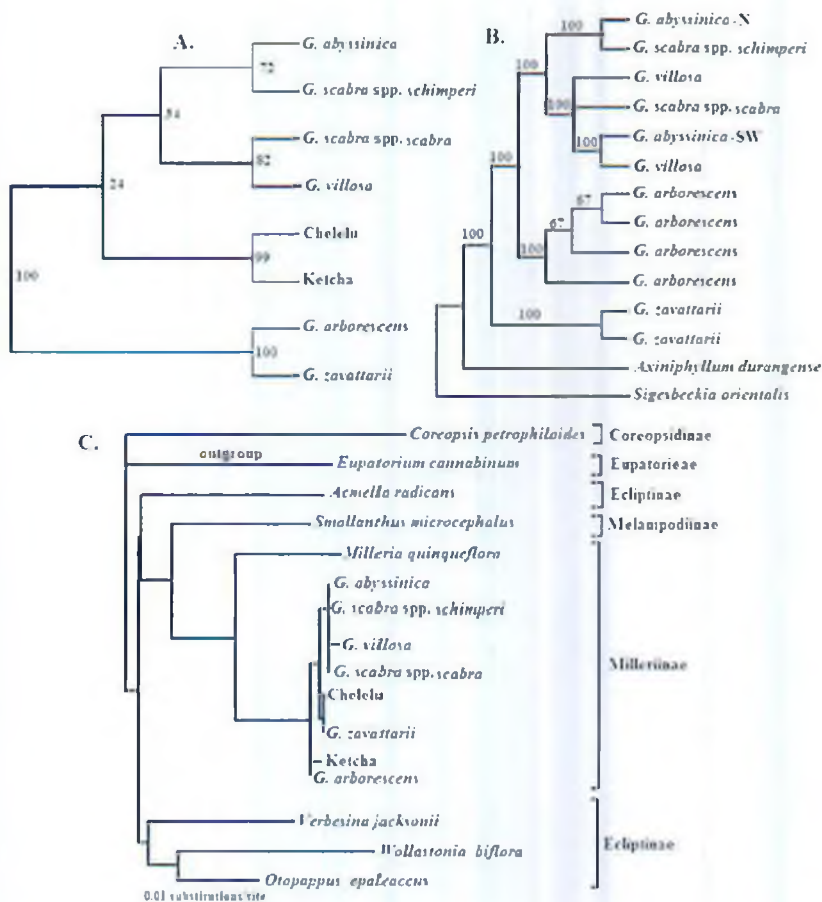


Figure 13. (A) Neighbor-joining tree from the combined data of AFLP and RAPD, (B) most parsimonious tree for the entire ITS region, and (C) neighbor-joining tree from the *trn*T/*trn*L and *trn*L/*trn*F intergenic spacers of cpDNA showing the phylogenetic relationship between guizotias and other closely related species.

G. arborescens and G. zavattarii

Dagne (1995, 2001), based on karyotype similarity and a high degree of chromosome pairing affinity in the F1 hybrid between *G. arborescens* and *G. zavattarii*, concluded that these species are more closely related to one another than to the other taxa. The degree of similarity between these species is relatively low when all data sets in this study are considered and thus do not support the previous grouping. Most putative taxon-specific markers generated both from AFLPs and RAPDs were specific to *G. arborescens* or *G. zavattarii* implying that these species are genetically distant from the other taxa and from each other. The phylogenetic analyses conducted in this study showed that *G. arborescens*, *G. zavattarii*, Ketcha and Chelelu evolved earlier than the rest of the taxa during the evolutionary history of the genus (Fig. 13A, 13C). Thus, Baagøe's (1974) suggestion regarding the derivation of *G. zavattarii* from *G. scabra* seems unlikely. The advancement versus primitiveness of morphological characters

suggested for *Guizotia* species by Baagoe (1974) is not in line with the results of phylogenetic analyses of this study.

Diagnostic markers

Twelve AFLP and four RAPD taxon specific markers were identified in this study. The identification of such taxon specific DNA markers in *Guizotia* has a potential use for conservation of these taxa and, most importantly, to improve *G. abyssinica*. In other words, these markers are potentially useful in *G. abyssinica* breeding, in species identification and determination of the degree of natural hybridization between *Guizotia* taxa that have overlapping geographic distribution. Such markers might be linked to desirable traits that are worth transferring to *G. abyssinica*, which needs to be one target of future studies. If the cross-incompatibility between *G. abyssinica* and *G. zavattarii* is postzygotic as suspected by Dagne (1994a), it can be overcome, for example, through ovule culture (*e.g.* Kumlehn & Nitzsche, 1996). Thus, *G. zavattarii* can be considered as a potential candidate as a source for some desirable traits, such as high OA seed oil (Dagne & Jonsson, 1997) to improve *G. abyssinica*. Conversion of such potential diagnostic markers to Sequence Characterized Amplified Region (SCAR) markers (Paran & Michelmore, 1993) improves the efficiency of their use as a molecular tool in marker assisted selection.

The position of *Guizotia* Cass. within Heliantheae (papers V &VI)

The genus *Guizotia* has been placed under different subtribes (Coreopsidinae, Verbesininae and Milleriinae) since 1873 (Bentham, 1873; Baagoe, 1974; Stuessy, 1977; Robinson, 1981; Karis, 1993). According to Stuessy's systematic review of Heliantheae (Stuessy, 1977), *Acmella radicans* (Jacq.) R. K. Jansen and *Verbesina jacksonii* B. L. Turner belong to subtribe Verbesininae, a subtribe that later was incorporated into Robinson's subtribe Ecliptinae (Robinson, 1981). *Coreopsis petrophiloides* belongs to the subtribe Coreopsidinae of the tribe Heliantheae, while *Eupatoria cannabinum* belongs to the tribe Eupatorieae (Stuessy, 1977; Robinson, 1981). The DNA sequence analysis in this study revealed that the nucleotide divergence between guizotias and *C. petrophiloides* was higher than the divergence between guizotias and *E. cannabinum*, which disagrees with the placement of the genus *Guizotia* under the subtribe Coreopsidinae. On the other hand, the ITS sequence of guizotias is more similar to various species of the subtribe Milleriinae (*e.g. Sigesbeckia* spp., *Axiniphyllum* spp. *Trigonospermum* spp.) than to species of other subtribes of the tribe Heliantheae (paper V). The phylogenetic analysis based on 32 parsimony informative characters from DNA sequence data of two intergenic spacers of cpDNA revealed that *Milleria quinqueflora* L. (subtribe Milleriinae; Fig. 13C) was the closest species to the genus *Guizotia*. The second closest species to the genus *Guizotia* was *Smallanthus microcephalus* (Hieron) H. Rob. Robinson (1981) placed *S. microcephalus* under the subtribe Melampodiinae. However, Panero *et al.* (1999) advised the transfer of this genus from this subtribe to the subtribe Milleriinae. Thus, this study strongly supports Robinson's (1981) placement of the genus *Guizotia* under the subtribe Milleriinae.

Conclusions

- *G. abyssinica* has a wide genetic basis that can be used for its improvement through breeding. The extent of niger genetic variation within populations is distributed regardless of the extent and altitude of cultivation, suggesting that all regions where niger is currently grown are equally important from a species conservation point of view. However, for the promotion of *in situ* conservation and enrichment of diversity through introgressive hybridization between both cultivated and wild forms some regions can be prioritized. Welo, where the highest within populations genetic variation was obtained and other regions such as parts of Jimma, Illubabor, Welega, Gojam and Tigray, where more than two *Guizotia* taxa overlap significantly, can be considered as priority regions.
- There is a significant "regional" differentiation between niger populations, and a significant positive correlation between genetic and geographic distance between populations.
- All wild and/or weedy guizotias possess a quite high level of genetic diversity.
- Higher genetic variation is found within populations than between populations in all guizotias, except in *G. zavattarii* that seems to have almost equal proportion. Significant population differentiation was obtained in all guizotias, with *G. zavattarii* exhibiting the highest differentiation.
- Chelelu and Ketcha acquired higher within-population genetic variation than the other guizotias while *G. arborescens* and *G. zavattarii* showed lower percentage of polymorphic loci.
- AFLP was superior to RAPD in various characteristics and thus has to be preferred to characterize accessions of niger held in gene bank. In comparison to RAPD, AFLP revealed genetic relationships between *Guizotia* taxa that is more inline with the DNA sequence data and the cytogenetic and hybridization studies and thus can be concluded that AFLPs give strong phylogenetic signals.
- *G. scabra* ssp. *schimperi* is the most likely progenitor of *G. abyssinica* although there are indications that *G. scabra* ssp. *scabra* and *G. villosa* might also have contributed successively at various stages to the evolution of *G. abyssinica*.
- The two subspecies of *G. scabra*, and Chelelu and Ketcha are best viewed at present as separate species of the genus *Guizotia*. However, a complete revision of the genus by considering those species taken into and out from the genus by different authors needs to be considered.
- Highly localized guizotias evolved first during the evolutionary history of the genus *Guizotia*.
- *Guizotia* Cass. belongs to the subtribe Milleriinae of tribe Heliantheae.

Recommendations and future prospects

G. abyssinica, just like many other domesticates (*e.g. Coffea arabica* L.), is Ethiopia's contribution to the world. Maintaining and promoting the genetic diversity of this crop and its wild relatives at their center of origin is highly valuable not only for Ethiopia but also for the rest of the world. This study assessed the level and pattern of genetic diversity in niger and its wild and/or weedy relatives and generated information of importance for conservation, breeding and utilization. In Ethiopia, niger germplasm collection has been conducted mainly in regions of its major cultivation and most of the accessions are from areas with better access to modern transportation and accordingly their representation of the existing genetic diversity of the crop within the country is questionable. The significant variation between populations and the regional differentiation of *G. abyssinica* populations obtained in this study has an important implication for its conservation and utilization. It indicates the need to conserve a large number of populations from all of its growing regions *ex situ* to prevent the loss of unique genetic variants. In other words, future germplasm collections should represent all growing regions by giving special emphasis to areas that are not yet represented or are underrepresented. In regions where niger production is not the priority, populations with unique genetic properties are at risk of being lost, as farmers may not continue growing them. Transferring representative samples of such populations to MaNPRs for *in situ* conservation, as a complementary approach to *ex situ* conservation, helps to maintain the existing genetic variation of the crop and increases the genetic variation within regions of major cultivation. Before characterizing gene bank collections at the molecular level, populations from remote and underrepresented areas should be collected first. This helps to increase the range of genetic variation among accessions and thus facilitates generation of a well represented core collection.

G. abyssinica has to be improved in order to make it a competitive oil crop worldwide. Niger breeding should focus primarily on increasing seed yield, as it is the major factor that hampers large-scale production and commercialization of this crop. Breeding for pest and disease resistance and oil content and quality is also of interest. The existing wide genetic basis in Ethiopian niger suggests that achieving these objectives may well be possible both through conventional breeding and marker assisted selection. Additionally, the considerable level of genetic variation in wild and/or weedy guizotias, especially in those that are cross-compatible with niger is a valuable resource to improve niger through interspecific hybridization in order to transfer desirable genes that may exist in these species into *G. abyssinica*. The diagnostic markers revealed in this study are useful in this regard. A loss of yield through shattering, for example, can be minimized by selecting non-shattering genotypes with determinate growth habit, provided that genes responsible for these traits exist in niger or in its wild and/or weedy relatives. Seed yield can be improved through stepwise selection of large seed size, a higher number of seeds/capitula, capitula/plant and branches/plant, as high variation in these traits exists in the Ethiopian niger gene pool. Problems associated to lodging can be overcome by looking for genotypes with short stature and strong stem.

Niger seed is well known for its high linoleic acid oil. Oleic acid accounts, commonly, for 5-13% of niger seed oil. However, there are indications that high OA (up to 40%) niger genotypes exist in the Ethiopian gene pool, which is a manifestation of the existing genetic diversity. OA has a higher oxidative stability than LA (Wanasundara & Shahidi, 1994), which makes high OA niger oil preferable to some applications in food industry. Therefore, developing niger varieties with a higher percentage of OA should be one direction of breeding in niger. The molecular marker techniques used in this study revealed high genetic polymorphism in guizotias. These and other molecular techniques could be extended to mapping and linkage analysis in *G. abyssinica*, as molecular markers linked to desirable traits are useful tools for the early generation selection of genotypes with desirable traits in niger breeding programs. Generally, in view of the existing genetic diversity, the potential for improving niger through breeding is promising.

G. zavattarii and *G. arborescens* are characterized by a highly localized geographic distribution, small population sizes, lower percentage of polymorphic loci and higher population differentiation in comparison with other taxonomically recognized guizotias. The phylogenetic evidence showed that they evolved earlier than most of the other guizotias. These conditions and the continuously increasing human activities in their natural habitats are signals of urgency to start their conservation immediately. Only single populations of Chelelu and Ketcha are known so far, which may indicate their localized distribution. It is of great interest to explore their geographic distribution and unravel the possible reasons behind their high genetic variation.

The cpDNA sequence data did not resolve the phylogenetic relationship between *G. abyssinica*, *G. villosa* and the two subspecies of *G. scabra*. It might be worthy to study more variable cpDNA regions and low-copy nuclear genes to resolve the phylogeny of this group and to further assess the results obtained based on the ITS data.

References

Abebe, M., Yermanos, D.M. & Bingham, F.T. 1978. The ecophysiology of noug (*Guizotia abyssinica* Cass.). *African Journal of Agricultural Science 5*, 55-66.

Adda, S., Reddy, T.P. & Kishor, P.B.K. 1993a. Embryogenesis and organogenesis in cultured anthers of an oil yielding crop niger (*Guizotia abyssinica* Cass). *Plant Cell Tissue and Organ Culture 35*, 75-80.

Adda, S., Reddy, T.P. & Kishor, P.B.K. 1993b. Plant regeneration from cotyledons of niger. *Plant Cell Tissue and Organ Culture 32*, 131-135.

Adda, S., Reddy, T.P. & Kishor, P.B.K. 1994a. Androclonal variation in niger (*Guizotia abyssinica* Cass.). *Euphytica 79*, 59-64.

Adda, S., Reddy, T.P. & Kishor, P.B.K. 1994b. Somatic embryogenesis and organogenesis in *Guizotia abyssinica*. *In Vitro Cellular and Developmental Biology-Plant 30*, 104-107.

Alonso-Blanco, C., Peeters, A.J.M., Koornneef, M., Lister, C., Dean, C., van den Bosch, N., Pot, J. & Kuiper, M.T.R. 1998. Development of an AFLP based linkage map of Ler, Col and Cvi *Arabidopsis thaliana* ecotypes and construction of a Ler/Cvi recombinant inbred line population. *The Plant Journal 14*, 259–271.

Alvarez, J.B., Moral, A. & Martin, L.M. 2006. Polymorphism and genetic diversity for the seed storage proteins in Spanish cultivated einkorn wheat (*Triticum monococcum* L. ssp. monococcum). *Genetic Resources and Crop Evolution 53*, 1061-1067.

Archak, S., Gaikwad, A.B., Gautam, D., Rao, E.V.V.B., Swamy, K.R.M. & Karihaloo, J.L. 2003. Comparative assessment of DNA fingerprinting techniques (RAPD, ISSR and AFLP) for genetic analysis of cashew (*Anacardium occidentale* L.) accessions of India. *Genome 46*, 362-369.

Archibald, J.K., Mort, M.E. & Crawford, D.J. 2003. Bayesian inference of phylogeny: a non-technical primer. *Taxon 52*, 187–191.

Asilbekova, D.T., Ul'chenko, N.T., Rakhimova, N.K., Nigmatullaev, A.M. & Glushenkoval, A.I. 2005. Seed lipids from *Crotalaria alata* and *Guizotia abyssinica*. *Chemistry of Natural Compounds 41*, 596-597.

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

Ayana, A. & Bekele, E. 1999. Multivariate analysis of morphological variation in sorghum (*Sorghum bicolor* (L.) Moench) germplasm from Ethiopia and Eritrea. *Genetic Resources and Crop Evolution 46*, 273-284.

Ayana, A., Bryngelsson, T. & Bekele, E. 2001. Geographic and altitudinal allozyme variation in sorghum (*Sorghum bicolor* (L.) Moench) landraces from Ethiopia and Eritrea. *Hereditas 135*, 1-12.

Ayres, D.R. & Ryan, F.J. 1997. The clonal and population structure of a rare endemic plant, *Wyethia reticulate* (Asteraceae): allozyme and RAPD analysis. *Molecular Ecology 6*, 761-772.

Baagoe, J. 1974. The genus *Guizotia* (Compositae). A taxonomic revision. *Botanisk Tidskrift 69*, 1-39.

Baagoe, J. 1977. Taxonomical application of the ligule microcharacters in Compositae, I: Anthemideae, Heliantheae, and Tageteae. *Botanisk Tidskrift 71*, 193-224.

Baldwin, B.G. 1992. Phylogenetic utility of the internal transcribed spacers of nuclear ribosomal DNA in plants: an example from the Compositae. *Molecular Phylogenetics and Evolution 1*, 3-16.

Baldwin, B.G., Sanderson, M.J., Porter, J.M., Wojciechowski, M.F., Campbell, C.S. & Donoghue, M.J. 1995. The ITS region of nuclear ribosomal DNA: a valuable source of evidence on angiosperm phylogeny. *Annals of the Missouri Botanical Garden 82*, 247–277.

Bänfer, G., Fiala, B. & Weising, K. 2004. AFLP analysis of phylogenetic relationships among myrmecophytic species of *Macaranga* (Euphorbiaceae) and their allies. *Plant Systematics and Evolution 249*, 213-231.

Barata, C. & Carena, M.J. 2006. Classification of North Dakota maize inbred lines into heterotic groups based on molecular and testcross data. *Euphytica 151*, 339-349.
Barrett, B.A. & Kidwell, K.K. 1998. AFLP-based genetic diversity assessment among wheat cultivars from the Pacific Northwest. *Crop Science 38*, 1261–1271.
Bayer, R.J., Greber, D.G. & Bagnall, N.H. 2002. Phylogeny of Australian Gnaphalieae (Asteraceae) based on chloroplast and nuclear sequences, the *trnL* intron, *trn*L/*trn*F intergenic spacer, *mat*K, and ETS. *Systematic Botany 27*, 801–814.
Bekele, E. 1985. The biology of cereal landrace populations: problems of gene conservation, plant-breeding selection schemes and sample size requirement. *Hereditas 103*, 119–134.
Bekele, E., Fido, R.J., Tatham, A.S. & Shewry, P.R. 1995. Heterogeneity and polymorphism of seed proteins in tef (*Eragrostis tef*). *Hereditas 122*, 67-72.
Bellon, M.R. 1996. The dynamics of crop infraspecific diversity: A conceptual framework at the farmer level. *Economic Bot*any *50*, 26-39.
Bennetzen, J.L., Ma, J. & Devos, K.M. 2005. Mechanisms of recent genome size variation in flowering plants. *Annals of Botany 95*, 127-132.
Bentham, G. 1873. Notes on the classification, history, and geographical distribution of Compositeae. *Botanical Journal of the Linnean Society 13*, 335- 577.
Birky, C.W. 1995. Uniparental inheritance of mitochondrial and chloroplast genes: mechanisms and evolution. *Proceedings of the National Academy of Sciences of the United States of America 92*, 11331-11338.
Blattner, F. R., Weising, K., Bänfer, G., Maschwitz, U. & Fiala, B. 2001. Molecular analysis of phylogenetic relationships among myrmecophytic *Macaranga* species (Euphorbiaceae). *Molecular Phylogenetics and Evolution 19*, 331–344.
Bossdorf, O., Auge, H., Lafuma, L., Rogers, W.E., Siemann, E. & Prati, D. 2005. Phenotypic and genetic differentiation between native and introduced plant populations. *Oecologia 144*, 1-11.
Boukar, O., Kong, L., Singh, B.B., Murdock, L. & Ohm, H.W. 2004. AFLP and AFLP-derived SCAR markers associated with *Striga gesnerioides* resistance in cowpea. *Crop Science 44*, 1259–1264.
Briquet, J. 1935. International rules of botanical nomenclature. Jena. 1-151 pp.
Brush, B. & Meng, E. 1998. Farmers' evaluation and conservation of crop genetic resources. *Genetic Resources and Crop Evolution 45*, 139-150
Brush, S.B. 1995. *In situ* conservation of landraces in centers of crop diversity. *Crop Science 35*, 346-354.
Bussell, J.D. 1999. The distribution of random amplified polymorphic DNA (RAPD) diversity amongst populations of *Isotoma petraea* (Lobeliaceae). *Molecular Ecology 8*, 775-789
Cavagnaro, P., Cavagnaro, J., Lemes, J. Masuelli, R. & Passera, C. 2006. Genetic diversity among varieties of the native forage grass *Trichloris crinita* based on AFLP markers, morphological characters, and quantitative agronomic traits. *Genome 49*, 906-918.
Clegg, M.T. 1993. Chloroplast gene sequences and the study of plant evolution. *Proceedings of the National Academy of Sciences of the United States of America 90*, 363-367.
Crawford, D.J. & Mort, M.E. 2005. Phylogeny of eastern North American Coreopsis (Asteraceae-Coreopsideae): insights from nuclear and plastid sequences and comments on character evolution. *American Journal of Botany 92*, 330–336.
Dagne, K. & Heneen, W.K. 1992. The karyotype and nucleoli of *Guizotia abyssinica* (Compositae). *Hereditas 117*, 73-83
Dagne, K. & 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.
Dagne, K. 1994a. Meiosis in interspecific hybrids and genomic interrelationships in *Guizotia* Cass. (Compositae). *Hereditas 121*, 119-129.
Dagne, K. 1994b. Mitotic and meiotic behavior of B-chromosomes of *Guizotia scabra* (Vis.) Chiov. ssp. *scabra* (Compositae). *Caryologia 48*, 35-45.
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.

Desprès, L., Gielly, L., Redoutet, B. & Taberlet, P. 2003. Using AFLP to resolve phylogenetic relationships in a morphologically diversified plant species complex when nuclear and chloroplast sequences fail to reveal variability. *Molecular Phylogenetics and Evolution 27*, 185-196.

Devi, N.A., Kumar, C.M., Naidu, Y.R. & Rao, M.N. 2006. Production and evaluation of biodiesel from sunflower (*Helianthus annuus*) and niger seed oil (*Guizotia abyssinica*). *Asian Journal of chemistry 4*, 2951-2958.

Dutta, P.C., Helmersson, S., Kebedu, E. & 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.

Friis, I. B. 1971. A new species of *Guizotia* (Compositae) from northeast tropical Africa. *Norwegian Journal of Botany 18*, 231-234.

Fujimoto, Y., Kakinuma, K., Eguchi, T., Ikekawa, N., Hirayama, N., Mbarushimana, A. & Ntahomvukiye, D. 1990. 12, 15-dihydroxylabda-8(17), 13-dien-19-oic acid from *Guizotia scabra*. *Phytochemistry 29*, 319-321.

Gebre-Medhin, T. & Mulatu, B. 1992. *Insect pests of noug, linseed and brassica*. In: Oilseeds research and development in Ethiopia. IAR, Addis Abeba- Ethiopia, pp. 174-177.

Geleta, M., Asfaw, Z., Bekele, E. & Teshome, A. 2002. Edible oil crops and their integration with the major cereals in North Shewa and South Wello, central highlands of Ethiopia: an ethnobotanical perspective. *Hereditas 137*, 29-40.

Genet, T. & Belete K. 2000. Phenotypic diversity in the Ethiopian noug germplasm. *African Crop Science Journal 8*, 137-143.

Gepts, P. 2002. A comparison between crop domestication, classical plant breeding, and genetic engineering. *Crop Science 42*, 1780-1790.

Getinet, A. & Sharma, S. M. 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.

Getinet, A. & Teklewold, A. 1995. An agronomic and seed-quality evaluation of niger (*Guizotia abyssinica* Cass.) germplasm grown in Ethiopia. *Plant Breeding 114*, 375-376.

Gupta, P.K., Varshney, R.K., Sharma, P.C. & Ramesh, B. 1999. Molecular markers and their applications in wheat breeding. *Plant Breeding 118*, 369-390.

Hall, B.G. 2001. *Phylogenetic trees made easy*. Sinauer Associates, Sunderland, MA.

Hall, T. 2005. BioEdit v. 7.0.5. Biological sequence alignment editor for Windows. Ibis Therapeutics a division of Isis pharmaceuticals. http://www.mbio.ncsu.edu/BioEdit/bioedit.html (last accessed 03-March- 2007).

Hamrick, J.L. & Godt, M.J.W. 1989. *Allozyme diversity in plant species*. In: Plant population genetics, breeding and genetic resources. (Eds. A.H.D. Brown, M.T. Clegg, A. L. Kahler & B.S. Weir). Sinauer, Sunderland MA, pp. 43-63.

Hamrick, J.L. & Godt, M.J.W. 1996. Effects of life history traits on genetic diversity in plant species. *Philosophical Transactions of the Royal Society-Biological Sciences 351*, 1291-1298.

Hamrick, J.L. & Godt, M.J.W. 1997. Allozyme diversity in cultivated crops. *Crop Science 37*, 26-30.

Harlan J.R. & de Wet J. 1971. Toward a rational classification of cultivated plants. *Taxon 20*, 509-517.

Harlan, J.R. 1969. Ethiopia: a centre of diversity. *Economic Botany 23*, 309-314.

Hawkes J.G. 1977. The importance of wild germplasm in plant breeding. *Euphytica 26*, 615-621.

Hawkes, J.G. 1983. *The diversity of crop plants*. Harvard University Press, Cambridge.

Heal, G., Walker, B., Levin, S., Arrow, K. Dasgupta, P. Daily, G., Ehrlich, P., Maler, K.-G., Kautsky, N., Lubchenco, J., Schneider, S. & Starrett, D. 2004. Genetic diversity and interdependent crop choices in agriculture. *Resource and Energy Economics 26*, 175-184

Hillis, D.M., Moritz, C. & Mable, B.K. 1996. *Molecular systematics*. 2nd Edition. Sinauer Associates. Sunderland, MA.
Hiremath, S.C. & Murthy, H.N. 1986. The structure, stability, and meiotic behavior of B-chromosomes in *Guizotia scabra* (Vis.) Chiov. ssp. *scabra* (Compositae). *Caryologia 39*, 397-402.
Hiremath, S.C. & Murthy, H.N. 1988. Domestication of niger (*Guizotia abyssinica*). *Euphytica 37*, 225-228.
Hiremath, S.C. & Murthy, H.N. 1992. Cytological studies in *Guizotia* (Asteraceae). *Caryologia 45*, 69-82.
Hiremath, S.C., Murthy, H.N. & Salimath, S.S. 1992. Quantitative nuclear DNA differences associated with genome evolution in *Guizotia* (Compositae). *Genetica 85*, 241-247.
Huelsenbeck, J.P. & Crandall, K.A. 1997. Phylogeny estimation and hypothesis testing using Maximum Likelihood. *Annual Review of Ecology and Systematics 28*, 437–466.
Humbles, J.E. 1972. Observations of genus *Seigesbeckia* L. *Ciencia Y Naturaleza 13*, 2-19.
Jaccard, P. 1908. Nouvelles recherches sur la distribution florale. *Bulletin de la Société Vaudoise des Sciences Naturelles 44*, 223-270.
Jadimath, V.G., Murthy, H.N., Pyati, A.N., Kumar, H.G.A. & Ravishankar, B.V. 1998. Plant regeneration from leaf cultures of *Guizotia abyssinica* (niger) and *Guizotia scabra*. *Phytomorphology 48*, 131-135.
Jana, S. 1999. Some recent issues on the conservation of crop genetic resources in developing countries. *Genome 42*, 562-569.
Johnson, E.L., Saunders, J.A., Mischke, S., Helling, C.S. & Emche, S.D. 2003. Identification of *Erythroxulum* taxa by AFLP DNA analysis. *Phytochemistry 64*, 187–197.
Johnson, L.A., & Soltis, E. 1994. *mat*K DNA sequences and phylogenetic reconstruction in Saxifragaceae s.str. *Systematic Botany 19*, 143-56.
Jones, C.J., Edwards, K.J., Castaglione, S., Winfield, M.O., Sala, F., van de Weil, C., Bredemeeijer, G., Vosman, B., Matthes, M., Daly, A., Brettschneider, R., Bettini, P., Buiatti, M., Maestri, E., Maleevschi, A., Marmiroli, N., Aert, R., Volckaert, G., Rueda, R., Linacero, R., Vazquez, A. & Karp, A. 1997. Reproducibility testing of RAPD, AFLP and SSR markers in plants by a network of European laboratories. *Molecular Breeding 3*, 381-390.
Kadereit, G. & Kadereit, J.W. 2005. Phylogenetic relationships, evolutionary origin, taxonomic status, and genetic structure of the endangered local lower Elbe river (Germany) endemic *Oenanthe conioides* (Nolte ex Rchb.f.) Lange (Apiaceae): ITS and AFLP evidence. *Flora 200*, 15-29.
Kandel, H. & Porter, P. (Eds.) 2002. *Niger (Guizotia abyssinica) (L. f.) Cass. production in northwest Minnesota*. University of Minnesota Extension Service.
Karis, P.O. 1993. Heliantheae sensu lato (Asteraceae) clades and classification. *Plant Systematics and Evolution 188*, 139-195.
Kelly, J.D. & Miklas, P.N. 1998. The role of RAPD markers in breeding for disease resistance in common bean. *Molecular Breeding 4*, 1-11.
Kumar, H.G.A., Murthy, H.N., Jadimath, V.G., Sheelavantmath, S.S., Pyati, A.N. & Ravishankar, B.V. 2000. Direct somatic embryogenesis and plantlet regeneration from explants of niger, *Guizotia abyssinica* (L.f.) Cass. *Indian Journal of Experimental Biology 38*, 1073-1075.
Kumlehn, J.J. & Nitzsche, W. 1996. Plant regeneration from ryegrass ovules cultivated on endosperm-derived feeder cells. *Plant Cell, Tissue and Organ Culture 44*, 235-241.
Kuzoff, R.K., Sweere, J.A., Soltis, D.E., Soltis, P.S. & Zimmer, E.A. 1998. The phylogenetic potential of entire 26S rDNA sequences in plants. *Molecular Biology and Evolution 15*, 251-263.
Landry, P.-A. & Lapointe, F.-J. 1996. RAPD problems in phylogenetics. *Zoologica Scripta 25*, 283-290.
Liang, H. & Hilu, K.W. 1996. Application of the *mat*K gene sequences to grass systematics. *Canadian Journal of Botany 74*, 125-134.
Loveless, M.D. & Hamrick, J.L. 1984. Ecological determinants of genetic structure in plant populations. *Annual Review of Ecology and Systematics 15*, 65-95.

Lynch, M. & Milligan, B.G. 1994. Analysis of population genetic structure with RAPD markers. *Molecular Ecology 3*, 91-99.

Ma, Y.S., Wang, W.H., Wang, L.X., Ma, F.M., Wang, P.W., Chang, R.Z. & Qiu, L.J. 2006. Genetic diversity of soybean and the establishment of a core collection focused on resistance to soybean cyst nematode. *Journal of Interactive Plant Biology 48*, 722-731.

Mantel, N.A. 1967. The detection of disease clustering and a generalized regression approach. *Nucleic Acids Research 21*, 1111-1115.

Marini, F., Magria, A.L., Marinib, D. & Balestrieric, F. 2003. Characterization of the lipid fraction of niger seeds (*Guizotia abyssinica* cass.) from different regions of Ethiopia and India and chemometric authentication of their geographical origin. *European Journal of Lipid Science and Technology 105*, 697–704

Meilleur, B.A. & Hodgkin, T. 2004. *In situ* conservation of crop wild relatives: status and trends. *Biodiversity and Conservation 13*, 663–684.

Meyerowitz, E.M. 1999. The first completely sequenced plant chromosomes, from the mustard *Arabidopsis thaliana*, reveal a dynamic genome that is constantly being rearranged. *Nature 402*,731-732.

Milbourne, D., Meyer, R., Bradshaw, J.E., Baird, E., Bonar, N., Provan, J., Powell, W. & Waugh, R. 1997. Comparison of PCR based marker systems for the analysis of genetic relationships in cultivated potato. *Molecular Breeding 3*, 127–136.

Murthy, H.N., Hiremath, S.C. & Salimath, S.S 1993. Origin, evolution and genome differentiation in *Guizotia abyssinica* and its wild species. *Theoretical and Applied Genetics 87*, 587-592.

Murthy, H.N., Jeong, J.H., Choi, Y.E. & Paek, K.Y. 2003. *Agrobacterium*-mediated transformation of niger [*Guizotia abyssinica* (L. f.) Cass.] using seedling explants. *Plant Cell Report 21*, 1183–1187.

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 of the United States of America 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. & Dagne, K. 1999. Self-incompatibility system in Ethiopian populations of *Guizotia abyssinica* (L.f.) Cass. (niger). *Sinet: Ethiopian Journal of Science 22*, 67–88.

Nikam, T.D. & Shitole, M.G. 1997. In vitro plant regeneration from callus of niger (*Guizotia abyssinica* Cass.) cv. Sahyadri. *Plant Cell Reports 17*, 155–158.

Nybom, H. & Bartish I.V. 2000. Effects of life history traits and sampling strategies on genetic diversity estimates obtained from RAPD markers in plants. *Perspectives in Plant Ecology, Evolution and Systematics 3*, 93-114.

Nybom, H. 2004. Comparison of different nuclear DNA markers for estimating intraspecific genetic diversity in plants. *Molecular Ecology 13*, 1143-1155.

Ohlrogge, J.B. & Jaworski, J.G. 1997. Regulation of fatty acid synthesis. *Annual Review of Plant Physiology and Plant Molecular Biology 48*, 109–136.

Ohta, T. 1994. Further examples of evolution by gene duplication revealed through DNA sequence comparisons. *Genetics 138*, 1331-1337.

Page, R.D.M., 1996. TreeView. An application to display phylogenetic trees on personal computers. *Computer Applications in the Biosciences 12*, 357-358.

Panero, J.L., Jansen, R.K., Jennifer, A. & Clevinger, J.A. 1999. Phylogenetic relationships of subtribe Ecliptinae (Asteraceae: Heliantheae) based on chloroplast DNA restriction site data. *American Journal of Botany 86*, 413–427.

Paran, I. & Michelmore, R.W. 1993. Development of reliable PCR-based markers linked to downy mildew resistance genes in lettuce. *Theoretical and Applied Genetics 85*, 985–93.

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 Biologica 45*, 97-99.

Powell, W., Morgante, M., Andre, C., Hanafey, M., Vogel, J., Tingey, S. & Rafalski, A. 1996. The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. *Molecular Breeding 2*, 225-238.

Prasad, V. 1990. Pollen tube growth and site of incompatibility reactions in niger (*Guizotia abyssinica* Cass.). *Current Science 59*, 466-468.

Raina, S.N., Rani, V., Kojima, T., Ogihara, Y., Singh, K.P. & Deyarumath, R.M. 2001. RAPD and ISSR fingerprints as useful genetic markers for analysis of genetic diversity, varietal identification and phylogenetic relationships in peanut (*Arachis hypogaea*) cultivars and wild species. *Genome 44*, 763–772.

Ramachandran, T.K. & Menon, P. 1979. Pollination mechanisms and inbreeding depression in niger (*Guizotia abyssinica* Cass.). *Madras Agricultural Journal 66*, 449-454.

Ramadan, M.F. & Mörsel, J-T (2003). Determination of the lipid classes and fatty acid profile of niger (*Guizotia abyssinica* Cass.) seed oil. *Phytochemical Analysis 14*, 366–370.

Ramadan, M.F. & Mörsel, J-T. 2004. Oxidative stability of black cumin (*Nigella sativa* L.), coriander (*Coriandrum sativum* L.) and niger (*Guizotia abyssinica* Cass.) crude seed oils upon stripping. *European Journal of Lipid Science and Technology 106*, 35–43.

Rieseberg, L.H. 1996. Homology among RAPD fragments in interspecific comparisons. *Molecular Ecology 5*, 99-105.

Riley, K.W. & Belayneh, H. 1989. *Niger*. In: Oil crops of the world. (Eds. G. Röbbelen, R.K. Downey & A. Shri). McGraw Hill Publishing Company. New York, pp. 394-403.

Robinson, H. 1981. A revision of the tribal and subtribal limits of the Heliantheae (Asteraceae). *Smithsonian Contributions to Botany 51*, 1-102.

Rohlf, F.J. 2000. *NTSYSpc, Numerical taxonomy and multivariate analysis system. Version 2.1.* Exeter Software, Setauket, New York, USA.

Saitou, N. & Nei, M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution 4*, 406-25.

Sang, T., Crawford, D. J. & Stuessy, T.F. 1997. Chloroplast DNA phylogeny, reticulate evolution, and biogeography of paeonia (Paeoniaceae). *American Journal of Botany 84*, 1120-1136.

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.

Schulz, D.L. 1990. Zur kenntnis der Gattung *Sigesbechia* L. in Africa. *Gleditschia, 46*, 613-628.

Seegeler, C.J.P. 1983. *Oil plants in Ethiopia their taxonomy and agricultural significance.* Center of Agricultural Publishing and Documentation. Wageningen. 122-146 pp.

Shaw, J. & Small, R.L. 2005. Chloroplast DNA phylogeny and phylogeography of the North American plums (*Prunus* subgenus *Prunus* section *Prunocerasus*, Rosaceae). *American Journal of Botany* 92, 2011–2030.

Shim, S.I. & Jorgensen, R.B. 2000. Genetic structure in cultivated and wild carrots (*Daucus carota* L.) revealed by AFLP analysis. *Theoretical and Applied Genetics 101*, 227–233.

Singh, B. & Patra, G.J. 1989. Character association and path coefficients of quantitative traits in niger (*Guizotia abyssinica*). *Indian Journal of Agricultural Sciences 59*, 442-445.

Sivolap, Y.M., Solodenko, A.E. & Burlov, V.V. 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. *Theoretical and Applied Genetics 91*, 1086–1091.

Slatkin, M. 1987. Gene flow and the geographic structure of natural populations. *Science 236*, 787-792.

Small, R.L., Cronn, R.C. & Wendel, J.F. 2004. Use of nuclear genes for phylogeny construction in plants. *Australian Systematic Botany 17*, 145-170.

Soleiman, V.D., Baum, B.R. & Johnson, D.A. 2007. Analysis of genetic diversity in barley cultivars reveals incongruence between S-SAP, SNP and pedigree data. *Genetic Resources and Crop Evolution 54*, 83-97.

Soltis, P.S., Soltis, D.E. & Chase, M.W. 1999. Angiosperm phylogeny inferred from multiple genes as a tool for comparative biology. *Nature 402*, 402-404.
Stuessy, T.F. 1977. *Heliantheae-Systematic review*. In: The biology and chemistry of the Compositae, Vol 1. (Eds. V.H. Heywood, J.B. Harborne & B.L. Turner). Academic Press. London, pp. 1106-1118.
Sujatha, M. 1997. In vitro adventitious shoot regeneration for effective maintenance of male sterile niger (*Guizotia abyssinica* (L.f.) Cass.). *Euphytica 93*, 89-95.
Swofford, D.L. 2000. *PAUP*: Phylogenetic analysis using parsimony, version 4.0, beta.* Sinauer associates inc. Sunderland, Massachusetts.
Taberlet, P., Gielly, L., Pautou, G. & Bouvet, J. 1991. Universal primers for amplification of three non-coding regions of chloroplast DNA. *Plant Molecular Biology 17*, 1105-1109.
Talhinhas, P., Leitao, J. & Neves-Martins, J. 2006. Collection of *Lupinus angustifolius* L. germplasm and characterization of morphological and molecular diversity. *Genetic Resources and Crop Evolution 53*, 563-578.
Thompson, J.A. & Nelson, R.L. 1998. Core set of primers to evaluate genetic diversity in soybean. *Crop Science. 38*, 1356–1362.
Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F. & Higgins, D.G. 1997. The CLUSTAL X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research 25*, 4876-4882.
Vavilov, N.I. 1951. The origin, variation, immunity and breeding of cultivated plants. *Chronica Botanica 13*, 1–364.
Vicient, C.M., Suoniemi, A., Anamthawat-Jónsson, K., Tanskanen, J. Beharav, A., Nevo, E. & Schulman, A.H. 1999. Retrotransposon BARE-1 and its role in genome evolution in the genus *Hordeum*. *Plant Cell 11*, 1769-1784.
Vles, R.O. & Gottelbos, J.J. 1989. *Nutritional characteristics and food uses of vegetable oils*. In: Oil crops of the world. (Eds. G. Röbbelen, R.K. Downy & A. Ashri). McGraw Hill, New York, USA, pp. 63-86.
Vos, P., Hogers, R., Bleeker, M., Reijans, M., van de Lee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J., Kuiper, M. & Zabeau, M. 1995. AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research 23*, 4407-4414.
Wanasundara, U.N. & Shahidi, F. 1994. Canola extract as an alternative natural antioxidant for canola oil. *Journal of the American Oil Chemists Society 71*, 817-822.
Weising, K., Nybom, H., Wolff, K. & Kahl, G. 2005. *DNA Fingerprinting in plants. Principles, Methods, and Applications*. 2nd edition. CRC Press, Tylor and Francis group. Boca Raton.
Weiss, E.A. 1983. *Oil seed crops*. Longman Inc. New York.
Welsh, J. & McClelland, M. 1990. Fingerprinting genomes using PCR with arbitrary primers. *Nucleic Acids Research 18*, 7213-7218.
Whitten, W.M., Williams, N.H. & Chase, M.W. 2000. Subtribal and generic relationships of Maxillarieae (Orchidaceae) with emphasis on Stanhopeinae: combined molecular evidence. *American Journal of Botany 87*, 1842-1856.
Williams, J.G.K., Kubelik, A.R., Livak, K.J., Rafalski, J.A. & Tingey, S.V. 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Research 18*, 6531-6535.
Winkworth, R.C., Grau, J., Robertson, A.W. & Lockhart, P.J. 2002. The origins and evolution of the genus *Myosotis* L. (Boraginaceae). *Molecular Phylogenetics and Evolution 24*, 180–193.
Wood, D. & Lennie, J.M. 1997. The conservation of agrobiodiversity on-farm: questioning the emerging paradigm. *Biodiversity and Conservation 6*, 109-129.
Yeh, F.C. & Boyle, T.J.B. 1997. Population genetic analysis of co-dominant and dominant markers and quantitative traits. *Belgian Journal of Botany 129*, 157.
Yin, T., Zhang, X., Huang, M., Wang, M., Zhuge, Q., Tu, S., Zhu, L. & Wu, R. 2002. Molecular linkage maps of the *Populus* genome. *Genome 45*, 541–555.
Zeven, A.C. & de Wet, J.M.J. 1982. *Dictionary of cultivated plants and their regions of diversity*. International Book Distributors, Dehra Dun, India.

Acknowledgements

First of all I wish to thank almighty God for keeping me alive and giving me the strength to overcome all challenges I have encountered during my studies. When I look back, I realize that this thesis is not only my work but is the effort and contribution of many people. I would like to convey my sincere gratitude to all of you who helped me accomplish this work.

My very sincere gratitude goes to my supervisors for accepting me as your student. You have been so kind and helpful to me and your coordinated arrangement made my studies smooth and straightforward. This thesis would not have been completed without your full and coordinated support. ***Professor Tomas Bryngelsson***, the first time I met you in Addis Ababa, I felt that you are very kind and helpful. Yes, I was right! You were always there to listen and to give advice. I provide my heartfelt regards to you as you guided me throughout my PhD training with your intelligent suggestions and ideas. Thank you very much for critical reading, constructive discussions and invaluable comments and suggestions on the separate papers and this thesis. I am very grateful to you for your concern about my future. ***Professor Endashaw Bekele***, you have been my model since I took your first course while I was undergraduate student. I am very grateful to you for your confidence in me, kindness and continuous encouragement. I always remember the warm and moral-boosting sentences in your e-mail as a prompt reply to my needs and questions. The discussions we have had throughout my study period were very helpful and made me love this study. Your critical reading of the manuscripts and this thesis at your highest priority and invaluable comments and suggestions were vitally important; thank you very much! ***Dr. Kifle Dagne***, I always remember the day you asked me if I am interested in this study. Thank you very much for that golden question. The initial discussion we had regarding *Guizotia* species was a key to the success of this study. I am greatly indebted to you for arranging my fieldwork while you were the head of the Department of Biology, AAU. Without your absolutely kind help, complete collection of those samples was unimaginable within that period of time. Your comments on each paper and this thesis were very helpful to improve their qualities. Your fatherly approach, concern and continuous encouragement to which I am highly thankful, made me happy, strong and successful.

My special thanks go to The Swedish International Development Agency (Sida/SAREC) for financing this study and the International Science Program (ISP) of Uppsala University for administrating and facilitating this Sida/SAREC supported program. I wish to express my sincere gratitude to ***Dr. Malin Åkerblom***, the Director of the International Program in Chemical Sciences (IPICS, which is part of ISP), for solving my financial problems during my study period in Ethiopia and for extending my stay in Sweden in order to complete this study. You were always there to help me. Thank you very much! I would also like to thank *Hossein Aminaey* and all ISP staff for your immediate response to everything I have asked.

I would like to express my special gratitude to all members of the Department of Crop Science (SLU) for making my stay in Sweden happy, comfortable and fruitful in many ways. *Britt Greén* and *Therese Bengtsson*: many thanks for your skilful assistance in molecular laboratory work, for the immediate replies to my requests and for many more reasons. I felt comfortable in the lab as I knew that you were there to help me. *Helén Lindgren* and *Annelie Ahlman*, thanks a lot for your help on several occasions in the absence of *Britt* and *Therese*. *Helen*, I am very grateful to you for your TV. I enjoyed the 2006 soccer world cup and other programs during my leisure time. *Monica Lotfinia*, you are very supportive to me and your care always reminds me of my mom. Whenever I talk to you, I feel happy and refreshed. Monica, you are great! Tack så mycket. *Susanne Hjerdin*, thank you for your hospitality and warm salutation whenever we meet. My sincere thanks go to *Lars Hagtorn* and *Goran Olsson* for fixing problems associated with my computer. *Lars*, I highly appreciate your effort to recover data from the crashed computer which happened twice during my studies. I am highly thankful to *Kristina Santén* (Tintin), *Dr. Helena Persson Hovmalm*, *Phuong Nguyen*, *Cuong Nhu*, *Alicia Sanchez*, *Åsa Ekman*, *Ann-Sofie Fält*, *Marissa Prieto-Linde*, *Gunilla Åhman*, *Svetlana Leonova*, *Thuy Doan*, *Dr. Beatrice Angio-Were*, *Anders Smolka*, *Dr. Shu-Chin Hysing*, *Toan Pham*, *Dennis Eriksson* and *Carlos Loaisiga* for your congenial company.

I express my heartfelt gratitude to *Professor Waheeb Heneen* for the hospitality, inspiration and encouragement I received from you. I am deeply indebted to *Professor Margareta Welander* for evaluating my review on plant transformation and for your help and encouragement during NOVA courses. My sincere thanks go to *Professor Arnulf Merker* for your interesting lectures during the last course of my PhD studies and for your appreciation and encouragement. *Dr. Anders Carlsson*, thank you very much for introducing me how to manipulate DNA sequence data and for valuable discussions on several occasions. I am also highly thankful to *Dr. Peter Olsson* for the valuable discussion we had on DNA sequencing data. It is my great pleasure to sincerely thank *Dr. Eva Johansson*, *Dr. Erland Liljeroth*, *Dr. Björn Salomon*, *Dr. Salla Marttila*, *Dr. Zhu Li-Hua*, *Professor Hilde Nybom*, *Professor Mats Gustafsson*, *Professor Sten Stymne*, *Dr. Ylva Hillbur* and many others for your encouragement and contribution to my success in different ways.

My sincere thanks go to *Pia Ohlsson* for sharing your experience on AFLP analysis with me, and *Dr. Agnese Kolodinska-Brantestam* and *Alfia Khairullina* for your help in solving some problems related to AFLP analysis. *Dr. Linus Masumbuko*, thanks a lot for discussions on data analysis and the great moments we shared. I convey my special regards to all of my housemates at Hovbeslagsskolan for making my life comfortable. You are affable sort of people, thanks.

I would like to thank the Department of Biology, Addis Ababa University for facilitating my involvement in this study and arranging a car for the fieldwork. I express my thanks to *Dr. Dawit Abate* for facilitating my fieldwork while you were working as a head of the Department. *Ejigu Belayneh*, you are an excellent driver. Without your absolute willingness, it was not possible to travel throughout

Ethiopia to collect the germplasm within that period of time. I am grateful to your unreserved assistance in all aspects during the fieldwork. *Gashaw Kedir*, thank you for making my fieldwork happy and successful both as a driver and as a friend. *Mulualem* and *Banchu*, I am thankful to you for your unreserved help under many circumstances. *Dr. Kassahun* and *Tamiru*, you are nice friends. We had a wonderful time and fruitful discussions while we were working in the genetics research laboratory, AAU. Thank you.

I am highly thankful to the Institute of Biodiversity Conservation (IBC), Addis Ababa, Ethiopia, for giving me the permission to take *Guizotia* germplasm to Sweden for this research work. I convey my special gratitude to all friends at IBC for encouraging me to start this study. *Adugna Abdi*, you have been very helpful in different ways. Thank you!

I would like to thank my Ethiopian compatriots and friends *Dr. Genet Birmeta*, *Dr. Kebebew Assefa*, *Dr. Tileye Feyissa*, *Dr. Esayas Aga*, *Yohannes Petros*, *Faris Hailu*, *Yitbarek W/Hawariat* and *Diriba Mulleta* for highly valuable discussions on academic matters and for your continuous encouragement while we were in Sweden. *Genet*, you have been like a sister to me; thank you so much. *Jone* and *Faris*, I always remember the great time we shared while we were living in one flat in Kornvägen 57 Arlöv. *Kebebew*, I am very grateful to you for taking care of my car at home during the last year of my studies in Sweden. You are really kind and helpful! *Tileye* and *Esayas*, you are specially remembered for all lunch we shared together in your house. *Yit*, thanks a lot for a benefit I got from you as a result of our close friendship. Tirunesh and Getachew (Lund), I would like to say a big thank you for your hospitality.

Lastly, I would like to express my deepest thanks and regards to my dear parents, *Geleta* and *Leakie*, for your continuous love and daily prayers for my wellbeing and better life. *Elfe*, *Workiye*, *Teshe*, *Defe*, *Eshe*, *Shewa*, *Bire*, *Chuche*, *Sintu*, *Nunish*, *Gishu*, *Burte* and other close relatives, thank you very much for your continuous encouragement and for keeping in touch while I was away. *Wube*, I feel wordless to be thankful for your love, patience and encouragement. I am longing for the day we will start to live together.

I

Genet Resour Crop Evol
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
© Springer Science+Business Media B.V. 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 $H'_{pop} = \sum H'_j/n$, 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 $H'_{loci} = \sum H'_j/L$, 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'_{pop}/H'_{sp} and $(1 - H'_{pop}/H'_{sp})$, 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 $H_j(i) = 2q_j(i)[1 - q_j(i)] + 2\mathrm{Var}[q_j(i)]$ 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 $q = x^{1/2}\left[1 - \frac{\mathrm{Var}(x)}{8x^2}\right]^{-1}$. $\mathrm{Var}(q)$ and $\mathrm{Var}(x)$ 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 $G_{ST} = (H_T - H_S)/H_T$, 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 $\overline{H'}_{sp}$ 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	**Wl**	**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[a]	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

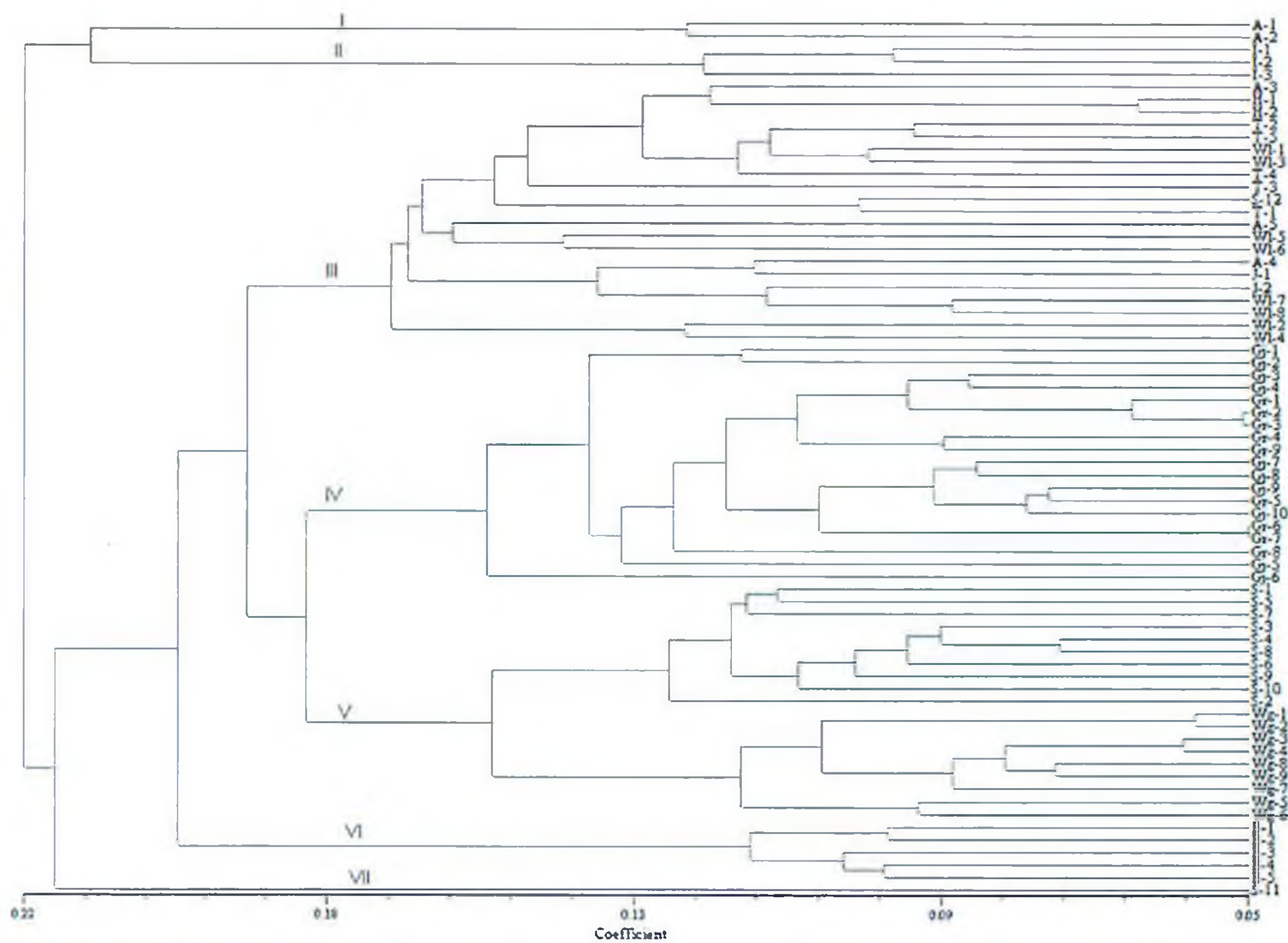


Fig. 2 UPGMA Phenogram of the 70 populations of *G. abyssinica* based on Nei's standard genetic distance estimated from RAPD band profiles

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.

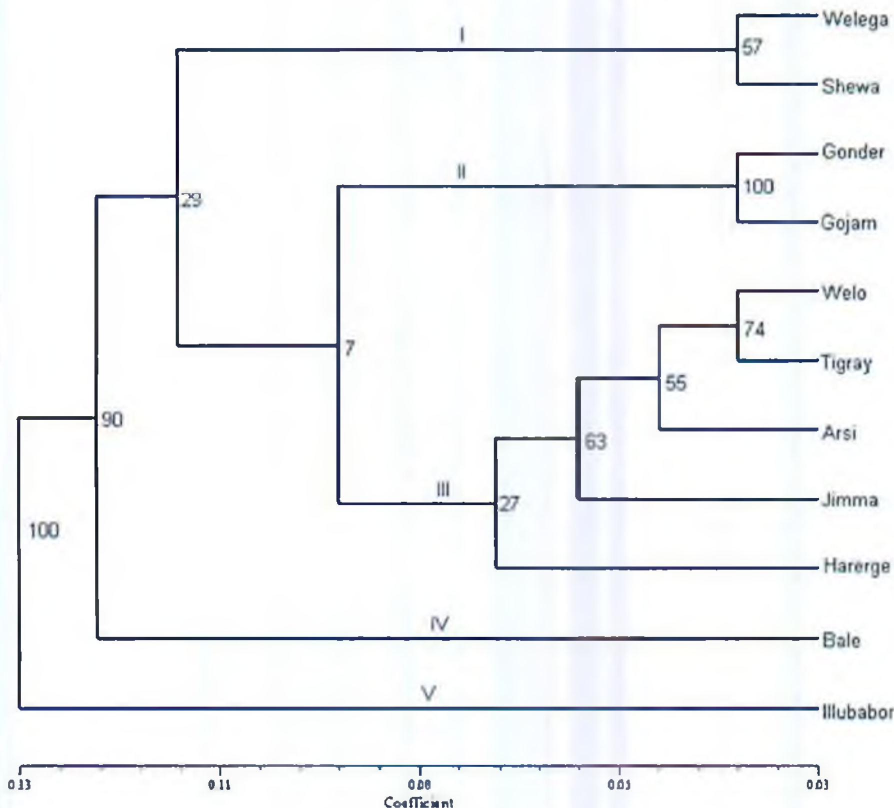


Fig. 3 UPGMA Phenogram of *G. abyssinica* of the 11 regions based on Nei's standard genetic distance. Numbers near branches are bootstrap values

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

Baagoe 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, Morsel 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 co-dominant 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

II

Assessment of genetic diversity of *Guizotia abyssinica* (L.F.) Cass. (Asteraceae) from Ethiopia using Amplified Fragment Length Polymorphism (AFLP)

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

[1]Department of Plant Protection Biology, Swedish University of Agricultural Sciences, Box 44,
SE-230 53 Alnarp, Sweden
[2]Department of Biology, Addis Ababa University, P. O. Box 1176, Addis Ababa, Ethiopia

Abstract

Seventeen populations of niger (*Guizotia abyssinica*) representing all regions in Ethiopia where this crop is grown was 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 (AMOVA; $P < 0.001$). There were no significant differences between populations in relation to the extent and altitude of cultivation. 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.

Key words: 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 in 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 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 - 2200 masl, 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 characterisation (Nayakar, 1976; Alemaw and Teklewold, 1995; Pradhan et al., 1995; Genet and Belete, 2000), and the only published marker based study used RAPD (Geleta et al., 2006).

The present study was undertaken to investigate the extent of genetic variation within and among populations of Ethiopian niger using the AFLP marker technique. It

also set out to identify any hotspots of diversity that may be important in optimising 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 (1640 - 2550 masl) representing all regions where niger is currently cultivated. Each region was represented by at least one population (Table 1; Figure 1). Seeds were grown in a greenhouse and fresh 15-30 day old leaves were used for genomic DNA extraction, using the modified CTAB procedure described in Assefa et al. (2003).

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

AFLP

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 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 ATP, 1 mM 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 (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 PCR buffer (as above), 25 ng *Eco*RI primer + 3 SNs, 30 ng *Mse*I primer + 3 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 microlitre 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: (i) 10% acetic acid was used as fixer solution and stopping solution and (ii) the concentration of sodium thiosulfate in 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. (2006), 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 PCs

A total of 539 fragments were detected among the 170 individual plants, of which 483 (over 89%) were polymorphic (Table 1). The number of polymorphic loci averaged 69 per PC. When each population was considered separately, the percentage of polymorphic loci ranged from 46% (T-1) to 60% (WI-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.

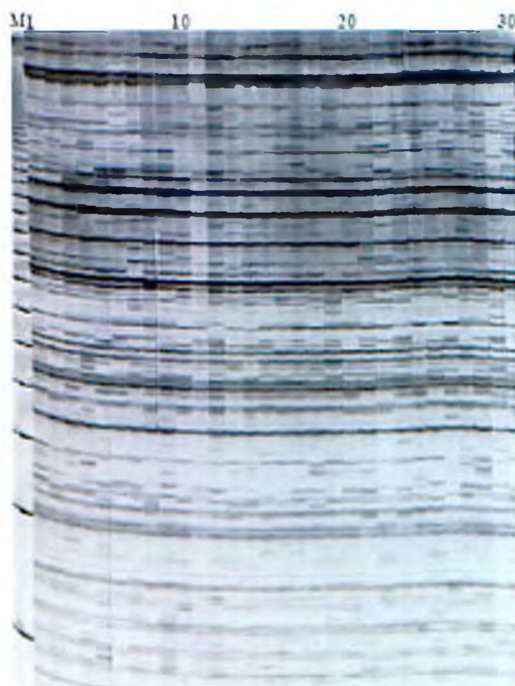


Figure 2. AFLP profiles generated by primer combination *E-ACA/M-CTA* for three populations. M: molecular size marker. Individuals of WI-1, WI-2 and I-1 populations represented in lanes 1-10, 11-20 and 21-30, respectively.

Table 1. *Population codes, geographic description, number of polymorphic loci (NPL), percent 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 parameters for the 17 populations (corresponding value of *PPL* is P_P); [b]corresponding values of each parameters when all individuals were considered together (corresponding value of *PPL* is P_S).

Table 2. *Summary of number of bands scored (NBS), number of polymorphic loci (NPL), percent 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
[cf]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
						$\overline{H'}_{POP}$	$\overline{H'}_{SP}$		$\overline{G'}_{ST}$	$\overline{H}_S$	$\overline{H}_T$		$\overline{G}_{ST}$	$\overline{F}_{ST}$
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												

[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. *Mean of values in column. [c]E = *Eco*RI primer (5'-GACTGCGTACCAATTC-3'); [d]M = *Mse*I primer (5'-GATGAGTCCTGAGTAA-3').

AFLP PCs used in this study were significantly different in 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 ($G'_{ST} = 0.415$; $G_{ST} = 0.354$; $F_{ST} = 0.357$).

Total and within population genetic variation

The overall genetic diversity estimated by Shannon diversity index as $\overline{H'_{SP}}$ and gene diversity estimate (Nei, 1978) as $\overline{H_T}$ was 0.628 and 0.320, respectively. Similarly, the overall within population variation estimated by Shannon diversity index ($\overline{H'_{POP}}$) and Nei's gene diversity estimate ($\overline{H_S}$) 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-1, 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 similar level of genetic variations 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 ($\overline{G'_{ST}}$), 0.269 ($\overline{G_{ST}}$) and 0.234 ($\overline{F_{ST}}$) (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 higher altitude group (> 2000 masl) and lower altitude group (< 2000 masl) 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 0.05-0.35 (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 analyzed 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 four-fold variation in genetic distance between pairs of populations that ranged from 0.040 (Gr-1 vs Gr-2; Gr-2 vs Wg-1) to 0.175 (H-1 vs T-1) was obtained, with the overall mean genetic distance of 0.118 (see appendix 1). 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 Figures 3A and 3B). Three clusters (Cluster I, II and III) were revealed in the UPGMA cluster analysis (Figure 3A), which was also clearly depicted in the PCoA (Figure 3B). Clusters I, II, and III consist of nine, three and five populations, respectively. The bootstrap value for branching of cluster 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; Figure 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 (Figure 3B). The second axis (PC-II) explained 20% of the total variation and was better than PC-I in discriminating populations within clusters.

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 populations from (i) MiNPR (ii) 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 of gene flow*

groups	Source of variation*	Degrees of freedom	Variance components	%age of variation	Fixation indices	Significant 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 (i) A-1, A-2, B-1, H-1, I-1, J-1, T-1, W-1, W-2 and (ii) Gj-1, Gj-2, Gr-1, Gr-2, Sh-1, Sh-2, Wg-1, Wg-2	AG	1	Va = 0.56	0.82	F_{ST} = 0.237	Va and F_{CT} = 0.195	
	APWG	15	Vb = 15.84	22.91	F_{SC} = 0.231	Vb and F_{SC} = 0.000	
					F_{CT} = 0.008	Vc and F_{ST} = 0.000	
	WP	153	Vc = 52.73	76.27			
	Total	169	69.13				
C. Grouping the populations into (i) A-1, A-2, B-1, Gj-1, Gr-1, Sh-1, Wg-1, Wl-2 and (ii) Gj-2, Gr-2, H-1, I-1, J-1, Sh-2, T-1, Wg-2, Wl-1	AG	1	Va = -0.96	-1.40	F_{ST} = 0.229	Va and F_{CT} = 0.923	
	APWG	15	Vb = 16.65	24.33	F_{SC} = 0.240	Vb and F_{SC} = 0.000	
	WP	153	Vc = 52.73	77.07	F_{CT} = -0.014	Vc and F_{ST} = 0.000	
	Total	169	68.42				
D. grouping the populations into (i) A-1, A-2, B-1, H-1; (ii) I-1, J-1, Sh-1, Wg-1, Wg-2; (iii) Sh-2, W-1, W-2; (iv) Gj-1, Gj-2, and (V) Gr-1, Gr-2, T	AG	4	Va = 5.22	7.48	F_{ST} = 0.244	Va and F_{CT} = 0.000	
	APWG	12	Vb = 11.80	16.92	F_{SC} = 0.183	Vb and F_{SC} = 0.000	
	WP	153	Vc = 52.73	75.59	F_{CT} = 0.075	Vc and F_{ST} = 0.000	
	Total	169	69.75				

*AP = Among populations; WP = Within populations; AG = Among groups; APWG = Among populations within groups.

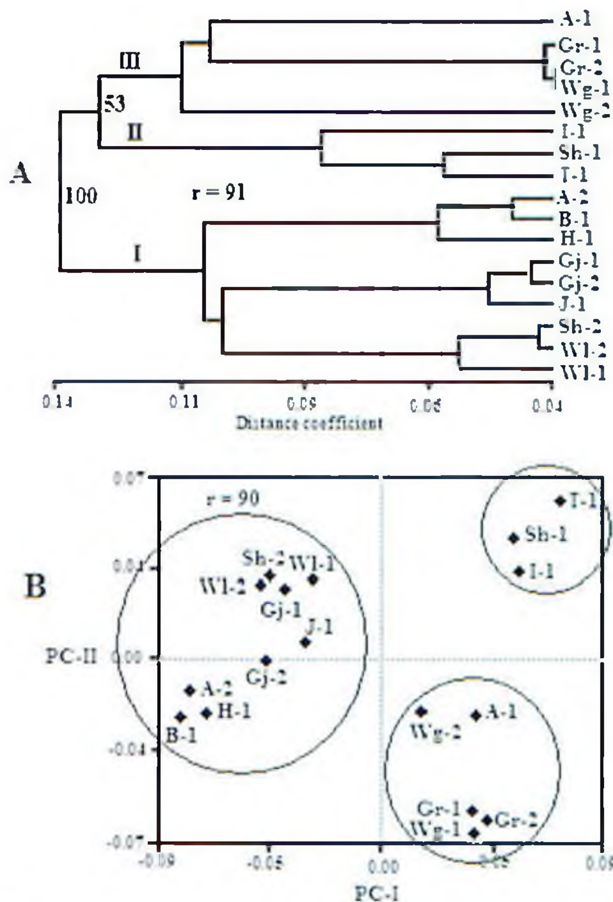


Figure 3. (A) UPGMA phenogram of the 17 populations of niger based on Nei's (1972) standard genetic distance, estimated from AFLP profiles; (B) two-dimensional plot generated by PCoA.

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. The AFLP PCs differed in their capability to detect polymorphic loci and unique alleles. *E-AGG/M-CTA* revealed the highest number of polymorphic loci that makes it preferable in order to increase the chance of getting polymorphic loci linked to economically important traits, provided that some of the loci are linked to such traits in niger. *E-ACA/M-CTA* seems to be preferable to search unique mutations in this species as it revealed the highest number of unique alleles specific to a given population or a group of populations. The highest estimate of population differentiation was obtained by *E-AAG/M-CTC*. Thus, this PC should

be preferred to identify populations that have lost amplifiable AFLP loci in an effort to find population specific markers linked to desirable traits.

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., 2006), 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. (2006) for AMOVA and Shannon diversity index was 65% and 57%, respectively, which is lower than that revealed in this 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 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., 2006) 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 this 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., 2006). Comparing the mean H'_{loci} (0.430) and H_j (0.200) obtained in this 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 based on RAPD based study (Geleta et al., 2006) was not supported by this 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) analyzed 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 are 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 large number of populations from all its growing regions *ex situ* in gene banks as a complement to on-farm conservation is the best opportunity to conserve 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, 1999).

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 (Rolf, 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 (Figures 3A and 3B) 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., 2006).

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

References

Abebe M, Yermanos DM and Bingham FT (1978) The ecophysiology of noog (*Guizotia abyssinica* Cass.). *Afr. J. Agric. Sci.* 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. *Mol. Ecol.* 6: 761-772.

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

Bossdorf O, Auge H, Lafuma L, Rogers WE, Siemann 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 Sci.* 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. *Proc. Natl. Acad. Sci. 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 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 and 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 and Alemaw G (1994) Variation in lipid composition of niger seed (*Guizotia abyssinica* Cass.) samples collected from different regions in Ethiopia. *J. Am. Oil Chem. Soc.* 71: 839-843.

Fichtl R and Adi A (1994) Honeybee flora of Ethiopia. Margraf Verlag, Germany.

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

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 (2006) Genetic diversity of *Guizotia abyssinica* (L. f.) Cass. (Asteraceae) from Ethiopia as revealed by Random Amplified Polymorphic DNA (RAPD). *Genet. Resour. Crop. Ev.* DOI 10.1007/s10722-006-0018-0.

Genet T and Belete K (2000) Phenotypic diversity in the Ethiopian noug germplasm. *Afr. Crop Sci. J.* 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. Institute of Plant Genetics and Crop Plant Research, Gatersleben/International Plant Genetic Resources Institute, Rome.

Hamrick JL and 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 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. *Syst. Zool.* 39: 242–252.

Loveless MD and Hamrick JL (1984) Ecological determinants of genetic structure in plant populations. *Ann. Rev. Ecol. Syst.* 15: 65-95.

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

Mantel NA (1967). The detection of disease clustering and a generalized regression approach. *Nucleic acids Res.* 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. *Theor. Appl. Genet.* 83: 1003–1012.

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

Murthy HN, Hiremath SC and 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. *P. Natl. Acad. Sci. 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). 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.

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 J. Genet. Pl. Br.* 55: 457-459.

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

Riley KW and Belayneh H (1989) Niger. In: Röbbelen G, Downey RK and 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.

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

Shim SI and Jorgensen RB (2000) Genetic structure in cultivated and wild carrots (*Daucus carota* L.) revealed by AFLP analysis. *Theor. Appl. Genet.* 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 Res.* 23: 4407-4414.

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

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

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

Appendix (Online Table)

Appendix 1. *Nei's standard genetic distance between 17 populations of G. abyssinica calculated based on AFLP data*

	A-1	A-2	B-1	Gj-1	Gj-2	Gr-1	Gr-2	H-1	I-1	J-2	Sh-1	Sh-2	T-1	Wg-1	Wg-2	Wl-1
A-2	0.144															
B-1	0.148	0.048														
Gj-1	0.141	0.105	0.106													
Gj-2	0.131	0.100	0.102	0.044												
Gr-1	0.101	0.133	0.130	0.127	0.121											
Gr-2	0.106	0.142	0.141	0.139	0.124	0.040										
H-1	0.148	0.066	0.059	0.114	0.110	0.128	0.139									
I-1	0.117	0.147	0.159	0.129	0.150	0.114	0.117	0.146								
J-1	0.119	0.096	0.099	0.056	0.050	0.109	0.114	0.104	0.126							
Sh-1	0.115	0.146	0.158	0.117	0.131	0.120	0.120	0.146	0.071	0.107						
Sh-2	0.136	0.107	0.109	0.107	0.107	0.135	0.138	0.110	0.134	0.094	0.131					
T-1	0.139	0.173	0.173	0.140	0.151	0.138	0.143	0.175	0.098	0.137	0.061	0.148				
Wg-1	0.113	0.139	0.137	0.139	0.126	0.043	0.040	0.133	0.132	0.111	0.123	0.137	0.147			
Wg-2	0.115	0.132	0.132	0.133	0.128	0.112	0.109	0.132	0.127	0.118	0.118	0.126	0.138	0.110		
Wl-1	0.126	0.109	0.110	0.107	0.100	0.118	0.125	0.108	0.120	0.086	0.108	0.062	0.131	0.127	0.116	
Wl-2	0.134	0.109	0.118	0.114	0.114	0.134	0.146	0.116	0.138	0.102	0.134	0.043	0.159	0.143	0.130	0.056

Note: The overall mean Nei's standard genetic distance was 0.118. Correlation between Nei's standard genetic distance and geographic distance was 0.258 (Mantel test $P < 0.01$).

Hurri Hills in northern Kenya (Hiremath and Murthy 1992; Dagne and Jonsson 1997). Chelelu and Ketcha populations do not exactly fit to any of the recognized taxa of the genus *Guizotia* and the taxonomic status of these populations has not yet been determined (Dagne 1995, 2001). We treated each of them as a separate "taxon" for the sake of simplicity. The geographic distribution of these populations has not been fully explored.

The economic use of the wild and/or weedy *Guizotia* taxa is not well known. However, *G. scabra* ssp. *scabra* was reported to be used for human consumption (seeds) in Nigeria, and to make fishing nets (stem) in Congo (Baagoe 1974). G. *scabra* ssp. *schimperi* was reported to be used as a substitute of *G. abyssinica* and also as herbal medicine in some places in Ethiopia (Baagoe 1974). The fatty acids of *G. scabra* ssp. *schimperi* are identical to that of *G. abyssinica* (Dagne and Jonsson 1997), thus the use of *G. scabra* ssp. *schimperi* seeds for human consumption is encouraging.

In spite of their biological and potential economic importance, almost no information is available on the extent and distribution of genetic variability in wild and/or weedy *Guizotia* species. In order to make reasonable decisions about sampling procedures and to develop conservation strategies that preserve maximum levels of genetic diversity one should know how genetic variation is distributed throughout each taxon and what their populations genetic structures look like. The use of amplified fragment length polymorphism (AFLP) and random amplified polymorphic DNA (RAPD) to assess genetic variability within and among populations and determining population structure is promising because many polymorphic loci can be obtained fairly easy without any prior knowledge of the genome of the species under study (Vos et al. 1995; Gupta et al. 1999; Hansen et al. 1999; Nybom and Bartish 2000; Nybom 2004). Hence, the present study was undertaken to investigate the extent of genetic variation within and between populations of all wild and/or weedy *Guizotia* taxa from Ethiopia, using AFLP and RAPD markers.

MATERIAL AND METHODS

Plant material and DNA extraction

Germplasm collection was conducted in Ethiopia from November to December 2003 (Table 1, Fig. 1). Plants of the same taxon found continuously without large gaps in between were considered as a population and each population was represented by 10 individual plants. Samples were collected randomly and systematically depending on the size of the populations. Different numbers of populations was studied per taxon based on the geographic range and abundance of the species within the country. Fresh leaves from 15 days to one month old plants grown in the greenhouse were used for genomic DNA extraction. DNA was extracted by a modified CTAB procedure as described in Assefa et al. (2003). Initially, this study was conducted using RAPD molecular marker technique. Thereafter, AFLP marker technique was applied on some of the populations previously studied using RAPDs in order to provide additional data and strengthen the conclusion that could be made from the study.

AFLP

The AFLP analysis was performed according to Vos et al. (1995) with some modifications. The procedures followed are as follows: genomic DNA (1 μg) was sequentially digested first with 5 units of *Mse*I at 65°C for 1 h and then with 5 units of *Eco*RI at 37°C for 90 min in a final volume of 50 μl, following the recommendations of the manufacturer of the enzymes (Fermenta Life Sciences), in a reaction buffer that contains 6.6 mM Tris-acetate pH 7.9, 2 mM magnesium acetate, 13.2 mM potassium acetate and 20 ng μl^{-1} BSA. To the double digested DNA sample, a total volume of 10 μl of ligation reaction mixture that contains 0.5 μmoles *Eco*RI adapter, 6 μmoles *Mse*I adapter, 50 mM Tris- HCl pH 7.6, 10 mM $MgCl_2$, 1 mM ATP, 1 mM DTT, 5% (w/v) polyethylene glycol-8000 and 1 unit T_4-DNA ligase was added, and then incubated for 3 h at 37°C in a thermocycler to ligate the resultant fragments to the corresponding adapters.

The double digested and ligated product was diluted 1:2.3 with $T_{10}E_{0.1}$ (10 mM Tris-HCl pH 8.0, 0.1 mM EDTA), and then 10 μl of each sample was used as a template to conduct the pre-amplification in a final volume of 20 μl that contains 1 × PCR buffer (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 dNTPs, 1.5 mM $MgCl_2$ and 0.5 U Thermowhite *Taq* DNA polymerase (Saveen Werner AB, Sweden). The preamplification reaction was carried out for 20 cycles of 1 min denaturing at 92°C, 30 s primers annealing at 60°C and 1 min primers extension at 72°C. Pre-amplification product was diluted 1:25 with $T_{10}E_{0.1}$, and then used as template for the selective PCR amplification to generate AFLPs.

Seven selective primer-pairs, out of the 56 initially screened, that detected high polymorphism and amplified the profiles that could be unambiguously and easily scored were used to genotype the entire populations (Table 2). This main DNA amplification was

Table 1. *Population codes, altitudes, regions and place of collections for seven Guizotia taxa used for this study.*

Species/population	Popn code	Region	Altitude (m a.s.l.)	Place of collection
Chelelu	Shewa-ch	Shewa	2475	About 20 km from Addis Ababa to Sendafa (Chelelu river)**
	Chida-ar-1	Kefa	2100	8 km from Chida to Ameya**
	Chida-ar-2	Kefa	2110	14 km from Chida to Ameya**
G. arborescens	Chida-ar-3	Kefa	2200	20.5 km from Ameya to Chida*
	O/Nadda-ar	Kefa	2382	The hill south of Omo-Nadda town (27 km to south of AA-Jimma road)**
	G/Gofa-sc	Gamu Gofa	2020	10 km from Sodo to Arba Minch**
	Illubabor-sc-1	Illbabor	1810	2 km north of Dedessa bridge*
	Illubabor-sc-2	Illubabor	1900	53 km from Gore to Bure**
	Keffa-sc-1	Keffa	2430	9.5 km from Chida to Ameya*
G. scabra ssp. *scabra*	Keffa-sc-2	Kefa	2245	43 km from Jimma to Bonga*
	Keffa-sc-3	Keffa	2192	11 km from Bonga to Menjo**
	Sidamo-sc	Sidamo	1890	3 km from Wadera to N/Borena*
	Tigray-sc	Tigray	2300	103 km from Maychew to Mekele**
	Welega-sc	Welega	1480	86 km from Nekemt to Gimbi*
	Bale-sh	Bale	2370	36 km east of Robe town*
	G/Gofa-sh	Gamu Gofa	2225	19.5 km from Sodo to Waka*
	Harerge-sh	Harerge	2200	44 km from Kobo to Hirna**
	Illubabor-sh	Illubabor	1900	47 km from Gore to Bure*
G. scabra ssp. *schimperi*	Keffa-sh	Kefa	1720	38 km from Jimma to Agaro**
	Shewa-sh	Shewa	2570	1 km from Sendafa to Chefe Donsa**
	Sidamo-sh	Sidamo	1770	2.4 km from K/Mengist to Shakiso**
	Tigray-sh	Tigray	1920	13 km from Abi Adi to Adwa*
	Welega-sh	Welega	2443	60 km from Bako to Shambu*
	Gojam-vl	Gojam	1920	16 km from B/Dar to Mota**
	Gonder-vl	Gondar	2220	6 km from Gondar to Azezo**
	Tigray-vl-1	Tigray	2008	20 km Slekleka to Shire*
G. villosa	Tigray-vl-2	Tigray	1920	15 km from Abi Adi to Adwa*
	Tigray-vl-3	Tigray	2410	22 km from Adi Gudem to Mekele*
	Tigray-vl-4	Tigray	2140	Aba Gerima hill (Adwa)**
	Arero-zv-1	Sidamo	1800	Arero town (Sillase church compound)*
	Arero-zv-2	Sidamo	1820	1.5 km from Arero to Babila**
	Mega-zv-1	Sidamo	1870	1 km north of Mega town *
G. zavattarii	Mega-zv-2	Sidamo	1725	3 km from Mega to Moyale**
	Yabelo-zv-1	Sidamo	1974	3 km from Yabelo to Konso*
	Yabelo-zv-2	Sidamo	2030	5.8 km from Yabelo to Konso**
Ketcha	Bale-kt	Bale	2380	About 64 km from Bale-Goba to Delo Mena**

*Populations used for RAPDs only; **Populations used for both AFLPs and RAPDs.

carried out using primers with three selective nucleotides (SNs) (Table 2) in a final volume of 20 µl containing 1 × PCR buffer (same as above), 25 ng *Eco*RI primer + 3 SNs, 30 ng *Mse*I primer + 3 SNs, 0.2 M dNTPs, 1.5 mM $MgCl_2$, 0.5 U *Taq* DNA polymerase and 10 µl pre-amplified and diluted DNA sample. The reaction was conducted through an initial 2 min denaturing at 94°C and a final 2 min extension at 72°C, with 36 intervening cycles as described in FERDINANDEZ and COULMAN (2004). GeneAmp® PCR system 9700 (Applied Biosystems) were used for incubation during restriction and ligation reactions and for preamplification and main amplification reactions.

The amplified product was denatured by adding 15 µl of formamide loading buffer (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. The AFLP electrophoresis was conducted using a Model S_2 Sequencing Gel Electrophoresis Apparatus (Whatman Biometra®). Seven l of the amplification products was loaded on 5% (w/v) polyacrylamide gels and separated at 90 W constant power until xylene cyanol FF runs two-third of the length of the plate. Before loading the samples, the gel was pre-ran for 45 min to warm the gel to 50–60°C. DNA size standard (100 bp ladder) from Amersham Biosciences AB was used to estimate the molecular size of the AFLP fragments. Following electrophoresis, DNA bands were visualized using the silver staining technique according to CAETANO-ANOLLES and GRESSHOFF (1994) with the following modifications: (1) 10% acetic

Fig. 1. Map of Ethiopia showing the locations of *Guizotia* populations studied. The big arrow indicate the capital city Addis Ababa.

acid was used as fixer solution and stopping solution and (2) the concentration of sodium thiosulfate in developing solution was 2 mg l^{-1}.

RAPD

Several protocols 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. Out of the 150 RAPD primers (from QIAGEN Operon GmbH, Germany) tested, ten primers (Table 2) that were highly reproducible, revealed high polymorphism, showed clearly resolvable banding patterns and amplified larger number of loci per sample were used for generating RAPD data. The DNA amplification, electrophoresis and staining

Table 2. *List of AFLP primer-pairs and RAPD primers used for this study.*

AFLP	RAPD	
Primer-pair	Primer	Base sequence
*E-AAG/**M-CTC	OPA-11	5′-CAATCGCCGT-3′
*E-ACA/**M-CTA	OPA-14	5′-TCTGTGCTGG-3′
*E-ACA/**M-CTG	OPB-18	5′-CCACAGCAGT-3′
*E-ACG/**M-CTC	OPB-20	5′-GGACCCTTAC-3′
*E-ACT/**M-CAC	OPD.20	5′-ACCCGGTCAC-3′
*E-AGG/**M-CAT	OPF-5	5′-CCGAATTCCC-3′
*E-AGG/** M-CTA	OPF-10	5′-GGAAGCTTGG-3′
	OPG-2	5′-GGCACTGAGG-3′
	OPG-16	5′-AGCGTCCTCC-3′
	OPG-17	5′-ACGACCGACA-3′

*E = *Eco* RI primer (5′-GACTGCGTACCAATTC-3′);
**M = *Mse* I primer (5′-GATGAGTCCTGAGTAA-3′).

of the amplified product was done as described in Geleta et al. (2006).

Data scoring and analysis

Both AFLPs and RAPDs were treated as dominant markers and each AFLP and 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 genotypes of each taxon. A locus was considered polymorphic when the frequency of present allele or null allele is less than 99% across the whole genotypes investigated for each taxon. Genetic diversity was calculated based on Nei's unbiased gene diversity (Nei 1978) with the modification provided by Lynch and Milligan (1994), using polymorphic loci only. The mean observed gene diversity within each population (H_j) and estimates of genetic differentiation between populations (G_{ST}) (Nei 1973) was also calculated as described in Geleta et al. (2006).

The NTSYSpc program (Rohlf 2000) was used to calculate Jaccard's similarity coefficient (Jaccard 1908) between populations of each taxon. POPGENE version 1.31 (Yeh and Boyle 1997) was used for analysis of percentage of polymorphic loci for each population of each taxon, while analysis of molecular variance (AMOVA) was done using Arlequin ver. 2 (Schneider et al. 2000).

RESULTS

Percentage of polymorphic loci and extent of genetic variation

Table 3 summarizes the percentage of polymorphic loci, H_j and Jaccard's similarity coefficient for each taxon of *Guizotia* studied. The mean percentage of polymorphic loci per population (P_p) was found to be 17%, 61%, 65% 62% and 24% in the case of AFLP and 38%, 51%, 51%, 57% and 35% in the case of RAPD, respectively for *G. arborescens*, *G. scabra* ssp. *scabra*, *G. scabra* ssp. *schimperi*, *G. villosa* and *G. zavattarii*. When all loci were considered for each taxon as a whole, the percentage of polymorphic loci (P_S) was 28.5%, 84.5%, 90.0%, 83.9% and 50.1% for AFLP and 86.6%, 99.6%, 99.6%, 98.5% and 92.7% for RAPD in that order for the above taxa. The percentage of polymorphic loci (P_P) for Chelelu and Ketcha was 41% and 81% in the case of AFLP and 75% and 92% in the case of RAPD, respectively. Thus, both marker systems revealed substantially lower polymorphism for *G. arborescens* and *G. zavattarii* than the rest of the taxa. Unlike, *G. arborescens* and *G. zavattarii* the percentage of polymorphic loci for Chelelu and

Table 3. *Percent polymorphic loci, H_j and Jaccard's similarity coefficient (JSC) generated from AFLPs and RAPDs for each population and the corresponding means for each taxon.*

Taxa	Population	Percent polymorphic loci				H^c_{loci}		JSC[c]	
		per population[a]		per taxon[b]		AFLP	RAPD	AFLP	RAPD
		AFLP	RAPD	AFLP	RAPD				
Chelelu	Shewa-ch	40.82	74.68	40.82	74.68	0.70	0.74		
	Chida-ar-1	17.30	39.78	28.54	85.64	0.50	0.37	0.89	0.53
	Chida-ar-2	18.65	41.99			0.44	0.38	0.89	0.51
G. arborescens	Chida-ar-3	nd	32.04			nd	0.33	nd	0.48
	O/Nadda-ar	16.18	36.46			0.43	0.29	0.90	0.48
	Mean	**17.38**	**37.57**			**0.46**	**0.34**	**0.89**	**0.50**
	G/Gofa-se	64.55	60.39	84.55	99.60	0.57	0.44	0.54	0.26
	Illubabor-se-1	nd	55.69			nd	0.47	nd	0.28
	Illubabor-se-2	61.82	44.71			0.54	0.37	0.54	0.26
	Keffa-se-1	nd	51.37			nd	0.41	nd	0.27
G. scabra ssp. *scabra*	Keffa-se-2	nd	47.45			nd	0.36	nd	0.28
	Keffa-se-3	63.27	56.47			0.48	0.43	0.54	0.26
	Sidamo-se	nd	45.10			nd	0.41	nd	0.28
	Tigray-se	54.73	50.59			0.55	0.41	0.54	0.26
	Welega-se	nd	49.02			nd	0.40	nd	0.24
	Mean	**61.09**	**51.20**			**0.54**	**0.41**	**0.54**	**0.26**
	Bale-sh	nd	48.73	90.02	99.58	nd	0.45	nd	0.22
	G/Gofa-sh	nd	58.47			nd	0.45	nd	0.22
	Harerge-sh	65.70	47.64			0.52	0.35	0.55	0.24
	Illubabor-sh	nd	61.02			nd	0.39	nd	0.23
G. scabra ssp. *schimperi*	Kefa-sh	61.34	62.97			0.58	0.41	0.55	0.24
	Shewa-sh	59.71	51.69			0.57	0.34	0.53	0.24
	Sidamo-sh	71.51	44.49			0.54	0.50	0.55	0.24
	Tigray-sh	nd	43.22			nd	0.31	nd	0.23
	Welega-sh	nd	41.95			nd	0.37	nd	0.24
	Mean	**64.57**	**51.13**			**0.55**	**0.40**	**0.55**	**0.23**
	Gojam-vl	64.48	56.93	83.93	98.50	0.58	0.47	0.49	0.24
	Gonder-vl	58.73	62.87			0.59	0.52	0.52	0.25
	Tigray-vl-1	nd	57.43			nd	0.46	nd	0.28
G. villosa	Tigray-vl-2	nd	55.94			nd	0.45	nd	0.28
	Tigray-vl-3	nd	56.93			nd	0.45	nd	0.28
	Tigray-vl-4	64.68	50.99			0.53	0.38	0.51	0.29
	Mean	**62.63**	**56.85**			**0.57**	**0.46**	**0.51**	**0.27**
	Arero-zv-1	nd	32.77	50.11	92.66	nd	0.35	nd	0.43
	Arero-zv-2	29.36	39.55			0.44	0.31	0.78	0.43
	Mega-zv-1	nd	33.33			nd	0.35	nd	0.47
G. zavattarii	Mega-zv-2	24.06	25.42			0.35	0.28	0.79	0.44
	Yabelo-zv-1	nd	38.42			nd	0.28	nd	0.47
	Yabelo-zv-2	17.66	39.55			0.26	0.22	0.79	0.45
	Mean	**23.69**	**34.84**			**0.35**	**0.30**	**0.79**	**0.45**
Ketcha	Bale-Kt	80.64	91.67	80.64	91.67	0.68	0.72		

nd = no data; [a]the mean of the populations per taxon is *Pp*; [b]Percent polymorphic loci when all individuals within a taxon are considered together across all loci (*Ps*); [c]The correlation between AFLP and RAPD is highly significant (P <0.001).

Ketcha was found to be high regardless of their small population sizes and being completely isolated populations. The degree of polymorphism (P_S) revealed by AFLPs is lower than that revealed by RAPDs for all the *Guizotia* taxa studied (Table 3). Conversely, the effective multiplex ratio (EMR) was higher for AFLPs (23.8) than for RAPDs (6.9) (Table 4).

This analysis revealed higher $\overline{H}_T$ for *G. arborescens* and *G. zavattarii* (AFLP = 0.37, RAPD = 0.28) as compared to the other taxa. The relatively higher overall gene diversity in these species was mainly due to relatively high population differentiation (Table 4). The total gene diversity estimates ($\overline{H}_T$) for *G. scabra* ssp. *scabra*, *G. scabra* ssp. *schimperi* and *G. villosa*

Table 4. *Summary of AFLP and RAPD based effective multiplex ratio (EMR), total gene diversity ($\overline{H}_T$), within population genetic diversity ($\overline{H}_S$) and population differentiation ($\overline{G}_{ST}$) for various Guizotia taxa.*

Taxa	EMR		Gene diversity		
	AFLP	RAPD	Parameter	AFLP	RAPD
Chelelu	12.8	8.8	$\overline{H}_S$	0.37	0.31
G. arborescens	1.9	2.6	$\overline{H}_S$	0.24	0.15
			$\overline{H}_T$	0.37	0.28
			$\overline{G}_{ST}$	0.30	0.34
G. scabra ssp. *scabra*	29.3	6.7	$\overline{H}_S$	0.24	0.17
			$\overline{H}_T$	0.32	0.23
			$\overline{G}_{ST}$	0.18	0.19
G. scabra ssp. *schimperi*	32.8	6.2	$\overline{H}_S$	0.24	0.17
			$\overline{H}_T$	0.32	0.22
			$\overline{G}_{ST}$	0.19	0.20
G. villosa	28.3	6.5	$\overline{H}_S$	0.24	0.19
			$\overline{H}_T$	0.33	0.23
			$\overline{G}_{ST}$	0.19	0.15
G. zavattarii	3.6	2.1	$\overline{H}_S$	0.17	0.14
			$\overline{H}_T$	0.37	0.28
			$\overline{G}_{ST}$	0.41	0.40
Ketcha	57.6	15.1	$\overline{H}_S$	0.33	0.30
All taxa	23.8	6.9			

were found to be almost similar (ranging from 0.32–0.33 (AFLP) and 0.22–0.23 (RAPD)). The lowest and highest overall within population gene diversity ($\overline{H}_S$) was recorded for *G. zavattarii* (AFLP = 0.17, RAPD = 0.14) and *G. villosa* (AFLP = 0.24, RAPD = 0.19), respectively (Table 4). Our analyses, based on the combined data of AFLP and RAPD, by grouping the taxa into annuals and perennials, demonstrated slightly higher within population diversity in perennials (H_S = 0.24) than in annuals (H_S = 0.20). This is mainly due higher within population variation (H_S) revealed in Chelelu (AFLP = 0.37; RAPD = 0.31) and Ketcha (AFLP = 0.33; RAPD = 0.30) as compared to the other taxa (Table 4).

The extent of genetic variation for each population of each taxon was estimated as H_j for both marker systems (Table 3). Pearson's correlation coefficient between AFLPs and RAPDs for this parameter was highly significant (r = 0.66, P < 0.001). On average, AFLPs revealed higher diversity in each population in all taxa as compared to RAPDs. Overall, the lowest and highest H_j value was recorded for Yabelo-zv-2 of *G. zavattarii* (AFLP = 0.13; RAPD = 0.11) and Chelelu (AFLP = 0.37; RAPD = 0.31) respectively (Table 3).

Population differentiation and genetic similarity between populations

Population differentiation was estimated using modified gene diversity estimate as G_{ST} and AMOVA as F_{ST}. The highest population differentiation was recorded for *G. zavattarii* as revealed by both AFLPs ($\overline{G}_{ST}$ = 0.41; F_{ST} = 0.50) and RAPDs ($\overline{G}_{ST}$ = 0.40, F_{ST} = 0.57). The two marker systems and the two population differentiation parameters were not in a complete agreement regarding the taxon with the lowest level of population differentiation. However, generally populations of *G. scabra* ssp. *scabra*, *G. scabra* ssp. *schimperi* and *G. villosa* were less differentiated as compared to populations of *G. arborescens* and *G. zavattarii* (Table 4, 5). The genetic variation between populations estimated as F_{ST} through analysis of molecular variance (AMOVA) was found to be significant (P < 0.0001) for all taxa in both marker systems (Table 5). Analysis of population differentiation based on combined data of AFLPs and RAPDs revealed significantly higher population differentiation (P < 0.001) in perennials (G_{ST} = 0.32; F_{ST} = 0.40) as compared to annuals (G_{ST} = 0.18; F_{ST} = 0.33).

The AFLP-based mean Jaccard's similarity coefficient between populations within taxon ranged from 0.51 (*G. villosa*) to 0.89 (*G. arborescens*), while the RAPD-based estimate ranged from 0.23 (*G. scabra* ssp. *schimperi*) to 0.50 (*G. arborescens*). Despite a relatively high similarity obtained using AFLPs, both marker systems agreed that the similarity between populations of *G. arborescens* and *G. zavattarii* is higher as compared to that of the other taxa (Table 3). The general agreement between the two marker systems was further supported by their significant positive correlation (r = 0.98, P < 0.001). In the case of RAPDs, of the five taxa that were represented by more than one populations, the highest genetic similarity

Table 5. *Analysis of molecular variance (AMOVA) for each of the five Guizotia taxa based on AFLP and RAPD data.*

Species		Source of variation	DF	Sum of squares	Variance components	% of variation	Fixation index (F_{ST})*
G. arborescens	AFLP	Among populations	2	156.27	6.29	29.31	0.29
		Within populations	27	409.90	15.18	70.69	
		Total	29	566.17	21.48		
	RAPD	Among populations	3	434.20	12.92	45.44	0.45
		Within populations	36	558.60	15.52	54.56	
		Total	39	992.80	28.44		
G. scabra ssp. *scabra*	AFLP	Among populations	3	760.40	18.85	22.49	0.22
		Within populations	36	2338.60	64.96	77.51	
		Total	39	3099.00	83.81		
	RAPD	Among populations	8	1413.60	15.03	36.30	0.36
		Within populations	81	2136.90	26.38	63.70	
		Total	89	3550.50	41.41		
G. scabra ssp. *schimperi*	AFLP	Among populations	3	633.17	14.29	17.33	0.17
		Within populations	36	2453.50	68.15	82.67	
		Total	39	3086.67	82.44		
	RAPD	Among populations	8	1362.73	14.63	37.80	0.38
		Within populations	81	1949.40	24.07	62.20	
		Total	89	3312.13	38.69		
G. villosa	AFLP	Among populations	2	558.07	21.70	25.90	0.26
		Within populations	27	1676.00	62.07	74.10	
		Total	29	2234.07	83.77		
	RAPD	Among populations	5	644.55	10.56	31.17	0.31
		Within populations	54	1259.10	23.32	68.83	
		Total	59	1903.65	33.88		
G. zavattarii	AFLP	Among populations	2	456.87	20.78	50.12	0.50
		Within populations	27	558.20	20.67	49.88	
		Total	29	1015.07	41.45		
	RAPD	Among populations	5	913.55	17.01	57.38	0.57
		Within populations	54	682.30	12.64	42.62	
		Total	59	1595.85	29.64		

*P-value < less than 0.0001.

was recorded for *G. arborescens*, between Chida-ar-2 and Chida-ar-3 (0.58). Geographically, these populations are located very close being separated by about six km only. While, the lowest genetic similarity was recorded for *G. scabra* ssp. *schimperi* between Tigray-sh and G/Gofa-sh (0.18) which are located north and south of the country (Fig. 1).

Guizotia scabra ssp. *scabra*, Welega-sc, being collected from the lowest altitude (1480 m a.s.l.), showed the lowest similarity (0.24) compared to the other populations within the taxon. Both marker systems agreed that the *G. villosa* population from Gojam (Gojam-vl) has the lowest similarity to the other populations of that species (AFLP = 0.49, RAPD = 0.24), especially to the populations from Tigray. Similarly, in the case of *G. zavattarii*, Arero populations (Arero-zv-1 and Arero-zv-2) showed lowest similarity to the rest populations (Table 3).

DISCUSSION

The results obtained from this study showed that the percentage of polymorphic loci revealed by AFLPs is lower than that revealed by RAPDs, which is in agreement with the reports of Milbourne et al. (1997) and Pejic et al. (1998). Conversely, the EMR calculated as the product of fraction of polymorphic loci and number of polymorphic loci per individual assay, was more than three fold higher for AFLPs than for RAPDs, which is also in a good agreement with previous reports (Powell et al. 1996; Milbourne 1997). This indicates that AFLPs detect higher numbers of polymorphism in a single assay, reflecting the higher efficiency of AFLPs to simultaneously analyze larger number of polymorphic loci as compared to RAPDs. Furthermore, AFLP revealed higher total gene diversity (H_T) and within population genetic variation (H_S), for all the *Guizotia* taxa

studied (Table 4), implying the superiority of AFLPs over RAPDs in detecting genetic variation in *Guizotia*. Nevertheless, both marker systems are in a good agreement on several important points and found to be very useful for genetic diversity study in *Guizotia*.

Both AFLPs and RAPDs revealed a relatively higher level of genetic polymorphism for those *Guizotia* taxa with relatively wider geographic distribution in Ethiopia (*G. scabra* ssp. *scabra* and *G. scabra* ssp. *schimperi*). Although restricted to the northern and northwestern part of the country, the geographic distribution of *G. villosa* is wider as compared to that of *G. arborescens* and *G. zavattarii*. The P_S value of *G. villosa* was higher than that of the latter taxa, thus demonstrating a direct relationship between the geographic range of a taxon and the extent of their genetic polymorphism. The two highly localized species with small population sizes (*G. arborescens* and *G. zavattarii*) showed lower within population diversity and reduced levels of polymorphism as compared to the other taxa. The level of genetic diversity in small and isolated populations is often low, mainly as a consequence of the pronounced effects of genetic drift (AYRES and RYAN 1997; PEDERSEN and LOESCHCKE 2001; COLE 2003; GRUBBS and CASE 2004). Therefore, higher proportions of monomorphic loci in these species, as compared to their relatively common congeners, could be partly explained by their highly localized distribution and small population sizes. Conversely, the weedy nature of the two subspecies of *G. scabra* and *G. villosa* might have contributed to their relatively higher diversity, as weediness encounters disturbed habitats that lead to further accumulation of genetic variation.

Unlike the case of *G. arborescens* and *G. zavattarii*, the within population genetic variation obtained for Chelelu and Ketcha was higher than that of the other taxa (Table 4). The implication is that, genetic drift could be counterbalanced by other factors that promote genetic variation even in small populations such as Chelelu and Ketcha. Factors that promote genetic variation, such as high rate of mutation and interspecific hybridization with other *Guizotia* taxa, might have played a significant role to maintain high genetic variation in these populations. Generally, the trend is that perennials have higher within population diversity as compared to annuals, provided that they have similar mating systems (HAMRICK and GODT 1989; NYBOM and BARTISH 2000). Thus, the higher diversity in Chelelu and Ketcha as compared to their annual congeners might be expected due to their perennial habit. However, low diversity in the other two perennial species (*G. arborescens* and *G. zavattarii*) is a counter-example to this trend. Therefore, different factors such as life forms and geographic distribution with differential intensity of selection contributed differently to the extent of within population diversity in different taxa of the genus *Guizotia*.

Analysis of the extent of genetic diversity in each population (H_j), and genetic similarity between populations of each taxon, is important to relate the genetic diversity to environmental factors, to identify populations with unique genetic constitution and to suggest genetic diversity centers for each taxon. In *G. arborescens*, the slightly lower genetic diversity of O/Nadda-ar population as compared to Chida populations suggests that the micro genetic diversity center for this species is the area between Chida and Ameya (Table 1, Fig. 1). On the other hand, the extent of genetic variability of the two subspecies of *G. scabra* seems to be distributed regardless of their geographic locations within the country.

Although *G. villosa* is relatively common in Tigray, populations from Gonder and Gojam seem to have higher genetic variation as compared to Tigray populations. Furthermore, Jaccard's similarity coefficient indicated that the Gojam population showed the lowest similarity to the other populations. Thus, future conservation activities including germplasm collection should also consider such population located at the periphery of the species distribution range, as they seem to have relatively high unique genetic constitution. The case of *G. zavattarii* is also interesting in that generally the two populations from Arero (Arero-1 and Arero-2) are genetically less similar to the other populations and seem to have slightly higher diversity. Arero is remote, with less human interference as compared to Mega and Yabelo, which may suggest that a stable community contributes to the genetic diversity of natural plant populations. This thus agrees with TILMAN and DOWNING (1994) who indicated the positive correlation between diversity and stability in a plant species.

Population differentiation

G_{ST} measure the proportion of variation among populations relative to the total species diversity estimated as H_T. F_{ST} expresses the proportion of variation explained by each population and is the average inter-population distance between any two populations (LI and NELSON 2002). In our analysis, the G_{ST} and F_{ST} values generated both from AFLPs and RAPDs suggest that all *Guizotia* taxa, except *G. zavattarii*, are more diverse within, with less genetic differentiation between, populations. In addition to inter-population pollen movement, interspecific gene

flow might have also contributed to reduced population differentiation in these species, as these *Guizotia* taxa co-exist in most of their natural habitats.

BUSSELL (1999) indicated that the G_{ST} values of outcrossing species is usually less than 20%, which agrees with our result when *G. scabra* ssp. *scabra*, *G. scabra* ssp. *schimperi* and *G. villosa* are considered (Table 4). The mean of G_{ST} values for *Guizotia* (AFLP = 0.25; RAPD = 0.26) was found to be higher than the average RAPD-based G_{ST} value (0.16) obtained for 27 outcrossing species by BUSSELL (1999), and G_{ST} value (0.18) obtained by (SHIM and JORGENSEN 2000) for wild carrot. Our result was also slightly higher than the mean G_{ST} values obtained from 15 RAPD-based studies (0.21) and eight AFLP-based studies (0.24) (NYBOM 2004). The mean F_{ST} values for our taxa (AFLP = 0.29, RAPD = 0.41) were also higher than that obtained by NYBOM (2004) for RAPD-based studies (0.24) and AFLP-based studies (0.23). Thus, this study revealed that the extent of population differentiation in *Guizotia* seems to be relatively high, specifically in *G. arborescens* and in *G. zavattarii*.

MURTHY et al. (1993) mentioned generally that *Guizotia* species are highly cross pollinated, and self-incompatible. The outcrossing nature of these species was also supported by the work of DAGNE (2001) who obtained viable and fertile hybrid plants from crosses between *G. arborescens* and *G. zavattarii*. However, there is no direct evidence of self-incompatibility of wild/weedy Guizotias. The extent of population differentiation in *G. arborescens* and *G. zavattarii* are higher than expected when we consider them as outcrossing species, suggesting that these species might have some degree of inbreeding. Their high population differentiation might be best explained by the fact that these species are highly localized, rare and have small population sizes, as population differentiation is argued to be higher in narrowly distributed species with small population sizes than in widespread species (LOVELESS and HAMRICK 1984; SLATKIN 1987).

Generally, although the extent of within and between population genetic variation varies among different *Guizotia* taxa studied, substantial amount of overall genetic diversity was held in all *Guizotia* taxa. Therefore, acknowledging the existing genetic variation within and between populations is important for conservation and breeding purposes. It has been suggested that *G. scabra* ssp. *schimperi* is the progenitor of *G. abyssinica* (BAAGOE 1974; HIREMATH and MURTHY 1988; MURTHY et al. 1993; DAGNE 1994, 1995, 2001). Crosses between *G. scabra* ssp. *schimperi* and *G. abyssinica* were highly successful as compared to most hybridization attempts within the genus *Guizotia* (MURTHY et al. 1993; DAGNE 1994). Correspondingly, our results revealed high genetic diversity in *G. scabra* ssp. *schimperi*, which seems to be a good opportunity for improvement of *G. abyssinica* as the chance to find desirable genes within *G. scabra* ssp. *schimperi* and transfer it to *G. abyssinica* through hybridization is high. Other *Guizotia* species could also be a good source of desirable gene for improvement of *G. abyssinica*.

Maintaining adequate levels of genetic diversity within and among populations is a critical aspect to consider (HAIG 1998). The case of *G. arborescens* and *G. zavattarii* needs special attention because these species are characterized by highly localized geographic distribution, small population sizes, lower percentage of polymorphic loci, higher genetic similarity between populations and higher population differentiation as compared to other taxonomically recognized Guizotias. If species specific requirements and environmental factors that immediately affect survival, reproduction and dispersal are critically determining the geographic range and abundance of these species, it would be of interest what specific requirements are restricting it. What ever the case is, unless immediate conservation activity is started, their highly restricted distribution, small population sizes and an increase in human activities in their natural habitat may lead these species to extinction. Threats to these species are mainly anthropogenic (fires and agricultural expansion). Finally, this study highlights the importance of molecular analysis in understanding the genetic diversity and population structure of wild and/or weedy *Guizotia* taxa and contributes to the knowledge of conservation of genetic resources in the Ethiopian flora.

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

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, D. R. and Ryan, F. J. 1997. The clonal and population structure of a rare endemic plant, *Wyethia reticulate* (Asteraceae): allozyme and RAPD analysis. – Mol. Ecol. 6: 761–772.

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

Bussell, J. D. 1999. The distribution of random amplified polymorphic DNA (RAPD) diversity amongst

populations of *Isotoma petraea* (Lobeliaceae). – Mol. Ecol. 8: 775–789.
Caetano-Anollés, G. and Gresshoff, P. M. 1994. Staining nucleic acid with silver: an alternative to radioisotopic and fluorescent labling. – Promega Notes 45: 13.
Cole, C. T. 2003. Genetic variation in rare and common plants. – Annu. Rev. Ecol. Syst. 34: 213–237.
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. and Jonsson, A. 1997. Oil content and fatty acid composition of seeds of *Guizotia* Cass. (Compositae). – J. Sci. Food Agric. 73: 274–278.
Ferdinandez, Y. S. N. and Coulman, B. E. 2004. Genetic relationships among smooth bromegrass cultivars of different ecotypes detected by AFLP markers. – Crop Sci. 44: 241–247.
Friis, I. B. 1971. A new species of *Guizotia* (Compositae) from north east tropical Africa. – J. Bot. 18: 231–234.
Geleta, M., Bryngelsson, T., Bekele E. et al. 2006. Genetic diversity of *Guizotia abyssinica* (L. F.) Cass. (Asteraceae) from Ethiopia as revealed by random amplified polymorphic DNA (RAPD). – Genet. Resour. Crop Evol. DOI 10.1007/s10722-006-0018-0.
Getinet, A. and Sharma, S. M. 1996. Niger (*Guizotia abyssinica* (L. f.) Cass. Promoting the conservation and use of underutilized and neglected crops. 5. – Inst. Plant Genet. Crop Plant Res., Gatersleben/Int. Plant Genet. Resour. Inst., Rome.
Grubbs, H. J. and Case, M. A. 2004. Allozyme variation in American ginseng (*Panax quinquefolius* L.): variation, breeding system, and implications for current conservation practice. – Conserv. Genet. 5: 13–23.
Gupta, P. K., Varshney, R. K., Sharma, P. C. et al. 1999. Molecular markers and their applications in wheat breeding. – Plant Breeding 118: 369–390.
Haig, S. M. 1998. Molecular contributions to conservation. – Ecology 79: 413–425.
Hansen, M., Kraft, T., Christansson, M. et al. 1999. Evaluation of AFLP in *Beta*. – Theor. Appl. Genet. 98: 845–852.
Hamrick, J. L. and Godt, M. J. W. 1989. Allozyme diversity in plant species. – In: Brown, A. H. D., Clegg, M. T., Kahler, A. L. et al. (eds), Plant population genetics, breeding and genetic resources. Sinauer, Sunderland MA, p. 43–63.
Hiremath, S. C. and Murthy, H. N. 1986. The structure, stability, and meiotic behavior of B-chromosome in *Guizotia scabra* (Vis.) Chiov.ssp. *scabra* (Compositae). – Caryologia 39: 397–402.
Hiremath, S. C. and Murthy, H. N. 1988. Domestication of niger (*Guizotia abyssinica*). – Euphytica 37: 225–228.
Hiremath, S. C. and Murthy, H. N. 1992. Cytological studies in *Guizotia* (Asteraceae). – Caryologia 45: 69–82.
Jaccard, P. 1908. Nauvelles researches sur la distribution florale. – Bull. Soc. Vaud. Sci. Nat. 44: 223–270.
Li, Z. and Nelson, R. L. 2002. RAPD marker diversity among cultivated and wild soybean accessions from four Chinese provinces. – Crop Sci. 42: 1737–1744.
Loveless, M. D. and Hamrick, J. L. 1984. Ecological determinants of genetic structure in plant populations. – Annu. Rev. Ecol. Syst. 15: 65–95.
Lynch, M. and Milligan, B. G. 1994. Analysis of population genetic structure with RAPD markers. – Mol. Ecol. 3: 91–99.
Milbourne, D., Meyer, R., Bradshaw, J. E. et al. 1997. Comparison of PCR-based marker systems for the analysis of genetic relationships in cultivated potato. – Mol. Breeding 3: 127–136.
Murthy, H. N., Hiremath, S. C. and Salimath, S. S. 1993. Origin evolution and genome differentiation in *Guizotia abyssinica* and its wild species. – Theor. Appl. Genet. 87: 587–592.
Nei, M. 1973. Analysis of gene diversity in subdivided populations. – Proc. Natl Acad. Sci. USA 70: 3321–3323.
Nei, M. 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. – Genetics 89: 583–590.
Nybom, H. 2004. Comparison of different nuclear DNA markers for estimating intraspecific genetic diversity in plants. – Mol. Ecol. 13: 1143–1155.
Nybom, H. and Bartish, I. V. 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.
Pedersen, A. A. and Loeschcke, V. 2001. Conservation genetics of peripheral populations of the mygalomorph spider *Atypus affinis* (Atypidae) in northern Europe. – Mol. Ecol. 10: 1133–1142.
Pejic, I., Ajmone-Marsan, P., Morgante, M. et al. 1998. Comparative analysis of genetic similarity among maize inbred lines detected by RFLPs, RAPDs, SSRs, and AFLPs. – Theor. Appl. Genet. 97: 1248–1255.
Powell, W., Morgante, M., Andre, C. et al. 1996. The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. – Mol. Breeding 2: 225–238.
Rohlf, F. J. 2000. NTSYSpc, Numerical taxonomy and multivariate analysis system, Ver. 2.1. Exeter Software. Setauket, NY, USA.
Schneider, S., Roessli, D. and Excoffier, L. 2000. Arlequin: a software for population genetics data analysis. Ver. 2.000. – Genetics and Biometry Lab, Dept. of Anthropology, Univ. of Geneva.
Shim, S. I. and Jorgensen, R. B. 2000. Genetic structure in cultivated and wild carrots (*Daucus carota* L.) revealed by AFLP analysis. – Theor. Appl. Genet. 101: 227–233.
Slatkin, M. 1987. Gene flow and the geographic structure of natural populations. – Science 236: 787–792.
Tilman, D. and Downing, J. A. 1994. Competition and biodiversity in spatially structured habitats. – Nature 367: 363–365.
Vos, P., Hogers, R., Bleeker, M. et al. 1995. AFLP: a new technique for DNA fingerprinting. – Nucleic Acids Res. 23: 4407–4414.
Yeh, F. C. and Boyle, T. J. B. 1997. Population genetic analysis of co-dominant and dominant markers and quantitative traits. – Belg. J. Bot. 129: 157.

Introduction

The genus *Guizotia* Cass. is a small Afro-montane endemic genus (Baagoe 1974) that belongs to the family Asteraceae, tribe Heliantheae. The genus comprises the following six recognized species, of which one was further classified into two subspecies (Baagøe 1974): *G. abyssinica* (L. f.) Cass., *G. arborescens* I. Friis, *G. jacksonii* (S. Moore) J. Baagøe, *G. scabra* (Vis.) Chiov. ssp. *scabra*, *G. scabra* (Vis.) Chiov. ssp. *schimperi* (Sch. Bip. in Walp.) J. Baagøe, *G. villosa* Sch. Bip. in Walp., *G. zavattarii* Lanza in Chiov. After two decades of the taxonomic revision by Baagøe (1974), two new populations of *Guizotia* were discovered by Dagne (1995) who called them "Chelelu" and "Ketcha", respectively, by the name of the place from where they were first collected. All these taxa are diploids with 2n = 2x = 30 chromosomes (Hiremath and Murthy 1986, 1992; Dagne 1995). Based on the fact that all taxa of the genus, except *G. jacksonii* are found in Ethiopia, Baagøe (1974) indicated that the genus has its center of distribution, diversity and perhaps its center of origin in Ethiopia.

All *Guizotia* species are wild and/or weedy, except *G. abyssinica*. *G. abyssinica* is cultivated mainly for its edible oil, particularly in Ethiopia but also in India and some other African and Asian countries (Baagøe 1974, Murthy et al. 1993, Getinet and Sharma 1996). G. *arborescens* is a shrubby perennial endemic to south west Ethiopia, and mountain Imatongs in Sudan and Uganda (Friis 1971, Baagøe 1974). The species is different from the other *Guizotia* taxa mainly in having an arborescent habit and petiolate leaves (Friis 1971, Baagøe 1974). *G. jacksonii*, which is not included in this study, is a rare creeping perennial and sparsely branching herb with restricted distribution in Kenya, Tanzania and Uganda (Baagøe 1974, Hiremath and Murthy 1992). *G. scabra* ssp. *scabra* is a perennial taxon of wide distribution in East Africa, Cameron and Nigerian highlands with a distributional gap in the Congolian rainforest area (Baagøe 1974; Hiremath and Murthy 1986, 1988). *G. scabra* ssp. *schimperi* is an erect, annual, terrestrial, moderately branching herb (Baagøe 1974) native to the Ethiopian highlands (Hiremath and Murthy 1992) and is a common weed of crops in mid and high altitude areas.

G. villosa is an erect, annual, and terrestrial much branching herb restricted in its distribution to the northern and northwestern part of the Ethiopian highland (Baagøe 1974, Hiremath and Murthy 1992). It could be easily distinguished from the other *Guizotia* species by its small heads and nearly square ray florets (Baagøe 1974). *G. zavattarii* is an erect, perennial, terrestrial plant restricted to southern Ethiopia around mount Mega, Arero and Yabelo, and the Hurri Hills in northern Kenya (Hiremath and Murthy 1992). Chelelu and Ketcha are perennial plants that were collected from areas located about 20 km on the Addis Ababa-Dessie road and 64 km on the Goba-Mena road, respectively. However, their geographic distribution has not been fully explored.

Several pieces of evidence presented by different researchers indicate that *G. abyssinica* might have originated from *G. scabra* ssp. *schimperi* through selection and further cultivation (Baagøe 1974; Hiremath and Murthy 1988; Murthy et al.

1993; Dagne 1994, 1995, 2001). However, differing from this, Bekele et al. (2006) based on DNA sequences derived from ITS region suggest that other *Guizotia* species might also have contributed to the origin of *G. abyssinica*. Baagoe (1974) united *G. scabra* and *G. schimperi* as two subspecies of *G. scabra* and named them *G. scabra* ssp. *scabra* and *G. scabra* ssp. *schimperi*, respectively. However, Hiremath and Murthy (1992) and Dagne (1995), based on the significant karyotypic differences between them, indicated that these taxa should be considered as separate species. Chelelu and Ketcha do not exactly fit to any of the recognized taxa of the genus *Guizotia* (Dagne 1995, 2001) and their taxonomic status has not been determined, although we treated each of them as a "taxon" for the sake of simplicity.

Amplified Fragment Length Polymorphism (AFLP) and Random Amplified Polymorphic DNA (RAPD) have been used to study genetic and phylogenetic relationship between plant species (*e.g.* Blattner et al. 2001, Bänfer et al. 2004, Kadereit and Kadereit 2005) and were reported to work well in closely related species (Landry and Lapointe 1996, Després et al. 2003). The use of these marker techniques to assess genetic relationship between different groups of taxa and to screen diagnostic markers is promising because many polymorphic loci can be obtained fairly easy, in a relatively short time and without any prior knowledge of the genome of the species under study (*e.g.* Vos et al. 1995, Gupta et al. 1999, Hansen et al. 1999, Nybom and Bartish 2000, Nybom 2004).

Hence, the present AFLP- and RAPD-based study was undertaken to: (i) determine the extent of genetic relationship among *Guizotia* taxa, (ii) to suggest the closest relative of *G. abyssinica*, (iii) to comment on the taxonomic status of the two subspecies of *G. scabra*, (iii) to suggest the taxonomic position of Chelelu and Ketcha populations within the genus, and (iv) to identify some diagnostic markers for different taxa in the genus *Guizotia*.

Materials and Methods

Plant material and DNA extraction

Germplasm was collected directly from farmers' fields and natural habitats in Ethiopia from November to December 2003 (Table 1, Fig. 1). In this study, each population of each taxon was represented by 10 individual plants. *G. abyssinica* populations were sampled at equidistant along the longest line found across the field. Each population of wild and/or weedy *Guizotia* taxa was sampled randomly and systematically depending on the size of the population. The number of populations studied per taxon varies depending on the geographic range and abundance of the taxon in Ethiopia (Table 1, Fig. 1). Seeds were grown in a greenhouse and fresh leaves from 15-30 days old were used for genomic DNA extraction. DNA was extracted by modified CTAB procedures as described in Assefa et al. (2003).

Table1. *Population codes, altitudes, regions and place of collections for the eight Guizotia taxa used in this study*

Taxa	Region	Population code	Altitude (m asl)	Place of collection
G. abyssinica	Arsi	Arsi-ab	2565	13 Km from Robe to Ticho
	Bale	Bale-ab	2450	2.7 Km from Adaba to Dodola
	Gojam	Gojam-ab	1890	35 Km from Amanuel to Bure
	Gonder	Gonder-ab	1590	9 Km from T/Dingay to Humera
	Hararge	Hararge-ab	1830	36 Km from A/Teferi to Gelemso
	Illubabor	Illubabor-ab	1865	3 Km from Metu to Gore
	Jimma	Jimma-ab	1860	25 km from Jimma to A/Ababa
	Shewa	Shewa-ab	1640	13 Km from Sh/Robit to D/Sina
	Tigray	Tigray-ab	1972	9.5 Km from Shire to Shiraro
	Welega	Welelga-ab	1940	21 Km from Nekemt to Ambo
	Welo	Welo-ab	2420	31 Km from W/Tena to Gashena
G. arborescens	Kefa	Chida-ar-1	2100	8 km from Chida to Ameya
	Kefa	O/Nadda-ar	2382	The hill south of Omo-Nadda town (27 km south of AA-Jimma road)
G. scabra ssp. *scabra*	G/Gofa	G/Gofa-sc	2020	10 Km from Sodo to Arba Minch
	Illubabor	Illubabor-sc	1900	53 Km from Gore to Bure
	Tigray	Tigray-sc	2300	103 Km from Maychew to Mekele
G. scabra ssp. *schimperi*	Hararge	Hararge-sh	2200	44 Km from Kobo to Hirna
	Kefa	Kefa-sh	1720	38 Km from Jimma to Agaro
	Sidamo	Sidamo-sh	1770	2.4 Km from K/Mengist to shakiso
G. villosa	Gondar	Gondar-vl	2220	6 Km from Gondar to Azezo
	Tigray	Tigray-vl-4	2140	Aba Gerima hill (Adwa)
G. zavattarii	Sidamo	Mega-zv-2	1725	3 Km from Mega to Moyale
	Sidamo	Arero-zv-2	1820	1.5 Km from Arero to Babila
Chelelu	Shewa	Shewa-ch	2475	about 20 km from Addis Ababa to Sendafa
Ketcha	Bale	Bale-ket	2380	about 64 km from Bale-Goba to Delo-Mena

PCR amplification and electrophoresis

RAPD

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. Out of 150 RAPD primers (from QIAGEN Operon GmbH, Germany) initially tested, ten primers (Table 3) that generated highly reproducible banding pattern, revealed greater polymorphism, showed clearly resolvable banding patterns and amplified larger number of loci per sample were used for generating RAPD data. The reproducibility of the result was also checked during data generation by replicating representative samples within a gel and across different gels, as suggested by Skroch and Neinhuis (1995). The DNA amplification, electrophoresis and staining of the amplified product were done as described in Geleta et al. (2006).

Figure 1. Map of Ethiopia showing locations of sample collection for eight *Guizotia* taxa.

AFLP

After optimizing AFLP protocol, a total of 56 AFLP primer combinations were tested for their capacity to detect polymorphisms and clarity of banding patterns in all the *Guizotia* taxa studied. Following this initial screening, seven primer combinations that amplified the profiles that could be unambiguously scored were used to genotype the entire population(s) of each taxon. Similar to RAPD, the reproducibility of the result was continuously checked during data generation by replicating representative samples within a gel and across different gels. The AFLP analysis was performed according to Vos et al. (1995) with some modifications. The procedures followed were as follows: Genomic DNA (1 μg) was sequentially digested first with 5 U of *Mse*I restriction enzyme at 65°C for 1 h and then with 5 U of *Eco*RI restriction enzyme at 37°C for 90 min in a final volume of 50 μl, following the recommendation of the manufacturer of the enzymes (Fermenta Life Sciences), in a reaction buffer that contains 6.6 mM Tris-acetate pH 7.9, 2 mM magnesium acetate, 13.2 mM potassium acetate and 20 ng/ul BSA. To the double digested DNA sample, a total volume of 10 μl of ligation reaction mixture containing 0.5 μmol *Eco*RI adapter, 6 μmol *Mse*I adapter, 1 x ligation reaction buffer (50 mM Tris-HCl, pH 7.6, 10 mM $MgCl_2$, 1 mM ATP, 1 mM DTT, 5% (w/v) polyethylene glycol-8000) and 1 unit T_4-DNA ligase was added, and then incubated for 3 h at 37°C in a thermocycler to ligate the resultant fragment to the corresponding adapters.

The double digested and ligated product was diluted 1:2.3 with $T_{10}E_{0.1}$ (10 mM Tris-HCl, pH 8, 0.1 mM EDTA), and then 10 µl of each sample was used as a template to conduct the pre-amplification in a final volume of 20 µl that contains 1x PCR buffer (20 mM Tris-HCl, pH 8.55, 16 mM $(NH_4)_2SO_4$, 0.01%Tween®20 and 2 mM $MgCl_2$), 30 ng each of *Eco*RI-A and *Mse*I-C primers, 0.2 mM dNTPs, 1.5 mM $MgCl_2$ and 0.5 U *Taq* DNA polymerase. The preamplification reaction was carried out for 20 cycles of 1 min denaturing at 92°C, 30 s primers annealing at 60°C and 1 min primers extension at 72°C. Pre-amplification product was diluted 1:25 with $T_{10}E_{0.1}$, and then used as a template for the selective PCR amplification to generate AFLPs.

The main DNA amplification was conducted using primers with three selective nucleotides (SNs) (Table 2). The reaction was carried out in a final volume of 20 µl containing 1x PCR buffer (same as above), 25 ng *Eco*RI primer + 3 SNs, 30 ng *Mse*I primer + 3 SNs, 0.2 M dNTPs, 1.5 mM $MgCl_2$, 0.5 U *Taq* DNA polymerase and 10 µl pre-amplified and diluted DNA sample. The main amplification was conducted using the following amplification profile: Initial 2 min denaturing at 94°C for 2 min and a final 2 min extension at 72°C, with an intervening 36 cycles. The first cycle was a denaturation step at 94°C for 30 s, an annealing step at 65°C for 30 s and primers extension step at 72°C for 60 s. The next 12 cycles emulated a touchdown PCR format decreasing the annealing temperature by 0.7°C each cycle to 56°C. The next 23 cycles were performed using the profile: 94°C for 30 s, 56°C for 30 s, and 72°C for 60 s. The amplified product was stored at 4°C until electrophoresis. GeneAmp® PCR system 9700 (Applied Biosystems) were used for incubation during restriction and ligation reactions and for preamplification and main amplification reactions.

The amplified product was denatured by adding 15 µl of formamide loading buffer (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. The denatured product was immediately placed on ice and then stored at -20°C until ready to be loaded. The AFLP electrophoresis was conducted using Model S_2 Sequencing Gel Electrophoresis Apparatus (Whatman Biometera®). Seven microliter of the amplification products were loaded on 5% (w/v) polyacrylamide gels with 1x TBE buffer and separated at 90 W constant power until xylene cyanol FF runs two-third of the length of the plate. Before loading the samples, the gel was pre-run at constant power of 90 W for 45 min to warm the gel to 50–60°C. DNA molecular weight marker V (8-587) (Roche Molecular Biochemicals, Germany) and 100 bp ladder (Amersham Biosciences AB) was loaded at flanking positions as molecular size standards and to facilitate data scoring within and among gels.

Following electrophoresis, DNA bands were visualized using silver staining technique according to Caetano-Anollés and Gresshoff (1994) with slight modifications. Gels were fixed in 10% acetic acid for 30 min and washed with distilled water three times (2 times for 4 min each and 1 time for 2 min). The fixed gel was impregnated with silver in the presence of 0.055% formaldehyde for 30 min, and then quickly rinsed in distilled water and transferred to developing solution (3% Na_2CO_3, 0.055% formaldehyde and 2 mg/l $Na_2S_2O_3$) until optimal

image contrast was obtained (2-3 min). Image development was stopped with the fixative (10% acetic acid), and the gel was washed with distilled water.

Table 2. *AFLP primer combinations and RAPD primers used for this study; and number of loci (NL) used to compare the Guizotia taxa, number of monomorphic loci (NML) and percent monomorphic loci (PML) for each AFLP primer combination and their corresponding mean/total*

AFLP				RAPD[a]		
Primer combination	NL	NML	PML	Primer	Base sequence	NL
*E-AAG/**M-CTC	111	10	0.09	OPA-11	5'-CAATCGCCGT-3'	28
*E-ACA/**M-CTA	95	4	0.04	OPA-14	5'-TCTGTGCTGG-3'	22
*E-ACA/**M-CTG	75	6	0.08	OPB-18	5'-CCACAGCAGT-3'	29
*E-ACG/**M-CTC	70	6	0.09	OPB-20	5'-GGACCCTTAC-3'	35
*E-ACT/**M-CAC	127	0	0.00	OPD.20	5'-ACCCGGTCAC-3'	42
*E-AGG/**M-CAT	90	4	0.04	OPF-5	5'-CCGAATTCCC-3'	49
*E-AGG/** M-CTA	90	0	0.00	OPF-10	5'-GGAAGCTTGG-3'	30
				OPG-2	5'-GGCACTGAGG-3'	42
				OPG-16	5'-AGCGTCCTCC-3'	44
				OPG-17	5'-ACGACCGACA-3'	32
Mean	94	4.3	0.05			35
Total	658	30	0.05			353

*E = *Eco*RI primer (5'-GACTGCGTACCAATTC-3'),
**M = *Mse*I primer (5'-GATGAGTCCTGAGTAA-3'). [a]all loci are polymorphic.

Data scoring and analysis

Each RAPD and AFLP band was considered as a single bi-allelic locus with one amplifiable and one null allele. Data were scored manually as 1 for the presence and 0 for the absence of a DNA band for each locus across the genotypes of each taxon. As a strategy to minimize scoring error both within and among gels, data scoring was performed twice, separately, and discrepancy was reconsidered and corrected. NTSYSpc program (Rohlf 2000) was used to calculate Nei's unbiased genetic distance (Nei 1978), Jaccard's similarity coefficient (Jaccard 1908) between each taxon and to calculate matrix correlation between the genetic similarities and their cophenetic value matrix and between similarity matrices generated from AFLPs and RAPDs. NTSYSpc program was also used for UPGMA (unweighted pair group method with arithmetic average) cluster analysis. Analysis of molecular variance (AMOVA) was done using Arlequin version 2 (Schneider et al. 2000) to calculate molecular variance and F_{ST} for each pair of taxa. FreeTree-Freeware program (Pavlicek et al. 1999) was used to generate neighbor joining tree of the genus *Guizotia* based on Ne's standard genetic distance (Nei 1972) and for bootstrap analysis.

Results and Discussion

AMOVA and genetic distance

Seven AFLP primer combinations and ten RAPD primers (Table 2) were used to examine the genetic relationship between *Guizotia* taxa and to suggest the taxonomic status of some of these taxa, as the information about genetic relationship among these cross compatible taxa has important implications for breeding programs. A total of 658 and 353 loci were scored from the seven AFLP primer combinations and the ten RAPD primers, respectively. The extent of molecular variance, genetic differentiation (F_{ST}) and genetic distance among each pair of *Guizotia* taxa, based on AFLPs and RAPDs, were calculated separately. The average molecular variance among the whole *Guizotia* taxa was also calculated for both AFLP and RAPD data.

In the case of AFLPs, the highest molecular variance (158.33), F_{ST} (0.95) and genetic distance (0.68) were recorded between *G. arborescens* and Chelelu, implying low genetic relationship between them (Tables 3, 4). The lowest values (46.83, 0.58 and 0.19, respectively as above) were recorded for *G. scabra* ssp. *scabra* vs. *G. villosa*, indicating that these two species are more related to one another as compared to the genetic relationship between any other pair of taxa included in this study (Tables 3, 4). The molecular variance and genetic distance calculated based on RAPD data ranged from 25.00 and 0.32 (between *G. scabra* ssp. *scabra* and *G. scabra* ssp. *schimperi*) to 51.50 and 0.54 (between *G. arborescens* and Chelelu), respectively (Tables 3, 4). The F_{ST} calculated based on RAPD data ranged from 0.43 (between *G. scabra* ssp. *scabra* and *G. scabra* ssp *schimperi*) to 0.77 (between *G. abyssinica* and *G. arborescens*) (Table 4). Thus, the range of all the three parameters generated from AFLPs was larger than that generated from RAPDs. For example, the extent of genetic distance among taxa revealed by AFLP were more than threefold, while that of RAPD was less than twofold. This implies that the discriminatory power of AFLP is better than that of RAPD, thus more suited to study the extent of genetic variation within and among *Guizotia* taxa. Other comparative studies involving RAPD and AFLP have shown that AFLP is a more efficient method to estimate genetic diversity because of their high reproducibility and high multiplex ratio (Powell et al. 1996, Russell et al. 1997, Pejic et al. 1998).

The mean molecular variance, F_{ST} and genetic distance among *Guizotia* taxa were 105.01, 0.81, and 0.44, respectively, for AFLP and 37.38, 0.59 and 0.42, in that order, for RAPD. The RAPD based mean genetic distance among taxa (0.42; see Table 4) was found to be more than twice of the overall mean of genetic distance within taxa among populations (0.20). Furthermore, considering all the taxa together, AMOVA revealed a highly significant genetic differentiation ($P < 0.0001$) between the taxa for both AFLPs and RAPDs (Table 3). Thus, both molecular marker techniques have proved a clear differentiation among the *Guizotia* taxa.

Table 3. *Degree of natural co-occurrence of Guizotia taxa (DNCO) and molecular variance among each pair of taxa as revealed based on AFLP and RAPD markers*

Pair of taxa	DNCO	AFLP		RAPD	
		VCAT[a]	TV[b]	VCAT[a]	TV[b]
G. arborescens vs. *G. abyssinica*	+	144.17	159.63	50.50	65.25
G. arborescens vs. *G. scabra* ssp. *scabra*	++	128.68	150.31	39.00	62.00
G. arborescens vs. *G. scabra* ssp. *schimperi*	+	145.23	161.69	41.00	61.75
G. arborescens vs. *G. villosa*	-	126.02	144.94	42.50	67.00
G. arborescens vs. *G. zavattarii*	-	127.31	134.81	44.50	60.75
G. arborescens vs. Chelelu	-	158.33	166.09	51.50	70.00
G. arborescens vs. Ketcha	-	147.66	162.76	45.50	64.25
G. abyssinica vs. *G. scabra* ssp. *scabra*	++	59.67	90.25	33.25	60.00
G. abyssinica vs. *G. scabra* ssp. *schimperi*	+++	55.83	81.25	30.00	54.50
G. abyssinica vs. *G.villosa*	++	74.94	102.81	43.50	71.75
G. abyssinica vs. *G. zavattarii*	-	132.67	149.13	43.75	63.75
G. abyssinica vs. Chelelu	+	106.10	124.61	34.50	56.75
G. abyssinica vs. Ketcha	-	82.93	108.78	38.00	60.50
G. scabra ssp. *scabra* vs. *G. scabra* ssp. *schimperi*	++	54.79	86.38	25.00	57.75
G. scabra ssp. *scabra* vs. *G. villosa*	+	46.83	80.88	36.25	72.75
G. scabra ssp. *scabra* vs. *G. zavattarii*	-	127.44	150.06	35.75	64.00
G. scabra ssp. *scabra* vs. Chelelu	-	101.40	127.31	33.50	64.00
G. scabra ssp. *scabra* vs. Ketcha	++	71.06	104.31	31.00	61.75
G. scabra ssp. *schimperi* vs. *G. villosa*	++	72.38	101.25	40.00	74.25
G. scabra ssp. *schimperi* vs. *G. zavattarii*	+	134.60	152.06	32.50	58.50
G. scabra ssp. *schimperi* vs. Chelelu	+	106.41	126.13	33.75	62.00
G. scabra ssp. *schimperi* vs. Ketcha	+	83.08	110.13	34.50	63.00
G. villosa vs. *G. zavattarii*	-	130.77	150.69	35.25	65.00
G. villosa vs. Chelelu	-	105.92	128.58	31.50	63.50
G. villosa vs. Ketcha	-	84.17	114.17	31.00	63.25
G. zavattarii vs. Chelelu	-	136.73	145.69	34.00	57.75
G. zavattarii vs. Ketcha	-	128.64	144.94	34.00	57.75
Chelelu vs. Ketcha	-	60.06	78.89	34.50	60.75
All species*		105.01	126.69	37.38	63.25

= no co-occurrence, + = low co-occurrence, ++ = medium co-occurrence, +++ = high co-occurrence. [a]Variance components among taxa. [b]Total variance. *genetic differentiation among the taxa is highly significant ($P < 0.0001$). **Note:** Molecular variance was calculated based on data from equal number of individuals from each taxon (Degrees of freedom are 1 and 19 for variance component among taxa and the total variance, respectively).

G. abyssinica versus other *Guizotia* taxa

G. abyssinica and *G. scabra* ssp. *schimperi* possess strong morphological resemblance to each other (Baagoe 1974). High percentage of crossability and genome homology between these taxa and high mean pollen fertility of their hybrid were evident (Hiremath and Murthy 1992, Murthy et al. 1993, Dagne 1994), which indicate the close relationship between them. Based on the existing evidences, several authors suggested that *G. abyssinica* might have originated from *G. scabra* ssp. *schimperi* through selection and cultivation (Baagoe 1974, Hiremath and Murthy 1988, Murthy et al. 1993, Dagne 1994). However, a firm exclusion of *G. scabra* ssp. *scabra* and *G. villosa* from being ancestor to *G. abyssinica* still demands more data (Bekele et al. 2006).

***Table 4**. Genetic differentiation (F_{ST}) and Nei's unbiased Genetic distance (Nei-UB) between pair of Guizotia taxa and their means generated based on (a) AFLPs (below the diagonal) and RAPDs (above the diagonal)*

			RAPD								
			Che	Aby	Arb	Sca	Sch	Vil	Zav	Ket	Mean
AFLP	Che	F_{ST}		0.61	0.74	0.52	0.54	0.50	0.59	0.57	0.58
		Nei-UB		0.39	0.54	0.44	0.43	0.44	0.40	0.42	0.41
	Aby	F_{ST}	0.85		0.77	0.55	0.55	0.61	0.69	0.63	0.63
		Nei-UB	0.43		0.52	0.38	0.33	0.53	0.48	0.43	0.42
	Arb	F_{ST}	0.95	0.90		0.63	0.66	0.63	0.73	0.71	0.70
		Nei-UB	0.68	0.62		0.45	0.45	0.50	0.46	0.49	0.47
	Sca	F_{ST}	0.80	0.66	0.86		0.43	0.50	0.56	0.50	0.53
		Nei-UB	0.42	0.24	0.55		0.32	0.53	0.45	0.38	0.40
	Sch	F_{ST}	0.84	0.69	0.90	0.63		0.54	0.56	0.55	0.55
		Nei-UB	0.43	0.22	0.63	0.22		0.53	0.40	0.44	0.40
	Vil	F_{ST}	0.82	0.73	0.87	0.58	0.71		0.54	0.49	0.54
		Nei-UB	0.44	0.30	0.53	0.19	0.29		0.46	0.43	0.46
	Zav	F_{ST}	0.94	0.89	0.94	0.85	0.89	0.87		0.63	0.61
		Nei-UB	0.56	0.56	0.51	0.55	0.57	0.56		0.47	0.42
	Ket	F_{ST}	0.76	0.76	0.91	0.68	0.75	0.74	0.89		0.58
		Nei-UB	0.23	0.33	0.65	0.30	0.34	0.35	0.54		0.41
	Mean	F_{ST}	0.85	0.78	0.90	0.72	0.77	0.76	0.90	0.78	
		Nei-UB	0.46	0.39	0.60	0.35	0.39	0.38	0.55	0.39	

Che = Chelelu, Aby = *G. abyssinica*, Arb = *G. arborescens*, Sca = *G. scabra* ssp. *scabra*, Sch = *G. scabra* ssp. *schimperi*, Vil = *G. villosa*, Zav = *G. zavattarii*, Ket = Ketcha.

We evaluated the genetic relationship between *G. abyssinica* and other guizotias based on molecular variance, F_{ST} and Nei's unbiased genetic distance. AFLP-based molecular variance, F_{ST} and genetic distance among *G. abyssinica* and *G. scabra* ssp. *schimperi* were 55.83, 0.69 and 0.22, respectively. On the other hand, RAPD-based analyses have shown that the values of these parameters were 30, 0.55 and 0.33, respectively. These values, both in the case of AFLPs and RAPDs, are the lowest when compared to corresponding values revealed between *G. abyssinica* and any other taxon, except that the AFLP-based F_{ST} between *G. abyssinica* and *G. scabra* ssp. *scabra* was lower (0.66). Therefore, both AFLPs and RAPDs revealed that the most closely related taxon to *G. abyssinica* is *G. scabra* ssp. *schimperi*. This finding is in complete agreement with most of the previous studies that suggested the likely evolution of *G. abyssinica* from *G. scabra* ssp. *schimperi* (see Figures 2, 3). Introgression is a likely event between these taxa due to the fact that *G. scabra* ssp. *schimperi* occurs as a weed in *G. abyssinica* fields, which might also have contributed to the higher genetic similarity between them.

Although there are some discrepancies between AFLPs and RAPDs, *G. scabra* ssp. *scabra* was generally the second most closely related taxa to *G. abyssinica*, with a molecular variance of 59.67 and 33.25, F_{ST} of 0.66 and 0.55, and genetic distance of 0.24 and 0.38 as generated from AFLP and RAPD data, respectively (Tables 3, 4). Dagne (1994) previously demonstrated that *G. abyssinica* and *G. scabra* ssp. *scabra* share high degree of homology between their chromosome complements, and sexual hybrids can be easily made between them. In addition, Murthy et al. (1993) indicated that *G. abyssinica* and *G. scabra* ssp. *schimperi* are

almost equivalent in their relationship to *G. scabra* ssp. *scabra*. These evidences indicate the close relationship between these three taxa, which is strongly supported by our finding (Tables 3, 4).

G. scabra ssp. *scabra* versus *G. scabra* ssp. *schimperi*

The molecular variance, genetic differentiation (F_{ST}), and genetic distance generated both from AFLPs and RAPDs have shown that *G. scabra* ssp. *scabra* and *G. scabra* ssp. *schimperi* are more closely related to each other as compared to most of the pairs of taxa compared in this study (Tables 3, 4). For example, AFLP based molecular variance and genetic distance among them were 54.79 and 0.22, respectively, which were the second smallest molecular variance and genetic distance observed in the whole data set, the smallest being 46.83 and 0.19, in that order, recorded for *G. scabra* ssp. *scabra* vs. *G. villosa*. Furthermore, RAPDs revealed that the molecular variance and genetic distance among these taxa were 25.00 and 0.32, respectively, which were the smallest corresponding values in the whole data set (Tables 3, 4). The second smallest molecular variance (30.00) and genetic distance (0.33) were recorded for *G. abyssinica* vs. *G. scabra* ssp. *schimperi*. According to Dagne (1994), the crossability between *G. scabra* ssp. *scabra* and *G. villosa* (82%) was higher than the crossability between *G. scabra* ssp. *scabra* and *G. scabra* ssp. *schimperi* (62%). Dagne (1994) further indicated that the pollen fertility of hybrids of *G. abyssinica* vs. *G. scabra* ssp. *schimperi* was much higher (81%) than that of *G. scabra* ssp. *scabra* vs. *G. scabra* ssp. *schimperi* (49%). Thus, our molecular markers based study is in good agreement with these findings.

The implication of taxonomic classification of *G. scabra* ssp. *scabra* and *G. scabra* ssp. *schimperi* as a subspecies of *G. scabra* by Baagoe (1974) is that they are more closely related to one another than to the other *Guizotia* species; and also that their similarity is more than the similarity between any other recognized taxa within the genus *Guizotia*. However, as we have already discussed above, our data and the report of Dagne (1994) showed that some pair of taxa are more similar or almost equally similar to one another as compared to the similarity between *G. scabra* ssp. *schimperi* and *G. scabra* ssp. *scabra*. Thus, these taxa are best viewed as separate species rather than as the subspecies of *G. scabra*. Hiremath and Murthy (1992) and Dagne (1995), based on the significant differences between the karyotypes of these taxa, also indicated that they should be considered as separate species. Morphologically, *G. scabra* ssp. *scabra* is distinguished from *G. scabra* ssp. *schimperi* by having (usually) more than 5 outer involucral leaves exceeding the disk centre, many-flowered heads, the perennial habit, and the usually more coarse general appearance (Baagøe, 1974). Generally, our finding and the results obtained from previous studies are strong evidences that suggest reconsideration of the taxonomic status of these taxa.

G. villosa

Based on morphological similarities, Baagoe (1974) suggested that *G. villosa* might have been derived from *G. scabra* (without mentioning the subspecies). In this study, AMOVA and genetic distance based on AFLP data revealed that the smallest molecular variance (46.83) and genetic distance (0.19) within the whole data set was for *G. villosa* vs. *G. scabra* ssp. *scabra* (Table 3, 4), implying that *G. villosa* is more closely related to *G. scabra* ssp. *scabra* than to any other taxa included in this study. Thus, these data support the grouping together of these taxa by Dagne (1995) and contradicts Hiremath and Murthy (1992) who speculated that *G. villosa* is evolved from *G. scabra* ssp. *schimperi*. On the other hand, RAPD revealed that *G. villosa* is more closely related to *G. zavattarii* as compared to other taxonomically recognized taxa (Table 3, 4). Therefore, the two marker systems are not in agreement regarding *G. villosa*. Nonetheless, both analyses agree in that *G. villosa* is more closely related to *G. scabra* ssp. *scabra* than to *G. scabra* ssp. *schimperi*.

Chelelu and Ketcha

The two molecular marker systems were not in agreement regarding the degree of genetic relationship between Chelelu and other *Guizotia* taxa. AFLPs showed that Chelelu is more closely related to *G. scabra* ssp. *scabra* (excluding Ketcha) with molecular variance of 101.40 (Table 3) and genetic distance of 0.42 (Table 4). AFLPs also revealed that the genetic distance between Chelelu and *G. abyssinica*; and Chelelu and *G. scabra* ssp. *schimperi* was equal (Table 4). On the other hand, RAPD data revealed different results when different parameters were used. RAPDs molecular variance (31.50) and F_{ST} (0.50) indicated that Chelelu is more closely related to *G. villosa*, while RAPD based genetic distance indicated that it is more closely but equally related to *G. abyssinica* and *G. zavattarii* (Table 4). The results from chromosome morphology and karyotype study (Dagne, 1995, 2001) indicated that Chelelu is more closely related to *G. abyssinica*. Thus, although there is no complete agreement between the AFLPs and RAPDs as well as the finding of Dagne (1995), when our data and the report of Dagne (1995) is taken into consideration, Chelelu seems to be more close to the group of taxa that contains *G. abyssinica*, *G. scabra* ssp. *schimperi* and *G. scabra* ssp. *scabra* than to the other taxonomically recognized taxa.

Morphologically, Ketcha seems to be more similar to *G. scabra* ssp. *scabra* than to any other *Guizotia* taxa. The current study suggested that Ketcha is more closely related to *G. scabra* ssp. *scabra* (excluding Chelelu) with molecular variance of 71.06 and 31.00 and genetic distance of 0.30 and 0.38 when calculated based on AFLPs and RAPDs, respectively (Tables 3, 4), which is in agreement with the work of Dagne (1995). RAPDs also showed that Ketcha is almost equally related to *G. scabra* ssp. *scabra* and *G. villosa*, which supports the grouping of Dagne (1995) as these three were grouped together. Thus, it could be generalized that the closest relative of Ketcha is *G. scabra* ssp. *scabra*.

AFLPs revealed that Chelelu and Ketcha are more closely related to one another than to any other taxa (Table 4) which was not in agreement with the result obtained from RAPD data and the report of Dagne (1995). Generally, all the analyses conducted based on both AFLPs and RAPDs suggest that these two populations are separate species as the molecular variance, F_{ST} and genetic distance between each of them and the other *Guizotia* taxa were equivalent to that revealed between taxonomically recognized species.

G. arborescens and *G. zavattarii*

Dagne (1995) grouped *G. arborescens* and *G. zavattarii* together based on karyotype similarity. The implication is that these two species are more closely related to one another than to the other *Guizotia* taxa. Additionally, based on crossing experiments, Dagne (2001) reported that *G. arborescens* and *G. zavattarii* showed a high degree of genomic affinity and thus a high degree of homology as reflected by the high mean values of bivalents in the F_1 hybrid. The extent of molecular variance, F_{ST} and genetic distance between these species, however, were high as revealed by our AFLP analysis (Tables 3, 4), indicating that the degree of genetic relationship these species have between themselves and with the other taxa is low. In the case of RAPD, the genetic similarity between these taxa was also low when compared to the similarities between most of other taxa. However, in the case of RAPD, the genetic similarities of these two taxa to the other taxa are comparable with the genetic similarities between several pairs of taxa (Tables 3, 4). Generally, our finding showed that the level of genetic relatedness between these taxa was relatively low and thus does not agree with the grouping of Dagne (1995). A study that compares the homologous state of the various AFLP and RAPD banding pattern by considering the sequences of the diagnostic markers given in Table 5 might shade more light to resolve the above differences between chromosomal and AFLP/RAPD data sets.

Cluster Analysis

Our UPGMA method of cluster analysis was based on the Jaccard's similarity coefficient for all three types of data sets (AFLPs, RAPDs and the combined data of AFLPs and RAPDs). Jaccard-based cluster analysis was reported to generate phenograms that are similar to the already established phylogenetic trees (Landry and Lapointe 1996). In our study, the bootstrap support for the clusters generated based on this coefficient was relatively higher as compared to those that based on other genetic distance and similarity coefficients tested. Bootstrap analysis is recommended to assess the internal consistency of the data and the strength of support for suggested relationships (Rieseberg 1996), and thus done with 500 repetitions in our case. The bootstrap analysis based on AFLPs generated three different trees with the original tree (being the most frequent tree) generated 497 times, while that conducted based on the combined data generated only two different trees with the original tree (being the most frequent tree) generated 304 times. On the other hand, RAPD based bootstrap analysis generated 234 different

trees with the original tree and the most frequent tree generated 12 and 17 times, respectively, implying internal inconsistency of the RAPD data.

The Jaccard-based cluster analysis was tested for its goodness of fit to the corresponding similarity coefficients. The cophenetic correlation coefficients between the genetic similarity matrices and their cophenetic value matrices calculated based on AFLPs, RAPDs and the combined data were 0.96, 0.71 and 0.95, respectively. This indicates a very good fit of respective clusters to the Jaccard similarity data generated from AFLPs and combined data, and a poor fit of RAPD based clusters to its genetic similarity data, as described in Rohlf (2000). The low cophenetic correlation coefficient in the case of RAPD does not mean that clustering is not possible, but only indicates that some distortion might have occurred (Mohammadi and Prasanna 2003). The correlation between the similarity matrices of AFLP and RAPD was low (r = 0.37). Such low correlation coefficient between AFLPs and RAPDs was also reported by other authors (*e.g.* Powell et al. 1996, Pejic et al. 1998).

In all three cases of cluster analyses, all intraspecific populations were clearly clustered together with higher genetic similarity as compared to the interspecific genetic similarity (data not shown), indicating that these taxa are well-separated. It may also indicate the monophyletic nature of each taxon. The phenograms generated based on AFLPs and RAPDs share some features in common. Both marker systems showed that *G. abyssinica* and *G. scabra* ssp. *schimperi* are very close to one another; and *G. arborescens* and *G. zavattarii* are the most divergent taxa in the genus *Guizotia*.

The difference in phenogram generated based on AFLPs and RAPDs in our study would seem to result from sampling of different regions of the genome by these marker systems. Other factors such as reproducibility problems in RAPDs (Jones et al. 1997), although we did our best to avoid problems of reproducibility, and co-migration of some non-homologous fragments (Rieseberg 1996) might also have some contribution. For example, about 9% of co-migrating RAPD bands generated for three wild species of sunflower were found to be non-homologous (Rieseberg 1996). The correlation coefficient between matrices of AFLPs and combined data was 0.99, which is a highly significant positive correlation. Additionally, the topology of phenogram generated based on the combined data is exactly identical to that generated based on AFLPs (Figs. 2A, 2C), implying that AFLP data is more influential and stronger than RAPD data in the combined data cluster analysis.

Taking the mean Jaccard's similarity coefficient as a reference, the cluster analysis based on AFLPs and the combined data revealed two clusters. Cluster I comprises all the studied taxa except *G. arborescens* and *G. zavattarii*, while cluster II comprises *G. arborescens* and *G. zavattarii* (Figs. 2A, 2C). On the other hand, cluster analysis based on RAPDs revealed that *G. arborescens* and *G. zavattarii* remained solitary without being grouped, while the rest were grouped together into cluster I (Fig. 2B). Dagne (1995) grouped the *Guizotia* taxa into three groups based on karyotypic similarity. According to Dagne (1995), the first group comprises *G. abyssinica*, *G. scabra* ssp. *schimperi* and Chelelu; the second

group comprises *G. arborescens* and *G. zavattarii*; and third group comprises *G. scabra* ssp. *scabra*, *G. villosa* and Ketcha. This grouping was also supported by crossing experiments and meiotic behaviors of the hybrids (Dagne 2001). Group I and group III of Dagne (1995) were clustered together into Cluster I of our cluster analysis in all the three sets of data.

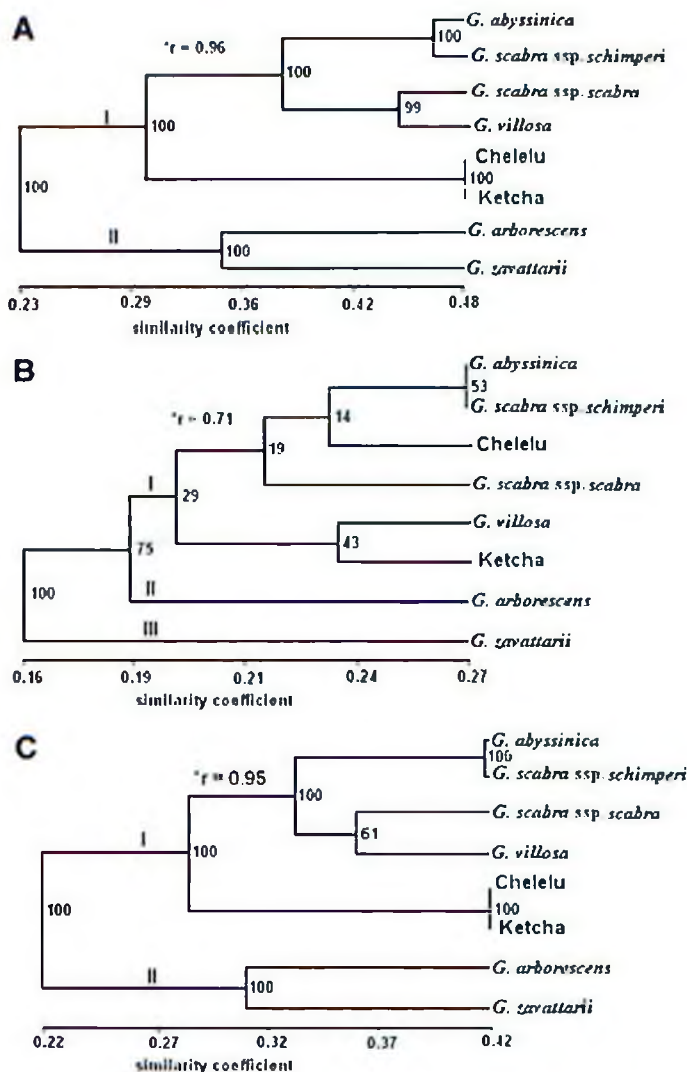


Figure 2. UPGMA Phenograms depicting genetic relationship between the eight *Guizotia* taxa based on Jaccard's similarity coefficients calculated from (A) AFLPs, (B) RAPDs and (C) combined data of AFLPs and RAPDs; with mean Jaccard's similarity coefficient of 0.30, 0.19 and 0.28, respectively. Numbers near branches are bootstrap values generated by 500 replications. *Correlation coefficient for Jaccard's similarity coefficient matrix vs. its cophenetic value matrix.

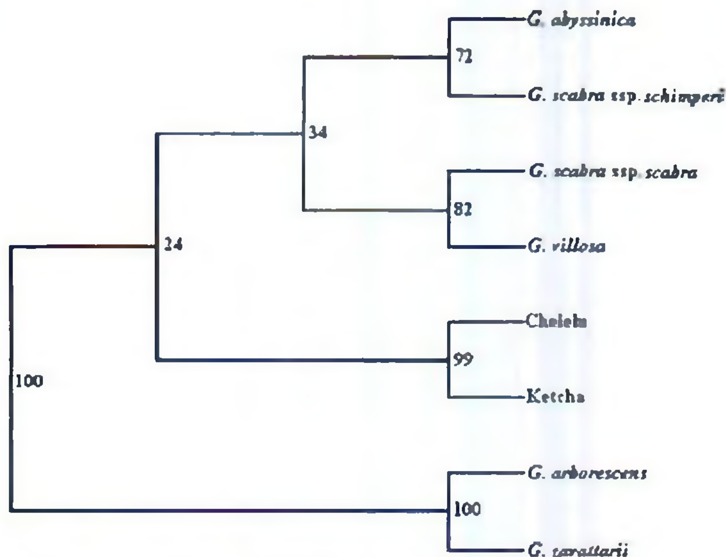


Figure 3. Neighbor-joining tree of eight *Guizotia* taxa generated based on Nei's standard genetic distance (Nei 1972) from combined data of AFLPs and RAPDs. Numbers near branches are bootstrap values generated by 500 replications.

AFLPs and combined data based clusters showed a sub-clustering of *G. abyssinica*, *G. scabra* ssp. *schimperi*, *G. scabra* ssp. *scabra*, and *G. villosa*, which is in agreement with Dagne (1994) who reported that these taxa are cross compatible; and that sexual hybrids could be easily produced among them. Group II of Dagne (1995) remained as a separate group (cluster II) in the case of AFLPs and combined data although they were not in the same cluster in the case of RAPDs. In spite of being clustered together in the case of AFLPs and combined data, the genetic similarity between *G. arborescens* and *G. zavattarii* and their genetic similarities to the other taxa were low and they are the most divergent species. Neighbor-joining tree was also constructed based on the combined data set of AFLPs and RAPDs to compare it with the UPGMA phenogram. The two cluster analysis generated similar phenograms except that the bootstrap support for neighbour-joining tree at some branches was lower than that of UPGMA phenogram (Fig. 3).

Diagnostic Markers

Several putative taxon specific markers were generated both from AFLPs and RAPDs. However, some of them are difficult to identify as they migrated almost the same distance with other bands and thus were not considered. Only those markers that were unambiguously and clearly separated from other bands were considered for this purpose. PCR-based genetic markers are now well-documented for species/cultivar identification (*e.g.* Raina et al. 2001, Johnson et al. 2003). The identification of taxon specific DNA markers in *Guizotia* would be of immense importance for conservation of these taxa and, most importantly, to improve *G.*

abyssinica through hybridization and introgression with the wild and/or weedy relatives.

Most putative taxon specific markers generated both from AFLPs and RAPDs were specific to *G. arborescens* or *G. zavattarii* (Table 5; Figs. 5, 6), implying that these two species are genetically distant from the other taxa. This is strongly supported by AMOVA and genetic distance analysis, which revealed that *G. arborescens* is most distantly related to the rest of the taxa with a mean F_{ST} and genetic distance of 0.90 and 0.60 (AFLP) and 0.70 and 0.47 (RAPD), respectively (Tables 3, 4). Similarly, *G. zavattarii* seems to be the second most distantly related taxon to the other taxa, although there is no complete agreement between the two marker systems (Tables 3, 4).

The RAPD marker with an approximate molecular size of 1630 bp and designated as OPB-18-1630 was generated by primer OPB-18. This marker was totally absent in *G. arborescens* and *G. zavattarii* but present in all other taxa at different frequencies, which makes it a putative marker to identify these two species from the rest. Though the previous crossing attempt resulted in only shrivelled seeds that failed to germinate (Dagne 1994), there is a possibility to produce hybrids between *G. abyssinica and G. zavattarii*. *G. zavattarii* can be a potential candidate as a source of some desirable traits to improve *G. abyssinica*, provided that germination problem of the hybrid seeds can be overcome. For example, the oleic acid content of *G. zavattarii* oil was reported to be about threefold higher than that of *G. abyssinica* (Dagne and Jonsson 1997). Therefore, this marker might be useful to screen the hybrid between these taxa starting from second (F_2) generation during breeding for niger varieties with high oleic acid and other desirable traits.

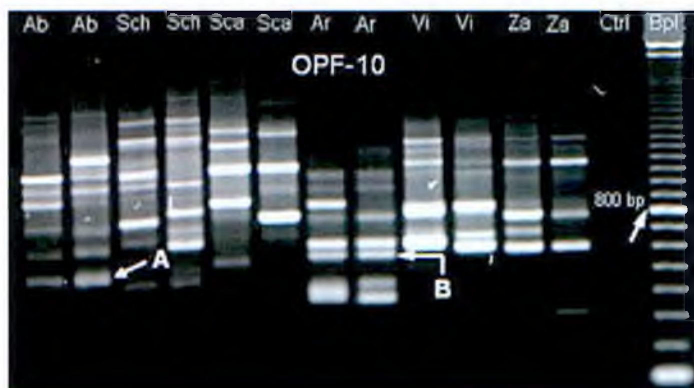


Figure 4. RAPD banding pattern amplified by primer OPF-10 showing diagnostic markers for *G. abyssinica* (arrow A) and *G. arborescens* (arrow B). Ab = *G. abyssinica*, Sch = *G. scabra* ssp. *schimperi*, Sca= *G. scabra* ssp. *scabra*, Ar = *G. arborescens*, Vi = *G. villosa*, Za = *G. zavattarii*, Ctrl = Control, Bpl = base pair ladder.

The RAPD marker designated OPF-10-510 was generated by primer OPF-10 and has a frequency of 97% in *G. abyssinica* but totally absent in *G. scabra* ssp. *schimperi* (Fig. 4). In addition, the marker was also recorded in *G. scabra* ssp. *scabra*, *G. villosa* and Ketcha at different frequencies. Therefore, it could be used

as a diagnostic marker for hybridization especially between *G. scabra* ssp. *schimperi* and *G. abyssinica*, and introgression of the marker to *G. scabra* ssp. *schimperi* together with other regions of the genome. OPF-10-590 RAPD marker was monomorphic in *G. arborescens* but totally absent in all other taxa (Fig. 4). Similarly, OPG-2-550 RAPD marker was found in *G. zavattarii* with 97% frequency but totally absent in all other taxa. Therefore, these two markers are potential candidates as species specific markers for *G. arborescens* and *G. zavattarii*, respectively (Table 5).

Table 5. *Diagnostic markers revealed by AFLPs and RAPDs and their estimated molecular size (EMS) in base pairs and description*

RAPD primer/ AFLP primer combination	Marker code	EMS	Description
OPB-18	OPB-18-1630	1630	Absent in *G. arborescens* and *G. zavattarii*, monomorphic in others
OPF-10 (see Fig. 4)	OPF-10-510	510	97% frequency in *G. abyssinica* , absent in *G. scabra* ssp. *schimperi*
	OPF-10-590	590	Monomorphic in *G. arborescens*, absent in all others
OPG-2	OPG-2-550	550	97% frequency in *G. zavattarii*, absent in all others
E-AAG/M-CTC	E-AAG/M-CTC-arb	220	Monomorphic in *G. arborescens*, absent in others
	E-AAG/M-CTC-zav	160	Monomorphic in *G. zavattarii*, absent in others
E-ACA/M-CTG	E-ACA/M-CTG-zav-310	310	Monomorphic in *G. zavattarii*, absent in others
	E-ACA/M-CTG-zav-230	230	Monomorphic in *G. zavattarii*, absent in others
E-ACG/M-CTC	E-ACG/M-CTC-arb-350	350	Monomorphic in *G. arborescens*, absent in others
	E-ACG/M-CTC-arb-240	240	Monomorphic in *G. arborescens*, absent in others
	E-ACG/M-CTC-arb-220	220	Monomorphic in *G. arborescens*, absent in others
	E-ACG/M-CTC-zav	180	Monomorphic in *G. zavattarii*, absent in others
E-ACT/M-CAC (see Fig. 5)	E-ACT/M-CAC-sch	350	Monomorphic in *G. scabra* ssp. *schimperi*, absent in *G. abyssinica* and *G. villosa*
	E-ACT/M-CAC-zav-340	340	Monomorphic in *G. zavattarii*, absent in others
	E-ACT/M-CAC-arb-320	320	Monomorphic in *G. arborescens*, absent in others
	E-ACT/M-CAC-zav-310	315	Monomorphic in *G. zavattarii*, absent in others
	E-ACT/M-CAC-arb-300	295	Monomorphic in *G. arborescens*, absent in others
	E-ACT/M-CAC-chel	255	Monomorphic in Chelelu, absent in others
E-AGG/M-CTA (see Fig. 6)	E-AGG/CTA-zav-450	450	Monomorphic in *G. zavattarii*, absent in others
	E-AGG/CTA-arb-390	390	Monomorphic in *G. arborescens*, absent in others
	E-AGG/CTA-zav-290	290	Monomorphic in *G. zavattarii*, absent in others
	E-AGG/CTA-arb-280	280	Monomorphic in *G. arborescens*, absent in others
	E-AGG/CTA-arb-230	230	Monomorphic in *G. arborescens*, absent in others
	E-AGG/CTA-arb-210	210	Monomorphic in *G. zavattarii*, absent in others
E-AGG/M-CAT	E-AGG/CTA-sch-210	210	Monomorphic in *G. scabra* ssp. schimperi, absent in *G. abyssinica*
	E-AGG/CTA-sch-160	160	Monomorphic in *G. scabra* ssp. schimperi, absent in *G. abyssinica*
	E-AGG/CTA-sch-150	150	Monomorphic in *G. scabra* ssp. schimperi, absent in *G. abyssinica*
	E-AGG/CTA-sch-110	110	Monomorphic in *G. scabra* ssp. schimperi, absent in *G. abyssinica*

Five AFLP markers that are interesting in terms of hybridization and introgression between *G. abyssinica* and *G. scabra* ssp. *schimperi* were revealed in our study. One of these markers, E-ACT/M-CAC-sch, was amplified by primer combination E-ACT/M-CAC (Fig. 5; Table 5), while the others, E-AGG/M-CAT-sch-210, E-AGG/M-CAT-sch-160, E-AGG/M-CAT-sch-150, E-AGG/M-CAT-sch-110, were amplified by primer combination E-AGG/M-CAT (Table 5). These markers were monomorphic in *G. scabra* ssp. *schimperi* and totally absent in *G. abyssinica*. Thus, they are possible diagnostic markers of *G. scabra* ssp. *schimperi* to identify it from *G. abyssinica*. *G. abyssinica* co-occurs with *G. scabra* ssp. *schimperi* more frequently than any other pair of *Guizotia* taxa (Table 3) due to the fact that *G. scabra* ssp. *schimperi* is a common weed in *G. abyssinica* fields. Therefore, hybridization and introgression between these taxa is likely as they are cross compatible and share the same pollinators, and fertile hybrids could be easily produced (Dagne 1994). Thus, this marker could be used to determine the extent of natural hybridization and introgression between these taxa. These markers are especially interesting from breeding point of view. One way to improve *G. abyssinica* is by transferring desirable genes from wild and/or weedy relatives to *G. abyssinica* through hybridization and introgression. Therefore, these markers may help to identify the hybrid between the two taxa and thus could be used as a molecular tool in marker assisted selection (MAS).

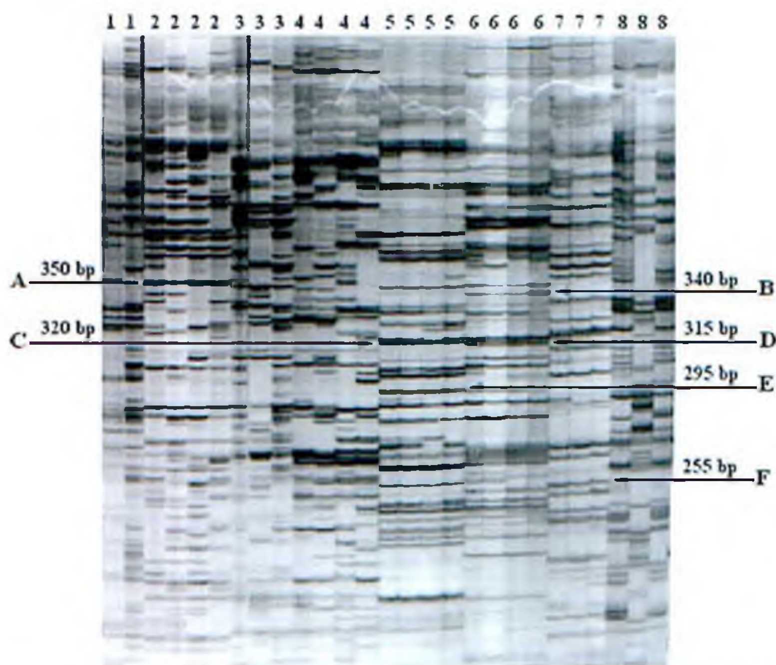


Figure 5. Banding pattern generated by AFLP primer combination *E-ACT/M-CAC* showing diagnostic markers for (i) *G. scabra* ssp. *schimperi* (arrow A), (ii) *G. zavattarii* (arrows B, D), (iii) *G. arborescens* (arrows C, E) and Chelelu (arrow F). (1) *G. abyssinica*, (2) *G. scabra* ssp. *schimperi*, (3) *G. scabra* ssp. *scabra*, (4) *G. villosa*, (5) *G. arborescens*, (6) *G. zavattarii*, (7) Chelelu, and (8) Ketcha.

To improve the efficiency of diagnostic markers revealed in this study, however, they need to be converted to Sequence Characterized Amplified Region (SCAR) markers (Paran and Michelmore 1993). In other words, these diagnostic markers need to be cloned and sequenced and then used to produce extended primers, as conversion of these markers into genetic probes could aid selection and identification of these taxa. SCAR markers are useful to plant breeding and genome analysis because of their simplicity, high reproducibility, low cost, and high efficiency as compared to polymorphic bands generated by AFLP and RAPD markers (Kelly and Miklas 1998, Boukar et al. 2004). For example, Boukar et al. (2004) identified an AFLP fragment from primer combination E-ACT/M-CAC linked to *Rsg*1, a gene that gives resistance to Nigerian strain of *Striga gesnerioides* (Willd.) Vatke. This AFLP fragment was cloned, sequenced, and converted into a co-dominant SCAR marker, which is useful in breeding programs (Boukar et al. 2004). SCAR markers also help to overcome the reproducibility problems associated with the RAPD technique (Paran and Michelmore 1993).

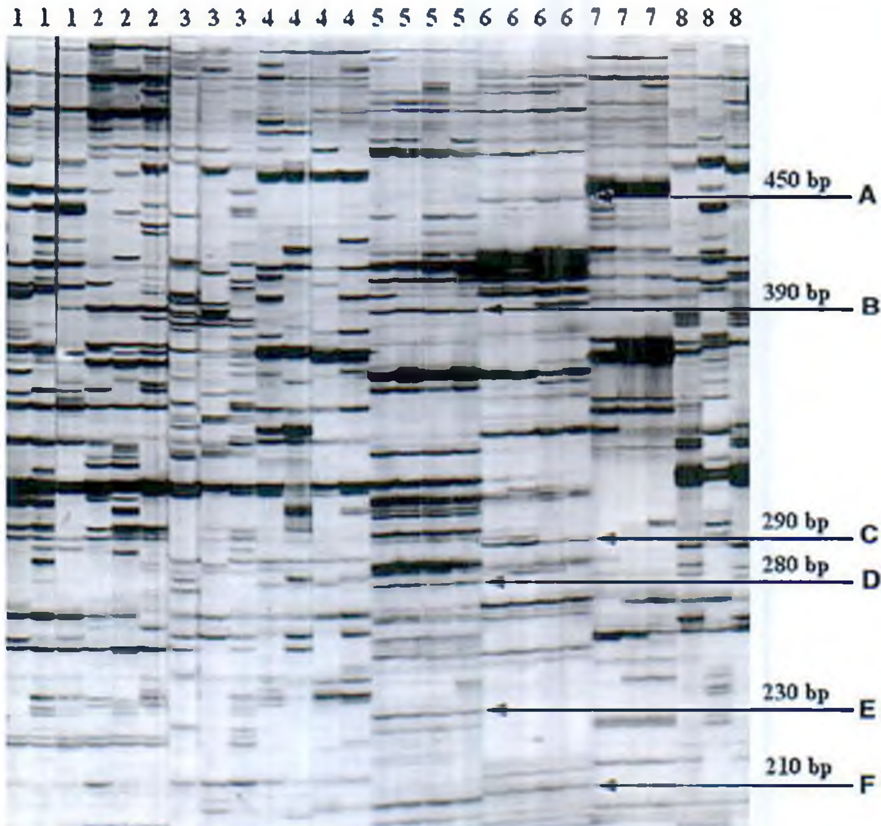


Figure 6. Banding pattern generated by AFLP primer combination *E-AGG/M-CTA* showing diagnostic markers for (i) *G. zavattarii* (arrows A, C, F), and (ii) *G. arborescens* (arrow B, D, E). (1) *G. abyssinica*, (2) *G. scabra* ssp. *schimperi*, (3) *G. scabra* ssp. *scabra*, (4) *G. villosa*, (5) *G. arborescens*, (6) *G. zavattarii*, (7) Chelelu, and (8) Ketcha.

Our analysis suggests that *G. scabra* ssp. *scabra*, *G. scabra* ssp. *schimperi*, Chelelu and Ketcha are separate species. In this study AFLP was found to be superior to RAPD in detecting genetic variation, in internal consistency of the data and the fitness of its clusters to the genetic similarity data and thus should be preferred. Nonetheless, in this study, RAPDs shared several common characteristics with AFLPs, which indicate that it can still be used as reliable markers. AFLP revealed genetic relationship between *Guizotia* taxa that was more inline with the cytogenetic and hybridization studies than that revealed by RAPDs and thus has a relatively strong phylogenetic signals.

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 in the laboratory work.

References

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

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

Bänfer G., Fiala B., Weising K. (2004) AFLP analysis of phylogenetic relationships among myrmecophytic species of Macaranga (Euphorbiaceae) and their allies. Pl. syst. Evol. 249: 213-231.

Bekele E., Geleta M., Dagne K., Jones A. L., Barnes I., Bradman N., Thomas M. G. (2006) Molecular phylogeny of genus *Guizotia* (Asteraceae) using DNA sequences derived from ITS. Genet. Resour. Crop Evol. DOI 10.1007/s10722-006-9126-0.

Blattner F. R., Weising K., Bänfer G., Maschwitz, U., Fiala, B. (2001) Molecular analysis of phylogenetic relationships among Myrmecophytic *Macaranga* Species (Euphorbiaceae). Mol. Phylogenet. Evol. 19: 331–344.

Boukar O., Kong L., Singh B. B., Murdock L. Ohm H. W. (2004) AFLP and AFLP-derived SCAR markers associated with *Striga gesnerioides* resistance in cowpea. Crop Sci. 44: 1259–1264.

Caetano-Anollés G., Gresshoff P. M. (1994). Staining nucleic acid with silver: an alternative to radioisotopic and fluorescent labling. Promega Notes Magazine 45: 13-20.

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). Pl. Syst. Evol. 195: 121-135.

Dagne K. (2001) Cytogenetics of new *Guizotia* Cass. (Compositae), interspecific hybrids pertaining to genomic and phylogenetic affinities. Pl. 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.

Després L., Gielly L., redoutet B., Taberlet P. (2003) Using AFLP to resolve phylogenetic relationships in a morphologically diversified plant species complex when nuclear and chloroplast sequences fail to reveal variability. Mol. Phylogenet. Evol. 27: 185-196.

Friis I. B. (1971) A new species of *Guizotia* (Compositae) from north east tropical Africa. J. Bot. 18: 231-234.

Geleta M., Bryngelsson T., Bekele E., Dagne K. (2006) Genetic diversity of *Guizotia abyssinica* (L. f.) Cass. (Asteraceae) from Ethiopia as revealed by Random Amplified Polymorphic DNA (RAPD). Genet. Resour. Crop Evol. DOI 10.1007/s10722-006-018-0

Getinet A., Sharma S. M. (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.

Gupta P. K., Varshney R. K., Sharma P. C., Ramesh B. (1999) Molecular markers and their applications in wheat breeding. Plant Breeding 118: 369-390.

Hansen M., Kraft T., Christansson M., Nilsson N. O. (1999) Evaluation of AFLP in *Beta*. Theor. Appl. Genet. 98: 845-852.

Hiremath S. C., Murthy H. N. (1986) The structure, stability, and meiotic behavior of B-chromosome in *Guizotia scabra* (Vis.) Chiov.ssp.*scabra* (Compositae). Caryologia 39: 397-402.

Hiremath S. C., Murthy H. N. (1988) Domestication of niger (*Guizotia abyssinica*). Euphytica 37: 225-228.

Hiremath S. C., Murthy H. N. (1992) Cytological studies in *Guizotia* (Asteraceae). Caryologia 45: 69-82.

Jaccard P. (1908) Nauvelles researches sur la distribution florale. Bull. Soc. Vaud. Sci. Nat. 44: 223-270.

Johnson E. L., Saunders J. A., Mischke S., Helling C. S., Emche S. D. (2003) Identification of *Erythroxulum* taxa by AFLP DNA analysis. Phytochemistry 64: 187–197.

Jones C. J., Edwards K. J., Castaglione S., Winfield M. O., Sala F., van de Weil C., Bredemeeijer G., Vosman B., Matthes M., Daly A., Brettschneider R., Bettini P., Buiatti M., Maestri E., Maleevschi A., Marmiroli N., Aert R., Volckaert G., Rueda R., Linacero R., Vazquez A., Karp A. (1997) Reproducibility testing of RAPD, AFLP and SSR markers in plants by a network of European laboratories. Mol. Breeding 3: 381-390.

Kadereit G., Kadereit J. W. (2005) Phylogenetic relationships, evolutionary origin, taxonomic status, and genetic structure of the endangered local lower elbe river (Germany) endemic *Oenanthe conioides* (Nolte ex Rchb.f.) Lange (Apiaceae): ITS and AFLP evidence. Flora 200: 15-29.

Kelly J. D., Miklas P. N. (1998) The role of RAPD markers in breeding for disease resistance in common bean. Mol. Breeding. 4: 1-11.

Landry P.-A., Lapointe F.-J. (1996) RAPD problems in phylogenetics. Zool. Ser. 25: 283-290.

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

Murthy H. N., Hiremath S. C., Salimath S. S. (1993) Origin, evolution and genome differentiation in *Guizotia abyssinica* and its wild species. Theor. Appl. Genet. 87: 587-592.

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

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

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

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

Paran I, Michelmore R. W. (1993) Development of reliable PCR-based markers linked to downy mildew resistance genes in lettuce. Theor. Appl. Genet. 85: 985–993.

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.

Pejic I., Ajmone-Marsan P., Morgante M., Kozumplick V., Castiglioni P., Taramino G., Motto M. (1998) Comparative analysis of genetic similarity among maize inbred lines detected by RFLPs, RAPDs, SSRs, and AFLPs. Theor. Appl. Genet. 97: 1248-1255.

Powell W, Morgante M, Andre C, Hanafey M, Vogel J., Tingey S., Rafalski A. (1996) The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. Mol. Breeding 2: 225-238.

Raina S. N., Rani V., Kojima T., Ogihara Y., Singh K. P., Devarumath R. M. (2001) RAPD and ISSR fingerprints as useful genetic markers for analysis of genetic diversity, varietal identification, and phylogenetic relationships in peanut (*Arachis hypogaea*) cultivars and wild species. Genome 44: 763–772.

Rieseberg L. H. (1996) Homology among RAPD fragments in interspecific comparisons. Molec. Ecol. 5: 99-105.

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

Russell R., Fuller J. D., Macaulay M., Hatz B. G., Jahoor A., Powell W., Waugh R. (1997) Direct comparison of levels of genetic variation among barley accessions detected by RFLPs, AFLPs, SSRs and RAPDs. Theor. Appl. Genet. 95: 714-722.

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.

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.

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

V

Genet Resour Crop Evol
DOI 10.1007/s10722-006-9126-0

RESEARCH PAPER

Molecular phylogeny of genus *Guizotia* (Asteraceae) using DNA sequences derived from ITS

Endashaw Bekele · Mulatu Geleta · Kifle Dagne · Abigail L. Jones · Ian Barnes · Neil Bradman · Mark G. Thomas

Received: 13 May 2006 / Accepted: 20 September 2006
© Springer Science+Business Media B.V. 2006

Abstract Complete sequences for the internal transcribed spacers of the 18s–26s nuclear ribosomal DNA were generated to establish phylogenetic relationships among five species of the genus *Guizotia*. Parsimony analysis and pairwise distance data produced a single tree with four clearly distinguished clades that accord with previously reported chromosomal data. The clades produced here have been discussed with reference to existing taxonomic treatments. It appears that *Guizotia scabra*, ssp. *scabra G. scabra* ssp. *schimperi* and *Guizotia villosa* have contributed to the origin of *Guizotia abyssinica*, the cultivated species of the genus. The present composition of the species of genus *Guizotia* and the subtribe the genus presently placed in are suggested to be redefined.

Keywords ITS · Phylogeny · *Guizotia*

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

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

A. L. Jones · I. Barnes · N. Bradman · M. G. Thomas
The Center for Genetic Anthropology (TCGA), University College of London, London, UK

Introduction

The genus *Guizotia* Cass. (nom. cons.) is a small but economically important genus that belongs to family Asteraceae, tribe Heliantheae and subtribe Coreopsidinae. Although Baagøe (1974) suggested transferring the genus to subtribe Verbesininae within the same tribe, Stuessy (1977) after revising the tribe Heliantheae maintained the genus in Coreopsidinae. On the other hand, Karis (1993) asserts that Robinson's (1981) placement of *Guizotia* in the Milleriinae seems more acceptable, even though the delimitation of this subtribe has to be emended. In addition to this, some African *Sigesbeckia* species treated by Humbles (1972) were transferred to genus *Guizotia* by Schulz (1990). According to Baagøe (1974), the distribution of *Guizotia* is typical of an afromontane, endemic genus. The genus is native to tropical Africa with most of the taxa restricted to East Africa, and with the highest concentration of species in Ethiopia.

Baagøe (1974) circumscribed the genus to six species viz.: *Guizotia abyssinica* (L. F) Cass., *Guizotia arborescens* I. Friis, *Guizotia jacksonii* (S. Moore) J. Baagøe *Guizotia scabra* (Vis.) Chiov. ssp. *scabra*, *G. scabra* (Vis.) Chiov. ssp. *schimperi* (Sch. Bip. in Walp.) J. Baagøe *Guizotia villosa* Sch. Bip. in Walp., and *Guizotia zavattarii* Lanza in Chiov. et al. Except *G. jacksonii* all species of *Guizotia* have been recorded in

Ethiopia. Baagøe (1974) gives an account on the endemism of *Guizotia* species and indicates that except for *G. scabra* ssp. *scabra* which has wider distribution, endemism is shown by *G. arborescens* (to the south west of Ethiopia and to the boarders of Sudan and Uganda), *G. villosa* (to the northern part of Ethiopian highlands), *G. scabra* ssp. *schimperi* as native to the Ethiopian highlands and *G. zavattarii* (to the Southern Ethiopia and North Kenya). *G. abyssinica*, *G. scabra* ssp. *schimperi* and *G. villosa* are annuals while the rest of the species of the genus are perennials.

Baagøe (1974), Hiremath and Murthy (1988) and Dagne and Heneen (1992) consider *G. scabra* ssp. *schimperi* to be the wild progenitor of *G. abyssinica*, the cultivated species. Hiremath and Murthy (1992) and Dagne (1995) concluded that *G. scabra* ssp. *scabra* and *G. scabra* ssp. *schimperi* reflect significant karyotypic differences, thus supporting the differences in morphology reported by Baagøe (1974). Hiremath et al. (1992) reported that *G. scabra* ssp. *scabra* and *G. scabra* ssp. *schimperi* differ distinctly in their genome size and based on this advised to treat these two taxa as independent species.

Baagøe also considered that *G. villosa* is derived from *G. scabra* and that *G. zavattarii*, derived from that of the same species or from an unknown ancestor common to the two. Later studies based on Karyology by Hiremath et al. (1992) suggested that *G. villosa* originated from species like *G. scabra* ssp. *schimperi*. The study on chromosome morphology of Dagne (1995) placed *G. villosa* to *G. scabra* ssp. *scabra* partly supporting Baagøe (1974) but still differing from her in the position and origin of *G. zavattarii*. Using chromosome morphology and giemsa C-banding Dagne and Heneen (1992) grouped the five *Guizotia* species into three:

Group I: *G. abyssinica* and *G. scabra* ssp. *schimperi*;
Group II: *G. zavattarii* and *G. arborescens*;
Group III:*G. villosa* and *G. scabra* ssp. *scabra*.

Baagøe (1974), using six presumed primitive and advanced morphological characters with in genus *Guizotia* concluded that *G. scabra* ssp. *scabra* has the highest number of primitive morphological characters while *G. arborescens* and *G. jacksonii* have advanced characters, the rest of the species were considered to be in the intermediate states.

Molecular techniques are being widely used in systematic and phylogenetic studies to provide a measure of genetic relatedness based on DNA sequences variation (Soltis et al. 1998). The internal transcribed spacers (ITS) of the nuclear ribosomal RNA genes have been among the most widely used sequences for DNA sequence variation studies. However, in spite of the small number of species in the genus *Guizotia*, there are no DNA sequences available for systematic purposes. The objective of this report is to make phylogenetic inferences using the information from ITS sequence generated for all species of *Guizotia* except for that of *G. jacksonii*. Here we report the results of comparison of their ITS sequences and relate this to the previous understanding and unresolved problems of the genus.

Materials and methods

Five *Guizotia* species, out of a total six (Baagøe 1974), were available for analysis from the samples collected (Table 1). The origin of these samples is given in Table 1. *Axiniphyllum durangense* B. L. Turner and *Sigesbeckia orientalis* L. were used as outgroups based on their ready availability of ITS sequences and their placement within the tribe Heliantheae. *Bidens setigera* (Sch. Bip.) Sherff, *Crepis aurea* (L.) Willd., *Helianthus annuus* L., *Senecio vulgaris* L. and *Vernonia noveboracensis* Cass. (L.) were also used as additional outgroups.

DNA was extracted from homogenized 300 mg of fine powder of leaf material in liquid nitrogen using 750 μl extraction buffer (pH 7.5) (consisting of 0.1 M Tris, 50 mM EDTA, 500 mM NaCl) and 100 μl of 10% SDS. The whole mix was incubated at 65°C for 20 min. Two hundred and fifty microliters of 5 M KAc was added to the mix and the samples were kept on ice for at least 30 min. The samples were later centrifuged at 14,000 rpm for 15 min and the supernatant was transferred to a new Eppendorf tube. To the supernatant equal volume of cold isopropanol was added and centrifuged for 10 min at 14,000 rpm. The supernatant

Table 1 *Guizotia* taxa used for ITS sequence analysis. Voucher specimens with their code are deposited in the National Herbarium of Addis Ababa University

	Species	Region	Altitude (m)	Location
1	*G. abyssinica*H15	Tigray	1,400	86 km from Shire to Adiarky
2	*G. abyssinica*J2	Illubabor	1,865	3 km from Metu to Gore
3	*G. arborescens*D170	Kelicho	2,110	14 km from Chida to Ameya
4	*G. arborescens*B200	Jimma	2,372	27 km from the Jimma-Addis Ababa road to Omo Nadda town
5	*G. arborescens*7	Kelicho	2,200	24 km from Chida to Ameya
6	*G. arborescens*8	Kelicho	2,200	24 km from Chida to Ameya
7	*G. scabra* ssp. *scabra*2	Sidamo	1,770	3.8 km from Kibre Mengist to Shakisso
8	*G. scabra* ssp. *schimperi*A23	Hararge	2,375	44 km from Kobo to Hirna
9	*G. villosa*A20	Tigray	2,410	102 km from Maychew to Mekele
10	*G. villosa*J15	Gondar	2,220	6 km from Gondar to Azezo
11	*G. zavattarii*K	Sidamo	1,820	2 km from Arero to Babila
12	*G. zavattarii*P	Sidamo	1,974	3 km from Yabelo to Konso

was poured off and the pellet was air-dried. Two hundred and fifty microliters of TE buffer (pH 7.6) consisting of 10 mM Tris and 1 mM EDTA was added to dissolve the pellet. After the pellet was dissolved, 250 μl of CTAB buffer (pH 7.5) composed of 0.2 M Tris, 50 mM EDTA, 2 M NaCl and 2% CTAB was added and the mix was incubated for 15 min at 65°C. Equal volume of chloroform was added to the mix and centrifuged for 5 min. The supernatant was recovered and chloroform extraction was repeated. The supernatant was taken into new Eppendorf tube and precipitated with equal volume of cold isopropanol and centrifuged for 15 min. The supernatant was poured off and the pellet was air-dried. One hundred microliters of TE buffer was added to dissolve the pellet overnight. Five microliters of RNase (1 mg/ml) was added the next morning and incubated at 37°C for 30 min. DNA quality and DNA concentration was determined using Shimadzu UV-240 Graphicord UV–visible light recording spectrophotometer.

The following primers were used:

18F 5′-GGAAGGAGAAGTCGTAACAAGG-3′
26R 5′-GCCGTTACTAAGGGAATCCTTGTTAG-3′.

Conditions for the thermal cycler were 95°C 4′, 95°C 30″, 54°C 30″, 72°C 1′ for 38 cycles. After 38 cycles, the PCR reactions were incubated at 72°C for 6 min. The PCR products were cleaned using Qiagen MinElute kit according to the manufacturer's instruction.

Automated sequencing of the purified double-stranded PCR products was carried out in both directions with sequencing conditions of 95°C 10 s, 54°C 1 s and 60°C 4 min for 25 cycles. The sequencing reaction was cleaned with 80 and 70% isopropanol at two different stages by centrifuging and the pellet was dried at 65°C for 5 min. To each of this pellet, 10 μl of high purity formamide was added and the two are mixed and incubated at 65°C for 5 min. The incubated samples were centrifuged at 1,000 rpm for 1 min. The samples were heated for 4 min at 96°C and then cooled on ice before it was loaded to 3700 genetic analyzer for sequencing.

Analysis of the ITS sequences were performed using PAUP 4.0 Beta 10 (Dave Swofford Sinauer Associates). Boundaries of the coding and spacer sequences were determined by comparison with published sequences in EMBL database. Sequences were aligned using Sequencher and manually edited. Forward and reverse sequences were compared, to check for consistency.

A pairwise distance value was calculated between taxa based on Kimura's 2-parameter correction method (Kimura 1980) with the estimated number of substitution per 1,000 bases. Parsimony analysis of the aligned sequences was

performed using PAUP 4.0 Beta 10. Trees were generated using the 'branch-and-bound' options. Bootstrap values were computed using PAUP. The consistency and homoplasy indices were calculated for all trees. A BLAST search was performed to search for homologous sequences of Asteraceae to compare with the sequences of *Guizotia* species reported here.

Results

Sequence analysis

Complete sequences of the ITS region were obtained for the five *Guizotia* taxa and the two outgroup species *S. orientalis* and *A. durangense* (GenBank accession number AF465890.1 and AF465846.1). The length of the entire ITS region was 645 alignment length of both the ingroup and *S. orientalis*, 644 for *G. villosa*A20 and 647 for *A. durangense*. The size of ITS1 is 263 bp while that of 5.8 s rRNA and ITS2 is 159 and 226, respectively. Forty parsimony informative characters (ingroups only) are present both in the whole fragment and ITS1 and ITS2 (excluding 5.8 s). There are additional four characters which are variable but not informative. Transition versus transversion ratios seem to be 1.3:1 excluding outgroup and 1.9:1 with the outgroup. These values are lower than the accepted 2:1 ratio required for Kimura's 2-parameter test when the outgroups are excluded. A total of four gaps were introduced into the alignment. Three gaps were introduced in the ingroup to improve fit with the outgroup. One gap was introduced into *S. orientalis* and *G. villosa*A20 to improve fit with the other taxa. The ratio of pyrimidines is slightly higher than purines (Table 2).

Sequence divergence comparison

Distances generated using Kimura's 2-parameter method are shown in Table 3. Divergence based on this method ranged from 0 to 5.5% between *Guizotia* species with the highest distance occurring between *G. villosa*J15 and *G. zavattarii*P. The divergence values between *Guizotia* species and outgroups vary from 2.7 to 7.2%. This shows that the sequences of both outgroups are not too different from those of *Guizotia*. Because of this they seem ideal as outgroup species due to the absence of hidden changes. Thus, exclusion of both outgroups can alter the topology of the trees produced. This is further noticed (Fig. 1) when seven species are used as outgroups. The neighbour-joining tree (Fig. 2), which shows branch lengths that are proportional to distances calculated with the method of Kimura (1980), indicates that *Guizotia* species are grouped into four groups and that there are two forms of *G. abyssinica* namely *G. abyssinica*J2 and *G. abyssinica*H15. The neighbour-joining tree (Fig. 3) produced similar results with that of Kimura's method except for the slight differences in the placement of the outgroups. On the other

Table 2 The proportions of the various bases and the number of ITS sites in the *Guizotia* species studied and the two outgroups considered

Taxon	A	C	G	T	No. of sites
*G. abyssinica*J2	0.22481	0.24651	0.27442	0.25426	645
G. scabra ssp. *schimperi*A23	0.22551	0.24496	0.27372	0.25581	645
*G. villosa*J15	0.22171	0.24883	0.27442	0.25505	645
*G. arborescens*B200	0.22016	0.25193	0.27287	0.25504	645
*G. arborescens*8	0.22016	0.25349	0.27287	0.25349	645
*G. arborescens*7	0.22016	0.25349	0.27287	0.25349	645
*G. arborescens*D170	0.22016	0.25271	0.27287	0.25426	645
G. scabra ssp. *scabra*2	0.22171	0.24883	0.26597	0.25350	645
*G. zavattarii*K	0.22171	0.25193	0.26822	0.25815	645
*G. zavattarii*P	0.22171	0.25116	0.26977	0.25736	645
G. orientalis—AF465890.1	0.21705	0.25581	0.27752	0.24961	645
*G. abyssinica*H15	0.22240	0.24492	0.27528	0.25740	645
A. durangense—AF465846.1	0.21638	0.26584	0.27666	0.24111	647
*G. villosa*A20	0.22205	0.24996	0.27566	0.25233	644
Mean	0.22112	0.25146	0.27379	0.25363	645.07

Table 3 Pairwise distances of ITS sequences in species of *Guizotia*, *Axiniphyllum durangense* and *Sigesbeckia orientalis*

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1														
2	0.0353													
3	0.0467	0.0270												
4	0.0467	0.0270	0.0000											
5	0.0450	0.0254	0.0000	0.0000										
6	0.0433	0.0254	0.0015	0.0016	0.0000									
7	**0.0720**	**0.0532**	0.0464	0.0466	0.4480	0.0448								
8	0.070	0.0482	0.0482	0.0483	0.0465	0.0465	0.0031							
9	0.0468	0.0253	0.0174	0.0173	0.0158	0.0158	**0.0550**	**0.0500**						
10	0.0467	0.0254	0.0142	0.0142	0.0126	0.0126	0.0515	**0.0499**	0.0031					
11	0.0636	0.0449	0.0383	0.0398	0.0383	0.0383	0.0351	0.0303	0.0382	0.0383				
12	0.0725	0.0533	**0.0500**	**0.0498**	**0.0484**	**0.0484**	0.0063	0.0032	**0.0515**	0.0400	**0.0383**			
13	0.0686	0.0499	0.0498	**0.0499**	0.0481	0.0481	0.0032	0.0000	**0.0516**	**0.0515**	0.0319	0.0032		
14	0.0654	0.0466	0.0417	0.0415	0.0400	0.0400	0.0319	0.0270	0.0399	0.0400	0.0078	0.0286	0.0286	

The highlighted values in the table are those with higher distance values in each column. Numbers *1–12* on first column/row are *Guizotia* taxa given in Table 1 in that order, while *13* and *14* are *S. orientalis* and *A. durangense*, respectively

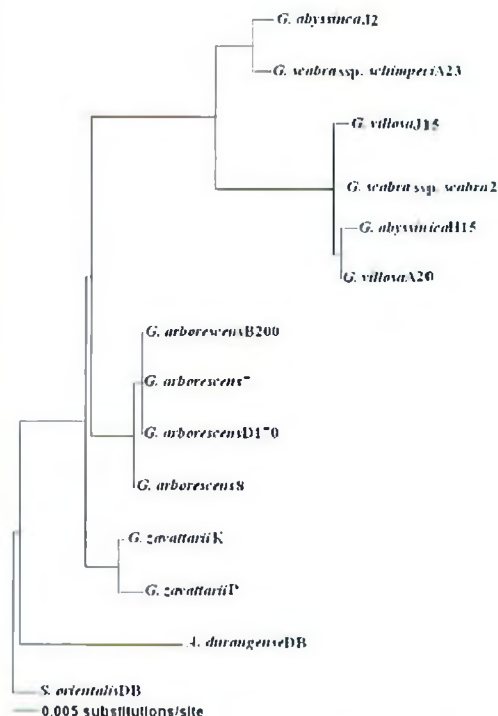


Fig. 1 Neighbour-joining tree obtained with ITS for five *Guizotia* species and five additional outgroups produced from pairwise distances calculated using Kimura's 2-parameter method

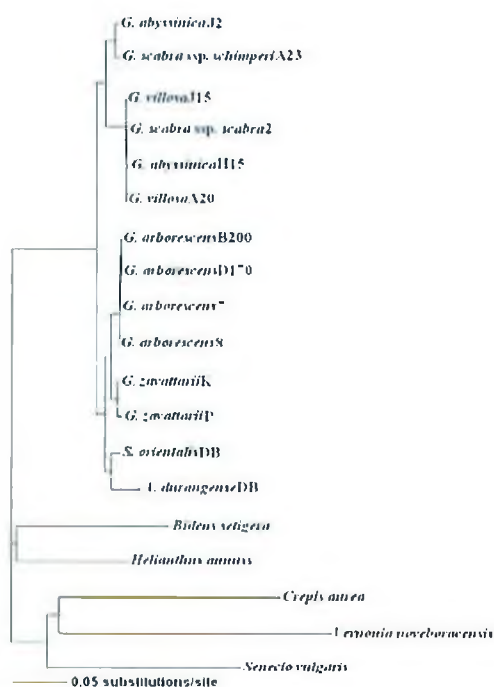


Fig. 2 Neighbour-joining tree obtained with ITS for five *Guizotia* species and two outgroups produced from pairwise distances calculated using Kimura's 2-parameter method

hand, the neighbour-joining tree (Fig. 4) constructed using ML-4KY-I-G resulted in similar clusters with that of the above two trees except for the slight differences in the placement of *G. abyssinica*H15 in the *G. villosa* and *G. scabra* ssp. *scabra* group.

Phylogenetic analysis of sequences

The 50% majority rule most parsimonious search performed on the entire ITS region produced one maximally parsimonious tree (Fig. 5). Bootstrapping showed that all the major branches were highly supported. The organization of taxa into clades is as follows: the two outgroups (*A. durangense* and *S. orientalis*) were separated from ingroups taxa in all trees, and the species of the genus *Guizotia*. Within the ingroups, four clades are distinguished. The first clade consists of *G. zavattarii*. The second clade consists of *G. arborescens*. The third clade consists of *G. villosa*, *G. scabra* ssp. *scabra* and *G. abyssinica*H15 (collected from Tigray). The fourth clade consists of *G. scabra* ssp. *schimperi* and *G. abyssinica*J2 (collected from Illubabor).

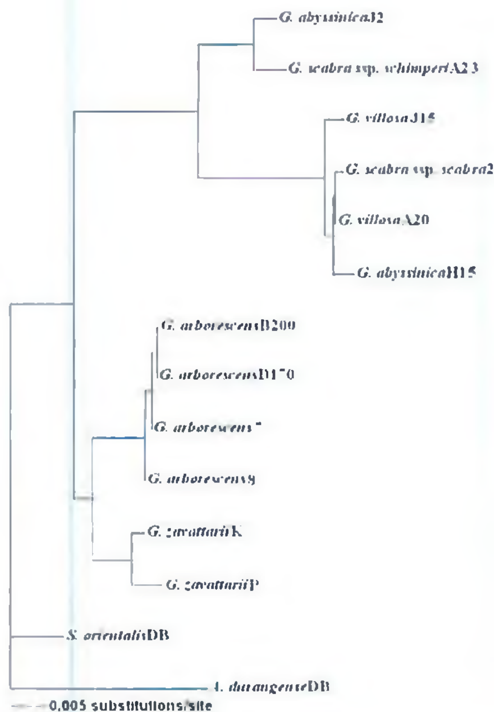


Fig. 3 Neighbour-joining tree obtained with ITS for five *Guizotia* species and two best outgroups produced from NJ-JC method

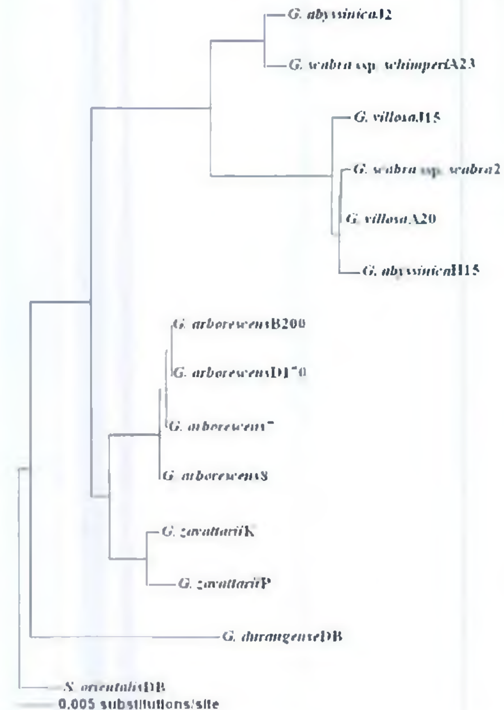


Fig. 4 Neighbour-joining tree obtained with ITS for five *Guizotia* species and two best outgroups produced from ML-Y-I-G method

Blast search to find homologous sequences

This search matched *G. abyssinica* sequences with those of other members of Asteraceae. Of these sequences, *G. abyssinica* was most similar to sequences generated from *S. orientalis*, *Sigesbeckia flosculosa* L'Her, *Sigesbeckia jorullensis* Kunth, *A. durangense*, *Trigonospermum annum* McVaugh et Lask., *Trigonospermum melampodioides* DC., *Milleria quinqueflora* L., *Coespeletia moritziana* (Sch. Bip. ex Wedd.) Cuatrec., *Rumfordia penninervis* S. F. Blake and *Espeletia pycnophylla* Cuatrec. On the other hand, the number of Asteraceae that showed similarity with *G. arborescens* and *G. zavattarii* is 27 and 23, respectively. The difference in number and extent of sequence similarities of species of various subtribes of tribe Heliantheae (Coreopsidinae, Milleriinae and Verbesininae) to various *Guizotia* species suggest that the present definition and

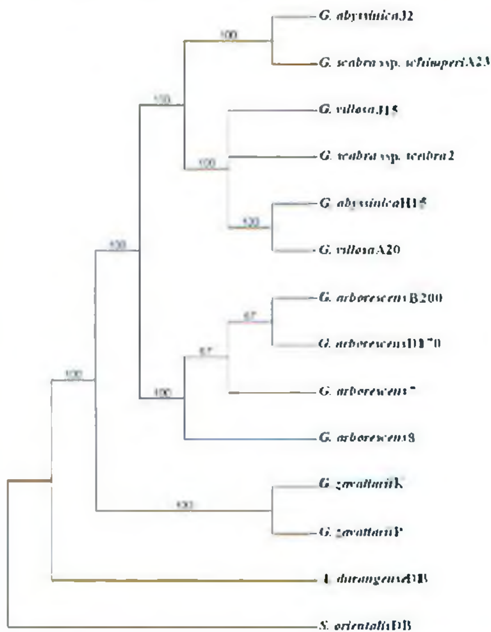


Fig. 5 Most parsimonious tree produced by the branch-and-bound search for the entire ITS region for five *Guizotia* species and two outgroups used. Bootstrap values are shown above the *line*

position of the genus *Guizotia* under subtribe Coreopsidinae has to be questioned, and need more data to resolve.

Discussion

A phylogeny tree has been constructed for genus *Guizotia* using ITS sequences. The total length of ITS region of *Guizotia* fall within the range reported for flowering plants (Baldwin et al. 1995). In all the species of *Guizotia* studied here, the length of ITS1 (159 bp) is higher than ITS2 (226 bp), a pattern which has also been found in other genera of Asteraceae (Baldwin 1992; Cerbah et al. 1998).

The most parsimonious tree produced separated the species of *Guizotia* into four clades. A truncated data set of ITS into ITS1 and ITS2 resulted retrieving the same topology (data not shown). The organization of taxa into clades produced by ITS is interestingly similar with that produced by Dagne and Heneen (1992) using chromosome morphology and Giemsa C-banding. The results of artificial inter-specific hybrids produced between *G. arborescens* and *G. zavattarii* through reciprocal crosses resulted in high percentage of seed setting and good pairing of chromosomes in the hybrids (Dagne 2001). Our ITS data and the clustering of *G. arborescens* and *G. zavattarii* corroborates with that of Dagne.

The advancement versus primitiveness of morphological characters suggested for species of genus *Guizotia* by Baagøe (1974) does not go with that of the ITS data clade formation and taxa organization.

Internal transcribed spacers data suggest that both *G. scabra* ssp. *scabra* and ssp. *schimperi* are the wild progenitor of *G. abyssinica*, the cultivated form. The cultivated *Guizotia* species collected from Northern Ethiopia seem to have evolved either from *G. scabra* ssp. *scabra* or *G. villosa* while that from the south might have evolved from *G. scabra* ssp. *schimperi*, thus suggesting that the significant karyotypic differences between *G. scabra* ssp. *scabra* and *G. scabra* ssp. *schimperi* and *G. villosa* reported by Dagne (2001) might exist between some of the cultivated forms. On the other hand, the karyotypes, C-banding and nucleolar numbers study made by Dagne (1995) agrees with the ITS data by placing *G. villosa* to *G. scabra* ssp. *scabra* than to *G. scabra* ssp. *schimperi* (see Figs. 3, 4, 5).

The diversity of plants, existence of complex home gardens and ethnic groups in Ethiopia has influenced the process of domestication (Bekele 1998). This is taken to promote diversification of many crops. Under such instances the two forms of *G. abyssinica* originating from two or even more wild progenitors of *Guizotia* species may take place.

The species of genus *Guizotia* seem to be natural. However, more data is required to suggest the ancestral species involved in the origin of the genus. The Ethiopian and Indian interface in the domestication processes is very interesting. For example, sesame (*Sesamum indicum* L.), which is the most ancient oil seed known and used by man (Purseglove 1976) is considered to have its primary centre of origin in Ethiopia or Sudan (Weiss 1971). It was taken at early date to India and eventually reached China. With

reference to *G. abyssinica* the case is clearer. Decandolle stated that the origin of cultivated plant could be found where it grows wild. This applies to the wild *Guizotia* species too. There is no need to think that these wild plants were brought from India as there are none. The presence of a wild progenitor of cultivated plants does not by itself bring about its domestication. For example, the wild African cotton [*Gossypium herbaceum* L. var. *africanum* (Watt) Mauer] which has remained in the wild state in South Africa has never been cultivated there at first. *Eragrostis pilosa* (L.) P. Beauv., which is a progenitor of *Eragrostis tef* (Zucc.) Trotter (Bekele and Lester 1981) is a cosmopolitan species but was domesticated only in Ethiopia. Since evolutionary events leading to man had occurred in and around Ethiopia, the hunter gatherer and the initial domestication stage must have been ancient in Ethiopia (Bekele 1998). The additional molecular data to probe such ancient period of domestication using model plants such as *Guizotia* species is highly desired.

Taxonomic considerations

So far there is no taxonomic treatment of Genus *Guizotia* into sections perhaps mainly because of the limited number of species in the genus. However, the distinct clades produced here imply such taxonomic delimitations. Dagne (2001) reported the successful production of artificial hybrids from all species within the genus although the chromosome behaviour of hybrids and degree of seed setting varies between the various species combinations used. Since species of *Guizotia* show self-incompatibility and are thus out breeders, the inter-specific hybrids production might also take place in nature provided the two species populations in the same locality flower at the same time.

Internal transcribed spacers data supports the group suggested by Dagne and Heneen (1992) and differ from that of Baagøe (1974) and Hiremath et al. (1992) on the origin of *G. villosa* and *G. zavattarii*. *G. villosa* is closer to *G. scabra* ssp. *scabra* than to *G. scabra* ssp. *schimperi*. *G. zavattarii* is quite different and distinct from all but closer to *G. arborescens*. The ITS data also supports the advice suggested by Hiremath et al. (1992) that the two subspecies of *G. scabra* be treated as independent species. In view of the unresolved sub-tribal definitions to accommodate *Guizotia*, more data is still required.

Acknowledgements The first author is very grateful for all the assistances provided to him by Centers for Genetic Anthropology of University College of London that kindly hosted the sequence study made. Part of the living cost for the present study was made by a grant from Sida/SAREC.

References

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

Baldwin B (1992) A phylogenetic utility of the internal transcribed spacers of nuclear ribosomal DNA in plants: an example from the Compositae. Mol Phylogenet Evol 1:3–16

Baldwin B, Snderson M, Porter J, Wojciechowski M, Campbell C, Donoghue M (1995) The ITS region of nuclear ribosomal DNA: a valuable source of evidence on angiosperm phylogeny. Ann Mol Bot Gard 82:247–277

Bekele E (1998) Historical and biological accounts of Ethiopia's unique position in the domestication and utilization of plant genetic resources: Ethiopia's gift to the world. Ethiopia's in broader perspective, vol. III. Nakanish Printing Co. Ltd, Kyoto, Japan, pp 775–789

Bekele E, Lester RN (1981) Biochemical assessment of the relationship of *Eragrostis tef* (Zucc.) Trotter with some wild *Eragrostis* species. Ann Bot 48:717–725

Cerbah M, Souza-Chies T, Jubier MF, Lejeune B, Siljak-Yakovlev S (1998) Molecular phylogeny of the Hypochaeris using internal transcribed spacers of nuclear rDNA: inference for Chromosoma Evolution. Mol Biol Evol 15:345–354

Dagne K (1995) Karyotypes, C-banding and nuclear 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, Heneen WK (1992) The karyotype and nucleoli of *Guizotia abyssinica* (Compositae). Hereditas 117:73–83

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

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

Hiremath SC, Murthy HN, Salimath SS (1992) Quantitative nuclear DNA differences associated with genome evolution in *Guizotia* (Compositae). Genetica 85:241–247

Humbles JE (1972) Observations of genus Seigesbeckia L. Ciene Naturaleza 13:2–19
Karis PO (1993) Heliantheae sensu lato (Asteraceae) Clades and classification. Plant Syst Evol 188:139–195
Kimura M (1980) A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J Mol Evol 16:111–120
Purseglove JW (1976) The origin and migration of crops in Tropical Africa. In: Harlan J, De Wet JMJ, Stemler ABI (eds) Origin of African plant domestication. Mountain Publishers, The Hague, Paris, pp 291–309
Robinson H (1981) A revision of the tribal and subtribal limits of the Heliantheae (Asteraceae). Smithson Contrib Bot 51:1–102
Schulz DL (1990) Zur Kenntnis der Gattung Sigesbechia L. in Africa. Gleditschia 46:613–628
Soltis DE, Soltis PS, Doyle JJ (1998) Molecular systematics of plants II: DNA sequencing. Kluwer, Dordrecht
Stuessy TF (1977) Heliantheae—systematic review. In: Heywood VH, Harborne JB, Turner BL (eds) The biology and chemistry of the Compositae, vol. 1. Academic, London, pp 1106–1118
Weiss EA (1971) Castor, sesame and safflower. Leonard Hill, London

VI

Molecular Phylogeny of the genus *Guizotia* (Asteraceae) based on sequences derived from the *trn*T/*trn*L and *trn*L/*trn*F intergenic spacers, *trn*L and 3'*trn*K/*mat*K introns and *mat*K gene

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

[1]Department of Plant Protection Biology, Swedish University of Agricultural Sciences, Box 44, SE-230 53 Alnarp, Sweden
[2]Department of Biology, Addis Ababa University, P. O. Box 1176, Addis Ababa, Ethiopia

Abstract

Phylogenetic analyses of the genus *Guizotia* (Asteraceae) were undertaken based on molecular sequence data from five chloroplast DNA regions. The genus is of particular interest for its single cultivated species, *G. abyssinica*. Parsimony and distance based phylogenetic analyses of molecular sequence data from the *trn*T/*trn*L and *trn*L/*trn*F intergenic spacers, *trn*L and 3'*trn*K/*mat*K introns and *mat*K gene revealed five clades, of which only one comprises more than one species. Of the five regions sequenced, the *trn*T/*trn*L intergenic spacer was the most variable within the genus. This phylogenetic analysis revealed a close phylogenetic relationship between *G. abyssinica*, *G. scabra* ssp. *scabra*, *G. scabra* ssp. *schimperi* and *G. villosa* although it fails to reveal the closest relative to *G. abyssinica*. Our analyses suggest that "Chelelu" and "Ketcha" are independent species within the genus *Guizotia*. This study supports the placement of the genus *Guizotia* within the subtribe Milleriinae. Those perennial *Guizotia* taxa with a highly localized geographic distribution seem to have evolved first during the evolutionary history of the genus.

Keywords: 3'*trn*K/*mat*K, *Guizotia*, Phylogeny, *trn*L, *trn*L/*trn*F, *trn*T/*trn*L, *mat*K

Introduction

The genus *Guizotia* Cass. is a small but economically important genus that belongs to the family Asteraceae, tribe Heliantheae. This genus has been placed under different subtribes of the tribe Heliantheae by different authors. Bentham (1873) placed the genus under subtribe Coreopsidinae. After one century, Baagoe (1974) suggested transferring of the genus to subtribe Verbesininae although Stuessy (1977), after revising the tribe Heliantheae, maintained the genus within the Coreopsidinae. Later, Robinson (1981) placed the genus under the subtribe Milleriinae. The placement of the genus under subtribe Milleriinae was also asserted by Karis (1993) who suggested that delimitation of Milleriinae has to be emended. The transferring of an African *Sigesbeckia* species (*S. somalensis* S. Moore) that belongs to the subtribe Milleriinae (Humbles, 1972) to the genus *Guizotia* by Schulz (1990) indirectly supports the placement of the genus *Guizotia* under this subtribe. The taxonomic position of the genus *Guizotia* is also recently questioned by Bekele et al. (2006) based on blast search to find homologous sequences between *Guizotia* species and various species from subtribes Coreopsidinae, Milleriinae and Verbesininae. According to Baagøe (1974), the genus *Guizotia* is an afromontane endemic genus, native to tropical Africa with most of the taxa restricted to East Africa, and with the highest concentration of species in Ethiopia.

Baagøe (1974) circumscribed the genus to six species viz.: *G. abyssinica* (L. F) Cass., *G. arborescens* I. Friis, *G. jacksonii* (S. Moore) J. Baagoe comb. nov., *G. scabra* (Vis.) Chiov. ssp. *scabra*, *G. scabra* (Vis.) Chiov. ssp. *schimperi* (Sch. Bip. In Walp) J. Baagoe stat. nov., *G. villosa* Sch. Bip. in Walp, and *G. zavattarii* Lanza in Chiov. & al. After Baagøe's (1974) taxonomic revision, two new populations of *Guizotia* were discovered in Ethiopia by K. Dagne who called them "Chelelu" and "Ketcha" (Dagne, 1995). These populations are distinctly different from each other and do not exactly fit to any of the recognized taxa of the genus *Guizotia* (Dagne, 1995, 2001). Their taxonomic status has not been determined, although we treated them as separate "taxa" for the sake of simplicity.

All the taxa of *Guizotia* are wild and/or weedy except *G. abyssinica,* which is cultivated mainly for its edible oil, particularly in Ethiopia and in India (Baagoe, 1974; Murthy et al., 1993). *G. abyssinica*, *G. scabra* ssp. *schimperi* and *G. villosa* are annuals while the rest of the taxa are perennials. Baagøe (1974) gives an account on the distribution of various *Guizotia* taxa. Except *Guizotia scabra* ssp. *scabra*, which extends from East Africa to Cameron and the Nigerian highlands with a distributional gap in the Congolian rainforest, the other taxa show endemism. This endemism is shown by *G. arborescens* (south west of Ethiopia and around the boarders of Sudan and Uganda), *G. scabra* ssp. *schimperi* (native to the Ethiopian highlands), *G. jacksonii* (Aberdares, Mt. Kenya and Mt. Elogon in Kenya and Uganda), *G. villosa* (northern and northwestern part of the Ethiopian highlands) and *G. zavattarii* (southern Ethiopia and northern Kenya).

Several pieces of evidence presented by different researchers indicate that *G. abyssinica* might have originated from *G. scabra* ssp. *schimperi* through selection

and further cultivation (Baagoe, 1974; Hiremath and Murthy, 1988; Murthy et al., 1993; Dagne 1994, 1995, 2001). Hiremath and Murthy (1992) and Dagne (1995) concluded that *G. scabra* ssp. *scabra* and ssp. *schimperi* display significant karyotypic differences and thus supporting the differences in morphology reported by Baagoe (1974). Hiremath et al. (1992) reported that the two subspecies of *G. scabra* differ distinctly in their genome size and, based on this, advised to treat them as independent species. Baagoe (1974), using six presumed primitive and advanced morphological characters within the genus *Guizotia* concluded that *G. scabra* ssp. *scabra* has the highest level of primitive morphological characters while *G. arborescens* and *G. Jacksonii* have the highest level of advanced characters, the rest taxa being at intermediate states.

Comparative DNA sequencing has become a widespread tool for systematic and phylogenetic studies as it is relatively fast and convenient, and offers large data set of discrete characters. Chloroplast DNA (cpDNA) sequence variations are being widely used for systematics and for phylogenetic reconstruction at different taxonomic levels (*e.g.* Taberlet et al., 1991; Johnson and Soltis, 1994; Liang and Hilu, 1996; Hilu and Liang, 1997; Bayer et al., 2002; Shaw and Small, 2005; Crawford and Mort, 2005). The sequences of the *trn*T/*trn*L and *trn*L/*trn*F intergenic spacers are useful for evolutionary studies of related species and probably also for populations of the same species (Taberlet et al., 1991). Taberlet et al. (1991) signify that the *trn*L intron is probably less variable as compared to the two intergenic spacers, and suggest that it could be more useful for evolutionary studies at higher taxonomic levels. *mat*K is a maturase encoding chloroplast gene of approximately 1500 bp located in the intron of the transfer RNA gene coding for lysine (*trn*K; tRNALysUUU) (Johnson and Soltis, 1994). The *mat*K gene has evolved at a higher rate than several other genes that have been used in systematic studies (Johnson and Soltis, 1994; Liang and Hilu, 1996) and was suggested to be appropriate for constructing infrafamilial phylogenies. The 3'*trnK/mat*K portion of the *trn*K intron has been used for systematics at the infrageneric level (*e.g.* Wang et al., 1999; Schultheis, 2001; Winkworth et al., 2002) and was shown to be more variable than the *mat*K/5'*trn*K portion of this intron (Shaw et al., 2005).

In spite of the small number of species in the genus *Guizotia*, there are no DNA sequences available for systematic purposes except the recent data from the internal transcribed spacers (ITS) of the nuclear ribosomal RNA gene (Bekele et al., 2006). The objectives of our work are (i) to attempt to construct the phylogeny of the genus *Guizotia* using DNA sequence data from four non-coding chloroplast DNA regions, namely: the *trn*T/*trn*L intergenic spacer, the *trn*L intron, the *trn*L/*trn*F intergenic spacer and the 3'*trn*K/*mat*K portion of *trn*K intron, as well as the *mat*K gene, and (ii) to comment on the subtribal placement of the genus *Guizotia* within the tribe Heliantheae. Herein we report the results of a comparison of the chloroplast DNA sequences for all *Guizotia* taxa except for *G. jacksonii* and relate it to the previous understanding and unresolved problems of the genus.

Material and Methods

Plant material and DNA extraction

Five *Guizotia* species, out of a total of six (Baagoe, 1974), and two yet taxonomically unclassified *Guizotia* populations (Dagne, 1995) were used in this study. The germplasm of all samples used in this study were collected from various regions in Ethiopia. The origin of these samples is given in Table 1. Each taxon was represented by two to four samples. Seeds were grown in a greenhouse and fresh leaves from 15-30 days old plants were used for genomic DNA extraction. DNA was extracted by a modified CTAB procedure as described in Assefa et al. (2003). DNA quality and concentration was determined using NanoDrop® ND-1000 spectrophotometer (Saveen Werner, Sweden).

Table 1. *Samples of various Guizotia taxa used for phylogenetic analysis*

Taxon	Altitude	Place of collection
G. abyssinica-1	1972	9.5 km shire to Shiraro
G. abyssinica-2	2372	78 km from Addis Ababa to Woliso
G. abyssinica-3	2425	63 km from Ginir to Gasera
G. abyssinica-4	1890	35 km from Amanuel to Bure
G. arborescens-1	2200	4.5 km from from Ameya to chida
G. arborescens-2	2382	The hill south of Omo-Nadda town[a]
G. arborescens-3	2100	8 km from Chida to Ameya
Chelelu-1	2475	About 20 km from Addis Ababa to Sendafa[b]
Chelelu-2	2475	About 20 km from Addis Ababa to Sendafa[b]
Ketcha-1	2380	About 64 km from Bale-Goba to Delo Mena
Ketcha-2	2380	About 64 km from Bale-Goba to Delo Mena
G. scabra ssp. *scabra*-1	1900	53 km from Gore to Bure
G. scabra ssp. *scabra*-2	2192	11 km from Bonga to Menjo[c]
G. scabra ssp. *scabra*-3	2020	10 km from Sodo to Arba Minch
G. scabra ssp. *schimperi*-1	2570	1 km from AA to D/Berehan road[d]
G. scabra ssp. *schimperi*-2	1720	38 km from Jimma to Agaro
G. scabra ssp. *schimperi*-3	2200	44 km from Kobo to Hirna
G. villosa-1	1920	16 km from B/Dar to Mota
G. villosa-2	2410	102 km from Maychew to Mekele
G. villosa-3	2220	6 km from Gondar to Azezo
G. zavattarii-1	1974	3 km from Yabelo to Konso
G. zavattarii-2	1820	1.5 km from Arero to Babila
G. zavattarii-3	1870	1 km North of Mega town

[a]27 km towards south of AA-Jimma road; [b]Chelelu river; [c]East of Bonga town; [d]East of Sendafa town.

PCR and Sequencing

Target cpDNA regions (Table 2) were amplified using a GeneAMP PCR system 9700 thermocycler with the following temperature profiles: initial 3 min denaturing at 94°C and final 7 min extension at 72°C with the intervening 30 cycles of 1 min denaturing at 94°C, 1 min of primer annealing at 48°C and 2 min

of primer extension at 72°C. The amplified product was stored at 4°C until purification. The whole *trn*K intron including the *mat*K gene was amplified using primers *MG1* and *MG15* while the *trn*T/*trn*L intergenic spacer was amplified using primers *a (B48557)* and *b (A4929)*. The *trn*L intron and the *trn*L/*trn*F intergenic spacer were amplified as a single fragment using primers *c (B49317)* and *f (A50272)* (Table 2). The PCR products were purified by QIAquick PCR purification kit (Qiagen GmbH, Germany) using microcentrifuge according to the manufacturer's instruction.

We employed cycle sequencing using the ABI PRISM® BigDye® Terminator v3.1 cycle sequencing kit (Applied Biosystems). Cycle sequencing was performed in a final volume of 10 µl containing 1x BigDye sequencing buffer, 2 µl of BigDye RR-100 cycle sequencing mix, 5 pmol of sequencing primer and 50-100 ng of purified double-stranded PCR product. The automated sequencing was carried out with sequencing conditions of 95°C for 30 sec, 50°C for 15 sec and 60°C for 4 min for 25 cycles. The *mat*K gene was sequenced from the 5'end to near the 3'end using sequencing primer *1110R*, *1240R*, *1408F*, *1541R* and *1694F*, respectively (Bayer et al., 2002; Table 2). Primer *1110R* sequenced the 5'most portion of the *mat*K gene and the 3'*trn*K/*mat*K portion of the *trn*K intron. A complete sequence of the *trn*T/*trn*F intergenic spacer was obtained by sequencing both strands using primer *a* (*B48557*) and primer *b* (*A4929*). The whole length of the *trn*L intron and the *trn*L/*trn*F intergenic spacer was sequenced using primer *c* (*B49317*) and primer *e* (*B49873*) respectively (Table 2). The product of the sequencing reaction was precipitated using a mixture of 29 µl of 96% ethanol and 1 µl of 3 M sodium acetate (pH 5.2) and centrifuged at 13200 rpm for 30 min. The precipitate was washed with 150 µl of 70% ethanol; air dried and submitted to BM labbet (www.BMlabbet.se) for sequencing on ABI PRISM® 3100 genetic analyzer (Applied Biosystems). The nucleotide sequences of representative samples of each taxon were submitted to EMBL nucleotide sequences database, where their gene bank accession numbers were assigned (Table 3).

Sequence alignment

Sequences were handled and manipulated using BIOEDIT version 7.0.5 (Hall, 2005). The quality of the sequences was visually inspected using SEQUNCE SCANNER version 1.0 (Applied Biosystems). Sequence alignments were created using CLUSTAL X version 1.81 (Thompson et al., 1997) and manually edited.

Sequence data analysis

Phylogenetic analysis was carried out using PAUP* 4.0 Beta 10 (Swofford, 2000). Gaps created during sequence alignment were included in the phylogenetic analysis and treated as (1) missing data and (2) fifth character. In the latter case, gaps of more than one bp long and shared by two or more taxa were treated as a single event. Phylogenetic analysis of both data sets produced nearly identical tree topology under different optimality criteria. Thus, only results from the first scheme are presented in this paper. In this phylogenetic analysis, it was assumed

that the first codon positions of the *mat*K gene weighs twice the weight of all other characters in the data set and that transversions cost twice transitions.

Trees were constructed using the maximum parsimony optimality criterion. Heuristic searches were performed through random sequence addition with 100 replicates using various branch swapping and branch length optimization options. From the resulting most parsimonious trees, a bootstrap 50% majority rule consensus tree was constructed. Clade support was estimated using bootstrap values (500 searches with 10 random additions, MAXTREES 1000). Using distance (weighted least squares with power = 2) as an optimality criterion Neighbor-joining trees were also constructed based on the Kimura-2-parameter method (Kimura, 1980).

Outgroup and additional ingroup taxa Selection

Outgroup and additional ingroup taxa were selected on the basis of (1) their taxonomic position within the Asteraceae, (2) ready availability of their sequences in the nucleotide gene bank and (3) their extent of sequence divergence from *Guizotia* taxa. All sequences of the outgroup and additional ingroup taxa for cpDNA regions used in this study was retrieved from National Center for Biotechnology Information (NCBI) databases through nucleotide-nucleotide BLAST (blastn) search. The gene bank accession numbers of these sequences are given in Table 3. The subtribal nomenclature of the taxa included in this study is based on the treatment of the tribe Heliantheae by Robinson (1981). Except *Eupatorium cannabinum* L., which belongs to the tribe Eupatorieae, all other taxa belong to the tribe Heliantheae. *E. cannabinum*, *Otopappus epaleaceus* Hemsl. (subtribe Ecliptinae) and *Helianthus annuus* L. (subtribe Helianthinae) were used as outgroups for phylogenetic analysis of the genus *Guizotia* based on the combined data from all regions sequenced.

For analysis of the phylogenetic relationship between the genus *Guizotia* and other genera that belong to various subtribes in the tribe Heliantheae, six species were used as additional ingroup taxa. These species are: *Acmella radicans* (Jacq.) R. K. Jansen, *Milleria quinqueflora* L., *Smallanthus microcephalus* (Hieron) H. Rob., *Verbesina jacksonii* B. L. Turner, *O. epaleaceus* and *Wollastonia biflora* (L.) DC. The subtribal placement of these species by Robinson (1981) is indicated in Figure 4. For this latter analysis, where *Coreopsis petrophiloides* B. L. Rob & Greenm and *E. cannabinum* were used as outgroup taxa, only sequencing data from the two intergenic spacers were used.

Table 2. *PCR and sequencing primers used for amplification and sequencing of various regions of chloroplast DNA*

Regions of chloroplast DNA	Primer name*	Primer sequence	Used for
*trn*K intron including *mat*K gene	MG1[a]	5-'CTA CTG CAG AAC TAG TCG GAT GGA GTA GAT-3'	Amplification
	MG15[a]	5'-ATC TGG GTT GCT AAC TCA ATG-3'	Amplification
	1110R[b]	5'-TAT TCT GTT GAT ACA TTC G -3'	sequencing
	1240R[b]	5'-CAG ATG AGC TGG GTA AGG T-3'	sequencing
	1408F[b]**	5'-CCT ATA TAC TTC TTA TGT ATG-3'	sequencing
	1541R[b]	5'-GCT CCA GAA GAT GTT GAT CG-3'	sequencing
	1694F[b]	5'-CTT TTG ATG AAT AAN TGG-3'	sequencing
*trn*T/*trn*L intergenic spacer	*a* (B48557)[c]	5'-CAT TAC AAA TGC GAT GCT CT-3'	Amplification, sequencing
	b (A4929)[c]	5'-TCT ACC GAT TTC GCC ATA TC-3'	Amplification, sequencing
*trn*L intron and *trn*L/*trn*F intergenic spacer	*c* (B49317)[c]	5'-CGA AAT CGG TAG ACG CTA CG-3'	Amplification, sequencing
	e (B49873)[c]	5'-GGT TCA AGT CCC TCT ATC CC-3'	sequencing
	f (A50272)[c]	5'-ATT TGA ACT GGT GAC ACG AG-3'	Amplification

*The superscript [a, b] and [c] refer to primers original reference Liang and Hilu (1996), Bayer *et al.* (2002) and Taberlet *et al.* (1991), respectively.
**the 12th and 20th nucleotides (indicated in bold) are different from Bayer *et al.* (2002).

Table 3. *Nucleotide sequence database accession numbers for the representative samples of Guizotia species and species used as additional ingroup taxa and outgroups in this study for five cpDNA regions*

Taxon	Gene bank accession number				
	3'*trn*K/*mat*K intron	*mat*K gene	*trn*L intron	*trn*T/*trn*L spacer	*trn*L/*trn*F spacer
Guizotia abyssinica	AM411125	AM411125	AM411136	AM411144-5	AM411162
Guizotia arborescens	AM411126-7	AM411126-7	AM411137-8	AM411146	AM411163
Chelelu	AM411128	AM411128	AM411139	AM411153	AM411164
Ketcha	AM411129	AM411129	AM412649	AM411154-5	AM411165
Guizotia scabra ssp. *scabra*	AM411130	AM411130	AM411140	AM411147-9	AM411166
Guizotia scabra ssp. *schimperi*	AM411131-2	AM411131-2	AM411141	AM411150-2	AM411167
Guizotia villosa	AM411133-4	AM411133-4	AM411142	AM411156-8	AM411168
Guizotia zavattarii	AM411135	AM411135	AM411143	AM411159-61	AM411169-70
Acmella radicans[a]	-	-	-	AY215885.1	AY216137.1
Coreopsis petrophiloides[b]	-	-	-	AY215903.1	AY216155.1
Eupatorium cannabinum[b]	AB217695.1	AB217695.1	AB217695.1	AB217695.1	AB217695.1
Helianthus annuus[b]	DQ383815.1	DQ383815.1	DQ383815.1	DQ383815.1	DQ383815.1
Milleria quinqueflora[a]	-	-	-	AY215954.1	AY216205.1
Otopappus epaleaceus[ab]	AY297651.1	AY297651.1	AY297673.1	AY297662.1	AY297684.1
Smallanthus microcephalus[a]	-	-	-	AY215986.1	AY216237.1
Verbesina jacksonii[a]	-	-	-	AY216002.1	AY216253.1
Wollastonia biflora[a]	-	-	-	AY297664.1	AY297686.1

[a]used as ingroup taxa. [b]used as outgroups. [ab]used as both an outgroup and as additional ingroup taxon. Note: In cases when more than one accession numbers represent a taxon, the accession numbers are given in range (*e.g.* AM411169-70 represents AM411169 and AM411169-70).

Results

Sequence characteristics of the Genus *Guizotia*

Significant length variations between taxa of the genus *Guizotia* were detected for the entire *trn*T/*trn*L intergenic spacer, ranging from 582 (*G. zavattarii*) to 634 (Ketcha) nucleotides long. The sequence length variation was mainly due to a variation in the number of tandem repeats of "*TATAGAAGATGAAAGAAGATAGA*", which were four, three and two for Ketcha, *G. arborescens* and the rest of the taxa, respectively. Gaps accounted for 7.8% of the aligned length of this spacer (Table 4). Considering indels as fifth characters, the highest proportion of DNA sequence divergence was recorded for the *trn*T/*trn*L intergenic spacer (9.3%; between *G. scabra* ssp. *scabra* and Ketcha). The sequence divergence for the other regions was less than 1% (Table 4). The G+C content of the *trn*T/*trn*L intergenic spacer averaged 28.7%. The complete sequence of the *trn*L intron was 433 bp for all guizotias. Of the regions sequenced, the G+C content was the highest (35.5%) in *trn*L intron. The *trn*L/F intergenic spacer was found to be shorter than both the *trn*T/*trn*L spacer and the *trn*L intron (ranging from 345-347 bp long) and has an average G+C content of 35.2%. A 0.5% gap was created during the alignment of *trn*L/*trn*F intergenic spacer due to two additional nucleotides in one sample of *G. zavattarii*. No indels were evident for the *trn*L intron, the *mat*K gene and the 3'*trn*K/*mat*K portion of the *trn*K intron.

In this study, 1255 bp long 5'-most portion of the *mat*K gene (G+C = 33.4%) and 301 bp long 3'*trn*K/*mat*K portion of the *trn*K intron (G+C = 28.5%) were sequenced (Table 4). Four variable sites were revealed within the *mat*K coding region among *Guizotia* taxa, two of which resulting in synonymous and the other two in non-synonymous amino acid substitutions. Transition and transversion events were equally frequent in the case of the *mat*K gene (Table 4). One transversion in the *mat*K coding region resulted in an exchange of asparagine and lysine, while one transition resulted in an exchange of valine and isoleucine amino acids. The proportion of sequence divergence of the combined data ranged from 0.1–2.2% between the *Guizotia* taxa (Table 4). Within-taxon sequence variation was revealed, at least for one of the taxa studied, in each region except in the *mat*K gene. The transition/transversion ratio (ns:nv) ranged from 0.50 (*trn*L intron)-2.00 (3'*trn*K/*mat*K portion of *trn*K intron; Table 4). The mean ns:nv ratio for the whole data set within the genus *Guizotia* was 0.86. Including corresponding sequences from the three outgroups (*E. cannabinum*, *O. epaleaceus* and *H. annuus*) resulted in a decrease of this ratio to 0.47, implying that transversion is more frequent than transition for these regions of chloroplast DNA within the tribe Heliantheae.

Phylogenetic Reconstruction

In the case of the whole data set, a total of 2997 aligned length (including outgroups) of nucleotides were used for phylogenetic analysis, of which 2823 characters are constant. Out of the 174 variable characters, 45 characters are parsimony informative. Interestingly, the *mat*K coding region contributed the

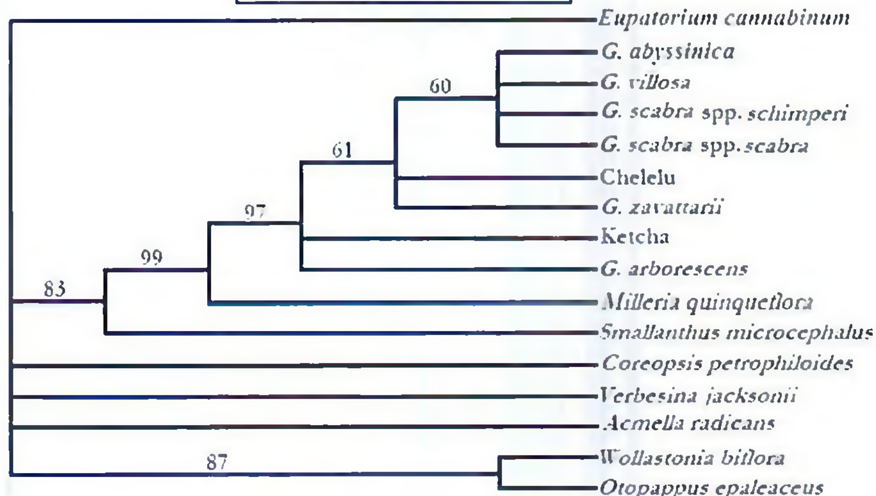


Figure 3. The 50% majority rule consensus tree of 227 equally parsimonious trees generated from phylogenetic analysis of DNA sequence data of the *trn*T/*trn*L and *trn*L/*trn*F intergenic spacers. *Eupatorium cannabinum* and *Coreopsis petrophiloides* were used as outgroups. Bootstrap values greater than 50% are indicated above the branches.

G. arborescens (clade B) and *G. zavattarii* (clade E) made their own separate clades, being positioned towards the base of the tree within the genus *Guizotia* (Figures 1 and 2). However, *G. zavattarii* appeared to be closer to the first clade than *G. arborescens* when distance was used as optimality criterion (Figure 2), which can also be observed in Figures 3 and 4. Based on their relatively large m type chromosomes, Dagne (1995) stated that these two species are more closely related to each other than to the other taxa. Our pair-wise distance calculated using the kimura-2-parameter method also revealed that they are more closely related to each other than to any of the taxa in the first clade (data not shown). The positioning of these species towards the base of a tree was also evident from the ITS data (Bekele et al., 2006). Thus, Baagoe's (1974) suggestion regarding the derivation of *G. zavattarii* from *G. scabra* seems unlikely. In this phylogenetic analysis it appeared that *G. arborescens*, *G. zavattarii*, Ketcha and Chelelu evolved earlier than the taxa under clade A during the evolutionary history of the genus. The conclusions of Baagøe (1974) to consider *G. arborescens* and *G. jacksonii* to have the highest level of advanced characters has to be questioned since it is not obvious as to how a given character set is termed as advanced, derived and primitive.

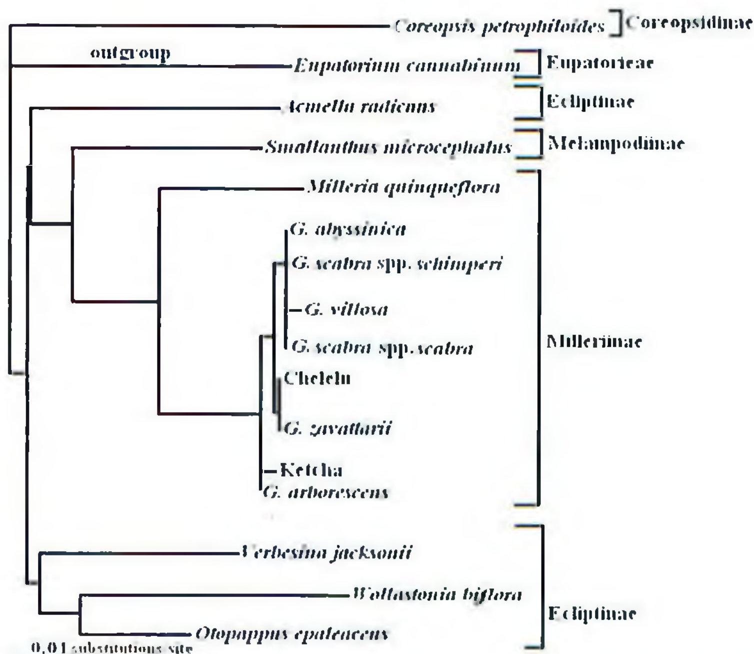


Figure 4. Neighbor-joining tree obtained with combined data from the *trn*T/*trn*L and *trn*L/*trn*F intergenic spacers for eight *Guizotia* taxa and seven additional species within the tribe Heliantheae produced from pair-wise distances calculated with the Kimura 2-parameter method. *E. cannabinum* and *C. petrophiloides* were used as outgroups.

The Chelelu and Ketcha populations were first reported to belong to the genus *Guizotia* by Dagne (1995). The study based on cross-compatibility and chromosome pairing of the hybrids between Chelelu and other recognized taxa also supports the placement of Chelelu in the genus *Guizotia* (Dagne, 2001). Our phylogenetic analysis confirms the placement of these populations in the genus *Guizotia*. In Figures 1 and 2, Chelelu and Ketcha appeared as if they are more closely related to each other, a pattern that was not obtained in Figures 3 and 4. The difference is due to the fact that Chelelu and Ketcha share complete sequence similarity for the *trn*L intron and the *mat*K gene, but showed some degree of sequence divergence in the other sequenced regions. The inclusion of data from the *trn*L intron and the *mat*K gene thus changed the tree topology. In spite of these differences, both Chelelu and Ketcha were clearly separated from the other taxa, which further strengthen the previous understanding that they do not fit to any of the recognized taxa.

Generally, the sequence divergence between Chelelu and Ketcha was found to be higher than the sequence divergence between any pair of taxa in the clade A. It is possible that these taxa share a common progenitor but have accumulated ample variation to the level that could make them separate species. The grouping of Chelelu

together with *G. abyssinica* and *G. scabra* ssp. *schimperi* and the grouping of Ketcha together with *G. villosa* and *G. scabra* ssp. *scabra* by Dagne (1995, 2001) based on chromosome morphology and meiotic behaviour of their hybrids were not observed in this analysis. All the data collected so far (both published and unpublished), including morphology, karyotype, crossing experiments, meiotic behaviour of hybrids, molecular marker data (RAPDs and AFLPs) and this particular study showed with certainty that Chelelu and Ketcha belong to the genus *Guizotia* and that they are distinct enough to be treated as separate species. Therefore, the next step should be assigning botanical names and publish them as new species.

The position of the genus *Guizotia* within the tribe Heliantheae

The genus *Guizotia* was placed under different subtribes (Coreopsidinae, Verbesininae and Milleriinae) of the tribe Heliantheae by different authors at different times (Baagoe, 1974; Stuessy, 1977; Robinson, 1981; Karis, 1993). One of the objectives of this study was to evaluate the placement of the genus *Guizotia* within Heliantheae by including sequences from typical species under the above subtribes into phylogenetic analysis and by using an outgroup from a tribe closer to Heliantheae. Bremer (1987) in his recircumscription of the tribe Heleniеae, based on information from Robinson's (1981) classification, indicated that the tribe Eupatorieae is very closely related to Heliantheae s. lato, which was also indicated by Bergqvist et al. (1992). Thus *E. cannabinum* was used as outgroup in this analysis. The attempt to use *C. petrophiloides* which belongs to the subtribe Coreopsidinae under the tribe Heliantheae (Stuessy, 1977; Robinson, 1981) as additional ingroup taxa was not successful due to the fact that its sequence divergence from the other ingroup taxa (including the genus *Guizotia*) was higher than the sequence divergence between *E. cannabinum* (the assigned outgroup) and the ingroup taxa. As a result, *C. petrophiloides* was treated as an additional outgroup taxon in this analysis. According to Stuessy's systematic review of Heliantheae (Stuessy, 1977), *A. radicans* and *V. jacksonii* belong to the subtribe Verbesininae, a subtribe that later was assimilated into Robinson's subtribe Ecliptinae (Robinson, 1981).

M. quinqueflora is a typical species of the subtribe Milleriinae. In our phylogenetic analysis, *M. quinqueflora* was the closest species to the genus *Guizotia*. Thus, our data supports the placement of the genus *Guizotia* under the subtribe Milleriinae (Robinson, 1981; Bergqvist et al., 992; Karis, 1993). The second closest species to the genus *Guizotia* was *S. microcephalus*. In Robinson's (1981) comprehensive taxonomic treatment of the tribe Heliantheae, *S. microcephalus* was placed under the subtribe Melampodiinae. However, Panero et al. (1999), in their phylogenetic analysis of the subtribe Ecliptinae, based on chloroplast restriction site data, indicated that the genus *Smallanthus* is closely related to the genus *Rumfordia* and collectively to other Milleriinae genera. These authors advised the transfer of this genus from Robinson's subtribe Melampodiinae to the subtribe Milleriinae, which further strengthens the placement of the genus *Guizotia* under the subtribe Milleriinae.

Acknowledgments

We thank Dr. Anders Carlsson and Dr. Peter Olsson for valuable discussions on the data analyses and Ms. Therése Bengtsson for her assistance in the laboratory work. This work was financed by the Swedish International Development Agency (SIDA/SAREC) through International Science Program (ISP).

References

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.

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

Bayer RJ, Greber DG and Bagnall NH (2002) Phylogeny of Australian Gnaphalieae (Asteraceae) based on chloroplast and nuclear sequences, the *trnL* Intron, *trnL/trnF* intergenic spacer, *matK*, and ETS. *Syst. Bot.* 27: 801–814.

Bekele E, Geleta M, Dagne K, Jones AL, Barnes I, Bradman N and Thomas MG. (2006). Molecular phylogeny of genus *Guizotia* (Asteraceae) using DNA sequences derived from ITS *Genet. Resour. Crop Ev.* DOI 10.1007/s10722-006-9126-0.

Bentham G (1873) Notes on the classification, history, and geographical distribution of Compositeae. *J. linn. Soc. Bot.* 13: 335- 577.

Bergqvist G, Bremer B and Bremer K (1992) Chloroplast DNA restriction site variation and phylogenetic interrelationships of some genera of the Heliantheae sensu lato (Asteraceae). *Nord. J. Bot.* 12: 149-154.

Bremer K (1987) Tribal interrelationships of Asteraceae. *Cladistics* 3: 210-253.

Crawford DJ and Mort ME (2005) Phylogeny of eastern North American coreopsis (Asteraceae-Coreopsideae): insights from nuclear and plastid sequences and comments on character evolution. *Am. J. Bot.* 92: 330–336.

Dagne K (1994) Meiosis in interspecific hybrids and genomic interrelationships in *Guizotia* Cass. (Compositae). *Hereditas* 12: 119-129.

Dagne K (1995) Karyotypes, C-banding and nuclear 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.

Hall T (2005) BioEdit v. 7.0.5. Biological sequence alignment editor for Windows. Ibis Therapeutics a division of Isis pharmaceuticals. http://www.mbio.ncsu.edu/BioEdit/bioedit.html (last accessed 03-March- 2007).

Hilu KW and Liang H (1997) The *mat*K Gene: sequence variation and application in plant systematics. *Am. J. Bot.* 84: 830-839.

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

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

Hiremath SC, Murthy HN and Salimath SS (1992) Quatitative nuclear DNA differences associated with genome evolution in *Guizotia* (Compositae). *Genetica* 85: 241-247.

Humbles JE (1972) Observations of genus *Seigesbeckia* L. *Ciencia Y Naturaleza* 13: 2-19.
Johnson LA and Soltis E (1994) matK DNA sequences and phylogenetic reconstruction in Saxifragaceae s.str. *Syst. Bot.* 19: 143-56.
Karis PO (1993) Heliantheae sensu lato (Asteraceae) Clades and classification. *Plant Syst. Evol.* 188: 139-195.
Kimura M (1980) A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.* 16: 111-120.
Liang H and Hilu KW (1996) Application of the *mat*K gene sequences to grass systematics. *Can. J. Bot.* 74: 125-134.
Murthy HN, Hiremath SC and Salimath SS (1993) Origin, evolution and genome differentiation in *Guizotia abyssinica* and its wild species. *Theor. Appl. Genet.* 87: 587-592.
Panero JL, Jansen RK, Jennifer A and Clevinger JA (1999) Phylogenetic relationships of subtribe Ecliptinae (Asteraceae: Heliantheae) based on chloroplast DNA restriction site data. *Am. J. Bot.* 86: 413–427.
Robinson H (1981). A revision of the tribal and subtribal limits of the Heliantheae (Asteraceae). *Smithsonian Contrib. Bot.* 51: 1-102.
Schultheis LM (2001) Systematics of *Downingia* (Campanulaceae) based on molecular sequence data: implications for floral and chromosome evolution. *Syst. Bot.* 26: 603–621.
Schulz DL (1990) Zur kenntnis der Gattung Sigesbechia L. in Africa. *Gleditschia* 46: 613-628.
Shaw J and Small RL (2005) Chloroplast DNA phylogeny and phylogeography of the northAmericans plums (*Prunus* subgenus *Prunus* section *prunocerasus*, Rosaceae). *Am. J. Bot.* 92: 2011–2030.
Shaw J, Lickey EB, Beck JT, Farmer SB, Liu W, Miller J, Siripun KC, Winder CT, Schilling EE and Small RL (2005) The tortoise and the hare II: Relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis. *Am. J. Bot.* 92: 142-166.
Stuessy TF (1977) Heliantheae-systematic review. In: Heywood VH, Harborne JB and Turner BL (Eds.) The biology and chemistry of the Compositae, Vol 1. Academic Press, London, pp. 1106-1118.
Swofford DL (2000) PAUP*: Phylogenetic analysis using parsimony, version 4.0, beta. Sinauer associates inc. Sunderland, Massachusetts.
Taberlet P, Gielly L, Pautou G and Bouvet J (1991) Universal primers for amplification of three non-coding regions of chloroplast DNA. *Plant Mol. Biol.* 17: 1105-1109.
Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F and Higgins DG (1997) The CLUSTAL X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucl. Acids. Res.* 25: 4876-4882.
Wang XR, Tsumura Y, Yoshimaru H, Nagasaka K and Szmidt AE (1999) Phylogenetic relationships of Eurasian pines (*Pinus*, Pinaceae) based on chloroplast *rbcL, matK, rpL20-rpS18* spacer, and *trnV* intron sequences. *Am. J. Bot.* 86: 1742–1753.
Winkworth RC, Grau J, Robertson AW and Lockhart PJ (2002) The origins and evolution of the genus *Myosotis* L. (Boraginaceae). *Mol. Phylogenet. Evol.* 24: 180–193.

Acta Universitatis Agriculturae Sueciae

Doctoral Thesis No. 2007:27

AFLPs, RAPDs and DNA sequencing were applied to study genetic diversity, phylogenetics and systematics of *Guizotia* Cass. A substantial level of genetic variation and a significant population differentiation were found in guizotias. G. *scabra* ssp. *schimperi* is the most likely progenitor of G. *abyssinica*. This study suggests that G. *scabra* ssp. *scabra*, G. *scabra* ssp. *schimperi*, and Chelelu and Ketcha populations are separate species, and that *Guizotia* Cass. belongs to the subtribe Milleriinae.

Mulatu Geleta, received his graduate education at the Department of Plant Protection Biology, SLU, Alnarp. He has BSc and MSc Degrees in Biology from Addis Ababa University, Addis Ababa, Ethiopia.

Online publication of thesis summary: http://epsilon.slu.se/eng/index.html

ISSN 1652-6880

ISBN 978-91-576-7326-8