Abayneh Derero

Genetic variation in *Cordia africana* Lam. in Ethiopia

Cuvillier Verlag Göttingen

Genetic variation in *Cordia africana* Lam. in Ethiopia

Dissertation

zur Erlangung des Doktorgrades
der Fakultät für Forstwissenschaften und Waldökologie
der Georg-August-Universität Göttingen

vorgelegt von

Abayneh Derero

Geboren in Addis Ababa, Äthiopien

Göttingen 2007

Bibliografische Information Der Deutschen Bibliothek
Die Deutsche Bibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über http://dnb.ddb.de abrufbar.
1. Aufl. - Göttingen : Cuvillier, 2007
Zugl.: Göttingen, Univ., Diss., 2007

978-3-86727-426-6

Berichterstatter: Prof. Dr. Reiner Finkeldey

Prof. Dr. Dirk Hölscher

Tag der mündlichen Prüfung: 21 August 2007

Gedruckt mit Unterstützung des DAAD

Nonnenstieg 8, 37075 Göttingen
Telefon: 0551-54724-0
Telefax: 0551-54724-21
www.cuvillier.de

1. Auflage, 2007
Gedruckt auf säurefreiem Papier

978-3-86727-426-5

To my family

Acknowledgments

I am deeply indebted to Prof. Dr. Reiner Finkeldey for accepting me as a PhD student at the Institute of Forest Genetics and Forest Tree Breeding, for his guidance, encouragement, and excellent supervision in my studies. Prof. Dr. Hans H. Hattemer has read part of my thesis, and he was helpful in other matters as well, and hence deserves great appreciation. I am also grateful to Dr. Demel Teketay, Dr. Girma Balcha and Dr. Tamrat Bekele for their kind help and collaboration in the beginning of my dissertation. I would like also to extend my sincere gratitude to Dr. Oliver Gailing for his supervision in both the molecular laboratory work and the writing up of the thesis. I am also grateful to Prof. Dr. Hans-Rolf Gregorius, Dr. Elizabeth Gillet and Prof. Dr. Martin Ziehe with whom I had series of fruitful discussions. I thank also Dr. Barbara Vornam and Dr. Ludger Leinemann for their collaboration whenever needed. I am also very grateful to Ms. Teresa Gatesman for proof reading my dissertation. I have many thanks to the secretary of the Institute, Mrs. Marita Schwahn, who was always kind and helpful in many official matters. My special thanks go to Olga Artes, August Capelle, Gerold Dinkel, Thomas Seliger, Oleksandra Dolynska and Christine Radler for their assistance in the laboratory and other activities. I have enjoyed the company and friendships of former and current PhD students in the Institute: Dr. Cui-Ping Cao, Dr. Hong T. Luu, Dr. Alexandru Lucian Curtu, Dr. Madhav Panday, Martin Mottura, Akindele Akinnagbe, Hani Siti Nuroniah, Valdir Marcus Stefenon, Nga Phi Nguyen, Nicolas G. Eliades, Marius Ekue, Yanti Rachmayanti, Taye Bekele, Sylvia Nascimento and Amaryllis Vidalis. I thank the coordinators in the PhD Program “Wood Biology and Technology”, Dr Kuersten and Dr. Buettner, and the fellow students for the nice academic atmosphere and experience sharing we had together. I would also like to extend my admiration to the people in the Harvest of the Lord for their kind fellowship. I extend my great thanks to my parents, my brothers, and to Mitiku for their invaluable support all along the way. My special acknowledgement goes to my wife Burte and my sons Yabets and Nati, for being the source of strength, comfort and joy for my life. DAAD offered me the scholarship and supported my fieldwork in Ethiopia. The Ethiopian Institute of Agricultural Research/Forestry Research Center provided me with vehicles for the fieldwork.

Table of contents

Chapter 1: General introduction

1.1 Investigation of genetic variation

Genetically, organisms are structured in the hierarchy of genes, genotypes, populations and species. The population genetic structure is the most fundamental piece of information for a species that requires genetic management (Brown, 1978, Yeh, 2000). To date, the description of population genetic structure and its dynamics has been based on allele and genotype frequencies in sampled populations with simply inherited traits whose transmissions follow Mendelian rules (Yeh, 2000). The assessment of genetic variation involves the determination of genetic multiplicity (i.e. proportion of polymorphic loci, number of alleles per locus and allelic richness) and genetic diversity (i.e. effective number of alleles, expected and observed heterozygosity) (Finkeldey and Hattemer, 2007). Its pattern is the result of the impact of evolutionary factors, which are mutation, gene flow and migration, mating system, genetic drift and selection (Wright, 1931; Finkeldey and Hattemer, 2007). The Hardy-Weinberg Principle, which states that in a large random-mating population with non-overlapping generations, the allele and genotype frequencies will remain constant from generation to generation when there is no mutation, migration and natural selection provides the foundation for all population genetic investigation (Yeh, 2000).

In plants, two very distinct vehicles mediate the dispersal function: the male gametophyte (pollen) and the young sporophyte (seed) (Bensch and Åkesson, 2005). Hence, in populations continuously distributed over a larger area, genetic isolation by distance is expected, since it is assumed that gene flow happens mainly between nearby locations (Rousset, 1977).

Genetic variation within and among populations can be investigated by employing biochemical markers (isozymes/allozymes), direct DNA sequencing and using molecular (DNA) markers (Weising *et al*., 2005). Until recently, research on the genetics of tropical trees was confined largely to allozyme studies of the genetic structure of adults in continuous forests (Hamrick and Murawski, 1991; Loveless, 1992). The first DNA marker exploited is referred to as Restriction Fragment Length Polymorphism (RFLPs; Botstein *et*

al., 1980) and involves the analysis of restriction-digested DNA in the so-called Southern blotting technique and hybridization with a sequence-specific probe (Weising *et al*., 2005). The recent molecular techniques such as Random Amplified Polymorphic DNA (RAPD; Williams *et al*., 1990), Intersimple Sequence Repeat Polymorphism (ISSR; Wolfe and Liston, 1998), microsatellites (also known as Simple Sequence Repeats, SSRs; Akkaya *et al*., 1992), Amplified Fragment Length Polymorphism (AFLP; Vos *et al*., 1995) as well as inverse sequence-tagged repeat (ISTR; Rohde, 1996) involve the Polymerase Chain Reaction (PCR), in which amplification of the fragments of genomic DNA is conducted using a heat-resistant DNA polymerase (Taq polymerase), primers and deoxyribonucleotide triphosphates at high temperatures (Saiki *et al*., 1988). The use of molecular markers in the investigation of genetic variation is getting a wide acceptance and broad application in fields such as phylogeny, taxonomy, ecology and genetics and breeding (Weising *et al*., 2005).

The DNA under investigation can be nuclear, extra-nuclear (organelle DNA) or the whole genome depending on the type of marker employed. The organelle and nuclear genes are inherited uniparentally and biparentally, respectively, and can be used to study the distribution of genetic diversity within and among populations and to infer the relative importance of seed and pollen dispersal (Petit *et al*., 1993, 2005). The nuclear DNA (nDNA) is subjected to intergenerational recombination whereas organelle genomes can exhibit intragenerational segregation (Murrey *et al*., 2000).

1.2 Physiography, climate and vegetation of Ethiopia

Ethiopia is located in the north-eastern part of Africa, popularly known as the Horn, between $3^0$24' and $14^0$53' N, and $32^0$42' and $48^0$12'E, and covers a total area of 1.13 million km^2 (CSA, 2001). It is bordered in the north and north-east by Eritrea, in the east by Djibouti and Somalia, in the south by Kenya, and in the west and southwest by Sudan. The population is estimated at more than 75 million, with 85% depending on agriculture, which is the mainstay of Ethiopian economy accounting for 54% of the GDP and about 90% of the exports (CSA, 2001).

The physiographic features include rugged mountains, deep gorges and river valleys, and rolling plains (Bekele, 1994). The very large dissected, dome-shaped mountain massif

consists of two plateaus, which are referred to as the North-Western (NW) Highlands and the South-Eastern (SE) Highlands (highlands being areas above 1500 m above sea level). Generally, the Ethiopian highlands are the biggest in tropical Africa (Friis *et al*., 2001). The altitude in Ethiopia ranges from 120 m below sea level at the Afar depression to 4370 m on Tullu Demtu (the highest peak in the SE) and 4620 m on Mount Ras-Dashen (the highest peak in the NW). The NW highlands are subdivided into a northern and a southern part by the extensive and up to 1000 m deep and 10-50 km wide Blue Nile (Abbay) Valley. The northern part is highly dissected, but with altitudes between ca. 2000 m and ca. 3000 m, they are generally higher than the southern highlands, which are less dissected and have extensive areas between ca. 1500 m and ca. 2000 m (Friis *et al*., 2001). Another classification has recognized three subunits in the NW highlands; namely, the north, central and the southwest (SW) (Friis, 1992). The geologically active Great Rift Valley extends from Lake Turkana in the SE to the Afar region in the north-east (NE) dissecting the two highland massifs in the Lake Region. In addition, lowlands, steppes and semi-desert areas (areas less than 1500 m in altitude mainly the Western Lowlands, Borana and Ogaden) stretch from the respective slopes and bound each of the massifs. This great terrain diversity is responsible for wide variations in climate, soil and natural vegetation (Friis, 2001).

The climate of Ethiopia is tropical monsoon with wide topographic-induced variations. The SW Highlands receive the highest annual rainfall (1400-2200 mm per annum) and the arid lowlands in Afar and in the SE Lowlands receive as little as 100 mm to 300 mm. The monthly average temperature shows very little seasonal variation, 2^{o}C in the south and 6^{o}C in the north, and the mean annual temperature decreases with altitude (Friis 1992). Generally, low temperatures are attained during the rainy months and high temperatures occur during the dry and sunny season. The climate of Ethiopia in terms of temperature variations can be broadly classified into three types (Breitenbach, 1963): Cool Zone (Dega)- cold mountains above 2400-2600 m, with average temperatures between 10 and 16°C; Temperate Zone (Woina-Dega)- temperate highlands between 1600-1900 m to 2400-2600 m, with mean temperatures between 16 to 20°C; and Hot Zone (Kolla)- hot lowlands below 1600-1900 m, with both tropical and arid conditions, and with mean temperatures between 20 and 29°C. A more complex climatic classification identifies three zones (Dry, Tropical and Temperate) with three subdivisions each (Anonymous, 1992).

The vegetation of Ethiopia has been surveyed quite extensively (e.g. von Breitenbach, 1963; Greenway 1973; Chaffey, 1979). The types and geographic distributions of the vegetation based on Friis (1992) can be summarized as follows: Deciduous woodland and wooded grassland cover the western lowlands, with the floristic composition varying considerably from the north to the south. The undifferentiated woodland, which is composed of species such as *Combretum collinum* and *Terminalia brownie*, is found on the escarpment of the NW highlands in the altitudinal range from 500 to 1500 m. Transitional rain forests exist on the SW slopes of the NW highlands. The Afromontane rain forests are located at the elevations between 1500 and 2000 m in the SW part of the NW highlands and in some parts of the SE highlands. Under drier conditions and at higher elevations, the forests in both plateaus tend to be dominated by the conifers *Podocarpus falcatus* and *Juniperus procera*, which are classified as undifferentiated Afromontane forests. The Rift Valley is mainly covered by *Acacia* woodland in the lakes area and with desert scrubland in the Afar area. The western lowlands are humid and are mainly characterized by a single semi-deciduous forest (Guineo-Congolian forest) and extensive broadleaved woodland. The SE lowlands are dry lowlands with *Acacia* and other semi-desert woody species.

A recent national document (IBC, 2005) identifies eight natural ecosystems in the country (1) Afroalpine and subafroalpine ecosystems (areas above 3200 m), (2) Dry evergreen montane forest and grassland complexes (an extensive area in both plateaus between [1600-] 1900-3300 m), (3) Moist evergreen montane forest ecosystems (situated in the SW and SE between 1500 and 2600m), (4) Lowland semi-evergreen forest ecosystems (found between 450 and 650 m), (5) *Acacia-commiphora* woodland ecosystems (occurs between 900 and 1900 m in the Rift valley and SE lowlands), (6) *Combretum-Terminalia* woodland ecosystems (found in the western escarpments between 500 and 1900 m), (7) Desert and semi-desert scrubland ecosystems (found in areas below 500 m), and (8) Aquatic ecosystems, which includes wetlands and lake areas.

However, the natural ecosystems and the natural vegetation have been greatly altered over the years. For example, the large areas of the forests over both the NW and the SE highlands have been replaced by secondary evergreen bush land, wooded grassland or farmland. The forests (not including the woodland and other vegetation cover) were reduced from the original 35% to 16% in 1952 (Sayer *et al.*, 1992) and to 3.6% currently (Anonymous, 2004). Remnants are left on the less accessible areas mainly in the SW and

SE. Isolated forest trees occur scattered throughout most of the highlands (Figure 1), suggesting the former presence of forest there. Patches of forest and woodland, with a species composition similar to that of the remaining natural forests, are very common around churchyards and graveyards, which are abundant in the highlands (Bekele, 1994). Deforestation has its longest history in the northern part of the country due to the farming activities for several centuries practiced there.

Figure 1 Contrasting view of continuous and scattered populations

1.3 *Cordia africana* Lam.

1.3.1 Taxonomy and botanic description

Cordia L. (generic name after Valeris Cordus, a German botanist) is a pantropical genus of about 250 species belonging to Boraginaceae Juss. (ICRAF, 1998), a plant family comprising about 100 genera and 2000 species that are characterized by flowers in helicoid cymes and by coarsely hairy herbage (Carr, 2006). *Cordia africana* Lam. (Synonym: *Cordia abyssinica* R. Br.) is a tree (rarely shrubby) species. Its English common names are East African cordia or large-leafed cordia or Sudan teak (ICRAF, 1998). On average, it attains a height between 14 and 21 m and a diameter at breast height (dbh) between 0.60 and 0.90 cm (chapter 5), and shows great morphological variation (Figure 2).

The botanical description of *Cordia africana* (Warfa 1988) is as follows: Crown spreading, umbrella-shaped or rounded. Bole typically curved or crooked. Bark greyish-brown to dark brown, smooth in young trees, but soon becoming rough and longitudinally fissured with

age; young branchlets with sparse long hairs. Leaves alternate, simple, ovate to subcircular, 7.5-17.5 (max. 30) cm long, 3.5-10.2 (max. 30) cm broad; thinly leathery; dark green above, paler green and velvety below, with prominent parallel tertiary net-nerves (about 7 pairs of lateral nerves); apex broadly tapering or rounded; base rounded to shallowly lobed; margin entire; petiole slender, 2.5-7.6 cm long. Buds oval, stalkless, pleated open into flowers that are bisexual, white, sweet scented, shortly pedicelate or subsessile, massed in compact panicles covering the crown, with a white mass of attractive flowers; calyx less than 1 cm long, strongly ribbed, back of lobes covered with short, soft, brown hairs; corolla lobes crinkled, white, long-exerted, funnel-shaped, about 2.5 cm long; cymes with many flowers. Fruit a drupe, smooth, spherical, oval tipped, fleshy, 1.3-1.5 cm long; green when young, yellow to orange when mature; with a sweet, mucilaginous pulp and short remains of the calyx at the base; contains 2-4 seeds, which lack endosperm.

Figure 2 Morphological variations and features in *Cordia africana*

1.3.2 Natural distribution and habitat

Cordia africana is native to Angola, the Democratic Republic of Congo, Djibouti, Eritrea, Ethiopia, Ghana, Guinea, Kenya, Malawi, Mozambique, South Africa, Sudan, Tanzania, Uganda, Zimbabwe, Saudi Arabia and Yemen (Warfa, 1988; Friis 1992). In Ethiopia, it is widespread in Broadleaved Afromontane Rain Forests, Undifferentiated (Dry) Afromontane Forests ('mixed *Podocarpus* forest') and in riverine forests as well as in the western lowlands (Friis, 1992; Figure 3). Generally, the species grows in areas with altitudes between 550 and 2600 m a.s.l. and annual rainfall of 700 to 2000 mm (Friis, 1992). It is an early colonizer in forest re-growth and is often found along forest margins, regenerates in clearings and forest gaps (Fichtl and Admasu, 1994; Derero *et al.*, 2003). It has a light quality sensing mechanism that hinders it from germinating beneath leaves (Yirdaw and Leinonen, 2002).

The natural distribution, the habitats and the populations of the species are very much affected, like other forest tree species, by deforestation, fragmentation and selective logging. Especially the northern part of the country represents an extreme case of deforestation in which the species is represented by scattered trees on farmlands, in and around homesteads, church compounds and graveyards; except at some spots, such as Zeghie Peninsula, where a relatively continuous forest exists.

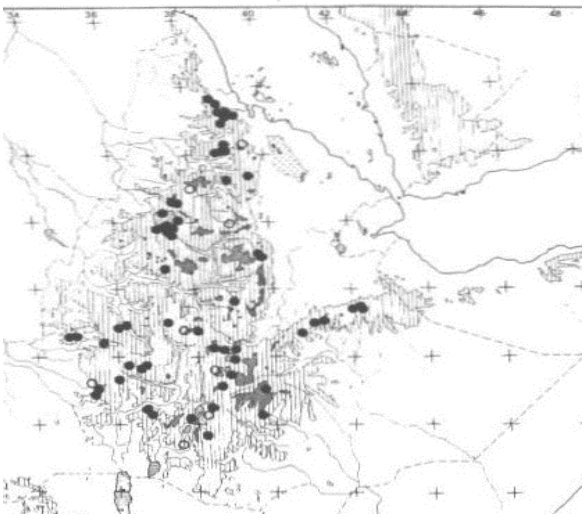


Figure 3 Natural distribution of *Cordia africana* in Ethiopia and Eritrea (from Friis, 1992).

1.3.3 Reproductive biology

Cordia africana begins flowering when a tree is 3-5 years old (ICRAF, 1988; Negash, 1995,). It is a monoecious species with complete flowers (hermaphrodite) and is known to be bee-pollinated (Warfa, 1988). The main flowering period of the species is from October to March (Fichtl and Admasu, 1994). After pollination by insects, fruit development takes a period of almost six months. The fruits are edible, and probably seed dispersal is carried out by mammals and birds (endozoochorus) (ICRAF, 1988).

1.3.4 Uses and economic importance

Cordia africana is a fast-growing and highly valued timber tree used for high-quality furniture, doors, windows, cabinet-making, drums, beehives, joinery, interior construction, mortars, panelling and veneering (ICRAF, 1998). In Ethiopia, it is one of the economically important tree species exploited commercially for timber (Abebe and Holm, 2003). Mixed planting and closer initial spacing in plantations have been recommended to improve the growth performance of the species and minimize problems posed by big branches on the quality of the wood (Mihretu, 1999; Mehari and Habte, 2005). The contribution of scattered trees of *C. africana* to various soil properties and its importance as a (coffee) shade tree in traditional agroforestry systems has been documented (Teketay and Tegineh, 1991; Yadessa *et al.*, 2001). The species fells its leaves heavily in the dry season, which not only serve as mulch but also contribute to nutrient cycling through decomposition (Negash, 1995; Teklay and Malmer, 2004). The species is also excellent honeybee forage since it supplies abundant pollen and copious nectar (Fichtl and Admasu, 1994), and beehives are often traditionally placed on the trees. The fruits are edible (Warfa, 1988) and serve as a (drought season) food that is marketed in some parts of the country. The seeds of this species are among the most highly demanded of the centrally distributed indigenous tree species for planting (Derero, 2002); however, insect infestation poses a problem in their utilization (Tigabu and Oden, 2002).

1.4 Objectives

1.4.1 General objective

The general objective of the study was to determine the genetic variation in populations of *C. africana* found at different altitudes and geographical locations ranging from scattered to continuous populations over a wide geographical distance, as well as to investigate populations that represent various levels of forest fragmentation in Ethiopia employing two DNA marker systems.

1.4.2 Specific objectives

The study was specifically aimed at the following five objectives:

- Investigate the genetic variation within various populations of the species in Ethiopia
- Determine the population structure of the species in Ethiopia
- Relate the genetic structure of the species to tree seed zones, natural ecosystems, altitudinal classes and geographical (physiographic) units in Ethiopia
- Investigate the genetic effects of forest fragmentation at a metapopulation scale in the transitional rainforest in south-west Ethiopia
- Identify species seed zones/tree breeding zones for *C. africana* in Ethiopia

Chapter 2: Variation at Amplified Fragment Length Polymorphisms

2.1 Introduction

Amplified Fragment Length Polymorphism (AFLP) is one of the DNA fingerprinting techniques widely applied in population genetic studies. DNA fingerprinting involves the display of a set of DNA fragments from a specific DNA sample, and AFLPs in this regard can be applied for DNAs of any origin or complexity without prior sequence knowledge (Vos *et al*., 1995, Weising *et al*., 2005).

The AFLP technique is based on the detection of genomic restriction fragments by PCR amplification using a limited set of generic primers. In AFLPs, polymorphisms between two or more genotypes arise from three sources: (1) sequence variation in one or both restriction sites flanking a particular fragment, (2) insertions or deletions within an amplified fragment, and (3) differences in nucleotide sequences immediately adjacent to the restriction sites (Weising *et al*., 2005). However, the polymorphism is displayed as only the presence and absence of equal-sized restriction fragments (Vos *et al*., 1995) that can be scored as 1 and 0, respectively, at such an AFLP locus. Hence, AFLP is a dominant marker technique and it is not possible to separate heterozygotes (1/0) from homozygotes (1/1). Despite this shortcoming, estimates of genetic diversity and genetic structure comparable to calculations from codominant markers can be made converting presence and absence data to expected heterozygosity by assuming HWE (Bensch and Åkesson, 2005). Nevertheless, populations may not be usually in HWE, and its dominant nature makes AFLP an inappropriate marker system in evaluating departures from the equilibrium. However, this shortcoming can also be overcome, and robust estimates of heterozygosity can be made by analyzing a large number of AFLP loci, which normally exhibit contrasting frequencies of null alleles (Kremer *et al*., 2005).

The possibility of generating enough polymorphic markers makes AFLPs suitable for individual identification, estimation of pairwise relatedness and parentage analyses to resolve individual differences at the DNA fingerprint level (Mueller and Wolfenbarger,

1999). Therefore, AFLPs can serve as alternative to microsatellites to assign individuals in investigating the affinity of each genotype to presumed populations of origin (Campbell *et al*., 2003). Furthermore, AFLPs can also be used to identify hybrid individuals (interspecific or intraspecific), even in systems where microsatellites may fail to do so (Bensch *et al*., 2002). AFLPs are also instrumental for gene mapping and linkage studies. Genetic linkage maps are essential in order to localize genetic regions that are associated with quantitative trait loci (QTL) (Bensch and Åkesson, 2005). In this regard, AFLPs, together with microsatellites, are the most widely used markers when developing new linkage maps (Erikson *et al*., 2004). Especially in genera where variation at other markers is low, AFLP can serve to solve species phylogenies (Cervera *et al*., 2005). A major concern with using AFLP data for phylogenetic reconstructions is that bands of the same length seen in two species may not be necessarily homologous (Bensch and Åkesson, 2005).

The aim of the present study was to investigate the genetic variation within and among populations of *C. africana* in Ethiopia. To this effect, the following hypotheses were tested: (1) there is genetic variation within and among populations, (2) the population structure of the species can be explained by the existing tree seed zones, natural ecosystems, altitudinal classes and physiographic features, and (3) the genetic variation of the species follows isolation-by-distance.

2.2 Materials and Methods

2.2.1 Sampling and plant material collection

A total of 22 different populations (sample size = 32) were sampled from virtually all the growing regions of *C. africana* in Ethiopia (Figure 4). Plant materials (leaves) were collected from adult trees keeping a minimum distance of 100 m from each other. The leaf samples were dried and preserved in silica gel for DNA isolation and further analyses.

The sampled populations are found distributed in the altitudinal range of 950 m to 2150 m a.s.l., and in the geographical range of 6°34’ to 13°28’N (about 865 km horizontal distance) and 35°14’ to 42°07’E (about 822 km horizontal distance) (Table 1, Appendix 1).

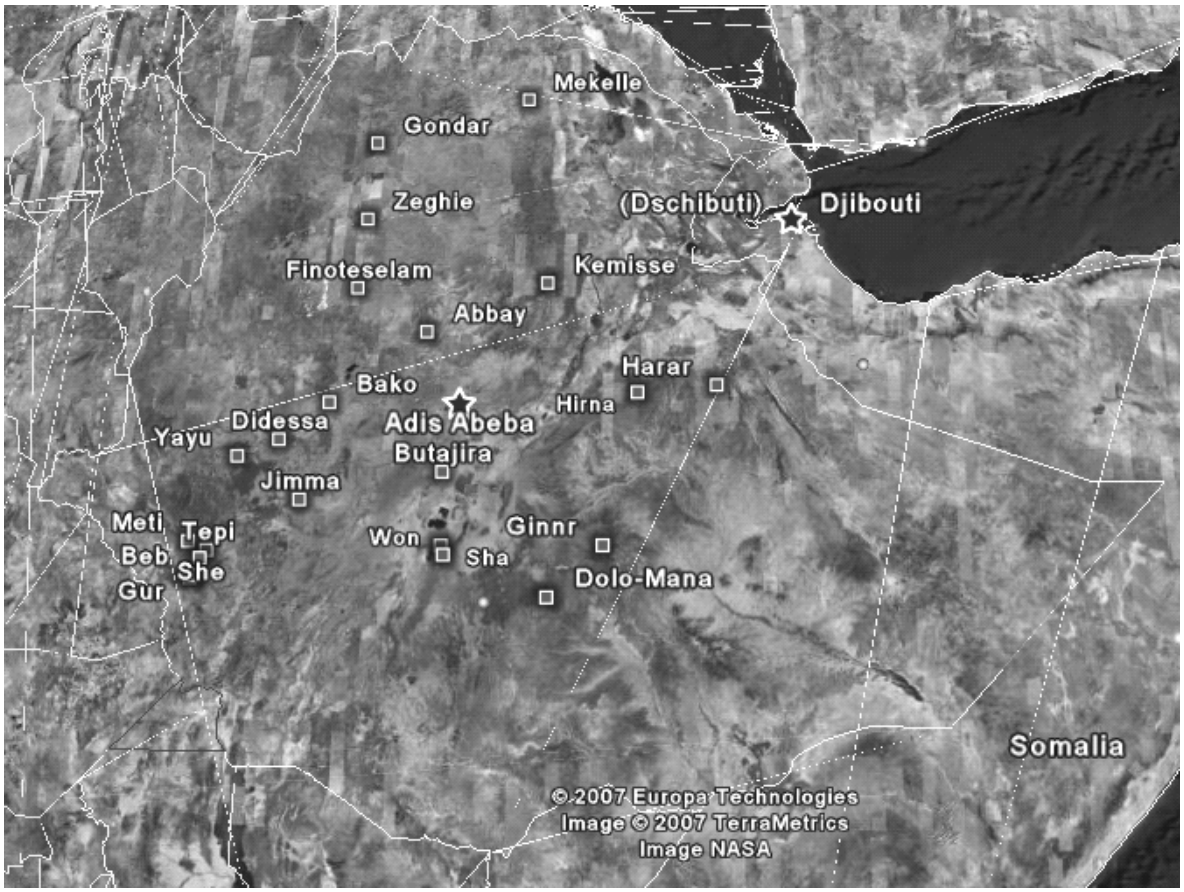


Figure 4 Map showing the locations of the 22 sampled populations of *Cordia africana* in Ethiopia

The populations were categorized arbitrarily into three geographical units following the physiographic units of the country, and referred to as the north, south-west (SW) and south-east (SE). The north was represented by populations entirely from the northern highlands, which is a highly denuded forest area. The SW (includes one population from the western part of the Central Plateau) was represented both by highland and lowland populations, and has the vast majority of the remnant high forests and represents relatively earlier deforestations. The SE (populations from the SE highlands + two populations from the Rift Valley) is more deforested than the SW but represents a much better forest condition compared to the north.

The populations can also be broadly classified as continuous populations with high tree density, (undisturbed, disturbed forests) and scattered populations with low tree density (remnant trees, regenerations and planted trees from wildings on traditional Agroforestry systems, in and around homesteads, in church compounds and in graveyards). The SW has

a third peculiar forest type referred to as coffee plantation, (natural forest converted to coffee plantation about three decades ago).

The populations also represent nine major tree seed zones and eighteen sub zones, which were delineated employing data on climate, altitudinal variation and natural vegetation type (Aalbæk, 1993). The names and designations of the nine major seed zones (given in

Table 1 List of populations of *Cordia africana* sampled from Ethiopia

No	Population	Code	Latitude (N)	Longitude (E)	Altitude	Altitudinal group	Seed Zone	Population Type*	Province Θ	Geo. unit	Natural ecosystem
1	Abbay[§]	Abb	10^{0}03'	38^{0}15'	1600-1900	Highland	11	Scattered	Gojjam/ Shewa	North	ACW
2	Bako	Bak	09^{0}06'	37^{0}00'	1500-1700	Highland	20	Scattered	Shewa	SW	DEMF
3	Bebeka	Beb	06^{0}51'	35^{0}21'	1050-1250	Lowland	23	Coffee	Kaffa	SW	MEMF
4	Butajira[+]	But	08^{0}10'	38^{0}34'	1850-1900	Highland	23	Scattered	Shewa	SE	ACW
5	Didessa	Did	08^{0}36'	36^{0}21'	1300-1650	Intermediate	11	Scattered	Illubabor	SW	ACW
6	Dolo-Mana	Del	06^{0}34'	39^{0}58'	1450-1600	Highland	24	Continuous	Bale	SE	MEMF
7	Finoteselam	Fin	10^{0}41'	37^{0}15'	1750-1950	Highland	20	Scattered	Gojjam	North	DEMF
8	Ginnr	Gin	07^{0}13'	40^{0}40'	2000	Highland	18	Scattered	Bale	SE	DEMF
9	Gondar[+]	Gon	12^{0}45'	37^{0}24'	1850-2000	Highland	19	Scattered	Gondar	North	DEMF
10	Guraferda	Gur	06^{0}49'	35^{0}18'	1000-1150	Lowland	23	Continuous	Kaffa	SW	MEMF
11	Harar	Har	09^{0}18'	42^{0}07'	1750-2050	Highland	21	Scattered	Hararghe	SE	DEMF
12	Hirna	Hir	09^{0}12'	41^{0}05'	1750-2000	Highland	21	Scattered	Hararghe	SE	DEMF
13	Jimma	Jim	07^{0}49'	36^{0}42'	1500-1800	Highland	23	Scattered	Kaffa	SW	MEMF
14	Kemisse	Kem	10^{0}43'	39^{0}51'	1500	Highland	15	Scattered	Wello	North	DEMF
15	Mekelle[§]	Mek	13^{0}28'	39^{0}28'	2000-2100	Highland	15	Scattered	Tigray	North	DEMF
16	Meti	Met	07^{0}17'	35^{0}14'	1150-1350	Lowland	23	Coffee	Illubabor	SW	MEMF
17	Shashemene	Sha	07^{0}13'	38^{0}36'	1950	Highland	23	Scattered	Shewa	SE	ACW
18	Sheko	She	07^{0}04'	35^{0}25'	1150-1400	Lowland	23	Continuous	Kaffa	SW	MEMF
19	Tepi	Tep	07^{0}10'	35^{0}30'	1000-1100	Lowland	23	Coffee	Illubabor	SW	MEMF
20	Wondogenet	Won	07^{0}06'	38^{0}38'	1800-2000	Highland	24	Continuous	Sidamo	SE	MEMF
21	Yayu	Yay	08^{0}23'	35^{0}48'	1250-1550	Intermediate	23	Continuous	Illubabor	SW	MEMF
22	Zeghie	Zeg	11^{0}41'	37^{0}18'	1800-1900	Highland	22	Continuous	Gojjam	North	DEMF

*Scattered = trees on farmlands, homesteads, church compounds, graveyards (remnants, regeneration & wildings), Coffee = natural forests converted to coffee plantations by systematic removal of individuals about three decades ago, continuous = natural forests (intact, disturbed, semi-forest); [+] Butajira = Sodo-Woreda, Gondar = Tikil-Dengay (Armacho), [§]Two sub populations at Abbay (Filikilk, Dejen) and Mekelle (Mekelle, Debir), Inter. = intermediate between lowland and highland, Θ = provinces of imperial Ethiopia, ACW = Acacia commiphora woodland ecosystem, DEMF = Dry evergreen montane forest and grassland complex, MEMF = Moist evergreen montane forest ecosystem

brackets) included in this study were: a) Broadleaved deciduous woodland (11), b) Western highlands dry *Juniperus* forest (15), c) Upper Wabe Juniperus forest (18), d)

Western highlands moist *Juniperus* forest (19), e) Western highlands undifferentiated Afromontane forest (20), f) SE highlands undifferentiated Afromontane forest (21), g) Lake Tana undifferentiated Afromontane forest (22), h) Western highlands broadleaved Afromontane rainforest (23) and i) SE highlands broadleaved Afromontane rainforest (24).

Furthermore, the populations represent three natural ecosystems of the country, namely the Acacia-commiphora Woodland Ecosystem, the Dry Evergreen Montane Forest and Grassland Complex, and the Moist Evergreen Montane Forest Ecosystem**.**

2.2.2 Laboratory investigations

2.2.2.1 DNA isolation

Total genomic DNA was extracted from dried leaves (taking about 20 mg dried leaves or 1 cm^2) using the DNeasyTM 96 Plant Kit (Qiagen, Hilden, Germany) following the manufacturer's instructions. The amount and quality of the DNA was checked on 1% agarose gel electrophoresis using Tri-acetate (1 x TAE) as running buffer. Molecular weight standards (Lambda DNA Marker, 100 bps ladders) were used to evaluate the quantity and quality of the DNA, which was stained with ethidium bromide solution and visualized by UV illumination.

2.2.2.2 Amplified Fragment Length Polymorphisms (AFLPs)

Probes from different populations were distributed randomly on the PCR plate. Genome-wide variation patterns were analysed with AFLPs following the protocol in Vos *et al*. (1995) with minor modifications (Appendix 2). The genomic DNA was digested with two different restriction enzymes, a six-cutter (*Eco*RI; 5'-G↓AATTC-3') and a four-cutter (*Mse*I; 5'-T↓TAA-3'). Then, double-stranded *Eco*RI and *Mse*I adaptors were ligated to the sticky ends created by the restriction enzymes to generate template DNA for PCR. The restriction-ligation was performed on the template DNA overnight at room temperature. Subsequently, the first PCR (preamplification) amplified the restriction-ligation fragments with the primer combination E01/M03, (5'-GACTGCGTACCAATTCA-3' and 5'-GATGAGTCCTGAGTAAG-3', respectively) (nomenclature according to www.keygene.com), which have a sequence corresponding to the adaptor and the restriction site plus one additional base, A and G, respectively. A small fraction of the

preamplification reaction was then used in a second (selective) PCR using two primer combinations that extend additionally three bases inward on each side (Figure 5). The selective primers used in the reactions were the fluorescent-dye-labelled (6-FAM) forward primer, E41, (5'-GACTGCGTACCAATTC AGG-3') and the reverse primer, M74, (5'-GATGAGTCCTGAGTAA GGT-3').

The AFLP products finally were resolved electrophoretically on the ABI 3100 Genetic Analyser with the Size Standard (GeneScan^{TM}-500, ROX^{TM}, Applied Biosystems, New York, USA). The length of electrophoresis products was measured with the help of the internal size standard GS-ROX500. Finally, individual alleles were analysed using Genescan © version 3.7 (Applied Biosystems) and genotyped using Genotyper © 3.7 NT (Applied Biosystems) (Figure 6).

Figure 5 Flowchart of steps required for an AFLP assay (Applied Biosystems)

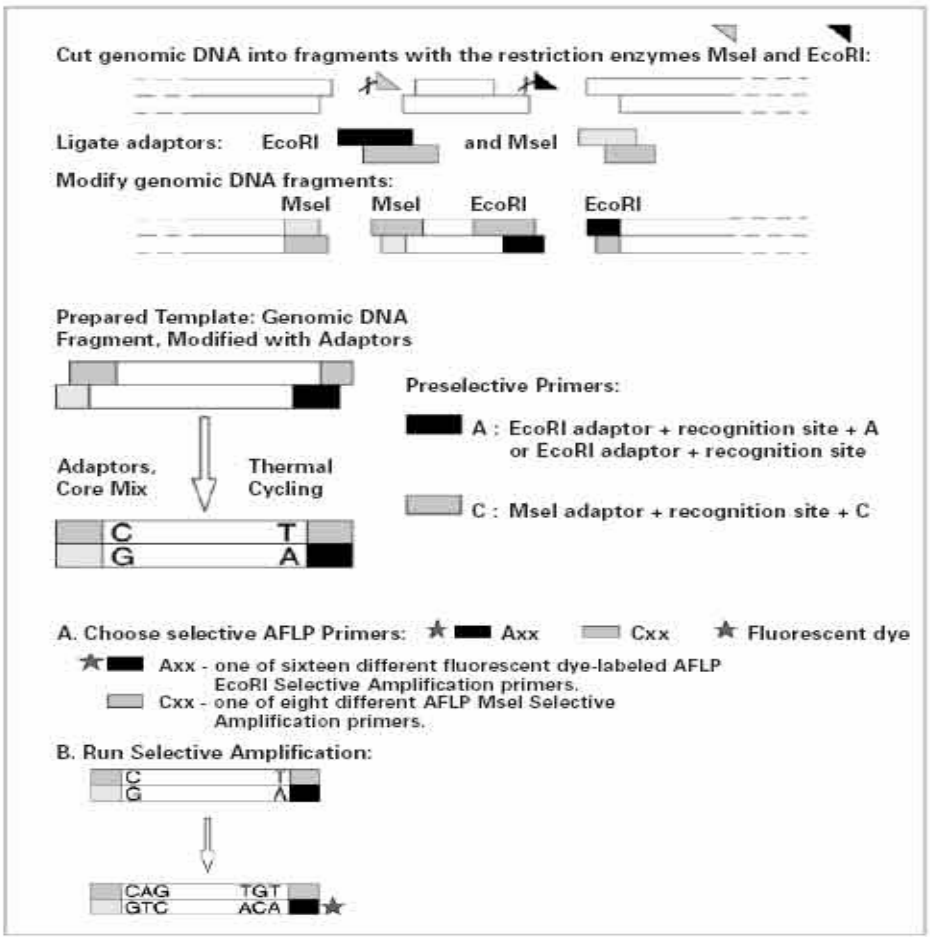
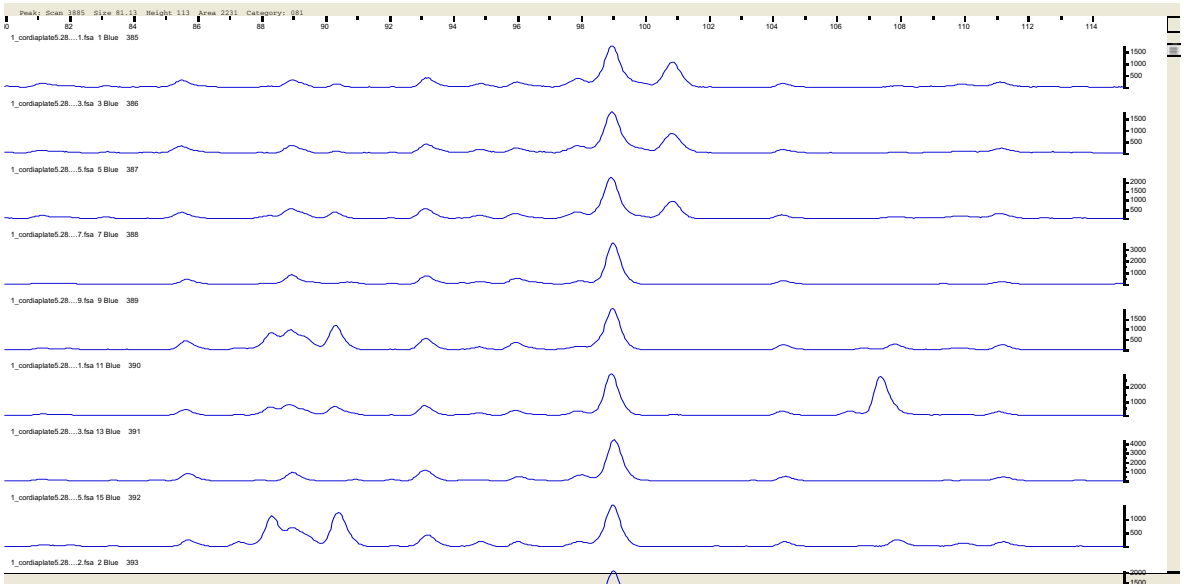


Figure 6 Electropherograms showing AFLP markers as visualised by Genescan 3.7 and Genotyper 3.7.



2.2.3 AFLP reproducibility

The AFLP procedure was repeated on 12 samples in order to test the reproducibility of the AFLP fragments. The number of cases in which a fragment was repeatable in the 12 pairs of samples was monitored. The size region below 80 base pairs (bp) generally represented low repeatability (ranging from 50 to 75%). For larger fragments, the repeatability varied between 75 and 100%. A total of 171 fragments revealed high repeatability values (individual values being above 90%) amounting to a mean repeatability of 98% (Appendix 3). However, fragments from a region lacking a clear banding pattern (120-129bp), loci and samples with too many missing data and fragments with unclear or stutter-band-like appearances were excluded from the further analyses. Consequently, a total of 90 bands with a mean repeatability of 97% were selected for further analysis. The mean size of the selected fragments was 203 bp (±83.50).

2.2.4 AFLP fragment scoring

The AFLP data were analysed with Genescan 3.7 and Genotyper 3.7 (Applied Biosystems). The scoring was carried out manually as “1” for fragments present, “0” for fragments absent and “?” for missing data (very weak banding pattern, impossible to distinguish between presence and absence of a fragment). The absence or presence of

fragments in defined size intervals (± 0.5 bp) was transformed into a binary 0/1 matrix. The AFLP bands were considered as putative loci and assumed to be dominant markers with two alleles (a dominant marker allele coding for the presence of a band at a given position, and a recessive null allele coding for the absence of the band).

2.2.5 AFLP data analyses

2.2.5.1 Genetic variation within populations

Genetic variation within populations was assessed from the binary AFLP matrix using AFLP-SURV 1.0 (Vekemans, 2002). The default method, a Bayesian approach with non-uniform prior distribution of allele frequencies (Zhivotovsky, 1999), was used for the computation of allelic frequencies from the observed frequencies of fragments at each marker locus in each population (Appendix 4) assuming HW genotypic proportions. Based on these estimates of allelic frequencies, the approach of Lynch and Milligan (1994) was used to estimate genetic variation. For each population, proportions of polymorphic loci (PPL) at the 5% level (i.e. loci with allelic frequencies lying within the range of 0.05 to 0.95) and expected heterozygosity or Nei's gene diversity (Hj) were computed. The Nei's gene diversity and the mean within-population gene diversity ($\hat{H}_W$), respectively, were estimated as:

$$\hat{H}_j = \frac{1}{L}\sum_{i=1}^{L}\hat{H}_j(i) \text{ ……. (Lynch and Milligan, 1994),}$$

where $\hat{H}_j$ is the mean observed gene diversity in the j_{th} population, L= number of observed loci, $\hat{H}_j$ (i) is the mean observed gene diversity in population j at the ith locus.

$$\hat{H}_W = \frac{1}{n}\sum_{j=1}^{n}\hat{H}_j \text{ ..... (Lynch and Milligan, 1994)}$$

2.2.5.2 Genetic variation among populations

The population genetic structure was investigated by the analysis of molecular variance (AMOVA) using the software Arlequin 3.1 (Excoffier *et al*., 2005). The multilocus AFLP phenotypes were considered as haplotypes. Genetic structure indices were estimated using

the information on the allelic content of the haplotypes, as well as their frequencies. The information on the differences in allelic contents between haplotypes was entered as a matrix of Euclidean squared distances.

Groupings of populations were conducted following altitudes, physiographic units, natural ecosystems, forest types (3 groups each), and seed zones (9 major seed zones). Furthermore, a combined grouping according to altitude and geographic location was done as follows: (1) SW lowlands, encompassing populations entirely from the transitional rain forest region (2) SW highland, including the intermediate (between lowland and highland) Yayu and Didessa populations (3) Northern highland, and finally (4) SE highland.

The significance of the covariance components associated with the different levels of genetic structure (within populations, within groups of populations, among groups) was tested using non-parametric permutation procedures (Excoffier, *et al*., 1992). The covariance components (σ_i^2 's) were used to compute fixation indices in terms of inbreeding coefficients (Slatkin, 1991) as:

$$F_{ST} = \frac{f_0 - f_1}{1 - f_1} = \qquad \text{(Slatkin, 1991),}$$

where f_0 is the probability of identity by descent of two different genes drawn from the same population, and f_1 is the probability of identity by descent of two genes drawn from two different populations. This fixation index computed in AMOVA setting is denoted as $\Phi_{ST.}$

Furthermore, a matrix depicting pairwise genetic distances between populations was computed using Nei's D (after Lynch and Milligan, 1994) for the overall analysis of genetic similarity among the 22 populations. A dendrogram was constructed based on this matrix employing the unweighted pair-group method with arithmetic means (UPGMA) in NTSYSpc2.0 (Rohlf, 1998). In addition, pairwise Φ_{ST} between populations were computed in Arlequin 3.1 (Excoffier *et al*., 2005).

2.2.5.3 Correlation between geographical and genetic distances

The matrix of geographical distances among populations (X-axis) was related using the matrix of Nei's genetic distances (Y-axis) in a spatial autocorrelation test in NTSYSpc2.0 (Rohlf, 1998). The two dissimilarity matrices were plotted, one matrix against the other element by element. The degree of relationship between the two matrices was analyzed by computing the product-moment correlation, r, and the Mantel test statistic (Mantel, 1967), Z.

$$Z = \sum_{i<j}^{n} X_{ij} Y_{ij} \text{ ,}$$

where X_{ij} and Y_{ij} are the off-diagonal elements of matrices X and Y. Permutational distribution was computed by comparing Matrix X with all possible matrices in which the order of the objects (or variables) in the Y matrix had been permuted. Then significance tests were performed by comparing the observed Z-value with its permutational distribution (10,000 permutations were used). Additional AMOVAs were computed taking subsets of the data (in SW-north, SW-SE, SW-north populations at a time and within SW, SE and north populations) to detect the finer geographical patterns.

2.3 Results

2.3.1 Genetic variation within populations

The mean percentage of polymorphic loci (PPL) was 85.7%, ranging from 62.2% at Dolo-Mana (a disturbed forest population SE) to 92.2% at Didessa (scattered population) and Meti and Wondogenet (relatively continuous populations in SW and SE, respectively) (Table 2). The mean within-population diversity (Hj) based on the 22 populations (mean sample size, n = 21.6) analyzed was 0.287. The population gene diversity (H) ranged from 0.220 at Dolo-Mana in the SE to 0.320 at Meti and Sheko in the SW.

The first three highest diversities were recorded in two continuous populations and on a farmland, and the lowest value in a disturbed forest population. However, there was no significant association between population type (continuous or scattered) and genetic

variation (Hj, PPL). The ranges of the rankings for the forest and the farmland populations overlapped for both PPL and diversity.

Table 2 List of *Cordia africana* populations and their genetic variation indices, listed in decreasing order of Hj

No	Population	Province	Geographical Unit	Type	N	PPL	Hj	S.E.(Hj)
1	Meti	Illubabor	SW	Coffee	26	92.2	0.320	0.016
2	Sheko	Kaffa	SW	Continuous	22	88.9	0.320	0.016
3	Harar	Haraghe	SE	Scattered	19	86.7	0.306	0.017
4	Didessa	Illubabor	SW	Scattered	23	92.2	0.304	0.016
5	Guraferda	Kaffa	SW	Continuous	20	91.1	0.303	0.015
6	Bebeka	Kaffa	SW	Coffee	23	87.8	0.301	0.017
7	Ginnr	Bale	SE	Scattered	18	84.4	0.300	0.017
8	Abbay	Showa/Gojjam	North	Scattered	21	86.7	0.299	0.018
9	Kemmise	Wello	North	Scattered	21	87.8	0.297	0.017
10	Jimma	Kaffa	SW	Scattered	22	87.8	0.296	0.015
11	Wondogenet	Sidamo	SE	Continuous	21	92.2	0.296	0.016
12	Yayu	Illubabor	SW	Continuous	19	87.8	0.295	0.017
13	Gondar	Gondar	North	Scattered	22	91.1	0.294	0.015
14	Bako	Shoa	SW	Scattered	18	83.3	0.294	0.017
15	Zeghie	Gojjam	North	Continuous	25	91.1	0.290	0.016
16	Tepi	Illubabor	SW	Coffee	22	88.9	0.290	0.015
17	Hirna	Hararghe	SE	Scattered	22	82.2	0.284	0.017
18	Finoteselam	Gojjam	North	Scattered	23	82.2	0.277	0.017
19	Shashemene	Shoa	SE	Scattered	23	83.3	0.271	0.018
20	Mekelle	Tigray	North	Scattered	17	72.2	0.268	0.019
21	Butajira	Shoa	SE	Scattered	23	82.2	0.259	0.017
22	Dolo-Mana	Bale	SE	Continuous	26	62.2	0.220	0.018

N = number of samples, PPL = percentage of polymorphic loci, Hj = genetic diversity (expected heterozygosity), S.E.(Hj) = standard error of Hj

2.3.2 Population genetic structure

2.3.2.1 Geographic patterns of the genetic variation

It is remarkable to note that the PPL were the highest in two populations from the SW and one population from the SE, and it was the lowest in two populations from the SE. Furthermore, the populations from the SW held 4 out of the first 5 positions in the ranking according to diversity, while two populations from the SE were the least diverse.

2.3.2.2 Population differentiation

The Analysis of Molecular Variance (AMOVA) revealed significant variation among the populations (Φ_{ST} = 0.072, $p < 0.001$), but 92.8 percent of the genetic variation was contained within the populations. The analyses of the populations according to altitudinal class and geographical units revealed also low but significant differentiation among the groups (Table 3). However, there was no significant differentiation among tree seed zones and among natural ecosystems.

Table 3 Genetic differentiations among 22 populations of *C. africana* employing AMOVA

Source of variation	df	SSD	MSD	Variance component	Percentage of variation	*p-value*
Among populations	21	610.74	29.08	0.85	7.27	*< 0.001*
Within populations	453	4895.83	10.81	10.81	92.73	*< 0.001*
Among altitudinal groups	2	104.06	52.09	0.22	1.87	*< 0.01*
Among populations	19	506.67	26.58	0.74	6.26	*< 0.001*
Within populations	453	4895.83	10.81	10.81	91.88	*< 0.001*
Among geographic units	2	113.24	56.62	0.19	1.66	*< 0.001*
Among populations	19	497.49	26.18	0.71	6.09	*< 0.001*
Within populations	453	4895.83	10.81	10.81	92.25	*< 0.001*
Among seed zones	9	294.98	32.78	0.15	1.24	*ns*
Among populations	12	314.17	26.18	0.71	6.17	*< 0.001*
Within populations	453	4895.83	10.81	10.81	92.89	*< 0.001*
Among natural ecosystems	2	74.43	37.22	0.06	0.51	*ns*
Among populations	19	536.31	28.23	0.09	6.93	*< 0.001*
Within populations	453	4895.83	10.81	10.81	92.56	*< 0.001*

df = degrees of freedom, SSD = sum of squared deviations, MSD = mean of squared deviations

2.3.2.3 Pairwise population comparisons

The Nei's genetic distance ranged from 0.000 to 0.046, the highest genetic distance being between the northernmost (Mekelle) and the southernmost populations (Dolo-Mana). The

UPGMA dendrogram (Figure 7) constructed using the matrix of pairwise genetic distances (Table 4) between populations revealed some tendency for geographical clustering. The upper cluster comprised mainly populations from the SW and the SE, which are separated from populations of the north. In general 18 out of the 22 populations showed geographical affinity. However, the topology of the dendrogram showed a weak resolution of clustering. On the other hand, the majority (206 out of 231) of the pairwise Φ_{ST} comparisons of the 22 populations revealed significant differentiations (87%, $p < 0.01$) (Table 5).

2.3.2.4 Correlation between geographical and genetic distance matrices

A significant positive correlation was obtained between the matrices of geographical (Appendix 1) and Nei's genetic distances (r = 0.35, $p < 0.01$), as well as between geographical and pairwise fixation indices (Φ_{ST}) (r = 0.34, $p < 0.01$) (Figure 8). Furthermore, Mantel tests on subsets of the data revealed significant correlations in the SW-north (r = 0.541, $p < 0.001$) and in the SE-north (r = 0.307, $p < 0.01$) populations (Figure 9). However, there was no such correlation in the SW-SE populations and within the north, the SW and the SE populations.

Table 4 Pairwise population comparisons in *Cordia africana* employing AFLP based on Nei's genetic distance

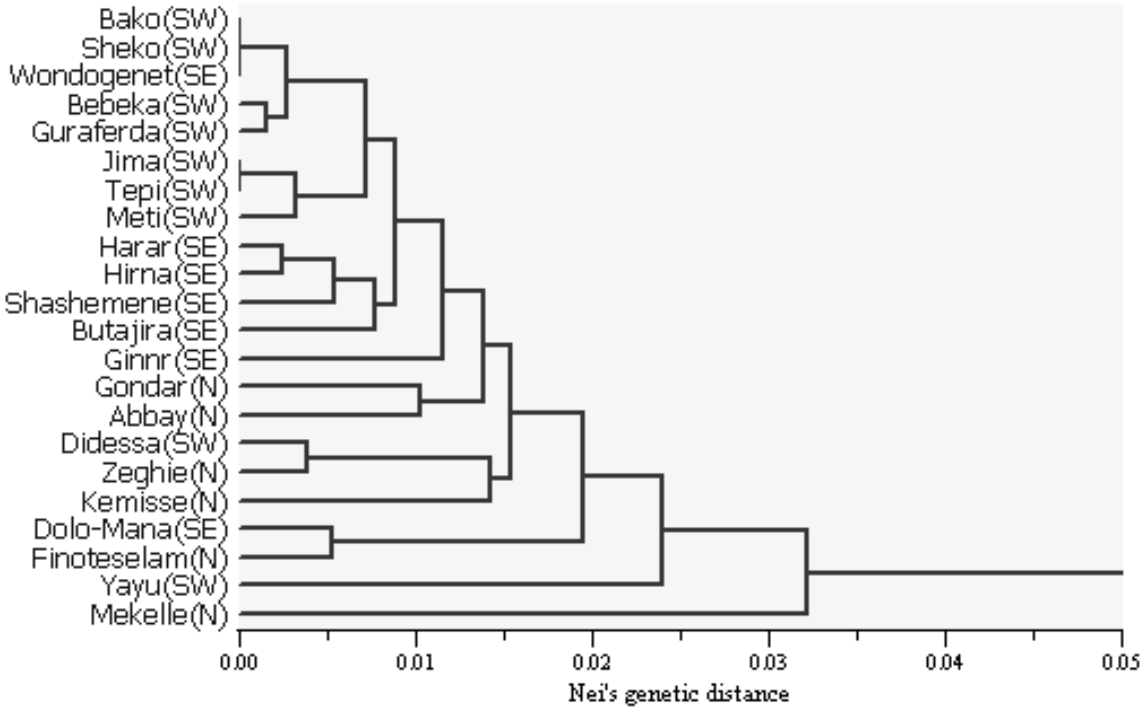
Popn*	Nei`s genetic distance																				
Bak	0.007																				
Beb	0.006	0.000																			
Del	0.038	0.017	0.034																		
Did	0.018	0.005	0.012	0.027																	
Fin	0.029	0.010	0.022	0.005	0.019																
Gur	0.016	0.001	0.002	0.026	0.006	0.019															
Gin	0.011	0.003	0.008	0.034	0.016	0.024	0.010														
Gon	0.010	0.010	0.021	0.027	0.010	0.021	0.011	0.019													
Har	0.004	0.000	0.003	0.018	0.007	0.016	0.003	0.004	0.011												
Hir	0.012	0.007	0.018	0.018	0.016	0.011	0.013	0.007	0.010	0.003											
Jim	0.022	0.000	0.012	0.014	0.012	0.005	0.008	0.013	0.013	0.010	0.004										
Kem	0.015	0.010	0.010	0.029	0.016	0.020	0.015	0.032	0.024	0.009	0.023	0.022									
Mek	0.019	0.029	0.042	0.046	0.036	0.036	0.034	0.024	0.013	0.026	0.015	0.031	0.041								
Met	0.013	0.000	0.006	0.017	0.008	0.008	0.004	0.012	0.016	0.004	0.006	0.006	0.018	0.034							
Sha	0.022	0.008	0.013	0.017	0.009	0.012	0.009	0.013	0.014	0.004	0.007	0.011	0.014	0.024	0.012						
She	0.008	0.000	0.001	0.029	0.007	0.023	0.005	0.018	0.013	0.007	0.020	0.013	0.014	0.036	0.006	0.016					
But	0.022	0.005	0.013	0.010	0.008	0.007	0.011	0.020	0.013	0.005	0.012	0.004	0.020	0.044	0.010	0.006	0.014				
Tep	0.024	0.008	0.011	0.016	0.015	0.007	0.007	0.019	0.017	0.010	0.006	0.000	0.024	0.036	0.001	0.012	0.010	0.008			
Won	0.014	0.000	0.004	0.016	0.005	0.011	0.006	0.011	0.013	0.003	0.014	0.006	0.014	0.035	0.008	0.003	0.000	0.003	0.008		
Yay	0.023	0.019	0.011	0.045	0.020	0.034	0.019	0.022	0.036	0.020	0.029	0.024	0.023	0.038	0.019	0.020	0.018	0.028	0.021	0.015	
Zeg	0.016	0.011	0.021	0.030	0.004	0.026	0.020	0.027	0.014	0.016	0.027	0.026	0.013	0.035	0.015	0.016	0.010	0.017	0.030	0.013	0.034
	Abb	Bak	Beb	Del	Did	Fin	Gur	Gin	Gon	Har	Hir	Jim	Kem	Mek	Met	Sha	She	But	Tep	Won	Yay

* for the full names of the populations see Table 1, ns = non significant, * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$

Table 5 Pairwise fixation indices (Φ_{ST}) and their level of significance

Bak*	0.036																				
	*																				
Beb	0.040	0.001																			
	*	ns																			
Del	0.191	0.095	0.162																		
	***	***	***																		
Did	0.079	0.050	0.057	0.152																	
	***	**	***	***																	
Fin	0.109	0.048	0.083	0.056	0.082																
	***	**	***	***	***																
Gur	0.089	0.022	0.028	0.131	0.044	0.078															
	***	ns	ns	***	*	***															
Gin	0.041	0.020	0.047	0.157	0.077	0.092	0.053														
	*	ns	*	***	***	***	**														
Gon	0.061	0.065	0.097	0.143	0.072	0.083	0.066	0.057													
	**	***	***	***	***	***	***	***													
Har	0.016	0.004	0.025	0.094	0.031	0.054	0.027	0.017	0.048												
	ns	ns	ns	***	*	**	ns	ns	**												
Hir	0.062	0.051	0.103	0.097	0.077	0.066	0.072	0.037	0.037	0.021											
	**	**	***	***	***	***	***	*	*	ns											
Jim	0.079	0.012	0.050	0.085	0.064	0.033	0.029	0.029	0.057	0.040	0.032										
	***	ns	**	***	***	*	*	*	***	**	ns										
Kem	0.082	0.060	0.069	0.146	0.073	0.064	0.070	0.133	0.112	0.053	0.110	0.080									
	***	***	**	***	***	***	**	***	***	**	***	***									
Mek	0.096	0.139	0.180	0.211	0.159	0.149	0.166	0.100	0.073	0.114	0.053	0.120	0.179								
	***	**	***	***	***	***	***	***	***	***	**	***	***								
Met	0.046	0.002	0.017	0.097	0.035	0.035	0.015	0.050	0.061	0.016	0.044	0.025	0.058	0.135							
	**	ns	ns	***	**	**	ns	**	***	ns	**	*	***	***							
Sha	0.104	0.051	0.074	0.080	0.050	0.056	0.057	0.077	0.071	0.027	0.042	0.056	0.061	0.115	0.055						
	***	**	***	***	***	***	**	***	***	*	**	***	**	***	***						
She	0.060	0.018	0.023	0.168	0.064	0.101	0.043	0.088	0.089	0.050	0.114	0.070	0.085	0.190	0.025	0.109					
	**	ns	ns	***	***	***	*	***	***	**	***	**	***	***	*	***					
But	0.101	0.027	0.049	0.068	0.043	0.048	0.050	0.065	0.074	0.018	0.058	0.030	0.084	0.172	0.043	0.027	0.089				
	***	*	**	***	**	***	**	**	***	ns	***	*	***	***	***	*	***				
Tep	0.089	0.034	0.049	0.098	0.072	0.043	0.027	0.065	0.062	0.043	0.040	0.005	0.095	0.145	0.006	0.069	0.050	0.048			
	***	*	**	***	***	**	*	**	***	**	**	*	***	***	ns	***	**	**			
Won	0.048	0.007	0.013	0.094	0.034	0.049	0.029	0.045	0.060	0.010	0.067	0.033	0.054	0.150	0.019	0.039	0.013	0.027	0.035		
	**	ns	ns	***	*	**	*	*	**	ns	***	*	**	***	ns	**	ns	*	*		
Yay	0.120	0.088	0.070	0.210	0.094	0.130	0.097	0.112	0.147	0.100	0.147	0.094	0.106	0.184	0.088	0.104	0.109	0.116	0.099	0.071	
	***	***	**	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	
Zeg	0.063	0.052	0.083	0.145	0.026	0.088	0.085	0.099	0.068	0.044	0.088	0.096	0.061	0.133	0.045	0.061	0.064	0.074	0.102	0.044	0.129
	***	**	***	***	*	***	***	***	***	**	***	***	***	***	***	***	***	***	***	**	***
	Abb	Bak	Beb	Del	Did	Fin	Gur	Gin	Gon	Har	Hir	Jim	Kem	Mek	Met	Sha	She	But	Tep	Won	Yay

* = for the full names of the populations see Table 1, ns = non significant, * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$



N = population from the northern unit, SW = population from the south-western unit, SE = populations from the south-eastern unit

Figure 7 An UPGMA dendrogram revealing genetic relatedness among 22 populations of *Cordia africana* in Ethiopia employing AFLP based on Nei's genetic distance matrix

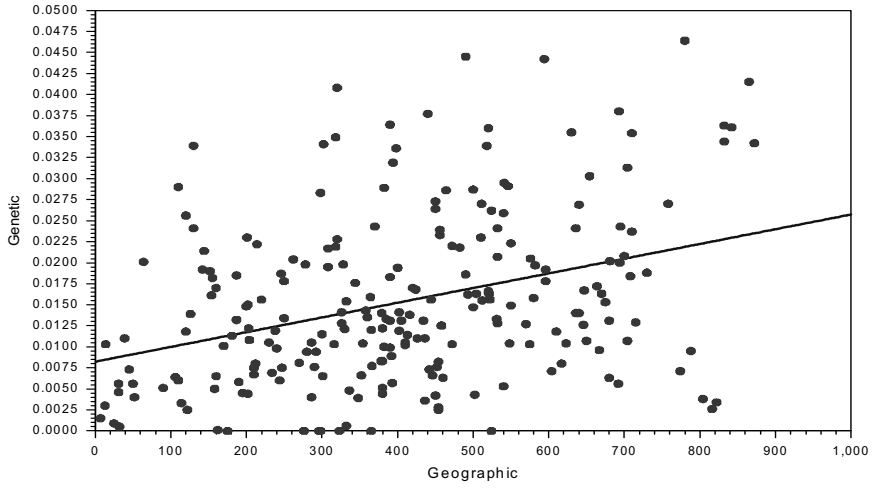


Figure 8 Correlation between geographic and genetic distance matrices of *Cordia africana* populations in Ethiopia

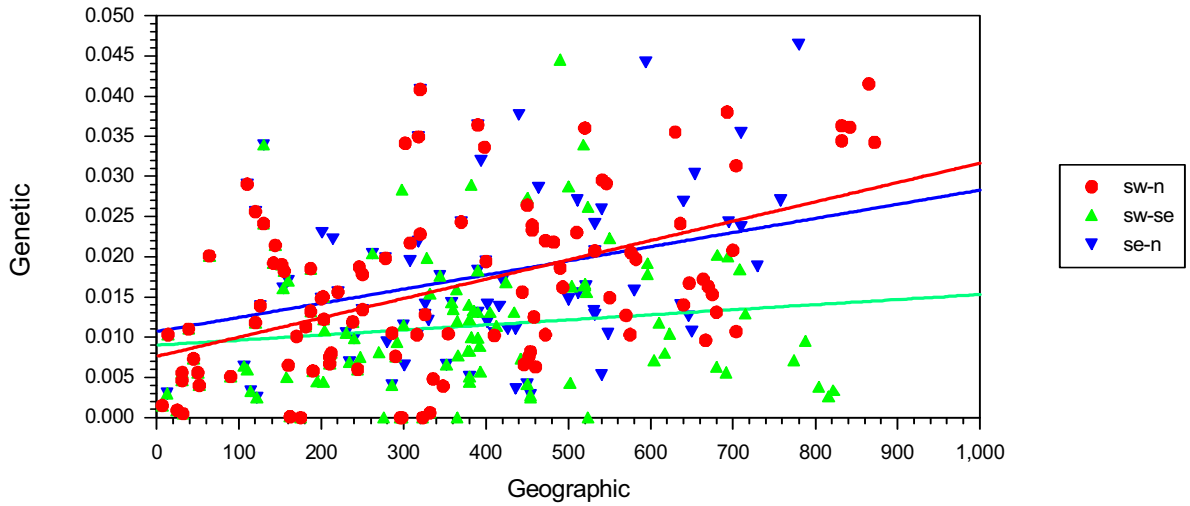


Figure 9 Correlation between geographic and genetic distance in SW-north, SE-north and SW-SE populations of *Cordia africana*

2.4 Discussion

2.4.1 Genetic variation within populations

Molecular markers have emerged as convenient methods for quantifying levels of genetic diversity in populations (Chase *et al*., 1996). The demography and evolutionary history of populations is often inferred from the variation at genetic markers that are assumed to be neutral. However, if a marker is actually subject to selection, conclusions based on patterns of genetic variation could be misleading (Rand 1995; Ford 2002) since neutral marker alleles could be linked to deleterious mutations or selectively favoured alleles, and genetic variation can erode faster than expected under neutral assumptions (van Oosterhout *et al.*, 2004). Gene diversity estimates based on dominant markers like RAPDs and AFLPs depends on the frequency of null homozygotes and the fixation index (F) of the populations (Kremer *et al*., 2005), and can either assume that populations are in Hardy-Weinberg equilibrium (HWE) (the inbreeding coefficient, F = 0) or use associated estimates of F from codominant markers (Nybom, 2004). There are different procedures for the estimation of null allele frequencies; namely, the square root method (Berstein, 1930), the Lynch and Milligan procedure (1994) and the Bayesian methods (Zhivotovsky, 1999), which may result in varying estimates. Krauss (2000) compared the three estimation procedures in *Personia mollis* and demonstrated that they give very similar estimates of heterozygosity. It is also argued that diversity estimates made with dominant markers such

as AFLP provide comparable data to surveys undertaken with codominant markers (Nybom and Bartish, 2000; Nybom, 2004). This is because, the larger number and wide distribution of AFLP markers throughout the genome compensates for the poor genetic information content at each locus (Mariette *et al*., 2002a, b).

Cordia africana under the assumption of HW exhibits relatively high mean diversity within populations (Hj = 0.287), which can be attributed to the wide ecological amplitude distribution of the species as has been noted for some other species (Hamrick *et al*., 1992). The mean diversity estimates of 10 Neotropical outcrossing tree species using AFLPs under the HW assumption (F = 0) was 0.23; values for individual species ranging approximately from 0.14 to 0.32. The analysis of the same set of species under the assumption of full selfing (F = 1) gave a mean diversity of about 0.24; values for individual species ranging from 0.17 to 0.38 (Kremer *et al*., 2005). In another survey, the overall mean value of the within population gene diversity (H_{pop}) summarized from 38 studies on outcrossing species employing RAPD was found to be 0.27 (Nybom, 2004). Some other tropical tree species investigated using AFLPs include *Moringa oleifera* with a mean diversity of 0.067 (Muluvi *et al*., 1999), *Calycophyllum spruceanum* with mean population diversities ranging from 0.249 to 0.349 (Russel *et al*., 1999), *Pterocarpus officinalis* with a population diversity ranging from 0.15 to 0.27 (Rivera-Ocasio *et al*., 2002), and *Shorea leprosula* and *S. parvifolia* with the mean population diversities of 0.168 and 0.138 each (Cao *et al*., 2006).

The sampling of the populations of *C. africana* in this study was designed to fully represent the growing regions of the species in Ethiopia, and therefore populations even from forest-denuded parts of the country were included. The samples also represented continuous forests to patchy and scattered populations. However, the genetic variation in terms of percentage of polymorphic loci (PPL) and population gene diversity (Hj) varied randomly among populations from different locations as well as among continuous and scattered populations; the PPL ranged from 62.2% at Dolo-Mana (Bale) to 92.2% at Meti and Didessa (Illubabor) and Wondogenet (Sidamo), while the Hj ranged from 0.220 at Dolo-Mana and 0.230 at Ginnr (both populations located in Bale) to 0.320 at Meti and Sheko (neighbouring populations located in Illubabor and Kaffa, respectively). Similarly, the lowest genetic diversities for *Hagenia abyssinica* (using ISSR markers, Feyissa *et al*., 2007) and *Coffea arabica* (using RAPD, Aga *et al*., 2003; and using ISTR, Aga and

Bryngelsson, 2006) were reported for populations from Bale province. In addition, the highest diversity for all three species (*C. africana, H. abyssinica* and *C. arabica*) was found in the SW populations.

Generally, populations of *C. africana* from the SW geographical unit harbour the highest diversity (Hj = 0.303), followed by the populations from the northern (Hj = 0.287) and the SE (Hj = 0.276) geographical units. In another study, it was found that the populations of *H. abyssinica* from the SE (Kersa, Adaba, Dinsho and Goba) were also characterized by the lowest genetic diversity (Feyissa *et al*., 2007). The analysis of diversity at species level (nearly 3000 plant taxa) in Ethiopia showed that the species richness follows a pattern with a marked centre in SE Ethiopia (Sidamo, Bale and Hararghe), a slightly less marked centre in the highlands from Shewa to Eritrea (central to north), and a poorly marked centre in SW and west Ethiopia (Friis *et al*., 2001); a converse trend to the pattern of genetic diversity detected in *C. africana*. Such a trend would imply the existence of trade-off between species and genetic diversity. Indeed, it is essential to assess the plant taxa diversity in each of the populations from where *C. africana* individuals were sampled for a stronger argument on trade-off between species and genetic diversity. Such a trade-off was observed in *Fagus sylvatica* L. in which the genetic diversity was strongly negatively correlated with the species diversity (Wehenkel *et al*., 2006).

In addition, the existence of significant differentiations between the lowland and highland provenances in *C. africana* magnify the importance of the other dimension of distance, the vertical distance, among the populations. In the analysis of nearly 3000 taxa in Ethiopia, diversity gradually increased with altitude from 350 taxa between 0 and 300 m to a maximum of 1600 taxa between 1200 and 1500 m a.s.l. The diversity gradually declines again to 400 taxa between 2700 and 3000 m, and to well below 100 on the highest mountain peaks (Friis *et al*., 2001).

The ranges of the rankings for the continuous and the scattered populations (most of which are in traditional agroforestry systems) overlapped for both PPL and diversity revealing that the latter group harboured substantial genetic diversity comparable to the former, and can be used as sources of genetic material for tree planting, tree improvement and conservation activities in areas where the natural forest has been denuded, such as in northern and eastern Ethiopia.

As a consequence of habitat loss and degradation, severe loss of genetic variation is predicted (Lowe *et al.*, 2005). Though the highly degraded regions and the highly fragmented populations were expected to have less diversity due to a less effective population size, the findings did not fully support this. For example, the effective population size in the scattered population, Abbay (located in Blue Nile Gorge), may be a few hundreds of individuals but it harbours very comparable genetic variation (Hj = 0.299) with - and even higher than - some of the populations from the continuous populations. Therefore, effective population size does not seem to affect genetic variation in *C. africana* strongly though tropical trees are thought to be particularly vulnerable to the effects of habitat degradation due to their demographic and reproductive characteristics, including low density of occurrence, complex self-incompatible breeding systems and high rates of outcrossing (Cascante *et al.*, 2002) as well as intimate interactions with pollinators and seed dispersers (Dick *et al.*, 2003; Ward *et al.*, 2005).

2.4.2 Population genetic structure

Deforestation in the Ethiopian plateaus and the escarpments has left the country with mosaics of forest patches, secondary evergreen bushland, wooded grassland, or farmland. Habitat fragmentation is expected to reduce gene flow by increasing spatial isolation between patches. Reductions in the proportions of immigrant genes, combined with allelic loss in small populations will increase genetic differentiation among disturbed remnants (Slatkin 1987; Young *et al.*, 1996). The level of the fixation index, which is a measure of genetic differentiation, in a population depends on the mating system and genetic structure (Wahlund effect; Hartl and Clark, 1997). Nei's genetic distance at AFLPs among the 22 populations ranged from 0.00 at some pairs to 0.05, the latter being between the northernmost and the southernmost populations. In the UPGMA dendrogram constructed using the Nei's genetic distance matrix at AFLP, 18 of the 22 populations showed some degree of geographical affinity in which most of the populations from the SW and the SE formed the upper cluster and the middle cluster, respectively, while the lower cluster was predominantly composed of the northern populations. However, the topology of the dendrogram did not show clear patterns of clustering (i.e. there were no three clear clusters as such) and generally revealed weak geographic structure.

The population differentiation among the populations at AFLP (Φ_{ST} = 0.072) was moderate. According to Wright (1978) and Hartl and Clark (1997), a value of F_{ST} lying in the range 0-0.05 indicates little genetic differentiation; a value between 0.05 and 0.15, moderate differentiation; a value between 0.15 and 0.25, great differentiation; and values above 0.25, very great differentiation. The moderate differentiation in *C. africana* at AFLP falls in the expected range of many of the outcrossing tropical species, which is generally less than 10% (Kremer *et al.*, 2005). The differentiation (G_{ST}) of 11 populations of *Cordia alliodora* at isozymes was 11.7% (Chase, *et al.*, 1995). In support of little differentiation, it is argued that if the average time that a population persists in one area is less than the time it takes for genetic drift to fix neutral alleles, which is of the same order of magnitude as the effective population size, there will be little differentiation of local populations due to genetic drift even if there is no exchange of individuals between established populations (Slatkin, 1987). For example, a very great population differentiation (Φ_{ST} = 0.52) was obtained at AFLPs for *Pterocarpus officinalis*, which involved four continental and four Caribbean populations, which apparently represents long-term differentiation (past colonization history) and genetic drift within the local populations but almost no possibility of gene flow between the two groups (Rivera-Ocasio *et al*., 2002).

The aforementioned argument of Slatkin (1987) can explain the low population differentiation and the weak geographical structuring in *C. africana*. The other possible argument may be related to the existence of high gene flow over long distances in the species. The pollination in *C. africana* is carried out by bees. The existence of scattered trees across the landscape, especially in the SW and SE, may facilitate pollen movement among distant populations, as the scattered remnants could act as stepping-stones for the pollinators. Direct measures of pollen flow in the neotropics (n = 11 studies) suggested that pollen dispersal was widespread among low-density tropical trees, ranging from a mean of 200 m to over 19 km for species pollinated by small insects or bats (Ward *et al.*, 2005). For example, for the predominantly outcrossing, insect-pollinated species, *Cordia alliodora*, most pollen originated from within 75 m of the mother tree, but a rare allele indicated a low but substantial proportion of pollen movement from as far as 280 m, and there was extensive overall gene flow with neighbourhood areas as large as 7 ha. (Boshier *et al*., 1995 a, b).

On top of the effective gene flow via pollen, it is very plausible to assume effective seed dispersal in *C. africana*. Edible fruits such as that of *C. africana* can be carried endozoochorously over long distances mainly by some migratory animals and birds. Loveless (1992) confirmed that zoochorous tropical tree species show on average a lower degree of differentiation (G_{ST} = 0.05) than species with abiotic means of seed dispersal (barochorous and anemochorous species; G_{ST} = 0.138).

Though the genetic differentiation in *C. africana* happens to be intermediate, the Analysis of Molecular Variance (AMOVA) at AFLP revealed significant differentiations (Φ_{ST}, 0.072, $p < 0.001$) among the populations. The majority (over 92% of the total variation detected) of the variation was contained within populations, which was due to differences among individuals within each population. This observation is in agreement with the argument that long-lived and outcrossing taxa retain most of their genetic variability within populations (Nybom and Bartish, 2000).

A further analysis of the structuring of the populations in terms of genetic distances over the geographical distances in a Mantel test revealed a significant correlation (r = 0.350, $p < 0.001$) between geographic and genetic distances. Furthermore, the highest genetic distance (Nei's genetic distance and Φ_{ST}) was recorded between the northernmost and the southernmost populations, and AMOVA revealed significant differentiation among the geographical units (the north, the SW and the SE). These findings imply the existence of isolation-by-distance among the populations of *C. africana*. In addition, there were significant correlations of the genetic and geographic distance matrices in the SW-north (r = 0.544, $p < 0.001$) and in the SE-north populations (r = 0.307, $p < 0.01$). However, there was no significant correlation among the SE-SW populations or within the north, within the SW or within the SE populations. The significant correlations of the geographic and genetic distance matrices in the SW-north and in the SE-north populations may depict that the genetic distance among the populations is mainly explained in the north-south axis. Population genetic studies covering the whole geographic ranges of *C. africana* in Tropical Africa and Tropical Arabia may reveal even stronger population differentiations and high correlation between the geographic and genetic distance matrices in the species. High population differentiation (Φ_{ST} = 0.73) was reported (short communication) for another Afromontane tree species, *Prunus africana*, in a study that covered all the geographical range (including 1 population from Ethiopia) of the species in six countries employing

RAPD (Dawson and Powel, 1999). Likewise, the isozyme analysis of 30 populations of *Faidherbia albida* sampled from the entire African continent revealed a very great differentiation (G_{ST} = 0.44), thereby asserting the argument of high differentiation with high geographical distance. The publication by Feyissa *et al*. (2007) on another Afromontane tree species, *Hagenia abyssinica* (Bruce) J. F. Gmel (dioecious, outcrossing, wind-dispersed and late successional tree species), in Ethiopia using ISSR revealed significant population differentiation (G_{ST} = 0.25; a comparatively higher value when compared to *C. africana*) and a significant correlation between genetic and geographic distances as well (r = 0.375, $p < 0.01$; which is a comparably identical value to *C. africana*). Even stronger differentiations were found in *Coffea arabica* using microsatellites (G_{ST} = 0.577; Silvestrini *et al*., 2006) and using RAPD, (Shannon-Weaver index of 0.35; Aga *et al*., 2003), which is typical for such a predominantly selfing species.

The hierarchical AMOVA also revealed significant differentiations between altitudinal groups (lowland, intermediate, and highlands) in the *C. africana* populations. The lowland populations (< 1500 m) were entirely from the transitional rainforest region in the SW, the intermediates were two populations (Didessa and Yayu in the SW) whose altitudes transgress the highland-lowland boundary, and all the remaining populations represented the highlands. Since the two intermediate populations are basically located on the SW highlands, AMOVA was run recognising only highland (hence including Didessa and Yayu in this group) and lowland provenances, with significant differentiation being obtained between these two groups, too (Φ = 0.0147). In Ethiopia, the highlands are categorized broadly as temperate (mean annual temperature = 16 to 20°C) and cool zones (mean annual temperature = 10 to 16°C); *C. africana* thrives best in the former zone. The temperature in the lowlands is hot (mean annual temperature = 20 to 29°C) and the areas are either dry or tropical, and the transitional rain forest region is located in the tropical part. Therefore, the significant differentiation of the lowland and the highland provenances may be related to such climatic differences.

In comparison, the AMOVA did not reveal significant differentiation among the existing tree seed zones. The existing tree seed zone of the country is a general zonation system covering all tree species delineated using data on landform, climate, and natural vegetation, and it was meant for setting guidelines for safe transfer of tree seeds (Aalbæk, 1993). The failure of the tree seed zone to reflect the population structure of *C. africana* can be related

to the fact that the seed zoning was only general and not species specific, and may as well indicate the existence of substantial gene flow through effective seed dispersal circumventing the local ecological differences. Therefore, it seems that the existing general seed zonation is too detailed for *C. africana*. According to the OCED (1974), the definition for seed zone of an individual species is: for a species, subspecies or distinct variety, the region of provenance is the area, or group of areas subject to sufficiently uniform ecological conditions on which are found stands showing similar phenotypic or genetic characters. However, due to the lack of or limited information on genetic variation and adaptability of individual species to non-native sites, general seed collection guidelines based on biological, climatological, and/or geographical criteria have been implemented (Aalbæk, 1993). In recent years, however, in an effort to be selective in seed zonation of tree species, Post *et al.* (2003) has developed hardwood seed zones using a geographic information system.

Like in the case of the existing seed zones of Ethiopia, no significant differentiation was found among populations that come from various natural ecosystems, the delineation of which was mainly based on the type of the natural vegetation. Of the eight natural ecosystems of the country, Dry Evergreen Montane Forest and Grassland complex and Moist Montane Forest ecosystem (from which most of the populations of *C. africana* were sampled) are the two major ecosystems that cover the plateaus of the country. The natural ecosystems were identified and described in the National Biodiversity Strategy and Action Plan document, which envisages the establishment of effective systems that ensure the conservation and sustainable use of the biodiversity in Ethiopia (IBC, 2005).

Chapter 3: Genetic effects of forest fragmentation in SW Ethiopia

3.1 Introduction

Habitat fragmentation of trees and other living organisms as an almost inevitable consequence of large-scale forest loss has emerged as a global concern (Finkeldey and Hattemer, 2007). The alarmingly high loss of tropical forests, which is estimated as 100,000 – 200,000 km^2 per year, has remained unabated (Katzman and Cale, 1990). The deforestation in many cases reduces and fragments the once continuous forests into mosaics of small habitat 'islands' embedded within a human-modified matrix (Bierregaard *et al*., 1992) often with little physical and limited functional resemblance to the original (Young and Boyle, 2000). The loss of the primary habitat, therefore, poses the single most important danger for biodiversity (Lira *et al*., 2003). For trees, degradation of primary habitat results from two main processes, fragmentation of forest into patches and disturbance of habitat following extraction processes (Lowe *et al*., 2005). Theoretically, fragmentation and disturbance lead to an erosion of genetic variation, reduced gene flow, elevated inbreeding and increased random genetic drift (Hattemer and Melchoir, 1993; Young *et al*., 1996). However, many tree species (White *et al*., 1999, Dick 2001) show evidence of substantial gene flow among fragmented populations. As a consequence, the genetic effects of fragmentation are apparently not so predictable as they are determined by the scale of fragmentation, the biology of the affected species and their associated pollinators and dispersers (Young and Boyle, 2000).

The objective of this study was to assess the genetic impacts of forest fragmentation into coffee plantations and farmlands in the transitional rain forest region of SW Ethiopia. The following hypotheses were tested: (1) genetic variants (alleles) are homogenously distributed among the subpopulations in the metapopulation, (2) there is a random association of loci in the subpopulations, and (3) subpopulations exhibit a similar level of genetic variation.

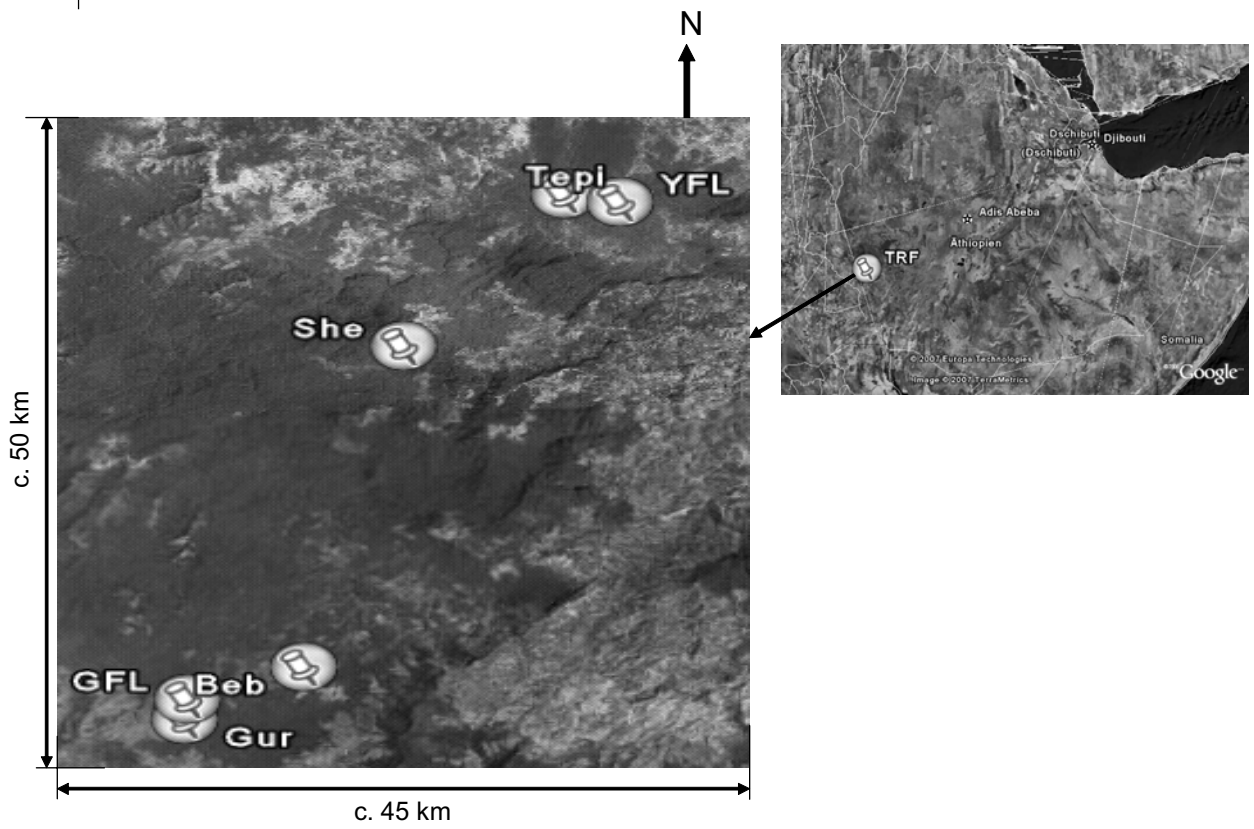
3.2 Materials and Methods

3.2.1 Sampling and plant material collection

Sampling was conducted in the Transitional Rain Forest area (Friis, 1992) in SW Ethiopia, which occurs in the altitudes between 500 and 1500 m. The transitional rain forest is a unique ecosystem that lies between the broadleaved Afromontane Rain Forest and the Guino-Conglian Type lowland forest. The forests Guraferda, Sheko and Yeki, which were demarcated and designated as state forests in 1988, form a forest continuum in the area. The three forests have a total area of about 3,400 km^2. The mean annual rainfall and the mean annual temperature in the area (at Yeki) are 2000 mm and 20-25^oC, respectively. The local people in the area have an ancient culture of non-destructive shifting cultivation in which they leave out the canopy trees and farm the forest for some years and then shift to another area (own assessment). However, the transitional rain forest area has been under intensive fragmentation following the establishment of state-owned coffee plantations Bebeka and Tepi, which have been operational since 1978 and 1979, respectivelly, and cover a total area of about 26,000 ha. Furthermore, a parallel timber extraction in the area began in 1985 with the introduction of Bebeka sawmill and three other mills, which intensified the forest disturbance and resource depletion. In addition, there has recently been more investment in coffee and tea plantations, and there are also new settlement enclaves (Abebe and Holm, 2003), which in general have shown an apparent booming of economic activities in the area in the last few decades. Therefore, it is of utmost importance to evaluate the effect of such forest fragmentation and disturbance on the genetic variation of economically important tree species such as *C. africana* in such a unique ecosystem in the country. To this effect, a total of six subpopulations were sampled from the area: forest, coffee plantation, farmland (two representatives each), i.e. with low, medium and high tree densities, respectively. They can be regarded as subpopulations forming a big metapopulation. Hence two farmland populations (designated as GFL and YFL) were sampled on top of the four populations included in the overall genetic variation study from this part of the country (namely, Guraferda, Bebeka, Sheko, and Tepi). The sampling and the plant material collection at GFL and YFL were carried out in a similar fashion as in the rest of the populations (see Chapter 2). The following picture (Figure 10) shows all the populations sampled from the transitional rain forest area. The geographical distance among the populations ranges from about 2 km to about 47 km (Table 6).

Table 6 Geographic distance (km) among subpopulations in the transitional rain forest region

GFL	4				
Gur	7	2			
She	25	27	31		
Tep	39	42	45	14	
YFL	42	45	47	17	3
	Beb	GFL	Gur	She	Tep



Gur & She = natural forests, Beb & Tepi = coffee plantations, GFL and YFL = farmlands, TRF = transitional rain forest

Figure 10 Location of populations sampled at a transitional rainforest in SW Ethiopia

3.2.2 Genetic Data Generation

The DNA isolation and the AFLP data generation were conducted as indicated in Chapter 2.

3.2.3 Analysis of variation within and among subpopulations

The percentage of polymorphic loci (PPL), Nei's diversity (Hj), Analysis of Molecular Variance (AMOVA), Nei's genetic distance and correlation of geographic and genetic distance (Mantel test) were conducted as described in Chapter 2.The homogeneity test of individual AFLP putative locus using X^2-tests at the 1% level of significance, computation of Nei's unbiased measure of genetic identity and genetic distance (Nei, 1978) and the construction of a UPGMA dendrogram was done in Popgene 1.31 (Yeh *et al.*, 1999). Ordination analysis was carried out using principal coordinate analysis (PCO) setting in GenAlEx (Peakall and Smouse, 2001).

Furtheremoe, a model-based clustering method was implemented in the Structure Program (Pritchard *et al*., 2000) for inferring population structure by assigning individuals to *a priori*-determined number of clusters. We considered AFLP phenotypes as genotypes by assuming the present band as genotype "1, -9" and the absent band as another genotype "2, -9"; in both cases, the genotypes are composed of a known and unknown allele, respectively. Missing data were also represented by –9. In addition, we assumed a No-Admixture Model, as is recommended for dominant markers in the Structure Program. A burn-in period of 100, 000 and 10,000 MCMC was used (Markov chain Monte Carlo). We varied the number of clusters (K) between 2 and 7, with 5 iterations each. After running the Structure Program, the optimum value of K was determined using an *ad hoc* statistic delta K based on the rate of change in the log probability of data between successive K values, as described in Evanno *et al*., (2005).

3.2.4 Association among AFLP loci

Since the dominant nature of AFLP does not allow gametic reconstruction, only the associations among AFLP loci were assessed. The analyses were done in exact tests of linkage disequilibrium setting in Arlequin 3.01 (Excoffier *et al.*, 2006) for all polymorphic pairs of loci in every population using a Marcov chain length of 10000 and dememorization step of 1000, taking a pair of loci at a time. The value of 'linkage disequilibrium' (non-random association among pairs of loci, %LD) for each locus in each population was then computed as:

$$LD(\%) = \frac{\sum_{i=1}^{n} Li}{n},$$

where n = the number of polymorphic loci within each population, L = number of 'linked loci' (non-randomly associated loci) per the ith polymorphic locus ($p < 0.05$). Then the resulting values at each locus were averaged over all the polymorphic loci in each population.

3.3 Results

3.3.1 Within subpopulation genetic variation and association among loci

The percentages of polymorphic loci (PPL) were the highest in the less affected subpopulations (forests and coffee plantations) and the lowest in the heavily disturbed populations (farmlands). The remnant natural forests, the coffee plantations and the farmlands harbour comparable levels of genetic diversity; the diversity at Tepi coffee plantation being the least and at Meti coffee plantation and Sheko forest being the highest. However, the analyses of the number of loci under significant non-random association revealed relatively elevated values in the farmlands (Table 7).

Table 7 Genetic variation and non-random associations among pairs of AFLP loci ('LD') in *Cordia africana* in the transitional rain forest, SW Ethiopia, in decreasing order of 'LD' ($p < 0.01$)

Population	Forest Type	n	PPL	Hj	S.E.(Hj)	N.L*	'LD'(%)*, $p < 0.05$	'LD'(%)*, $p < 0.01$
YFL	Farm	22	86.7	0.313	0.017	64	17.5	11.3
GFL	Farm	21	86.7	0.304	0.017	69	15.1	7.5
Bebeka	Coffee	23	87.8	0.301	0.017	69	12.4	7.4
Tepi	Coffee	22	88.9	0.290	0.015	72	8.1	6.2
Sheko	Forest	22	88.9	0.320	0.016	70	11.3	6.1
Guraferda	Forest	20	91.1	0.303	0.015	67	11.2	5.9

*N.L. number of pairs of loci under significant non-random associations; LD (%), percentage of loci under significant non-random association

3.3.2 Genetic variation among subpopulations

The homogeneity test revealed that the majority of the loci (85.6%) were distributed homogeneously in the subpopulations, only 13 putative AFLP loci (86, 89, 118, 128, 136, 142, 144, 177, 178, 195, 207, 208 and 322) showed significant deviations ($p < 0.01$) from homogeneity (Appendix 5). The Nei's genetic distance varied from 0.010 to 0.041 (Table 8). The UPGMA dendrogram (Figure 11) constructed in POPGENE 3.1 grouped the farmland populations together and separated them from the other natural forest and coffee populations. The principal coordinate analysis (PCO) seemed to support the clustering as the clouds of the two farmlands were concentrated at two locations in the first principal axis, which explained 27% of the variation. The second axis explained 10% of the variation; hence, the two axes explain 35% of the variation among the individuals in the metapopulation (Figure 12, Appendix 6).

Table 8 Genetic identity and genetic distance among the subpopulations in SW Ethiopia

Population	Bebeka	GFL	Guraferda	Sheko	Tepi	YFL
Bebeka		0.989	0.990	0.991	0.982	0.989
GFL	0.011		0.980	0.976	0.960	0.988
Guraferda	0.011	0.020		0.984	0.984	0.978
Sheko	0.010	0.024	0.016		0.981	0.975
Tepi	0.018	0.041	0.016	0.019		0.965
YFL	0.011	0.012	0.022	0.026	0.035	

Nei's genetic identity (above diagonal) and genetic distance (below diagonal)

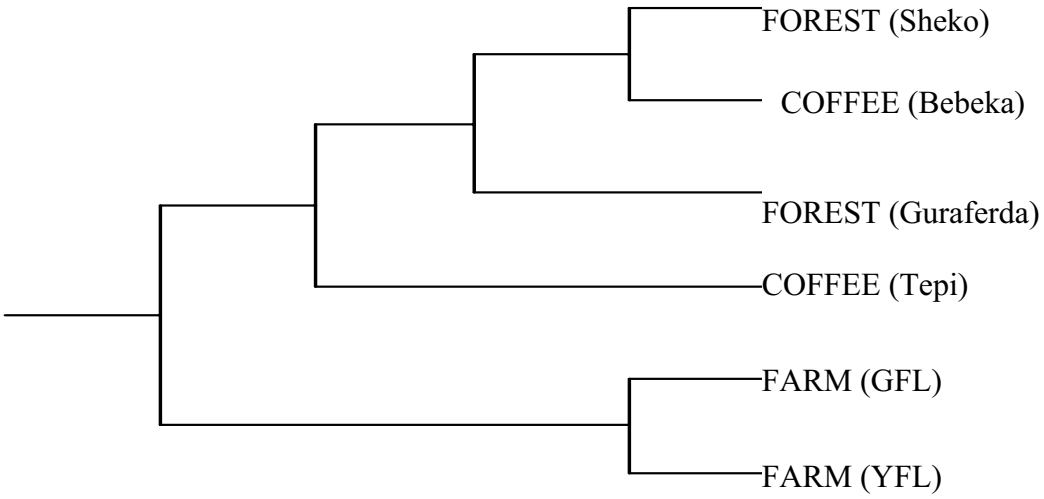
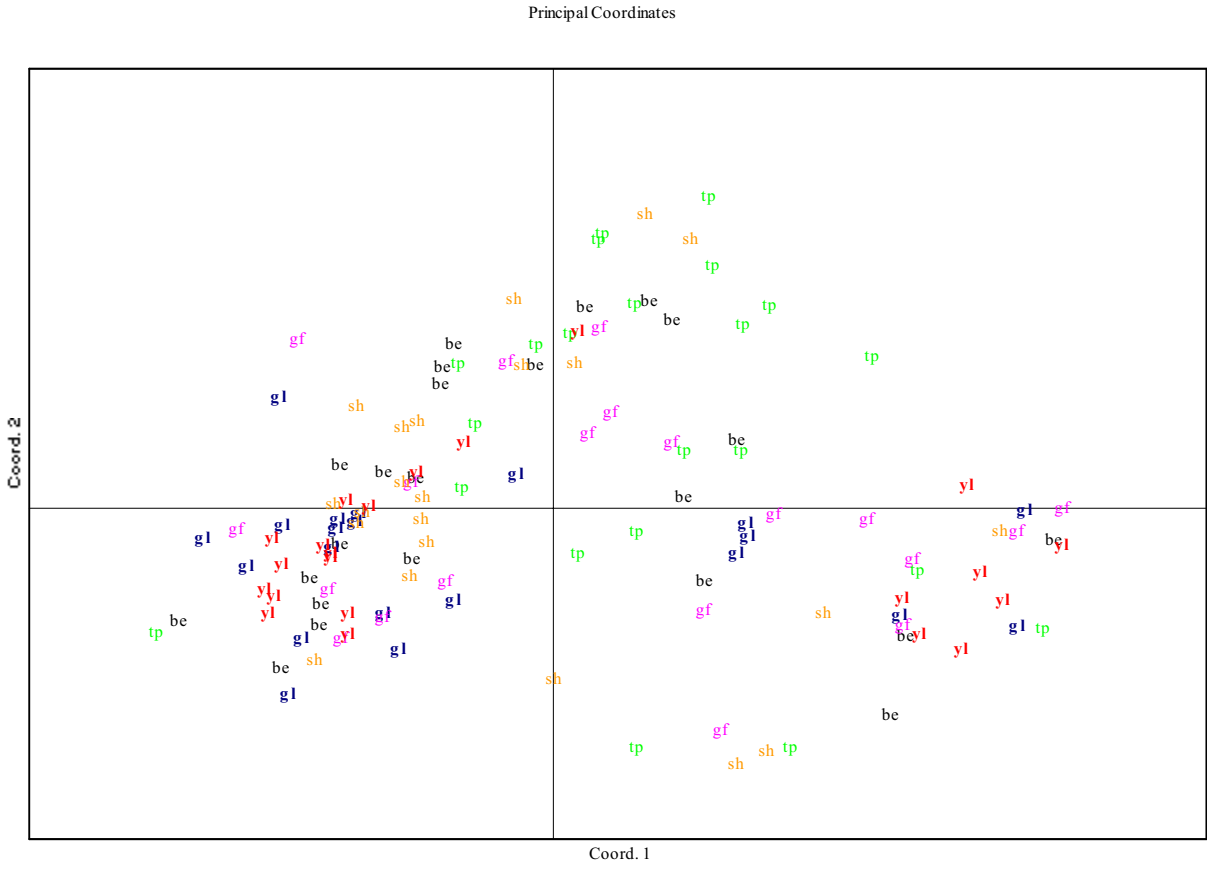


Figure 11 UPGMA dendrogram showing the genetic relatedness of sub-populations in the transitional rain forest, SW Ethiopia



tp = tepi, be = bebeka, sh= sheko, gl = GFL, yl = YFL and gf = Guraferda

Figure 12 Principal coordinates analysis among the six subpopulations in SW Ethiopia

There was a generally very low but significant genetic differentiation among the subpopulations (Φ_{ST} = 0.04, p < 0.001). However, there was no significant differentiation among fragmentation levels (forest, coffee plantation and forest). Thus, AMOVA did not reveal any significant differentiation either between the groups (the two farms and the rest) or between the two locations (Tepi, Bebeka), though it revealed significant differentiations among the subpopulations at Tepi (Table 9). The population pair-wise comparisons using fixation indices (Φ_{ST}) revealed significant differentiations (*p < 0.01*) in only 33% of the cases (Table 10). The Mantel test in NTSYSpc2.0 showed no correlation between the geographic and genetic distances.

Table 9 Analysis of molecular variance among subpopulations in the transitional rainforest area, SW Ethiopia

Source of variation	df	SSD	MSD	Variance component	Percentage of variation	*p-value*
Among groups (Forests, Coffee plantations, Farms)	2	48.88	24.44	0.07	0.54	*ns*
Among populations	3	64.78	21.60	0.46	3.84	*< 0.01*
Within populations	124	1433.07	11.56	11.56	95.62	*< 0.001*
Among clusters (Forests + Coffee plantations, Farms)	1	31.15	31.15	0.18	1.51	*ns*
Among populations	4	82.53	20.63	0.42	3.44	*< 0.001*
Within populations	124	1433.07	11.56	11.56	95.05	*< 0.001*
Among locations (Bebeka & Tepi)	1	21.13	21.13	-0.03	-0.26	*ns*
Among populations	4	92.55	23.14	0.53	4.44	*< 0.001*
Within populations	124	1433.07	11.56	11.56	95.82	*< 0.001*
Among populations around Bebeka	2	32.42	16.21	0.23	2.03	*ns*
Within populations	61	686.02	11.25	11.25	97.97	*< 0.001*
Among populations around Tepi	2	60.14	30.07	0.83	6.53	*< 0.001*
Within populations	63	747.05	11.86	11.86	93.47	*< 0.001*

Table 10 Pairwise population differentiation (Φ_{ST}) and their respective significance level in SW Ethiopia

GFL	0.005 ns				
Guraferda	0.027 ns	0.039 ns			
Sheko	0.023 ns	0.040 *	0.041 *		
Tepi	0.050 **	0.092 ***	0.033 *	0.052 **	
YFL	0.016 ns	0.022 ns	0.053 *	0.062 **	0.081 **
Population	Bebeka	GFL	Guraferda	Sheko	Tepi

ns = not significant, *, **, *** = significant at 5%, 1% and 0.1% level, respectively

3.3.3 Population structure based on assignment of individuals

The optimum value for the number of inferred clusters (k) was found to be 2; most of the individuals could be assigned to either one of these clusters accurately (Figure 13). However, the assignment of individual genotypes to the two clusters did not follow the *a priori* defined subpopulations; individuals from all the subpopulations were represented in both clusters (Table 11).

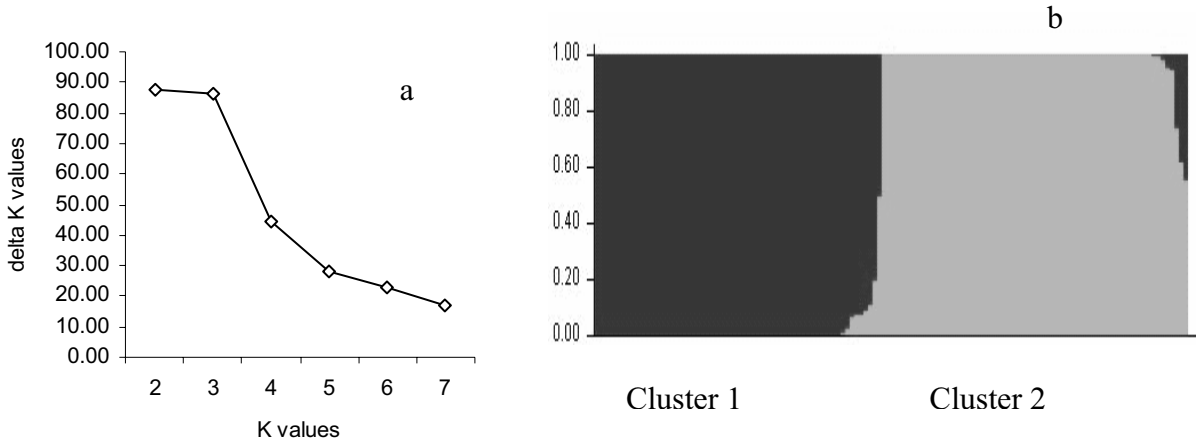


Figure 13 The optimum number of clusters as indicated by the highest peak (a) and membership of individuals to each of the two clusters (b)

Table 11 Proportion of each a priori defined subpopulation in each of the inferred clusters

No	A priori defined subpopulations	Number of individuals	Inferred cluster 1	Inferred cluster 2
1	Bebeka	23	0.41	0.59
2	GFL	21	0.29	0.71
3	Guraferda	20	0.64	0.56
4	Sheko	22	0.40	0.60
5	Tepi	22	0.82	0.18
6	YFL	22	0.36	0.64

3.4 Discussion

The transitional rainforest is situated in the SW escarpments between 500 and 1500 m and receives about 2000 mm annual rainfall. In this forest area, two state-owned coffee plantations exist and small-scale commercial selective logging and sawmilling have been practiced (Abebe and Holm, 2003). For trees, degradation of their primary habitat results from two main processes: fragmentation of forest into patches following clearance, and disturbance of habitat following extraction processes, such as selective logging (Lowe *et al*., 2005). Though severe impact on genetic structure and genetic variation is predicted as a consequence of habitat loss, degradation and silvicultural management practices (Barrett and Kohn, 1991; Finkeldey and Ziehe, 2004), particular life history characteristics of some species (e.g. mating system, pollen and seed dispersal mechanisms) may help to mitigate against genetic diversity loss in post-fragmentation and disturbance landscapes (Dick, 2001; White *et al*., 2002) as well as in manipulated forest stand structures.

The gene diversity observed in the six subpopulations (two remnant forests, two coffee plantations which were derived from part of the natural forest by systematic removal of trees some three decades ago, and two scattered populations on farmlands) of *C. africana* in the rain forest region did not reveal any trend of loss of diversity with the degree of fragmentation and level of disturbance. The finding on the diversity of *C. africana* conforms to theoretical and simulated predictions where genetic diversity will be lost slowly in post-impact generations, which in the case of trees may take decades or centuries, but the impact on inbreeding may be observed immediately after disturbance. The majority of studies carried out in the Neotropics, which examined genetic diversity

consequences (nine out of thirteen studies) also found no significant differences between impacted and control populations (Lowe *et al*., 2005). No clear short-term impact of selective logging in terms of levels of diversity, allelic richness, outcrossing rates, biparental inbreeding, pollen pool differentiation, pollen dispersal distance, allele frequencies in adults and the pollen pool, or spatial genetic structure (SGS) among adult trees was noticed on an Amazonian tree population of *Carapa guianensis* Aubl. (Cloutier *et al*., 2007). However some studies that assessed progeny inbreeding (six out of eight), reproductive output (seven out of ten), and fitness (all six) highlighted significant impacts (Lowe *et al*., 2005).

The trend of percentage of polymorphic loci (PPL) (though there were no pronounced differences) reflected the severity of fragmentation to some extent as the farmlands exhibited the least PPL, which is in agreement with the argument that measures of allelic richness are suitable for assessing short-term diversity loss, but measures of diversity are more appropriate for assessment of longer-term loss (Lowe *et al*., 2005). In an investigation on the impact of fragmentation on populations of mahogany, *Swietenia humilis*, using 10 microsatellite loci, the distribution of the genetic variation detected within the fragmented population was largely a reflection of the previous more continuous forest, than indicative of the impact of more recent forest disturbance. However, the effect of fragmentation was evident through a loss in the fragments of the low-frequency alleles detected in the continuous stand. Furthermore, a high proportion of the microsatellite loci in this species, *S. humilis*, did not lie within the HW expectations due to an apparent non-random mating/inbreeding within the fragments (White *et al.*, 1999). The dominant nature of the marker employed in the present study did not allow the carrying out of such a HW test on the populations of *C. africana*. In another study on hermaphroditic tropical tree species, equivalent numbers of alleles per ha were observed between continuous and fragmented forests; however, significant inbreeding was most associated with the latter (Aldrich *et al.*, 1998). A transient reduction in population size could result in a substantial loss of alleles due to a reduction in the representation of the original gene pool (Frankel and Soulé, 1981). In a liverwort, *Radula flaccida*, the effect of habitat fragmentation was reflected in neither neutral genetic diversity nor in population structure, but was rather associated with an increase in linkage disequilibrium (LD) (Zartman, *et al.*, 2006). The 'LD' (non-random association among the polymorphic loci) in each population of *C. africana* showed that the two farmland-populations had relatively the highest percentage of

significant 'LD'. Elevated linkage disequilibrium among pairs of polymorphic loci may indicate inbreeding (Ohta, 1982), recent population bottlenecks or admixture at subpopulation level, and high population turnover at a metapopulation level (Tero *et al*., 2003). The detection of significant LD among several pairs of loci coupled with the existence of only two alleles per locus at an intermediate frequency in *Pinus maximartinezii* using isozymes was also related to an extreme bottleneck (Ledig *et al.,* 1999). However, in view of the current study, the highest 'LD' in the farmlands (mainly in YFL) may not be completely related to inbreeding since it seems that the trees in both the continuous and scattered situations are remnants and do not reflect the generation after fragmentation. Nevertheless, such a trend can be considered as an indicator of a non-random selection of preferred genotypes or could indicate future negative genetic impacts of the fragmentation and forest disturbance.

Founder effects, small population size and migration can all create and maintain linkage disequilibria (Lynch and Walsh, 1998). In addition to the elevated non-random association among AFLP loci in farmlands, the genetic similarity of the two farmlands in comparison to the remaining coffee and forest populations, as depicted on the UPGMA dendrogram, might be related to some kind of preferential removal of some genotypes, which apparently took place when transforming the forest to agricultural land use. However, the removal of individuals, while the coffee plantations were being established was less severe since it took place systematically and there still exists a high number of trees. As a consequence, the coffee plantations still showed genetic identity to the natural forests. In addition, a coffee plantation gives protection to the remnant trees and hence happens to be a sound land use system for the preservation of the gene pool of *C. africana* and other tree species, even though natural regeneration is hampered in such a land-use system. Previously, the forests in this part of the country used to be either intact or were being utilized by a non-destructive method of shifting cultivation. None the less, the establishment of the state-owned coffee plantations has marked an era of high forest depletion in the area, due to the accompanied increase in the number of inhabitants and economic activities, which has made the remnants even more susceptible to further depletion.

Hierarchical AMOVAs (1) recognizing three groups; namely farmlands, coffee plantations and forest, (2) comparing the three subpopulations around Bebeka with the three subpopulations around Tepi, which are within 47 km horizontal distance, (3) three

subpopulations around Bebeka that are within 7 km among each other, and (4) three subpopulations around Tepi that are within 14 km among each other, gave a significant differentiation only in the last one. Pairwise comparison revealed significant differentiations ($p < 0.01$) in five of the fifteen cases, four of which were attributed to the pairwise comparisons in which Tepi was involved, and the remaining one was between Sheko and YFL. Therefore, most of the pairwise comparisons (66%) did not reveal any significant differentiation among the subpopulations of this specific metapopulation.

The chi square test proved that most of the loci analyzed were distributed homogenously among the subpopulations. The No-Admixture-model-based analysis of the population structure assuming correlated allele frequencies among the subpopulations also indicated a weak population genetic structure, since the assignment of individuals to the *a posteriori* defined populations did not follow the *a priori* defined subpopulations.

The low differentiation among the subpopulations can be related to the existence of a wide-scale gene flow prior (and/or post) habitat fragmentation in the transitional rainforest area as a whole. Despite the typically low population densities and animal-mediated pollination of tropical forest trees, outcrossing and long-distance pollen dispersal is the norm (Ward *et al*., 2005). Cascante *et al.* (2001) discovered that genetic diversities of progenies from continuous and isolated trees were comparable; however, these authors identified a higher effective self-fertilization rate and inbreeding coefficient in the latter. The low differentiation among the subpopulations in the transitional rain forest (for example, when compared to the differentiation at the national level indicated in Chapter 2) may be due to high gene flow (Burczyk *et al*., 2004), which may require complex models to quantify it (Adams, 1992, Burczyk *et al*., 2006)..

Chapter 4: Variation at chloroplast microsatellites

4.1 Introduction

Two organelle genomes coexist in the cytoplasm of plant cells: the mitochondrial genome and the chloroplast genome. Both these two genomes undergo virtually no recombination and have usually different modes of inheritance varying from strictly maternal to strictly paternal (Harris and Ingram, 1991; Röhr *et al*., 1998). The mode of inheritance of the chloroplast genome, cpDNA, is usually maternal in angiosperms (Reboud and Zeyl, 1994), and hence, it is transmitted only through seeds. However, there is evidence for biparental inheritance in about one-third of the genera that have been investigated (Birky, 1995) and there are also some examples of paternal inheritance in some species (Chat *et al*., 1999). The maternally inherited cpDNA has less potential for gene flow than nuclear genes, which can also move by pollen dispersal. As a result of this, genetic variation in the chloroplast genome is often more geographically structured than that of the nuclear genome (Cavers *et al*., 2003). Furthermore, as the rate of cpDNA sequence evolution is slow (Wolfe *et al*., 1987), the observed patterns of genetic variation reflect the outcome of processes over long time scales (Ennos *et al*., 1999), which makes cpDNA ideal for studying historical patterns of gene flow, in particular migration and colonization (Cavers *et al*., 2003), and also hybridization and introgression in related plant taxa (Petit *et al*., 2002; Curtu *et al*., 2007). The vast majority of plant phylogeographic studies, which aim at studying the principles and historical processes governing the geographical distributions of genealogical lineages, employ cpDNA markers (Weising *et al*., 2005).

The size of cpDNA generally ranges from 120 kilobases (kb) to 180 kb. The genomes of cpDNA typically have between 100 to 120 genes, many of which code for components of photosynthesis (Murray *et al*., 2000). One of the methods to investigate cpDNA variation is through the use and analysis of the so-called microsatellites, also known as “simple sequence repeats.” Microsatellites are abundant polymorphic elements of eukaryotic nuclear as well as organelle genomes and consist of short DNA sequence motifs reiterated in tandem (Wang *et al*., 1994). Size variation of the microsatellites is due to a variable number of repeat motifs, and can be visualized by PCR with pairs of flanking primers and electrophoretic separation of the amplification products (Weising *et al*., 1998).

Microsatellites are comparatively rare in organelle DNA (Wang *et al*., 1994) and are normally composed of short poly (A) or poly (T) tracts, with maximum sizes of about 20 bp (Weising *et al*., 1998). The chloroplast genome of a wide range of plant species can be analyzed using some sets of universal chloroplast microsatellite primers (Weising and Gardner, 1998) and/or other sets of universal primers (Taberlet *et al*., 1991; Demesure *et al*., 1995).

The analyses of cpDNA data involves combining the polymorphisms observed at various loci into distinct haplotypes, followed by the analysis of haplotype distribution and frequencies across different geographical regions, quantification of the genetic divergence, and evaluation of the genetic relationships between haplotypes (Weising *et al*., 2005).

The analysis and understanding of the evolution of chloroplast microsatellites also involves different models, as is the case for nuclear microsatellites. According to the infinite alleles model (IAM; Kimura and Crow, 1964), each mutation creates a novel allele at a given rate. In effect, this model does not allow for homoplasy, implying alleles are identical by descent (Balloux and Lugon-Moulin, 2002). The other extreme model is the stepwise mutation model (SMM; Kimura and Ohta, 1978) according to which each mutation creates a novel allele either by adding or deleting a single repeated unit of the microsatellite with equal probabilities in both directions. Hence, alleles of very different size will be more distantly related than alleles of similar sizes (Balloux and Lugon-Moulin, 2002). Estimates of both population differentiations and the proportion of migrants can vary depending on the model chosen for the analyses.

The aim of the current study was to determine the population structure of *C. africana* in Ethiopia using universal chloroplast microsatellites (cpSSRs).

4.2 Materials and Methods

4.2.1 Chloroplast Microsatellites

4.2.1.1 Primers employed and samples analysed

All ten fluorescence-labelled consensus chloroplast microsatellite primers (*ccmp*1, *ccmp*2, *ccmp*3, *ccmp*4, *ccmp*5, *ccmp*6, *ccmp*7, *ccmp*8, *ccmp*9 and *ccmp*10, developed from tobacco

cpDNA by Weising and Gardner, 1998) were tested on eight individuals to amplify DNA in *C. africana*. Those primers that gave amplification products were then tested on 60 samples (three samples per population) to detect polymorphisms. Finally, only the polymorphic primers were employed to analyze all 22 *C. africana* populations (mean sample size, n = 9.2), and the resulting variants were combined to form haplotypes for further data analysis.

4.2.1.2 Polymerase chain reaction (PCR) and genotyping

The 15 µl PCR mixture contained 5-10 ng template DNA, 10 x PCR buffer (containing 15 mM $MgCl_2$), 10 pmol forward and reverse primers each, 10 mM dNTPs, 1U of Taq polymerase (Qiagen, Hot Star Master Mix, Hilden, Germany) and 1.5 µl Q-solution (Qiagen). The PCR was carried out with an initial activation (95°C for 15'), thirty-five cycles of denaturation (94°C for 1'), annealing (50°C for 1') and elongation (72°C for 1'), and final extension (72°C for 10'). The PCR products were electrophoresed on the automated capillary sequencer ABI Prism 3100® Genetic Analyser with polymer 3100 POP-6™ (Applied Biosystems, New York, USA). The length of electrophoresis products was measured with the help of the internal size standard (GS ROX-500). Finally, individual alleles were analysed using Genescan © version 3.7 (Applied Biosystems) and genotyped using Genotyper © 3.7 NT (Applied Biosystems) (Figure 14). The repeatability of the observed patterns was checked over 12 samples.

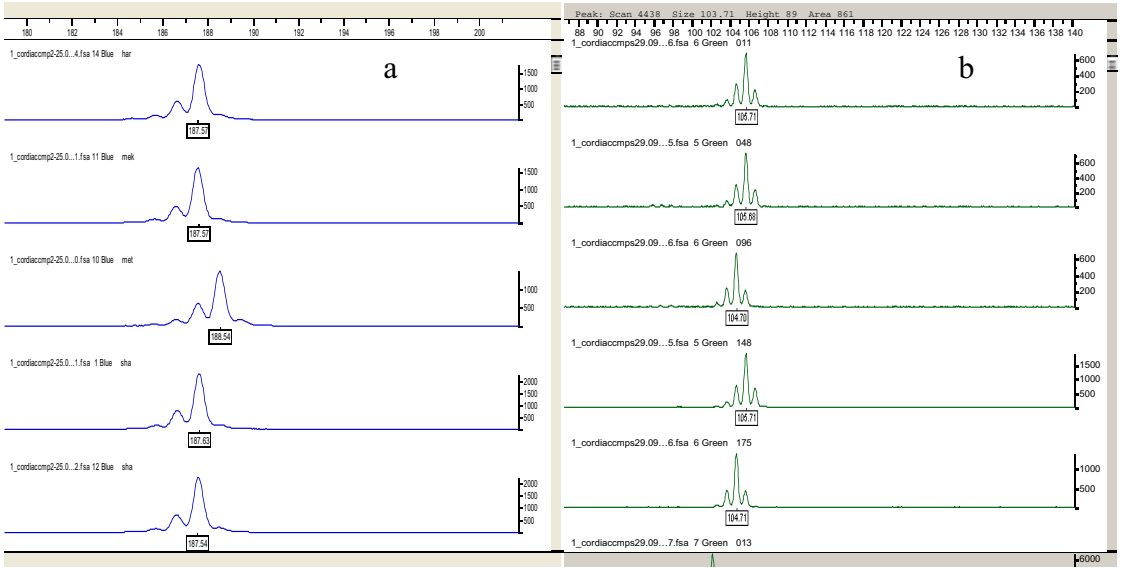


Figure 14 Electropherograms showing size variation in chloroplast microsatellite *ccmp*2 (a) and *ccmp*10 (b) as visualised by Genscan 3.7 and Genotyper 3.7.

4.2.2 Genetic diversity and population differentiation

The four variants identified at the two cpSSR loci were combined to three haplotypes (Table 12). Pie charts showing the frequency of haplotypes in each population were drawn and plotted on the map of the country to show the geographic pattern of distribution of haplotypes. Genetic diversity within populations (H_S) and total diversity (H_T), and population differentiation values were computed using the Permut-cpSSR programme (Pons and Petit, 1996) with weighted distance between haplotypes performing 10,000 permutations. The coefficient of gene differentiation (G_{ST}; Nei, 1973) is given as:

$$G_{ST} = \frac{D_{ST}}{H_T} = \frac{H_T - H_S}{H_T}, \qquad \text{(Nei, 1973)}$$

where D_{ST} is the average gene diversity between subpopulations.

Since microsatellites appear to follow a stepwise mutation model (SMM; Kimura and Otha, 1978), the differentiation measure, R_{ST}, was also used, devised explicitly on this mutation model:

$$R_{ST} = \frac{S - S_w}{S},$$

where S is the average squared difference in allele size between all pairs of alleles, and S_W, the average sum of squares of the differences in allele size within each subpopulations.

These two quantities (S and S_w), and hence R_{ST}, can be calculated from the variance of allele sizes, whereas F_{ST} will typically be derived from the allele frequencies (Balloux and Lugon-Moulin, 2002).

A third type of measure of differentiation for microsatellite data (N_{ST}) was used, which takes into account the similarities between the haplotypes as opposed to G_{ST} (given above), which makes use only of the allelic frequencies (Pons and Petit, 1996), and is given as:

$N_{ST} = V_T\text{-}V_S/V_T$

where V_T is the total diversity and V_S is within population diversity.

The unbiased Nei's genetic distance was computed among the populations in Popgene 1.31 (Yeh *et al.*, 1999), and Mantel test was done between the geographical and genetic distance matrices.

4.3 Results

4.3.1 Variation at cpSSRs

Out of the ten consensus chloroplast microsatellite primers, seven amplified DNA in *C. africana* (*ccmp*2, *ccmp*3, *ccmp*4, *ccmp*6, *ccmp*7, *ccmp*8 and *ccmp*10). However, only two of the primers (*ccmp*2 and *ccmp*10) were polymorphic (Table 12).

4.3.2 Haplotype variation

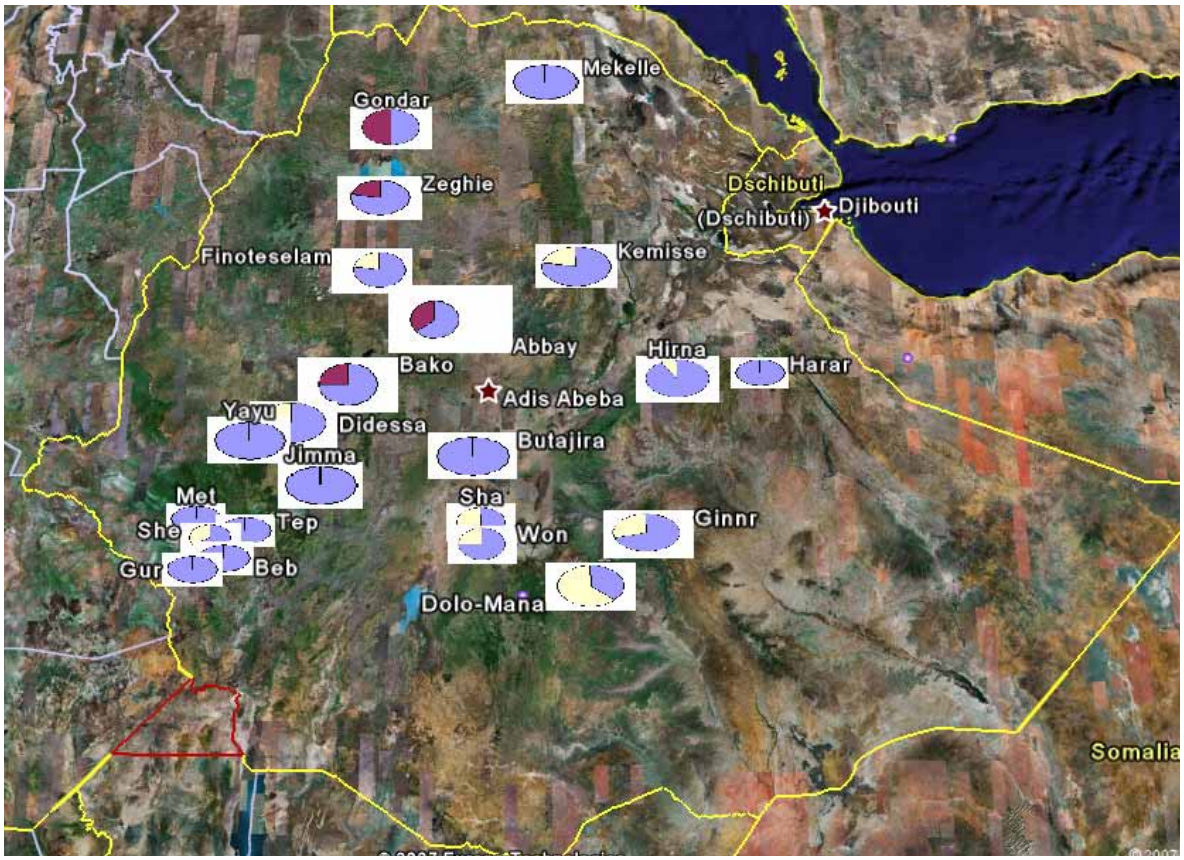
A total of three haplotypes were characterised from the two polymorphic cpSSR loci (*ccmp2* and *ccmp10*) (Table 13, Appendix 7).

Table 12 List of consensus chloroplast primer pairs and the respective sizes of the amplification products (in base pairs, bp) in *Cordia africana*

Code	Sequence alignment	Location	Repeat in tobacco	Size (bp)
*ccmp*2	5'-GATCCGGACGTAATCCTG-3'	5' to trnS	$(A)_{11}$	187/188
	5'-ATCGTACCGAGCGTTCGAAT-3'			
*ccmp*3	5'-CAGACCAAAAGCTGACATAG-3'	trnG intron	$(T)_{11}$	102
	5'-GTTTCATTCGGCTCCTTTAT-3'			
*ccmp*4	5'-AATGCTGAATCGAYGACCTA-3'	atpF intron	$(T)_{13}$	127
	5'-CCAAAATATTBGGAGGACTCT-3'			
*ccmp*6	5'-CGATGCATATGTAGAAAGCC-3'	ORF 77-ORF 82	$(T)_5C(T)_{17}$	101
	5'-CATTACGTGCGACTATCTCC-3'	intergenic		
*ccmp*7	5'-CAACATATACCACTGTCTAAG-3'	atpB-rbcL	$(A)_{13}$	123
	5'-ACATCATTATTGTATACTCTTTC-3'			
*ccmp*8	5'-TTGGCTACTCTAACCTTCCC-3'	rpl20-rps12	$(T)_6C(T)_{14}$	60
	5'-TTCTTTCTTATTTCGCAGDGAA- 3'	intergenic		
*ccmp*10	5'-TTTTTTTTTAGTGAACGTGTCA-3 '	rpl2-rps19	$(T)_{14}$	105/106
	5'-TTCGTCGDCGTAGTAAATAG-3 '	intergenic		

Table 13 List of haplotypes in *Cordia africana* based on chloroplast microsatellites

No	*ccmp*2 (bp)	*ccmp*10 (bp)	Haplotype	% proportion
1	187	105	Hap-A	13
2	187	106	Hap-B	81
3	188	106	Hap-C	6



Met = Meti, She = Sheko, Tep = Tepi, Beb = Bebeka, Gur = Guraferda, Won = Wondogenet, Sha = Shashemene

Figure 15 Geographic distributions of haplotypes of *Cordia africana* and their frequency distribution in Ethiopia

The majority (81%) of the individuals belonged to Haplotype-B, which was detected in all the populations analysed. Eight of the populations were fixed for Haplotype-B. Haplotype-A was limited to a few (18%) populations in the NW mountain massif (3 populations belonging to the northern unit, 1 population belonging to the SW unit), and was found

always in association with Haplotype-B. Haplotype-C had a wider distribution (41 % of the populations), and was also found in association with Haplotype-B (Figure 15). No population was fixed for either Haplotype-A or Haplotype C. The diversity within populations (H_S) and the total diversity (H_T) were 0.273 and 0.334, respectively. The population differentiation values (G_{ST}, N_{ST}, R_{ST}) were 0.182, 0.22 and 0.22, respectively.

4.4 Discussion

The assessment of cpDNA variation in several populations in the present study over a wide geographical range was presumed to be ideal for revealing some historical patterns of gene flow via seed in populations of *C. africana* in Ethiopia. For example, cpDNA has been shown to be a valuable tool for the identification of postglacial colonization routes of oaks in Europe (Petit *et al*., 2002a, b). Such kinds of phylogeographical studies would benefit by sampling as many populations as possible to ensure the inclusion of individuals from the overall natural distribution range of the target species (e.g. 2600 populations in Petit *et al*., 2002). In the present study, only three haplotypes were characterized in *C. africana* in Ethiopia in the 22 populations (every population had one or two haplotypes), which were located in an area covering a geographical distance of 800 km (in both the north-south and east-west axes), and in an altitudinal range of 1000 m. Exactly the same number of cpDNA haplotypes with a very similar distribution fashion (i.e. total of three haplotypes; one or two haplotypes in each population) has been reported for *Abies kawakamii* elsewhere (Shih *et al*., 2006).

Although most reported genetic surveys indicate a trend towards a fixation of cpDNA within populations in angiosperms (e.g. *Quercus* spp., Dumolin-Lapègue *et al*., 1997; *Corytophora alta*, Hamilton, 1999; *Cedrela odorata*, Cavers *et al*., 2003), this was not the case in *C. africana* (64% of the populations had mixed haplotypes in the current study), in *Abies kawakamii* (Shih *et al*., 2007), Dalbergia sissoo (two to three haplotypes per population; Pandey *et al.*, 2004) or in *Dicorynia guianensis* (only a single population was fixed, Caron *et al*., 2000). The chloroplast genome is also expected to show a high geographical structuring of populations within a given species. However, the distribution of the three haplotypes of *C. africana* over the country revealed a weak geographical pattern in contrast to other tropical outcrossing species (Cavers *et al*., 2003; Fontaine *et al*., 2004). Haplotype A (187/105) had a wider distribution in the north, SW and SE, with a

relatively high frequency in the SE populations. Haplotype B (187/106) was found in all 22 populations, and 8 of the populations were fixed for it. Haplotype C (188/106) was limited to the NW highlands.

No pattern of isolation-by-distance among populations of *C. africana* was observed at the cpDNA in the Mantel test. Generally the population differentiation among the populations of *C. africana* was very low (G_{ST} = 0.182) when compared to other tree species such as *Cedrela odorata* (G_{ST} = 0.96; Cavers *et al.*, 2003), various oak species (G_{ST} = 0.60 to 0.90; Cottrell *et al.*, 2002) and the endangered tropical species, *Caesalpinia echinata* (Φ_{ST} = 0.91; Lira *et al.*, 2003). In another study a moderate between-population differentiation (G_{ST} = 0.41), associated strikingly with the maintenance of high within-population diversity, was found in *Dicorynia guianensis* (Caron *et al*., 2000). Comparison of the differentiation indices (G_{ST}, N_{ST} and R_{ST}) at cpSSR of populations of *C. africana* with other broad-leaved woody species reveals that *C. africana* is one of the species with the lowest population differentiation (Table 14).

Table 14 Comparative population differentiations at cpSSR of *C. africana* and other four species

Species	**No. of loci**	**G_{ST} (F_{ST})**	**N_{ST}**	**R_{ST}**	**Source**
Magnolia stellata	3	0.137	n. a. *	0.080	Setsuko *et al*., 2007
Cordia africana	2	0.182	0.220	0.220	Current study
Vitis vinifera	5	0.188	0.220	0.282	Petit *et al*., 2005
Caesalpinia echinata	7	0.911	0.980	0.997	Petit *et al*., 2005

* n. a. = data not available

The low population differentiation and the weak geographical structuring in *C. africana* may indicate the dispersal of seeds over long distances. The fruits of *C. africana* are indehiscent and are eaten by a wide range of mammals and birds, some of which may migrate and travel long distances making the seed dispersal effective enough to result in low population differentiation. The ability of the species to grow in both the mountain areas and the Rift Valley, in the lowlands as well as in river gorges (i.e. low level of habitat discontinuity) could also have contributed to the weak differentiation among the populations of *C. africana.*

Chapter 5: Marker assisted designation of tree seed zones

5.1 Introduction

Seed collection zones (regions of provenance) are subdivisions of land areas established to identify seed sources and to control the movement of seeds and planting stock. Seed zones are needed for many species because of the genetic variation associated with their geographic distribution. Zone boundaries may be delineated from experimental data that identify genetic variation, or by analysis of the environmental factors that have most likely acted as selective forces in creating such genotypic variations (Cunningham, 1975).

Zone delineation depends on the assumptions behind the choice of optimal source populations. These assumptions in turn determine procedures that are appropriate for sampling, data analysis, and subdivision of source populations and planting sites into zones (Westfall and Conkle, 1992). In Ethiopia, a general tree seed zonation system was introduced in 1993 employing data on landform, climate, vegetation and altitude to set guidelines for the safe transfer of tree seeds (Aalbæk, 1993). The Ethiopian seed zonation system recognised a total of 45 zones for all the species, of which 25 are the major seed zones and the remaining 20 are sub seed zones under 10 of the major seed zones. The seed zones were meant to reflect the genetic constitution of the species. However, the general zonation system failed to reflect population genetic structure in *C. africana* and appears to be too detailed for this species (Chapter 2).

Elsewhere, for example in Germany, the delineation of regions of proveneance is species specific, and is based on forest ecological regions (growth regions), which have been combined to form 46 horizontal “Basic Ecological Units” covering the whole territory of Germany. The basic ecological units were joined to form regions of proveneance taking into consideration the phenotypical and genetic characteristics of the tree species in the different regions. The system also recognised altitudinal zones as distinct provenance regions in the mountainous regions. Consequently, there are 178 regions of provenance for all the species, ranging from 7 to 30 for the indigenous species, from 2 to 6 for the foreign species and one for the genus Populus (Anonymous, 1998). In the USA, species-specific zonation systems have been elaborated; for example, breeding zones for Douglas fir were developed from allozyme patterns, ecological and geoclimatic information, and seed

production needs (Westfall and Conkle, 1992). The aim of this chapter is to identify species-specific seed zones based on broad physiographic features (geographical and altitudinal variations) in the country, consolidate the zonation with patterns of variation at the neutral DNA markers (AFLPs), and describe and analyse the existing morphometric differences among and within the new zones.

5.2 Materials and Methods

5.2.1 Identification of seed zones

Following the physiographic variations in the major natural distribution areas of *C. africana* in Ethiopia, four zones can be identified; namely, 1. Northern highlands (NHL), 2. SW highlands (SWHL), 3. SW lowlands (SWLL) and 4. SE highlands (SEHL). The 22 populations sampled can accordingly be classified as belonging to the NHL (Abbay, Finoteselam, Zeghie, Gondar, Kemisse and Mekelle), SWHL (Bako, Jima, Yayu, Didessa), SWLL (Bebeka, Guraferda, Sheko, Tepi and Meti), and SEHL (Butajira, Shashemene, Wondogenet, Dolo-Mana, Ginnr, Hirna and Harar - the first two populations are located in the Rift Valley in the proximity of the SEHL).

5.2.2 Analysis of molecular variance at AFLPs among the species seed zones

AFLP data generated in chapter 2 were used to compute AMOVA among the newly identified species seed zones/tree breeding zones.

5.2.3 Analysis of variance of morphological traits among the species seed zones

Morphometric analysis was conducted on the randomly sampled trees across all the species seed zones for descriptive and comparative purposes. Tree diameter, height, volume and tree form were assessed in each seed zone. These measurements reflect the existing tree growth conditions in the seed zones (though individuals might greatly vary in age and there can be demographic manipulations as well). Measurements on dbh and total tree height were conducted using diameter tape and a Haga Hypsometer, respectively. Visual

assessment of bole shape and forking properties were conducted as well as the measurement of forking heights. In addition, the tree volume and the tree form were assessed. The estimation of tree volume (V) was conducted using all 32 samples in each population as: $V = (\pi * d^2 * h * f * 0.0001)/ 4$, where d = dbh (cm); h = height (m); f = form factor. The form factor was taken arbitrarily as 0.33, am empirical value which had already been applied in some rain forest timber species (Lamb and Borschmann, 1998). The tree form (TFS) was computed as a product of forking height and bole shape, [both scored from 1-6 (Appendix 8)] taken from 10 individuals chosen randomly from each population. Mean comparisons of tree diameter, height, volume and tree form were conducted in a one-way ANOVA setting in SPSS 12.0.

5.3 Results

5.3.1 AFLP population structure along the species seed zones

AMOVA revealed a significant but weak differentiation among the four species seed zones ($\Phi = 0.015$, $p < 0.001$). The populations in the NHL were the most differentiated ones ($\Phi_{ST} = 0.093$) when compared to the among population differentiation in the other SSZ (Table 15). A closer look at the populations from the SEHL in a separate UPGMA dendrogram (Figure 16) revealed the grouping of three geographically closer populations in one cluster (no such topology was found in the other species seed zones). However, AMOVA did not reveal any significant differentiation between the two groups (the southern cluster and the rest).

5.3.2 Morphometric variations along the species seed zones

With regard to the morphometric measurements, the sampled individuals had a mean diameter of 58.7 cm ($\pm$ 33.4 cm), and a mean height of 14.3 m ($\pm$ 7.1 m). The mean comparisons of tree size among the species seed zones revealed that the NHL had the lowest and the SWLL had the highest values for all the metric parameters. The SWHL and SEHL had equivalent sized trees and were intermediate sized when compared to the individuals from the other two species seed zones (SSZ) (Figure 17). Furthermore, one-way analysis of variance (ANOVA) revealed significant differences ($p < 0.001$) among the species seed zones for most of the parameters except for tree form (Table 16).

Table 15 Analysis of molecular variation among the species seed zones/tree breeding zones

Source of variation	df	SSD	MSD	Variance component	Percentage of variation	*p-value*
Among species seed zones	3	140.57	46.86	0.18	1.51	*< 0.001*
Among populations	18	470.17	26.12	0.718	6.08	*< 0.001*
Within populations	453	4895.83	10.81	10.81	92.42	*< 0.001*
Among populations in NHL	5	166.09	32.22	1.07	9.26	*< 0.001*
Within populations	122	1278.08	10.45	10.48	90.74	*< 0.001*
Among populations in SWHL	3	83.56	27.85	0.818	6.76	*< 0.001*
Within populations	78	875.83	11.23	11.23	93.24	*< 0.001*
Among populations in SWLL	3	66.10	22.03	0.47	3.82	*< 0.001*
Within populations	83	980.96	11.82	11.82	96.18	*< 0.001*
Among populations in SEHL	6	142.15	23.69	0.62	5.72	*< 0.001*
Among clusters (the south and the rest)	1	24.29	24.29	0.01	0.05	*ns*
Among populations in the clusters	5	117.87	23.57	0.62	5.69	*< 0.001*
Within populations	145	1483.55	10.23	10.23	94.28	*< 0.001*

df = degrees of freedom, SSD = sum of squared deviations, MSD = mean of squared deviations

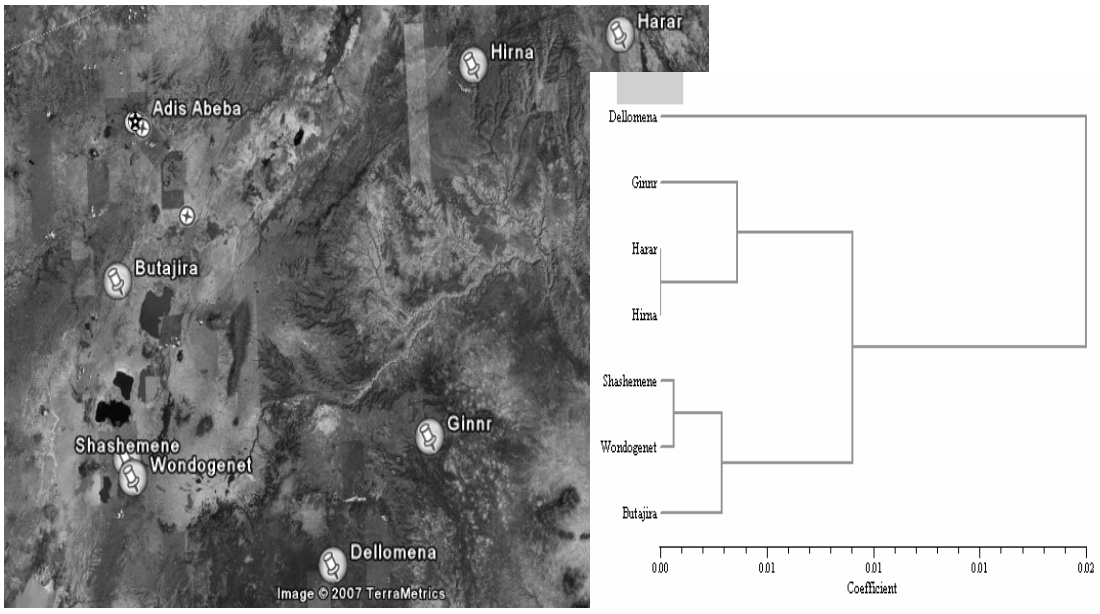
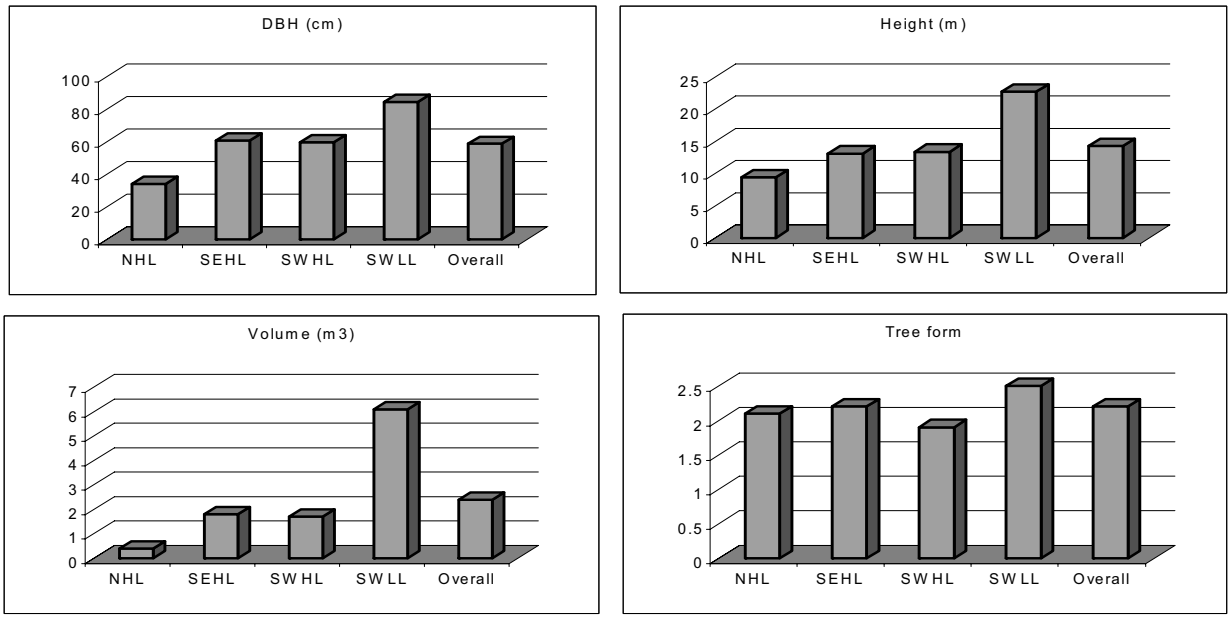


Figure 16 UPGMA dendrogram of populations of the SE based on Nei's genetic distance matrix at AFLP

Figure 17 Mean tree-: diameter, height and volume among *Cordia* africana seed zones



DBH = diameter at breast height, NHL = northern highlands, SEHL = southeastern highlands, SWHL = southwestern highlands

Table 16 One-way Analysis of Variance (ANOVA) for adult tree- height, diameter, volume and form among *Cordia africana* seed zones

Source of variation	df	Sum of squares	Mean square	F	p-value
Height					
Between SSZ	3	16467	5489	205	< 0.001
Within SSZ	700	18705	27		
Total	703	35173			
Diameter					
Between SSZ	3	224231	74744	93	< 0.001
Within SSZ	700	561012	801		
Total	703	785249			
Volume					
Between SSZ	3	3082	1027	49	< 0.001
Within SSZ	700	14616	21		
Total	703	17698			
Tree Form					
Between SSZ	3	9	3.0	1.2	ns
Within SSZ	216	532.5	2.5		
Total	219	540.6			

5.4 Discussion

Three types of investigation are used to generate information about genetic differences among provenances: genetic marker studies, short-term seedling experiments and long-term provenance trials (White *et al*., 2007). In the publication by Loha *et al.*, (2006), analysis of the effect of provenance on seed morphometric traits, germination and seedling growth of *C. africana* taking four populations from the SE (Alemaya, Degaga, Dilla and Wondogenet), one from the SW (Bedelle) and one from the north (Zeghie) of Ethiopia gave significant provenance effects on most of the parameters. The provenance effect was 71-98% for seed morphometric traits, 80% for germination capacity, 57-58% for seedling height and 3-13% for root collar diameter (Loha *et al.*, 2006). In the present study, the relic adult trees found in the scattered and continuous populations in the four newly defined seed zones (tree breeding zones) of *C. africana* (NHL, SWHL, SWLL and SEHL) showed significant differentiation in both morphometrical (diameter, height, volume) and genetic data (AFLPs). However, the differentiation between seed zones at AFLPs was lower than the differentiations among populations within the seed zones. The NHL are the plateaus in Gondar, Gojam, Wello and Tigray (in Amhara and Tigray Regional States), the SWHL are the western parts of Shoa, Wellega, Kaffa and Illubabor (Oromia and the Southern Peoples State), the SWLL are the escarpments in Kaffa, Illubabor and Wellega (Gambella, the Southern Peoples State and Oromia) and finally; the SEHL are the plateaus that extend from the southern parts of Shoa, Arsi, Sidammo, Bale to Hararghe (Oromia and the Southern Peoples State). The significant morphometric differences observed among the species seed zones (tree breeding zones) could be the result of genetic (G) and environmental (E) differences and their interaction (GxE), and the result of demographic manipulations and tree age differences as well.

The use of allozymes in breeding zone designation has been argued since multi-locus analyses of allozymes indicate significant geographic variation in widespread forest trees (Westfall and Conkle, 1992). The final selection of clones from plus-tree collections was carried out on the basis of alloenzyme markers for establishment of seed orchards (Hosius *et al*., 2000). However, a careful interpretation is needed for other marker-based studies since they have limitations regarding their representation of the genome, and the assessment of the adaptive potentials of tree species (Finkeldey and Mátyás, 1999; Szmidt and Wang, 1999). In view of this, provenance/progeny genetic testing of plantations may

be unavoidable for the refining of the species seed zonation proposed for *C. africana*. Zone designations for breeding programs are made on the basis of two perspectives: the source populations and the distribution of planting sites. When local populations are confirmed or assumed to be optimal, then the source populations, which define the seed zone, and the breeding zone, defined by the planting environments, occupy the same zone. But when non-local zones are optimal, planting site zones and the commercially important traits determine the appropriate zones in the source populations (Namkoong, 1988). Breeding zones for Douglas fir were developed from patterns of allozyme variation, ecological and geoclimatic information, and seed production needs (Westfall and Conkle, 1991). Another study on the same species, compared tree seed zone models (general tree seed zonation introduced to Oregon in 1966), soil models and physiographic models (considering elevation, latitude, longitude) and found that the latter fitted best to the species genetic structure based on the assessment of 2-year-old seedlings in a common garden experiment (Campbell, 1991). In the current seed zonation, we also have proved that such a physiographic model (by employing molecular techniques instead of genetic testing plantations) could be an option for setting seed collection guidelines for *C. africana* in Ethiopia.

However, it has to be underlined that individual populations should remain as the focus of interest rather than provenance regions. For example, a closer look at the findings of the comparison of seed morphometric traits, germination, and seedling growth by Loha *et al*. (2006) (recognizing that the population analyzed belong to the three newly defined provenance regions) substantiates this argument since the populations behaved independently of their identity with regard to the provenance regions. The pattern of genetic variation observed at the AFLPs found as described in the present study was also similar in this regard, since there was stronger differentiation among populations within the seed zones than among the seed zones themselves. Nevertheless, this should not undermine the necessity of seed zone delineation following some ecological features, such as physiographic differences (considering both the geographical location and altitude, which apparently dictate the climate and hence physiological adaptations), in order to avoid the risk of maladaptations of non-local sources. It is believed that any marker-assisted definition of seed zones has an advantage in comparison with zone delineation based solely on ecological conditions or phenotypic traits of trees. Since the common possession of one or a few variants of genetic markers in natural populations indicates a common descent of

the trees, the presence of a similar genetic make-up in populations at loci that are of adaptational relevance is highly probable.

Chapter 6: General Discussion and Conclusions

In Chapter 2, it was indicated that the two highest and the two lowest estimates of Nei's genetic diversity in *C. africana* populations were found in the SW (neighbouring populations in the Illubabor and Kaffa provinces, respectively) and in the SE (Bale province) of Ethiopia. The lowest within-population diversities in two other plant taxa (*Hagenia abyssinica* and *Coffea arabica*) were recorded also in one of these areas (Bale province) (Aga *et al.*, 2003, 2006; Feyissa *et al*., 2007). Furthermore, for both *H. abyssinica* and *C. arabica*, the highest values were recorded in the SW/western provinces as was found in *C. africana*. Considering the fact that the SE represents the highest and the SW the lowest diversity in terms of species richness of plant taxa (Friis *et al*., 2001), it can be speculated that some kind of trade-off exists between species and genetic diversity.

The ranges of the rankings of the populations of *C. africana* in terms of diversity overlapped for continuous and scattered (mainly trees from the traditional agroforestry system) populations, which implies that the latter still have a role to play as important elements of the forest genetic resources of the country, especially in forest-denuded areas.

Relatively lower population differentiation was obtained for *C. africana* (Φ_{ST} = 0.07, moderate differentiation) as compared to the differentiation in *Hagenia abyssinica* (G_{ST} = 0.25; Feyissa, *et al*., 2007) and *Coffea arabica* (Shannon-Weaver index of 0.35; Aga *et al*., 2003) in Ethiopia. The lower population differentiation in *C. africana* can be explained by the fact that the species has a wide ecological amplitude (it grows in the altitudinal range of 900 to 2300 m) when compared to *H. abyssinica*, which is restricted to the cool mountain regions (2450-3250 m; Hedberg, 1989) and hence has habitats (unlike *C. africana*) intersected by the Rift Valley, extensive woodlands and river valley gorges. The existence of comparatively effective seed dispersal (by animals) in *C. africana* as compared to *H. abyssinica* can further explain the relatively weak population structure in the former. Whereas sparse gene flow between most *H. abyssinica* populations has been suggested (Feyissa *et al*., 2007), the gene flow in *C. africana* is further facilitated by the relatively high abundance of scattered trees of *C. africana* in many parts of the Ethiopian landscape, as the scattered and solitary-looking trees can act as stepping stones for the gene flow. The very high differentiation in *Coffea arabica,* on the other hand, is predominantly due to its

selfing ability (Meyer, 1965) as opposed to the predominantly outcrossing nature of *C. africana*.

Dendrograms revealed some tendency of geographical structuring in both *H. abyssinica* (Feyissa *et al*., 2007) and *C. arabica* (Aga *et al*., 2003) as was the case in *C. africana* in the current study. The populations of *C. africana* show a tendency towards isolation-by-distance as witnessed by the significant correlations between geographical and genetic distances ($r = 0.35$, $p < 0.01$; with a stronger tendency in the north-south axis) coupled with the significant differentiation between the lowland and highland provenances and the three geographical units as well.

As presented in Chapter 3, the short-term genetic effects of fragmentation (Lowe *et al*., 2004) were observed by the PPL (increases from scattered to relatively continuous population, though not at a pronounced level) and by the elevated levels of non-random association between pairs of putative AFLP loci (linkage disequilibrium) in the farmland populations as compared to the relatively continuous coffee plantations and natural forests. However, genetic diversity did not show a clear trend towards loss with the severity of the fragmentation and forest disturbance in the present dissertation, and this mitigation against the loss of genetic diversity could have been due to effective systems of cross-mating and gene flow (Dick, 2001; White *et al*., 2002). Similarly no effects of fragmentation on genetic diversity were observed in some studies done elsewhere in the tropics (Lowe et al., 2005; Zartman *et al*, 2006; Cloutier *et al*., 2007). Nevertheless, the genetic effects of fragmentation were observed with (progeny) inbreeding (Aldrich *et al*., 1998; White *et al*., 1999; Cascante *et al*., 2001; Lowe *et al*., 2005) and at linkage disequilibrium (Ledig *et al*., 1999; Zartman *et al*., 2006), indicating recent bottlenecks (Tero *et al*., 2003). The elevated values of non-random association among AFLP loci (especially in one of the farmlands, YFL) in the present study may indicate the preferential removal of individuals in the past and a possible risk of genetic bottlenecks in future.

In Chapter 4, it is indicated that the population structure of *C. africana* at cpSSR in Ethiopia showed low differentiation (for the maternally inherited genome, $G_{ST} = 0.18$) and weak geographical structure, which is in contrast to some outcrossing tropical trees (Cavers *et al*., 2003; Lira *et al*., 2003). The low population differentiation may be attributed to a common descent and high gene flow via seeds.

The two marker systems (AFLPs and cpSSRs) were employed to study the population genetic structure of *C. africana* Lam. with the premises of unravelling complementary patterns. AFLPs are believed to enable the assessment of genome-wide variations (mainly the biparentally inherited nuclear genome, migration by both pollen and seed), while cpSSRs represent the uniparentally inherited chloroplast genome (predominantly maternally in angiosperms, gene flow only via seed). Furthermore, the relatively low rate of mutations in the chloroplast genome makes them ideal for studies of population history.

The population differentiation detected for AFLPs as well as cpSSRs was comparatively low. Furthermore, weak geographical structuring was revealed in both of the markers. In this regard, the two marker systems showed congruence in *C. africana*. In another study, congruence between the two marker systems in differentiating a subspecies from the rest of the species, *Fagus sylvatica*, was reported (Gailing and Wuehlisch, 2004). In addition, Ribeiro *et al*. (2002) reported a strong correlation between genetic distances computed for the two marker systems in *Pinus pinaster*, a species in which the chloroplast is inherited paternally and represents gene flow via pollen. However, the similar test in *C. africana* between Nei's genetic distances based on AFLPs (Table 4) and cpSSRs (Appendix 9) did not reveal a significant correspondence, which possibly reflects the different gene dispersal patterns by pollen and seeds.

Since the existing general seed zonation (Aalbæk, 1993) failed to reflect the species population structure at AFLPs, a new seed zonation for *C. africana* has been introduced in Chapter 5. In the new system, four broad species seed zones are identified: 1. Northern highlands (NHL), 2. South-west highlands (SWHL), 3. South-west lowlands (SWLL) and 4. South-east highlands (SEHL) based on the natural distribution of the *C. africana* in Ethiopia and following the existing physiographic features (i.e. considering both geographical and altitudinal variation). ANOVA of metric differences (disregarding age and environmental differences) and AMOVA of AFLPs revealed a significant overall variation among the seed zones, which can be implemented in practical forestry. Such a physiographic-based seed zoning has been practiced elsewhere (Campbell, 1991.)

Given the fact that populations representing virtually all the natural populations of *C. africana* in Ethiopia were sampled and analysed with biparentally and unipaternally inherited DNA markers, some relevant conclusions can be drawn. It is believed that the high reproducibility of the two marker systems (AFLPs and cpSSR) allows a precise

assessment of within- and among-population variation patterns of the allelic structure of various populations of *C. africana*. Genetic variation at the AFLP markers in *C. africana* is very much dependent on individual populations, and hence, populations are the most important units when it comes to the issue of genetic resource utilization and conservation. The positive correlation between genetic and geographical distance among populations supports the use of populations as units of conservation. Accordingly, populations from extreme geographical locations and altitudinal gradients need to be included. The north-south axis seems to be more important than the east-west axis; hence, the relic populations from the forest-denuded areas of northern Ethiopia should not be neglected. For example, a rare chloroplast haplotype was found mainly in the northern populations, which even underlines the need to include the northern provenances in future research and conservation endeavours. In summary, the systematic inclusion of populations in any efforts should incorporate populations from the NHL, SWHL, SWLL and SEHL for reasons of both allelic structure (AFLP and cpSSR data) and for physiological and adaptational characteristics.

Chapter 7: Summary

Biodiversity is an issue of global concern. It is organized in ecosystems, species and genetic diversity within species. Among the various elements of biodiversity, forests represent the biologically most diverse terrestrial ecosystems. However, deforestation poses one of the greatest risks to the maintenance and preservation of biodiversity. Especially deforestation in the tropics is alarming, and it seems hardly achievable to halt it in the near future. It is of utmost urgency that countries should embark on both *in-situ* and *ex-situ* conservation activities. Such efforts need primarily to be based on characterizing and understanding the genetic organization of the target species. One of the species that deserves attention in Ethiopia is the broadleaved tropical tree, *Cordia africana* Lam. (Boraginaceae).

The present dissertation aims at determining the genetic variation within and among populations of *C. africana* in Ethiopia. The variation in the DNA of *C. africana* was assessed employing two marker systems; namely, amplified fragment length polymorphisms (AFLPs) and chloroplast microsatellites (cpSSR). AFLPs represent DNA markers that are randomly distributed across the genome and are generally considered as dominant markers. The chloroplast genome is a non-recombining genome and is generally inherited maternally in angiosperms. The analysis of chloroplast microsatellites in *C. africana* was presumed to detect genetic structures reflecting efficient gene flow via seeds, as the fruits of the species are indehiscent and edible.

The populations were sampled from various geographical locations and altitudinal gradients covering the total distribution range of the species. Furthermore, the populations represented scattered (mainly trees from traditional agroforestry systems) and continuous forest conditions, various regions of provenance (seed zones), natural ecosystems and different levels of fragmentation. The populations can be categorized as belonging to the northern highlands (NHL), the south-west highlands (SWHL), the south-west lowlands (SWLL) or the south-east highlands (SEHL). Genetic variation was assessed in each population. It is assumed that different evolutionary forces (mating system, genetic drift, selection and mutation) acting on each of those populations in varying intensities shape their genetic variation patterns. Although molecular markers *per se* are not appropriate to assess adaptive traits, some of the gene loci may be in linkage disequilibrium with the

genes undergoing selection, and hence may also reveal variation patterns related to adaptation.

The percentage of polymorphic loci (PPL) for individual populations ranged from 62.2% at Dolo-Mena (Bale, SE) to 92.2% at Meti, Didessa (Illubabor, SW) and Wondogenet (Sidamo, SE). AFLP analysis revealed that *C. africana* harbours relatively high average within-population genetic diversity ($H_S = 0.29$). Parameter values of individual populations ranged from 0.22 at Dolo-Mena (Bale, SE) to 0.32 at Meti and Sheko (Illubabor, SW).

No significant difference in genetic diversity was detected between continuous populations and populations under traditional agroforestry systems with comparatively low tree density. This indicates the importance also of scattered populations as essential genetic resources, especially in areas where the natural forests are already gone.

In comparison, the 22 populations of *C. africana* exhibited moderate levels of differentiation ($\Phi_{ST} = 0.072$, $p < 0.001$), which is also in agreement with the observations in many other outcrossing tropical tree species. However, the majority of pairwise comparisons also revealed significant differentiation (87%, $p < 0.01$) indicating the importance of including as many populations as possible in conservation and tree improvement programs. In addition, in a hierarchical AMOVA depicting the importance of geographical locations and altitudinal gradients (i.e. the two dimensions of distance as components of the isolation-by-distance principle) in allelic distributions, significant differentiation was found also among populations in the NHL, SWHL, SWLL and SEHL. Likewise, the Mantel test showed a significant correlation ($r = 0.34$, $p < 0.01$) between geographic and genetic distance matrices indicating isolation-by-distance, which supports the principle conservation strategy outlined above.

Differentiation among geographical regions is only partly reflected in the UPGMA dendrogram using the AFLP genetic distance matrix. Eighteen of the 22 populations showed a certain degree of geographical affinity. Nevertheless, there was no significant differentiation among the existing provenance regions (seed zones), natural ecosystems or forest types (scattered and continuous populations). This failure to reveal regional genetic structure may be related to the neutrality of markers and/or to the presence of effective gene flow blurring area-specific genetic differences and compensating for demographic manipulations as well.

In the study of genetic effects of forest fragmentation in the transitional rain forest in south-west Ethiopia, three levels of fragmentation and disturbance were distinguished, i.e. natural forest, coffee plantations that originally were part of the continuous forests and were converted to plantations by the systematic removal of individuals about three decades ago, and scattered populations. Two representatives of each level of disturbance were used in the study. The analysis revealed strikingly high levels of significant non-random association among AFLP loci (linkage disequilibrium) mainly in one of the scattered (farmland) populations implying the preferential removal of individuals in the past from this population and indicating a possible risk of genetic bottlenecks in the future. Furthermore, the two scattered populations had the least PPL (although the range was rather low) and formed a separate cluster in the UPGMA tree.

The analysis of chloroplast microsatellites (cpSSR) revealed the existence of three haplotypes: Haplotype-A, Haplotype-B and Haplotype-C in proportions of 13, 81 and 6 percent, respectively. The most abundant haplotype (Hap-B) was found in all of the populations, and 8 of the 22 populations were fixed for it. Whereas Hap-C was detected only on the north-western (NW) highlands, the other haplotype (Hap-A) was found on either side of the Rift Valley (which dissects the NW and the SE plateaus). Haplotype-A and Haplotype-C never co-occurred in any given population. The differentiation among the populations was relatively low (G_{ST} = 0.18) for elements of the maternally inherited genome indicating an effective gene flow via seeds.

In conclusion, although the species, *C. africana*, in general does not show strong differentiation and geographical structuring in either of the marker systems, conservation and tree improvement activities would still benefit from including populations that are distributed over a wide geographical and altitudinal range.

Chapter 8: Zusammenfassung

Biodiversität ist von globaler Bedeutung und ist auf der Ebene von Ökosystemen und Arten sowie in der innerartlichen genetischen Variation organisiert. Wälder stellen die biologisch vielfältigsten terrestrischen Ökosysteme dar. Jedoch birgt Abholzung die größten Risiken für die Erhaltung der Biodiversität in Waldökosystemen. Besonders alarmierend ist die Entwaldung in tropischen Regionen, wobei es kaum erreichbar erscheint, diesem Prozess in naher Zukunft Einhalt zu gebieten. Es ist von äußerster Dringlichkeit, dass die betroffenen Länder Strategien der Erhaltung *in situ* und *ex situ* entwickeln. Solche Bemühungen setzen die Kenntnis der Struktur genetischer Variation der entsprechenden Arten voraus. Besondere Aufmerksamkeit verdient darunter die in Äthiopien beheimatete tropische Laubbaumart *Cordia africana* Lam. (Boraginaceae).

Die vorliegende Arbeit hat die Untersuchung der genetischen Variation in und zwischen Populationen von *Cordia africana* in Äthiopien zum Ziel. Dabei wurden zwei molekulare Markersysteme verwendet, nämlich amplified fragment length polymorphisms (AFLPs) und Chloroplastenmikrosatelliten (cpSSRs). AFLPs sind zufallsmäßig über das gesamte Genom verteilt und werden allgemein als dominante Marker angesehen. Das Chloroplastengenom weist keine Rekombination auf und wird bei Angiospermen im allgemeinen vom Samenelter (maternal) vererbt. Demnach war zu erwarten, dass die Untersuchungsergebnisse der genetischen Struktur an Chloroplastenmikrosatelliten Muster der Samenverbreitung bei *Cordia africana* widerspiegeln. Die Platzfrüchte sind überdies eßbar.

Populationen aus unterschiedlichen geografischen Regionen und Höhenlagen des gesamten Verbreitungsgebiets der Art wurden beprobt. Außerdem repräsentierten die Populationen unterschiedlich fragmentierte (hauptsächlich aus traditionellen agroforestry-Systemen stammende) und kontinuierliche geschlossene Wälder, sowie verschiedene Herkunftsregionen (seed zones) und natürliche Ökosysteme. Geographisch gehören die Populationen zu den nördlichen, südwestlichen und südöstlichen Hochlandgebieten und zum südwestlichen Tiefland. Die genetische Variation wurde in jeder Population bestimmt. Es wurde angenommen, dass unterschiedliche evolutionäre Faktoren (Paarungssystem, genetische Drift, Selektion und Mutation) auf die genetischen Variationsmuster der

einzelnen Populationen eingewirkt haben. Obwohl die verwendeten neutralen molekularen Marker *per se* nicht geeignet sind, die Variation in adaptiven Merkmalen abzuschätzen, können einige dieser zahlreichen Marker mit Genen assoziiert sein, die der Selektion unterliegen und somit Muster adaptiver genetischer Variation wiedergeben.

Der Prozentsatz polymorpher Genorte (PPL) reichte von 62.2% für die Population Dolo-Mena (Bale, SE) bis zu 92.2% für die Populationen Meti, Didessa (Illubabor, SW) und Wondogenet (Sidamo, SE). Die AFLP-Analyse ergab durchschnittlich vergleichsweise hohe genetische Variation innerhalb der Populationen (H_S = 0.29). Die Parameterwerte reichten von 0.22 in Dolo-Mena und 0.230 in Ginnr (beide in Bale, SE) bis zu 0.32 in Meti und Sheko (SW). Ein ähnliches Muster geringer genetischer Diversität in Südwesten des Landes wurde auch bei zwei anderen Holzpflanzenarten beobachtet.

Dagegen wurden keine bedeutsame Unterschiede in der genetischen Diversität zwischen Populationen aus kontinuierlichen, geschlossenen Wäldern und solchen geringer Dichte in traditionellen agroforestry-Systeme gefunden. Diese Tatsache weist auf die Bedeutung auch stark fragmentierter Populationen als wichtige genetische Ressourcen besonders in solchen Gebieten hin, in denen natürliche Wälder bereits verschwunden sind.

Andererseits zeigten die Populationen von *Cordia africana* in Übereinstimmung mit anderen fremdbefruchteten tropischen Arten eine mäßige Differenzierung zwischen Populationen (Φ = 0.072, $p < 0.001$). Die meisten paarweisen Vergleiche zwischen den Populationen erbrachten allerdings signifikante Differenzierung (87%, $p < 0.01$), so dass es wünschenswert ist, möglichst viele Populationen in Erhaltungs- und Züchtungsprogramme aufzunehmen. Außerdem wurde eine signifikante Differenzierung zwischen Populationen auch innerhalb geografischer Regionen und aus unterschiedlichen Höhenstufen gefunden. Diese Ergebnisse wurden durch einen statistischen Test (Mantel-Test) gestützt, mit dem eine signifikante Korrelation (r = 0.34, $p < 0.01$) zwischen den genetischen und geografischen Distanzmatrizen nachzuweisen war (isolation by distance). Zudem spiegelt sich die geografische Differenzierung zwischen Regionen im Baumdiagramm (UPGMA) wider, in dem die meisten südwestlichen Populationen aufgrund ihrer genetischen Ähnlichkeit von den anderen Populationen getrennt sind.

Im Gegensatz zur geografischen Differenzierung zeigt sich keine signifikante Differenzierung zwischen den Herkunftsregionen (seed zones), natürlichen Ökosystemen

und Wäldern mit unterschiedlicher Fragmentierung. Die fehlende Differenzierung kann durch Genfluss zwischen den Populationen und durch die Selektionsneutralität der verwendeten Marker erklärt werden.

Bei der Untersuchung genetischer Effekte der Fragmentierung im Regenwald Südwest-Äthiopiens wurden Populationen mit unterschiedlichem Grad der Störung und Fragmentierung betrachtet, d.h. 2 natürliche Populationen, 2 Kaffeeplantagen (ursprünglich Naturwald, der 30 Jahre vorher durch die systematische Entnahme von Bäumen umgewandelt worden war) und 2 stark verstreute Populationen. Die Analyse ergab auffallende Assoziationsungleichgewichte, welche auf selektive Entnahmen in der Vergangenheit und auf eventuelle künftige Risiken hinwiesen. Die beiden stark verstreuten Populationen wiesen zudem vergleichsweise geringe Werte des Parameters PPL auf und bildeten im UPGMA-Baum eine eigene Gruppe.

Durch Analyse von Chloroplastenmikrosatelliten (cpSSRs) Honntez drei Haplotypen Hap-A, Hap-B und Hap-C mit Häufigkeitsanteilen von 13, 81 bzw. 6 Prozent erkannt wereden. Der häufigste Typ Hap-B trat in allen Populationen auf; 8 der 22 Populationen waren sogar auf diesen Typ fixiert. Während sich Hap-C nur in den nordwestlichen Hochlagen vorfand, wurde Hap-A zu beiden Seiten des Rift-Tals beobachtet (welches das nordwestliche Plateau vom südöstlichen abgrenzt). Die Haplotypen Hap-A und Hap-C traten nie in ein und derselben Population auf. Die Differenzierung zwischen den Populationen war relativ schwach (G_{ST}= 0.18) im vergleich zu maternal vererbte Elementen des Genoms bei anderen Arten.

Abschließend ist festzustellen, dass die Baumart *C. africana* nach den Untersuchungen mit beiden Markersystemen nur gering differenzierte geografische Struktur aufweist. In Maßnahmen zur Erhaltung genetischer Ressourcen und zur Züchtung sollten daher möglichst viele Populationen aus einem weiten geografischen und Hochen Bereich einbezogen werden.

References

Aalbæk, A., 1994. Tree Seed Zones for Ethiopia. Forestry Reaserch Center/ National Tree Seed Project, Addis Ababa, 120 pp.

Abebe, T. and Holm, S., 2003. Estimation of wood residues from small-scale commercial selective logging and sawmilling in tropical rain forests of south-western Ethiopia. *International Forestry Review* 5 (1): 45-52.

Adams, W.T., 1992. Gene dispersal within forest tree populations. *New Forests* 6: 217-240.

Aga, E. and Bryngelsson, T., 2006. Inverse sequence-tagged repeat (ISTR) analysis of genetic variability in forest coffee (*Coffea arabica* L.) from Ethiopia. *Genetic Resources and Crop Evolution* 53: 721-728.

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

Akkaya, M.S., Bhagwat, A. A. and Cregan, P. B., 1992. Length polymorphisms of simple sequence repeat DNA in soybean. *Genetics* 132: 1131-1139.

Aldrich, R. P., Hamrick, J. L., Chavarriaga, P. and Kochert, G., 1998. Microsatellite analysis of demographic genetic structure in fragmented populations of the tropical tree *Symphonia globulifera*. *Molecular Ecology* 7: 933-944.

Anonymous, 1992. Ethiopian national report on environment and development: A report prepared for the United Nations Conference on Environment and Development, Rio de Janeiro, Addis Ababa, 116 pp.

Anonymous, 1998. Rules governing forest reproductive material in Germany. Federal Ministry of Food, Agriculture and Forestry (BMELF), Bonn, 57 pp.

Anonymous, 2004. Forest Resources of Ethiopia. Woody Biomass Inventory and Strategic Planning Project, Addis Ababa, 20 pp.

Balloux, F. and Lugon-Moulin, N., 2002. The estimation of population differentiation with microsatellite markers. *Molecular Ecology* 11: 155-165.

Barrett, S. C. H. and Kohn, J. R., 1991. Genetic and evolutionary consequences of small population size in plants: implications for conservation. In: *Genetics and Conservation of Rare Plants* (eds. Falk, D. A. and Holsinger, K. E.), pp. 3-30. Oxford University Press, Oxford.

Bekele, T., 1994. Studies on remnant Afromontane forests on the central plateau of Shoa, Ethiopia. PhD Dissertation, Uppsala University, Uppsala.

Bensch, S. and Åkesson, M., 2005. Ten years of AFLP in ecology and evolution: why so few animals? *Molecular Ecology:* 14: 2899-2914.

Bensch, S., Helbig A. J. Salomon, M. and Seibold I., 2002. Amplified Fragment Length Polymorphism analysis identifies hybrids between two species of warbles. *Molecular Ecology* 11: 473-481.

Berstein, F., 1930. Uber die Erblichkeit der Blutgruppen. *Zeitschriftfür Abstammung und Vererbungslehre* 54: 400-426.

Bierregaard, R. O. Jr., Lovejoy T. E., Kapos, V., dos Santos, A. A. and Hutchings, R. W., 1992. The biological dynamics of tropical rain forest fragments. *Bioscience* 42: 859-866.

Birky, C. W., 1995. Uniparental inheritance of mitochondrial and chloroplast genes: mechanisms and evolution. *Proceedings of the National Academy of Sciences,* USA 92: 11331-11338.

Boshier, D. H., Chase, M. R. and Bawa, K. S., 1995. Population genetics of *Cordia alliodora* (Boraginaceae), a Neotropical tree. 2. Mating system. *American Journal of Botany* 82 (4): 476-483.

Boshier, D. H., Chase, M. R. and Bawa, K. S., 1995. Population genetics of *Cordia alliodora* (Boraginaceae), a Neotropical tree. 3. Gene flow, neighbourhood, and population substructure. *American Journal of Botany* 82 (4): 484-490.

Botstein, D., White, R. L., Skolnick, M. and Davis, R.W., 1980. Construction of a genetic linkage map in man using restriction fragment length polymorphism. *American Journal of Human Genetics* 32 (3): 314-331.

Brown, A. H. D., 1978. Isozymes, plant population genetic structure and genetic conservation. *Theoretical and Applied Genetics* 52: 145-157.

Burczyc, J., Adams, W. T., Birkes, D. S. and Chybicki, I. J., 2006. Using genetic markers to directly estimate gene flow and reproductive success parameters in plants on the basis of naturally regenerated seedlings. *Genetics* 173 (1): 363-372.

Burczyc, J., DiFazio, S. P. and Adams, W. T., 2004. Gene flow in forest trees: how far do genes really travel. *Forest Genetics* 11 (2-3): 1-14.

Campbell, D., Duchesne, P. and Bernatchez, L., 2003. AFLP utility for population assignment studies: analytical investigation and empirical comparison with microsatellites. *Molecular Ecology* 12: 1979-1991.

Campbell, R. K., 1991. Soils, seed zone maps, and physiography: Guidelines for seed transfer of Douglas-fir in Southwestern Oregon. *Forest Science* 37 (4): 973-986.

Cao, C. P., Finkeldey, R., Siregar, I. Z., Siregar, U. J. S. and Gailing O., 2006. Genetic diversity within and among populations of *Shorea leprosula* Miq. and *Shorea parvifolia* Dyer (Dipterocarpaceae) in Indonesia detected by AFLPs. *Tree Genetics and Genomes* 2: 225-239.

Caron, H., Dumas, S., Marque, G., Messier, C., Bandou, E., Petit, R. J. and Kremer, A., 2000. Spatial and temporal distribution of chloroplast DNA polymorphism in a tropical tree species. *Molecular Ecology* 9: 1089-1098.

Carr, G. D., 2006. Boraginaceae. www.botany.hawaii.edu/faculty/carr/boragin.htm

Cascante, A., Quesada, M., Lobo, J. J. and Fuchs, E. A., 2001. Effects of dry tropical forest fragmentation on the reproductive success and genetic structure of the tree *Samanea saman*. *Conservation Biology* 16 (1): 137-147.

Cavers, S., Navarro, C. and Lowe, A. J., 2003. Chloroplast DNA phylogeography reveals colonization history of a Neotropical tree, *Cedrela odorata* L., in Mesoamerica. *Molecular Ecology*12: 1451-1460.

Cervera, M. T., Storme, V., Soto, A., Ivens, B., Van Montagu, M., Rajora, O. P. and Boerjan, W., 2005. Intraspecific and interspecific genetic and phylogenetic relationships in the genus *Populus* based on AFLP markers. *Theoretical and Applied Genetics* 111: 1440-1456.

Chaffey, D. R., 1979. SW Ethiopia forest inventory project. A reconnaissance inventory of forest in SW Ethiopia. Ministry of Overseas Development. Land Resources Development Center. Project Report 31: pp 316.

Chase, M. R., Boshier, D. H. and Bawa, K. S., 1995. Population genetics of *Cordia alliodora* (Boraginaceae), a Neotropical tree. 1. Genetic variation in natural populations. *American Journal of Botany* 82 (4): 468-475.

Chat, J., Chalak, L. and Petit, R. J., 1999. Strict paternal inheritance of mitochondrial DNA in intraspecific crosses of kiwifruit. *Theoretical and Applied Genetics* 99: 314-322.

Cloutier, D., Kanashiro, M., Ciampi, Y. A. and Schoen, D. J., 2007. Impact of selective logging on inbreeding and gene dispersal in an Amazonian tree population of *Carapa guianensis* Aubl. *Molecular Ecology* 16: 797-809.

Cottrell, J. E., Munro, R. C., Tabbener, H. E., Gillies, A. C. M., Forrest, G. I., Deans, J. D. and Lowe, A. J., 2002. Distribution of chloroplast DNA variation in British Oaks

(*Quercus robur* and *Q. petrea*): the influence of postglacial colonization and human management. *Forest Ecology and Management* 156: 181-195.

CSA, 2001. Statistical Abstracts. Central Statistical Agency, Addis Ababa.

Cunningham, R. A., 1975. Provisional tree and shrub seed zones for the Great Plains. *USDA Forest Service Research Paper* RM-150.

Curtu, A. L., Gailing, O., Leinemann, L. and Finkeldey, R., 2007. Genetic variation and differentiation within a natural community of five oak species (*Quercus* spp). *Plant Biology* 9: 116-126.

Dawson, I. K., and Powell, W., 1999. Genetic variation in the Afromontane tree *Prunus africana*, an endangered medicinal species (short communication). *Molecular Ecology* 8: 151-156.

Demesure B., Sodzi N. and Petit R. J., 1995. A set of universal primers for amplification of polymorphic non-coding regions of mitochondrial and chloroplast DNA in plants. *Molecular Ecology* 4: 129-131.

Derero A., Bekele, T. and Näslund, B-Å., 2003. Population structure and regeneration of woody species in a broadleaved Afromontane rain forest, South-west Ethiopia. *Ethiopian Journal of Natural Resources* 5 (2): 255-280.

Derero, A., 2002. Prospects of forest genetic resources conservation and tree seed provision in Ethiopia. In: Proceedings of a national conference on forest resources of Ethiopia: status, challenges and opportunities (eds. Balcha, G., Yeshitela, K. and Bekele, T.), pp 139-149, Institute of Biodiversity Conservation (IBC), Addis Ababa.

Dick C., 2001. Genetic rescue of remnant tropical trees by an alien pollinator. *Proceedings of the Royal Society of London. Series B, Biological Sciences* 268: 2391-2396, London.

Dick, C. W., Etchelecu, G. and Austerlitz, F., 2003. Pollen dispersal of tropical trees (*Dinizia excelsa*: Fabaceae) by native insects and African honeybees in pristine and fragmented Amazonian rainforest. *Molecular Ecology* 12: 753-764.

Dumolin-Lapègue, S., Demesure, B., Fineschi, S., Le Corre, V. and Petit, R. J., 1997. Phylogeographic structure of white oaks throughout the European continent. *Genetics* 146: 1475-1487.

Ennos, R. A., Sinclair, W. T., Hu, X-S, and Langdon, A., 1999. Using organelle markers to elucidate the history, ecology and evolution of plant populations. In: *Molecular Systematics and Plant Evolution* (eds. Hollingsworth, P. M., Bateman, R. M and Gornall, R. J.), pp 1-19, Taylor and Francis Ltd, London.

Evanno, G., Regnaut S. and Goudet, J., 2005. Detecting the number of clusters of individuals using the software structure: a simulation study. *Molecular Ecology* 14 (8): 2611-2620.

Excoffier, L., Laval G., and Schneider S., 2005. Arlequin ver. 3.1: An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online* 1: 47-50.

Feyissa, T., Nybom, H., Bartish, I. V. and Welander, M., 2007. Analysis of genetic diversity in the endangered tropical tree species *Hagenia abyssinica* using ISSR markers. *Genetic Resources and Crop Evolution* 54: 947-958.

Fichtl, R., and Admasu A., 1994. Honeybee Flora of Ethiopia. Magraf Verlag, Weikersheim, 510 pp.

Finkeldey, R. and Hattemer, H. H., 2007. Tropical Forest Genetics. Springer-Verlag, Berlin, Heidelberg, 315 pp.

Finkeldey, R. and Mátyás, G., 1999. Assessment of population history and adaptive potential by means of gene markers. In: Forest Genetics and Sustainability (ed. Mátyás, C.), pp. 91-104. Kluwer Academic Publishers, Dordrecht, Boston, London.

Finkeldey, R. and Ziehe, M., 2004. Genetic implications of silvicultural regimes. *Forest Ecology and Management* 197: 231-244.

Fontaine, C., Lovett, P. N., Sanou, H., Maley, J. and Bouvet, J-M., 2004. Genetic diversity of the shea tree (*Vitellaria paradoxa* C. F. Gaertn), detected by RAPD and chloroplast microsatellite markers. *Heredity* 93: 639-648.

Ford, M. J., 2002. Applications of selective neutrality tests to molecular ecology. *Molecular Ecology* 11: 1245-1262.

Frankel, O. H. and Soulé, M. E., 1981. Conservation and Evolution. Cambridge University Press, Cambridge, 366 pp.

Friis, I., 1992. Forests and Forest Trees of Northeast Tropical Africa. HMSO, Kew Bulletin Additional Series XV, 396 pp.

Friis, I., Edwards, S., Kelbessa, E. and Demissew, S., 2001. Diversity and endemism in the flora of Ethiopia and Eritrea- what do the published Flora volumes tell us? In: Biodiversity Research in the Horn of Africa Region (eds. Friis, I. and Ryding, O.), pp173-193. *Proceedings of the Third International Symposium on the Flora of Ethiopia and Eritrea at the Carlsberg Academy, Copenhagen, August 25-27, 1999*.

Gailing, O. and Wuehlisch, G. V., 2004. Nuclear markers (AFLPs) and chloroplast microsatellites differ between *Fagus sylvatica* and *F. orientalis*. *Silvae Genetica* 53 (3): 105-110.

Greenway, P. J., 1973. A classification of the vegetation of East Africa. *Kirkia* 9: 1-68.

Hamilton, M. B., 1999. Tropical gene flow and seed dispersal. *Nature* 401: 129-130.

Hamrick, J. L. Godt, M. J. W. and Sherman-Broyles, S. L., 1992. Factors influencing levels of genetic diversity in woody plant species. *New Forests* 6: 95-124.

Hamrick, J. L., and Murawski, D. A., 1991. Levels of allozymes diversity in populations of uncommon Neotropical tree species. *Journal of Tropical Ecology* 7: 395-399.

Harris, S. A. and Ingram, R., 1991. Chloroplast DNA and biosystematics: The effects of intraspecific diversity and plastid transmission. *Taxon* 40: 393-412.

Hartl, D. L. and Clark, A. G., 1997. Principles of population genetics, 3rd Ed. Sinauer Associates: Sunderland, Massachusetts, 542 pp.

Hattemer, H. H., and Melchior, G. H., 1993. Genetics and its application to tropical forestry. In: Tropical Forestry Handbook, Vol. I (ed. Pancel L.), pp 330-380, Springer, Berlin, Heidelberg, New York.

Hedberg, O., 1989. Flora of Ethiopia, Pittosporaceae to Araliaceae. In: Rosaceae, Vol. 3 (eds. Hedberg I. and Edwards, S.), Addis Ababa University, Addis Ababa, 659 pp.

Hosius, B., Bergmann, F., Konnert, M. and Henkel, W., 2000. A concept for seed orchards based on isoenzyme gene markers. *Forest Ecology and Management* 131: 143-152.

IBC, 2005. National Biodiversity Strategy and Action Plan. Institute of Biodiversity Conservation, Addis Ababa, pp 115.

ICRAF, 1998. Agroforestry Data Base: *Cordia africana* Lam. ICRAF.

Katzman, M. T. and Cale, W. G., 1990. Tropical forest preservation using economic incentives. *Bioscience* 40: 827-833.

Kimura, M. and Crow, J. F., 1964. The number of alleles that can be maintained in a finite population. *Genetics* 49: 725-738.

Kimura, M. and Otha, T., 1978. Stepwise mutation model and distribution of allelic frequencies in a finite populations. *Proceedings of the National Academy of Sciences,* USA 75: 2868-2872.

Krauss, L. S., 2000. Accurate gene diversity estimates from amplified fragment length polymorphism (AFLP) markers. *Molecular Ecology* 9: 1241-1245.

Kremer, A., Caron, H., Cavers, S., Colpaert, N., Gheysen, G., Gribel, R., Lemes, M., Lowe, A. J., Margis, R., Navarro, C. and Salgueiro, F., 2005. Monitoring genetic diversity in tropical trees with multilocus dominant markers. *Heredity* 95:274-280.

Lamb, D. and Borschmann, G., 1998. Agroforestry with high value trees. RIDROC publication No 98/142.

Ledig, F. T., Conkle, M. T., Bermejo-Velazquez, B., Eguiluz-Piedra, T., Hodgskiss, P. D., Johnson, D. R. and Dvorak, W. S., 1999. Evidence for an extreme bottleneck in a rare Mexican Pinyon: Genetic diversity, disequilibrium and the mating system in *Pinus maximartinezii*. *Evolution* 53 (1): 91-99.

Lira, C. F., Cardoso, S. R. C., Ferreira, P. C. G., Cardoso, M. A. and Provan, J., 2003. Long-term population isolation in the endangered tropical tree species *Caesalpinia echinat*a Lam. revealed by chloroplast microsatellites. *Molecular Ecology* 12: 3219-3225.

Loha, A., Tigabu M., Teketay, D., Lundkvist, K. and Fries, A., 2006. Provenance variation in seed morphometric traits, germination, and seedling growth of *Cordia africana* Lam. *New Forests* 32: 71-86.

Loveless, M. D., 1992. Isozyme variation in tropical trees: patterns of genetic organization. *New Forests* 6: 67-94.

Lowe, A. J., Boshier, D., Ward, M., Bacles, C. F. E. and Navarro, C., 2005. Genetic resource impacts of habitat loss and degradation; reconciling empirical evidence and predicted theory for Neotropical trees. *Heredity* 95: 255-273.

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

Lynch, M. and Walsh, J. B., 1998. Genetics and Analysis of Quantitative Traits. Sinauer Asscos. Inc. Sunderland, MA. 980 pp.

Mantel, N. A., 1967. The detection of disease clustering and a generalized regression approach. *Cancer Research* 27:209-220.

Mariette, S., Cottrell, J. Csaikl, U., Goikoetxea, P., Konig, A., Lowe, A. J., Van Dam B.C., Barreneche T., Bodenes C., Streiff R., Burg K., Groppe K., Munro R.C., Tabbener H. and Kremer A., 2002. Comparison of levels of genetic diversity detected by AFLP and microsatellite markers within and among mixed Q. petrea (Matt.) Liebl. and Q. robur L. stands. *Silvae Genetica* 51: 72-80.

Mariette, S., Le Corre V., Austerlitz, F. and Kremer, A., 2002b. Sampling within the genome for measuring within population diversity: trade-offs between markers. *Molecular Ecology* 11: 1145-1156.

Mehari, A. and Habte, B., 2006. Influence of initial spacing on growth and branching characteristics of *Cordia africana* trees established on Eritrean highland. *New Forests* 31: 185-193.

Meyer, G. F., 1965. Notes on wild *Coffea arabica* from Southwestern Ethiopia, with some historical considerations. *Economical Botany* 19:136-151.

Mihretu M., 1999. The combined effect of mixed planting and spacing on the growth performance of *Cordia africana*. *Ethiopian Journal of Natural Resources* 1 (2): 201-214.

Mueller, K. E. and Wolfenbarger L. L., 1999. AFLP genotyping and fingerprinting. *Trends in Ecology and Evolution* 14: 389-395.

Muluvi, G. M., Sprent, J. I., Soranzo, N., Provan, J., Odee, D., Folkard, G., McNicol, J. W. and Powell, W., 1999. Amplified fragment length polymorphism (AFLP) analysis of genetic variation in *Moringa oleifera* Lam. *Molecular Ecology* 8: 463-470.

Murray, B. G., Young, A. G. and Boyle, T. J., 2000. Basic Genetics. In: Forest Conservation Genetics: principles and practice (eds. Young, A., Boshier, D. and Boyle, T.), pp 7-19. Commenwealth Scientific and Industrial Research Organization (CSIRO) Publishing, Victoria, Australia.

Namkoong, G., Kang, H. C. and Brouard, J. S., 1988. Tree Breeding: principles and strategies. Springer-Verlag, New York.

Negash, L., 1995. Indigenous Trees of Ethiopia: biology, uses and propagation techniques, Addis Ababa University, 285 pp.

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

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

Nybom, H. and Bartish I., 2000. Effects of life history traits and sampling strategies on genetic diversity estimates obtained with RAPD markers in plants. *Perspectives in Plant Ecology, Evolution and Systematics* 3/2: 93-114.

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

OECD, 1974. OECD Scheme for the Control of Forest Reproductive Material Moving in International Trade. OECD, Paris.

Ohta, T., 1982. Linkage disequilbrium due to random genetic drift in finite subdivided populations. *Proceedings of the National Academy of Sciences, USA* 79: 1940-1944.

Pandey, M., Gailing O., Leinemann, L. and Finkeldey, R., 2004. Molecular markers provide evidence for long-distance planting material transfer during plantation establishment of *Dalbergia sisso* Roxb. in Nepal. *Annals of Forest Science* 61: 603-606.

Peakall, R. And Smouse, P. E., 2001. GenAlEx V5:Genetic Analysis in Excel. Population genetics software for teaching and research. Australian National University, Canberra.

Petit, R. J., Brewer, S. Bordács, S., Burg, K., Cheddadi, R., Coart, E., Cottrell, J., Csaikl, U. M., van Dam, B. C., Deans, J. D., Espinel, S., Fineschi, S., Finkeldey, R., Glaz, I., Goicoechea, P. G., Jensen, J. S., König, A. O., Lowe, A. J., Madsen, S. F., Mátyás, G, Munro, R. C., Popescu, F., Slade, D., Tabbener, H., de Vries, S. M. G., Ziegenhagen, B., de Beaulieu, J.-L. and Kremer, A., 2002a. Identification of refugia and post-glacial colonisation routes of European white oaks based on chloroplast DNA and fossil pollen evidence. *Forest Ecology and Management* 156 (1–3): 49–74.

Petit, R. J., Csaikl, U. M., Bordács, S. Burg, K., Coart, E., Cottrell, J., van Dam, B., Deans, J. D., Dumolin-Lapègue, S., Fineschi, S., Finkeldey, R., Gillies, A., Glaz, I., Goicoechea, P. G., Jensen, J. S., König, A. O., Lowe, A. J., Madsen, S. F., Mátyás, G., Munro, R. C., Olalde, M., Pemonge, M. -H., Popescu, F., Slade, D., Tabbener, H., Taurchini, D, de Vries, S. G. M., Ziegenhagen, B. and Kremer, A., 2002b. Chloroplast DNA variation in European white oaks phylogeography and patterns of diversity based on data from over 2600 populations. *Forest Ecology and Management* 156 (1–3): 5–2

Petit, R. J., Duminil, J., Fineschi, S., Hampe, A., Salvini, D. and Vendramin, G. G., 2005. Comparative organization of chloroplast, mitochondrial and nuclear diversity in plant populations. *Molecular Ecology* 14:689-701.

Petit, R. J., Kremer, A. and Wagner, D. B., 1993. Finite island model for organelle and nuclear genes in plants. *Heredity* 71: 630-641.

Pons, O. and Petit, R. J., 1996. Measuring and testing genetic differentiation with ordered and unordered alleles. *Genetics* 144: 1237-1245.

Post, L. S., Schlarbaum, S. E.,.van Manen, F., Cecich, R. A., Saxton, A. M. and Schneider, J. F., 2003. Development of hardwood seed zones for Tennessee Using a Geographic Information System. *Southern Journal of Applied Forestry* 27 (3): 172-175.

Pritchard, J. K., Stephens M., and Donnelly P., 2000. Inference of population structure using multilocus genotype data. *Genetics* 155: 945-959.

Rand, D. M., 1995. Neutrality tests for molecular markers and the connection between DNA polymorphism, demography, and conservation biology. *Conservation Biology* 10: 665-671.

Reboud, X. and Zely, C., 1994. Organelle inheritance in plants. *Heredity* 72:132-140.

Ribeiro, M. M., Mariette, S. Vendramin, G. G., Szmidt, A. E., Plomion, C. and Kremer, A., 2002. Comparison of genetic diversity estimates within and among populations of maritime pine using chloroplast simple-sequence repeat and amplified fragment length polymorphism data. *Molecular Ecology* 11: 869-877.

Rivera-Ocasio, E., Aide, T. M. and McMillan W. O., 2002. Patterns of genetic diversity and biogeographical history on the tropical wetland tree, *Pterocarpus officinalis* (Jack.), in the Caribbean basin.

Rohde, W., 1996. Inverse sequence-tagged repeat (ISTR) analysis, a novel and universal PCR-based technique for genome analysis, a novel and universal PCR-based technique for genome analysis in the plant and animal kingdom. *Journal of Genetics and Breeding* 50:249-261.

Rohlf, F. J., 1998. NTSYSpc version 2.0. Numerical taxonomy and multivariate analysis system. Exeter Software, Setauket, New York.

Röhr H., Kües, U. and Stahl, U., 1998. Organelle DNA of plants and fungi: inheritance and recombination. *Progress in Botany* 60: 39-87.

Rousset, F., 1997. Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. *Genetics* 145: 1219-1228.

Russell, J. R., Weber, J. C., Booth, A., Powell, W., Sotelo-Montes, C. and Dawson, I. K., 1999. Genetic variation of *Calycophyllum spruceanum* in the Peruvian Amazon Basin, revealed by amplified fragment length polymorphism (AFLP) analysis. *Molecular Ecology* 8: 199-204.

Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf S. J., Higuchi R., Horn G. T., Mullis K. B., Erlich H. A., 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* 239: 487-491.

Sayer, J. A., Harcourt, C. S. and Collins, N. M., 1992. The Conservation Atlas of Tropical Forests. Africa. Macmillan, Basingstroke, 288 pp.

Setsuko, S., Ishida, K., Ueno, S., Tsumura, Y. and Tomaru, N., 2007. Population differentiation and gene flow within a metapopulation of a threatened tree *Magnolia stellata* (Magnoliaceae). *American Journal of Botany* 94(1): 128-136.

Shih, F. L., Hwang, S.Y., Cheng, Y. -P., Lee, P. -F. and Lin, T. -P., 2007. Uniform genetic diversity, low differentiation, and neutral evolution characterize contemporary refuge populations of Taiwan fir (*Abies kawakamii*, Pinaceae). *American Journal of Botany* 94: 194-202.

Silvestrini, M., Junqueira, M. G., Favarin, A., Guerreiro-Filho, O., Maluf, M. P., Silvarolla, M. B., and Colombo, C. A., 2007. Genetic diversity and structure of Ethiopian, Yemen and Brazilian *Coffea arabica* L. accessions using microsatellite markers. *Genetic Resources and Crop Evolution* 54 (6): 1367-1379.

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

Slatkin, M., 1991. Inbreeding coefficients and coalescence times. *Genetical Researc*h 58: 167-175.

Szmidt, A. and Wang, X.-R., 1999. Genetic markers in forest genetics-the tunnel remains dark. In: Forest Genetics and Sustainability (ed. Mátyás, C.), pp.31-48. Kluwer Academic Publishers, Dordrecht, Boston, London.

Taberlet, P., Gielly, L., Pautou, G. and Bouvet, J., 1991. Universal primers for amplification of three non-coding regions of chloroplast DNA. *Plant Molecular Biology* 17: 1105-1109.

Teketay, D. and Tegineh, A., 1991. Shade trees of coffee in Hararghe, Eastern Ethiopia. *International Tree Crops Journal* 7: 17-27.

Teklay T. and Malmer A., 2004. Decomposition of leaves from two indigenous trees of contrasting qualities under shaded-coffee and agricultural land uses during the dry season at Wondo Genet, Ethiopia. *Soil Biology and Biochemistry* 36: 777–786.

Tero, N., Aspi, J., Siikamäki, P., Jäkäläniem, A. and Tuomi, J., 2003. Genetic structure and gene flow in a metapopulation of endangered plant species, *Silene tatarica*. *Molecular Ecology* 12: 2073-2085.

Tigabu, M. and Oden, P. C., 2002. Multivariate classification of sound and insect-infested seeds of a tropical multipurpose tree, *Cordia africana*, with near infrared reflectance spectroscopy. *Journal of Near Infrared Spectroscopy* 10: 45-51.

Tsumura, Y., Kawahara, T., Wickneswari, R. and Yoshimura, K. 1996. Molecular phylogeny of Dipterocarpaceae in SE Asia using RFLP of PCR-amplified chloroplast genes. *Theoretical and Applied Genetics* 93: 22-29.

Van Oosterhout, C., Van Heuven, M. K. and Brakefield, P. M., 2004. On the neutrality of molecular genetic markers: pedigree analysis of genetic variation in fragmented populations. *Molecular Ecology* 13: 1025-1034.

Vekemans, X., 2002. AFLP-SURV version 1.0. Distributed by the author. Laboratoire de Genetique et Ecologie Vegetale, Universite Libre de Bruxelles, Belgium.

Von Breitenbach, F., 1963. The indigenous trees of Ethiopia. Ethiopian Forestry Association , Addis Ababa, 305 pp.

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

Wang, Z., Weber, J. L., Zhong, G. and Tanksley, S. D., 1994. Survey of plant short tandem DNA repeats. *Theoretical and Applied Genetics* 88: 1-6.

Ward, M., Dick, C. W., Gribel, R. and Lowe, A. J., 2005. To self, or not to self…A review of outcrossing and pollen-mediated gene flow in Neotropical trees. *Heredity* 95: 246-254.

Warfa, A. M., 1988. *Cordia* (Boraginaceae) in NE tropical Africa and tropical Arabia. PhD Dissertation, Uppsala University, Uppsala, pp 78.

Wehenkel, C., Bergmann, F. and Gregorius, H. R., 2006. Is there a trade-off between species diversity and genetic diversity in forest tree communities? *Plant Ecology* 185: 151-161.

Weising, K. and Gardner, R. C., 1998. A set of conserved PCR primers for the analysis of simple sequence repeat polymorphisms in chloroplast genomes of dicotyledonous angiosperms. *Genome* 42 (1): 9-19.

Weising, K., Nybom, H., Wolff, K. and Kahl, G., 2005. DNA Fingerprinting in Plants: Principles, Methods, and Applications. 2nd Ed, Tayler and Francis Group, USA, 144 pp.

Westfall, R. D. and Conkle, M. T., 1991. Allozyme markers in breeding zone designation. *New Forests* 6: 279-309.

White, G. M., Boshier, D. H. and Powell, W., 1999. Genetic variation within a fragmented population of *Swietenia humilis* Zucc. *Molecular Ecology* 8: 1899-1909.

White, G. M., Boshier, D. H. and Powell, W., 2002. Increased pollen flow counteracts fragmentation in a tropical dry forest: an example from *Swietenia humilis* Zuccarini. *Proceeding of the National Academy of Sciences*, USA 99: 2038-2042.

Williams, J. G. K., Kubelik, A. R, Livak K. J., Rafalski, J. A. and Tingey S. V., 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Research* 18: 6531-6535.

Wolfe, A. D. and Liston, A., 1998. Contribution of PCR-based methods to plant systematics and evolutionary biology. In: Molecular systematics of plants II: DNA sequencing (ed. Soltis, D. E. Soltis, P. S. and Doyle, J. J), pp 43-86, Kluwer Academic, Boston.

Wolfe, K. H., Li, W.-H. and Sharp, P. M., 1987. Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast and nuclear DNAs. *Proceedings of the National Academy of Sciences*, USA 84: 9054-9058.

Wright, S., 1931. Evolution in Mendelian populations. *Genetics* 16:97-159.

Wright, S., 1978. Evolution and the genetics of population, variability within and among natural populations. The University of Chicago Press, Chicago.

Yadessa, A., Itanna, F. and Olsson, M., 2001. Contribution of indigenous trees to soil properties: The case of scattered trees of *Cordia africana* Lam. in croplands of western Oromia. *Ethiopian Journal of Natural Resources* 3 (2): 245-270.

Yeh, F. C., 2000. Population genetics. In: Forest Conservation Genetics: principles and practice (eds. Young, A., Boshier, D. and Boyle, T.), pp 21-38. Commenwealth Scientific and Industrial Research Organization (CSIRO) Publishing, Victoria, Australia.

Yeh, F. C., Yang, R.-C. and Boyle, T., 1999. POPGENE version 1.31, free software. University of Alberta, Alberta.

Yirdaw, E. and Leinonen, K., 2002. Seed germination responses of four Afromontane tree species to red/far-red ratio and temperature. *Forest Ecology and Management* 168: 53-61.

Young, A. G. and Boyle, T. J., 2000. Forest fragmentation. In: Forest Conservation Genetics: principles and practice (eds. Young, A., Boshier, D. and Boyle, T.), pp 123-134. Commenwealth Scientific and Industrial Research Organization (CSIRO) Publishing, Victoria, Australia.

Young, A., Boyle, T. and Brown, T., 1996. The population genetic consequences of habitat fragmentation for plants. *Trends in Ecology and Evolution* 43: 413-418.

Zartman, C. E., Mcdaniel, S. F. and Jonathan S. A., 2006. Experimental habitat fragmentation increases linkage disequilibrium but does not affect genetic diversity or population structure in the Amazonian liverwort *Radula flaccida*. *Molecular Ecology* 15: 2305-2315.

Zhivotovsky, L. A., 1999. Estimating population structure in diploids with multilocus dominant DNA markers. *Molecular Ecology* 8: 907-913.

Appendices

Appendix 1 Matrix of pair wise geographical distances (km) between the 22 populations of *Cordia africana* sampled from Ethiopia

Popn	Geographical distances (km)																					
Bak	160	0																				
Beb	460	323	0																			
Del	440	424	518	0																		
Did	250	90	238	450	0																	
Fin	110	170	482	540	246	0																
Gur	493	332	7	524	244	490	0															
Gin	426	454	617	130	504	532	623	0														
Gon	316	410	700	758	472	222	704	730	0													
Har	436	524	816	390	604	512	822	286	650	0												
Hir	330	442	708	344	522	436	715	234	548	122	0											
Jim	308	162	203	360	120	336	212	434	570	382	502	0										
Kem	200	354	667	464	444	278	675	394	370	280	201	472	0									
Mek	400	546	865	780	630	390	872	710	250	540	500	704	320	0								
Met	458	298	50	520	210	452	52	610	670	804	692	190	647	832	0							
Sha	318	270	385	160	292	402	392	390	636	450	352	204	402	695	380	0						
She	454	296	25	500	210	456	31	596	680	774	681	187	640	842	31	364	0					
But	214	195	405	240	247	301	413	262	531	380	300	202	308	594	390	106	379	0				
Tep	456	290	39	520	202	446	45	596	664	788	680	175	636	832	32	365	14	378	0			
Won	326	276	380	154	158	410	393	230	646	454	358	110	416	710	380	13	365	114	366	0		
Yay	320	152	181	490	64	302	187	550	520	694	382	130	510	693	142	328	155	298	144	332	0	
Zeg	220	286	576	654	348	120	582	640	126	580	511	450	326	318	550	521	575	420	541	532	398	
	Abb	Bak	Beb	Del	Did	Fin	Gur	Gin	Gon	Har	Hir	Jim	Kem	Mek	Met	Sha	She	But	Tep	Won	Yay	

Appendix 2: Protocol of AFLP

3.1 Restriction-ligation Mix
Prepare the RLR and RLM as follows:

A. Restriction Mix, RLR (20 µl for 10 probes)

T4 DNA Ligase buffer 10X	2.0 µl
0.5 M Nacl	2.0 µl
BSA (1 mg/ml)	1.0 µl
*Mse*I (10 u/µl)	0.8 µl
*Eco*RI (10 u/µl)	4.0 µl
T4 DNA Ligase (4 u/µl)	1.9 µl
HPLC H_2O	8.3 µl

B. Ligation Mix, RLM (60 µl for 10 probes)

T4 DNA Ligase buffer 10X	10.0 µl
0.5 M Nacl	10.0 µl
BSA (1mg/ml)	5.0 µl
*Mse*I Adapter pair	6.0 µl
*Eco*RI Adapter pair	6.0 µl
HPLC H_2O	23.0 µl

Finally, for each probe (4 µl DNA) add 6 µl RLR + 2 µl RLM and let the reaction take place at room temperature overnight.

3.2. Preselective amplification

A. PCR Mix (25 µl)

PCR-buffer (10X)	2.5 µl
dNTPs (10mM)	0.38 µl
Primer M03	0.25 µl
Primer E01	0.25 µl
Taq-polymerase	0.1 µl
HPLC H_2O	8.0 µl
Restriction-Ligation-Reaction product	0.8 µl

B. PCR programme
- Initiation 72 ^oC for 2 min
- 20 cycles of denaturation (94 ^oC for 10 sec), annealing (56 ^oC for 30 sec) and extension (72 ^oC for 2 min).
- Final extension at 60 ^oC for 30 min.

3.3 Selective amplification

A. PCR Mix (15µl)

PCR-buffer (10X)	1.67 µl
dNTPs (10mM)	0.25 µl
*Eco*RI primer (E41)	0.25 µl
*Mse*I primer (M74)	0.25 µl
Taq-polymerase	0.07 µl
HPLC H_2O	9.25 µl
Preamplification product	4.0 µl

B. PCR Programme
- Initiation at 94 ^oC for 2 min
- 33 cycles of denaturation (94 ^oC for 10 sec), annealing (65 ^oC for 30 sec, the annealing deceasing from 65 to 56, 1 ^oC at every cycle up to the 10^{th} cycle) and extension (72 ^oC for 2 min.
- Final extension at 60 ^oC for 30 min.

Appendix 3: List of loci and their respective repeatability

No	Sam.	PN.	81	84	85	86	88	89	90	91	92	93	94	95	96	98	99	101	102	103	104	105	106	107	108	109	110	111	113	114	116	117	118	119	120	121	122	123
1	100	2	1	1	0	1	1	1	1	0	1	1	1	1	1	1	1	0	1	0	1	0	0	0	1	1	1	1	1	0	1	1	1	1	0	0	1	1
2	100R	2	1	1	0	1	1	1	1	0	1	1	1	1	1	1	1	0	1	0	1	0	0	0	1	1	1	1	0	1	?	1	1	1	0	0	1	1
			r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	n	?	?	r	r	r	r	r	r	r
3	102	2	1	1	0	1	1	1	1	0	1	1	0	1	1	1	1	0	1	0	1	0	0	0	1	0	1	1	1	0	0	1	1	1	0	1	0	1
4	102R	2	1	1	0	1	1	1	1	0	1	1	0	1	1	1	1	0	1	0	1	0	0	0	1	0	1	1	1	0	0	1	1	1	0	0	0	1
			r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	?	r	r
5	201	3	0	0	1	1	1	1	1	0	1	1	1	0	1	1	1	0	0	1	1	0	0	0	0	0	1	1	0	0	0	0	1	1	0	1	0	1
6	201R	3	0	0	1	1	1	1	1	0	1	1	1	0	1	1	1	0	0	1	1	0	0	0	0	0	1	1	0	0	0	1	1	1	0	1	0	1
			r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r
7	202	3	1	1	1	1	1	1	1	0	1	1	1	0	1	1	1	0	1	1	1	0	0	0	0	0	1	1	1	0	0	0	1	1	1	1	1	0
8	202R	3	1	1	1	1	1	1	1	0	1	1	1	0	1	1	1	0	1	1	1	0	0	0	0	0	1	1	1	1	0	0	1	1	1	1	1	0
			r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	n	r	r	r	r	r	r	r	r
9	204	3	1	0	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	1	1	0	0	0	0	1	1	1	1	0	0	0	1	1	0	1	0	1
10	204R	3	1	0	1	1	1	1	0	0	1	1	1	1	1	1	1	0	1	1	1	0	0	0	0	1	1	1	1	0	1	0	1	1	0	1	0	1
			r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r
11	207	3	1	0	0	1	1	1	0	0	1	1	1	0	1	1	1	1	0	0	1	0	0	0	0	0	0	1	0	0	0	1	1	1	1	0	1	1
12	207R	3	0	0	0	1	1	1	0	0	1	1	1	0	1	1	1	1	0	0	1	0	0	0	0	0	0	1	0	0	0	1	1	1	1	0	0	0
			n	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r
13	208	3	1	0	1	1	1	1	0	0	1	1	1	1	1	1	1	0	0	0	1	0	0	0	1	0	1	1	0	0	0	0	1	1	0	1	0	0
14	208R	3	1	0	1	1	1	1	0	0	1	1	1	0	1	1	1	0	0	0	1	0	0	0	1	0	1	1	0	0	0	0	1	1	0	1	0	0
			r	r	r	r	r	r	r	r	r	r	r	n	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r
15	292	4	1	0	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	1	1	0	0	0	1	1	1	1	1	0	0	0	1	1	0	1	0	1
16	292R	4	1	0	0	1	0	1	1	0	1	1	0	1	1	1	1	0	1	1	1	0	0	0	1	1	1	1	1	0	0	0	1	1	0	1	0	1
			r	r	n	r	n	r	r	r	r	r	n	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r
17	386	5	1	0	?	0	1	1	1	0	1	1	0	1	1	1	1	1	0	0	1	0	0	0	0	0	1	1	1	0	0	1	1	1	1	1	0	1
18	386R	5	1	1	1	0	1	1	1	0	1	1	0	1	1	1	1	1	1	0	1	0	0	0	0	0	1	1	1	0	1	0	1	1	1	1	0	1
			r	n	?	r	r	r	r	r	r	r	r	r	r	r	r	r	n	r	r	r	r	r	r	r	r	r	r	r	n	n	r	r	r	r	r	r
19	482	6	0	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	0	0	1	0	1	1	1	0	0	0	1	1	1	0	1	0
20	482R	6	0	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	0	0	0	0	1	1	1	0	0	0	1	1	1	0	1	0
			r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	n	r	r	r	r	r	r	r	r	r	r	r	r	r
21	488	6	1	0	1	1	1	1	0	1	0	1	0	1	1	1	1	1	1	0	1	0	0	0	1	0	1	1	0	1	0	0	1	1	0	0	1	1
22	488R	6	1	0	1	1	1	1	0	1	0	1	0	1	1	1	1	1	1	0	1	0	0	0	0	0	1	1	0	1	0	0	1	1	0	0	1	1
			r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	n	r	r	r	r	r	r	r	r	r	r	r	r	r
23	497	6	1	1	1	0	0	1	0	0	0	1	1	1	1	1	1	0	1	0	1	1	1	1	0	0	0	1	0	0	0	0	1	1	0	1	1	1
24	497R	6	1	1	1	0	1	1	1	0	1	1	1	1	1	1	1	0	1	0	1	1	1	1	0	0	0	1	0	0	0	0	1	1	0	1	1	1
			r	r	r	r	?	r	n	r	?	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r
sum r			11	11	10	12	10	12	11	12	11	12	11	11	12	12	12	12	11	12	12	12	12	12	11	12	12	12	11	10	10	11	12	12	12	11	12	12
%			92	92	91	100	91	100	92	100	100	100	92	92	100	100	100	100	92	100	100	100	100	100	92	100	100	100	92	91	91	92	100	100	100	100	100	100

No	Sam.	PN.	124	125	126	127	128	129	130	132	134	136	137	139	140	141	142	143	144	147	148	149	150	151	153	154	159	160	161	162	163	165	166	168	169	170	171	173
1	100	2	1	1	1	1	1	0	0	1	0	1	1	1	1	1	1	1	1	0	0	0	0	0	0	1	0	1	1	1	1	0	0	0	0	1	0	1
2	100R	2	1	0	1	1	1	0	0	1	0	1	0	1	1	1	1	1	1	0	0	0	0	0	0	1	0	1	1	1	1	0	0	0	0	1	0	1
			r	?	r	r	r	r	r	r	r	r		r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r
3	102	2	1	1	0	0	1	0	1	1	0	1	1	1	1	1	1	1	1	1	0	1	1	0	0	1	0	1	1	1	1	1	1	0	0	1	0	0
4	102R	2	1	0	0	0	0	0	0	1	0	1	1	1	1	1	1	1	1	1	0	1	1	1	0	1	0	1	1	1	1	1	1	0	0	1	0	0
			r	?	r	r	n	r	n	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r
5	201	3	1	1	1	1	1	0	0	1	0	1	0	0	0	1	1	1	1	0	0	0	1	0	0	1	0	0	0	1	1	0	0	0	0	1	0	1
6	201R	3	1	1	0	1	1	0	0	1	0	1	0	0	0	1	1	1	1	0	0	0	1	0	0	1	0	0	0	1	1	0	0	0	0	1	0	1
			r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r
7	202	3	1	1	1	1	0	0	1	0	1	0	1	1	0	1	1	1	1	0	0	0	1	0	0	1	0	1	1	1	1	0	1	0	0	1	0	1
8	202R	3	1	1	1	1	0	0	1	1	0	1	0	1	0	1	1	1	1	0	0	0	1	0	0	1	0	1	1	1	1	0	0	1	1	1	0	1
			r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r
9	204	3	1	1	1	1	1	0	0	1	0	0	0	1	0	1	1	1	1	0	0	0	0	0	0	1	0	1	0	1	1	0	0	0	n	r	r	r
10	204R	3	1	1	1	1	1	0	0	1	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0	1	0	1	0	1	1	0	0	0	0	1	0	1
			r	r	r	r	r	r	r	r	r	r	r	n	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	n
11	207	3	0	0	1	0	1	0	0	0	0	1	1	1	0	1	1	1	1	0	0	0	0	0	0	1	1	1	0	1	1	0	0	0	0	0	0	0
12	207R	3	0	0	1	1	1	0	0	1	0	1	1	1	0	1	1	1	1	0	0	0	0	0	0	1	0	0	0	1	1	0	0	0	0	0	0	0
			r	r	r	n	r	r	r	?	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	n	n	n	r	r	r	r	r	r	r	r	r
13	208	3	1	1	1	1	1	0	0	1	0	1	0	0	1	1	1	1	1	1	1	0	0	0	0	1	0	0	0	1	1	0	1	0	0	0	1	1
14	208R	3	1	1	1	1	1	0	0	0	0	1	0	0	1	1	1	1	1	1	1	0	0	0	0	1	0	0	0	1	1	0	1	0	0	0	1	1
			r	r	r	r	r	r	r	?	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r
15	292	4	1	1	1	1	1	0	0	1	0	1	0	1	1	1	1	1	1	0	0	0	0	1	0	1	0	1	1	1	1	0	0	0	0	1	0	1
16	292R	4	1	0	1	0	1	0	0	1	0	1	0	1	1	1	1	1	1	0	0	0	0	1	0	1	0	1	1	1	1	0	0	0	0	1	0	1
			r	?	r	?	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r
17	386	5	0	0	1	1	1	0	0	1	0	1	1	0	0	1	1	1	1	0	0	0	1	0	0	1	1	0	1	1	1	0	0	0	0	1	0	1
18	386R	5	1	0	1	1	1	0	0	1	0	1	1	0	0	1	1	1	1	0	0	0	1	0	0	1	1	0	1	1	1	0	0	0	0	1	0	1
			r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r
19	482	6	1	1	0	0	1	1	1	0	0	1	0	1	0	1	1	1	1	0	0	0	1	0	0	1	1	1	1	1	1	0	0	1	1	1	0	1
20	482R	6	1	1	0	0	1	1	1	0	0	1	0	1	0	1	1	1	1	0	0	0	1	0	0	1	1	1	1	1	1	0	0	1	1	1	0	1
			r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r
21	488	6	0	0	1	1	1	0	1	1	0	1	0	1	1	1	1	1	1	0	0	0	1	1	0	1	0	1	1	1	1	0	0	0	0	1	1	1
22	488R	6	0	0	1	1	1	0	1	1	0	1	0	1	1	1	1	1	1	0	0	0	1	1	0	1	0	1	1	1	1	0	0	0	0	1	1	1
			r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r
23	497	6	1	0	1	0	1	0	0	0	0	1	0	1	1	1	1	1	1	0	0	0	0	1	1	0	0	1	1	1	1	0	0	0	0	1	0	0
24	497R	6	1	0	1	0	1	0	0	1	0	1	0	1	1	1	1	1	1	0	0	0	1	1	1	0	0	1	1	1	1	0	0	0	0	1	0	1
			r	r	r	r	r	r	r	n	r	r	r	r	r	r	r	r	r	r	r	r	n	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r
	sumr		12	12	11	10	11	12	11	9	12	12	11	11	12	12	12	12	12	12	12	12	11	12	12	12	11	11	11	12	12	12	11	11	11	12	12	11
	%		100	100	92	91	92	100	92	90	100	100	92	92	100	100	100	100	100	100	100	100	92	100	100	100	92	92	92	100	100	100	92	92	92	100	100	92

No	Sam.	PN.	174	175	177	178	179	181	182	184	185	186	187	189	190	192	194	195	198	201	202	204	207	208	210	214	221	223	224	225	227	229	232	233	235	238	242	246
1	100	2	1	1	1	1	0	1	1	0	0	0	0	1	1	0	1	1	0	1	1	0	1	1	0	1	1	0	1	0	1	0	1	1	1	1	0	1
2	100R	2	1	1	1	1	0	0	1	0	0	0	0	1	1	0	0	0	0	?	1	0	1	1	0	0	0	0	0	1	0	0	0	0	1	1	0	1
			r	r	r	r	r	?	r	r	r	r	r	r	r	r	n	n	r	?	r	r	r	r	r	?	?	?	?	?	?	r	?	?	r	r	?	r
3	102	2	0	1	1	1	0	1	1	0	0	0	0	1	1	0	0	0	0	1	1	0	1	1	1	1	1	0	1	1	0	0	0	0	1	0	0	1
4	102R	2	0	1	1	1	0	1	1	0	0	0	0	1	1	0	0	0	0	1	1	0	1	1	1	1	0	0	1	0	0	0	0	0	1	0	0	1
			r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	?	r	r	?	r	r	r	r	r	r	?	r	r
5	201	3	1	1	1	1	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	1	1	1	0	0	1	0	0	0	0	0	0	0	0	0
6	201R	3	1	1	1	1	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	1	1	0	0	0	1	0	0	0	0	0	0	0	0	0
			r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r
7	202	3	1	1	0	1	0	0	0	1	1	0	1	1	1	0	0	0	0	0	0	0	1	1	1	0	0	0	1	1	0	0	0	1	1	1	0	1
8	202R	3	1	1	0	1	0	0	0	1	1	0	1	1	1	0	0	0	0	0	0	0	1	1	0	0	0	0	1	1	1	0	0	0	1	1	0	1
			r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	?	r	r	r	r	r	r	r	r	r	r	r	r	r
9	204	3	1	1	1	1	0	0	0	0	0	0	0	1	1	0	0	0	1	1	0	0	1	1	1	0	0	0	1	1	1	0	1	0	1	0	0	1
10	204R	3	1	1	1	1	0	0	0	0	0	0	0	1	1	0	0	0	1	1	0	0	1	1	1	0	0	0	1	1	1	0	1	0	1	0	0	1
			r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r
11	207	3	1	1	1	1	1	1	1	0	0	0	0	1	0	0	0	0	1	1	1	0	1	1	0	0	1	0	0	0	0	0	1	0	1	1	0	0
12	207R	3	1	0	1	1	0	1	1	0	0	0	0	1	0	0	0	0	0	1	1	0	1	1	0	0	1	0	0	0	0	0	1	0	0	0	0	0
			r		r	r		r	r	r	r	r	r	r	r	r	r	r	n	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	n	?	r	r
13	208	3	1	1	1	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0
14	208R	3	1	1	1	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0
			r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	?	?	r	r	r	r	r	r	r
15	292	4	1	1	1	1	0	0	0	0	0	0	0	1	0	0	1	1	0	0	1	0	1	1	0	0	0	0	0	1	1	0	0	0	1	0	0	1
16	292R	4	1	0	1	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	1	0	1	1	0	0	0	0	0	1	1	0	0	0	1	0	0	1
			r	r	?	r	r	r	r	r	r	r	r	r	r	r	?	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r
17	386	5	1	1	1	1	1	0	1	0	0	0	0	1	1	0	1	1	0	0	1	0	1	1	1	1	0	0	1	0	0	0	1	1	1	0	0	0
18	386R	5	1	1	1	1	1	0	0	0	0	0	0	1	1	0	1	1	0	0	1	0	1	1	1	1	0	0	1	1	1	0	0	0	1	0	0	1
			r	r	r	r	r	r	n	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	n	?	n	r	?	?	r	?	?	r	r	r	?
19	482	6	1	1	1	1	0	0	0	1	1	0	0	1	1	0	0	0	0	0	0	0	1	1	0	0	1	1	1	1	0	0	1	0	1	0	0	1
20	482R	6	1	1	1	1	0	0	0	1	1	0	0	1	1	0	0	0	0	0	0	0	1	1	0	0	0	1	1	1	0	0	1	0	1	0	0	1
			r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	?	r	r	r	r	r	r	r	r	r	r	r
21	488	6	1	1	0	1	1	0	0	0	0	1	1	1	1	0	1	0	1	1	1	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	1
22	488R	6	1	1	0	1	1	0	0	0	0	1	1	1	1	0	1	0	1	1	1	0	1	1	1	0	0	0	0	0	0	0	0	0	1	1	0	1
			r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	?	?	r	?	?	?	?	r	r
23	497	6	1	1	1	1	0	0	1	0	0	0	0	1	1	0	0	1	0	0	1	0	1	1	1	0	0	0	0	1	0	0	0	0	1	0	1	1
24	497R	6	1	1	1	1	0	0	1	0	0	0	0	1	1	0	0	1	0	0	1	0	1	1	1	0	0	0	0	0	0	0	0	0	1	0	0	1
			r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	?	?	?	?	r	?	?	r	?	n	r
	sum r		12	11	12	12	11	11	11	12	12	12	12	12	12	12	10	11	11	11	11	12	12	12	11	10	8	9	10	6	6	12	7	7	10	9	10	11
	%		100	92	100	100	92	100	92	100	100	100	100	100	100	100	91	92	92	100	100	100	100	100	100	91	100	90	100	100	100	100	100	100	91	90	91	100

No	Sam.	PN.	252	255	256	258	261	262	263	265	271	272	274	275	279	283	286	289	291	292	293	296	301	306	309	312	314	315	316	320	322	324	325	326	332	335	339	344
1	100	2	1	0	0	1	0	0	0	1	1	0	1	1	0	0	0	1	1	1	1	1	0	0	1	1	0	0	0	0	1	0	0	0	1	1	1	1
2	100R	2	?	0	0	?	0	0	0	1	1	0	1	1	0	0	0	?	1	1	1	1	0	0	0	1	0	0	0	0	1	0	0	0	0	1	1	1
			?	r	r	?	r	r	r	r	r	r	r	r	r	r	r	?	r	r	r	r	r	r	n	r	r	r	r	r	r	r	r	r		r	r	r
3	102	2	0	0	0	0	0	0	0	1	0	0	1	1	0	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0	1	0	0	0	0	1	1	0
4	102R	2	0	0	0	0	0	0	0	1	1	0	1	1	0	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0	1	0	0	0	0	1	1	1
			r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	
5	201	3	0	0	0	0	0	0	0	0	1	0	1	1	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	1	1
6	201R	3	0	0	0	0	0	0	0	0	1	0	1	1	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	1	1
			r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r
7	202	3	0	0	0	0	0	0	0	1	0	0	1	1	0	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0	1	0	0	0	0	1	1	0
8	202R	3	0	0	0	0	0	0	0	1	0	0	?	1	0	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0	1	0	0	0	0	1	1	0
			r	r	r	r	r	r	r	r	r	r	?	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r
9	204	3	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	1	1	1	1	0	0	1	0	0	0	0	0	1	0	0	0	0	1	1	1
10	204R	3	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	1	1	1	1	0	0	1	0	0	0	0	0	1	0	0	0	0	1	1	1
			r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r
11	207	3	0	0	0	?	0	0	0	0	0	0	1	1	0	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0	1	0	0	0	0	1	1	0
12	207R	3	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0	1	0	0	0	0	1	1	0
			r	r	r	?	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r
13	208	3	0	0	0	0	0	0	0	0	0	1	1	1	0	1	0	0	1	1	1	1	0	0	0	0	0	0	0	0	1	1	0	0	0	1	1	0
14	208R	3	0	0	0	0	0	0	0	0	0	1	1	1	0	1	0	0	1	1	1	1	0	0	0	0	0	0	0	0	1	1	0	0	0	1	1	0
			r	r	r	r	r	r	r	r	r	r	r	r	0	r	r	r	r	r	r	r	r	r	0	r	r	r	r	r	r	r	r	r	r	r	r	r	r
15	292	4	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0	1	0	1	0	0	1	1	1
16	292R	4	0	0	0	1	0	0	0	0	0	0	?	1	0	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0	?	0	1	0	0	1	1	1
			r	r	r	n	r	r	r	r	r	r	?	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	?	r	r	r	r	r	r	r
17	386	5	1	0	0	0	0	0	0	1	1	0	1	1	0	0	0	0	1	1	1	0	1	0	0	0	0	0	0	0	1	0	1	1	0	0	0	0
18	386R	5	?	0	0	0	0	0	0	1	0	0	1	1	0	0	0	0	1	1	1	0	1	0	0	0	0	0	0	0	1	0	1	1	0	0	0	0
			?	r	r	r	r	r	r	r	n	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r
19	482	6	0	0	1	0	1	1	1	1	0	0	0	1	0	0	1	0	1	1	1	1	1	0	0	0	0	0	0	0	1	0	1	1	1	1	1	0
20	482R	6	0	0	1	0	1	1	1	1	0	0	0	1	0	0	1	0	1	1	1	1	1	0	0	0	0	0	0	0	1	0	1	1	1	1	1	0
			r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r
21	488	6	0	0	1	0	0	0	0	1	0	0	1	1	0	0	0	1	1	1	1	0	0	1	0	0	1	1	0	0	1	1	1	1	0	0	0	1
22	488R	6	0	0	1	0	0	0	0	1	0	0	1	1	0	0	0	1	1	1	1	0	0	1	0	0	1	1	0	0	1	1	1	1	0	0	0	1
			r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r
23	497	6	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	1	1	1	0	0	0	?	1	0	1	0	0	1	0	0	0	0	1	1	0
24	497R	6	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	1	1	1	0	0	0	?	0	0	1	0	0	1	0	0	0	0	1	1	0
			r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	?	r	r	r	r	r	r	?	n	r	r	r	r	r	r	r	r	r	r	r	r
	sum r		10	12	12	9	12	12	12	12	11	12	9	12	12	12	12	11	12	12	12	12	12	12	10	11	12	12	12	12	11	12	12	12	11	12	12	11
	%		100	120	100	90	100	100	100	100	92	100	100	100	100	100	100	100	100	100	100	100	100	100	91	92	100	100	100	100	100	100	100	100	92	100	100	92

No	Sam.	PN.	353	354	356	358	361	362	371	374	378	380	381	384	391	394	395	396	407	410	425	435	445	446	462
1	100	2	1	1	0	1	0	1	1	0	1	1	1	0	0	0	1	0	?	0	0	1	1	1	1
2	100R	2	?	0	0	?	0	?	?	0	1	1	1	0	0	0	1	0	?	0	?	1	1	1	1
			?	?	r	?	r	?	?	r	r	r	r	r	r	r	r	0	?	r	?	r	r	r	r
3	102	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	?	0	?	1	0	1	1
4	102R	2	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	?	0	?	1	0	1	1
			r	r	r	r	r	r	r	r	r	n	r	r	r	r	r	r	?	r	?	r	r	r	r
5	201	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	?	0	0	0	0	0	1
6	201R	3	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	?	0	0	0	0	0	?
			r	r	r	r	r	r	r	r	r	n	r	r	r	r	r	r	?	r	r	r	r	r	?
7	202	3	0	0	0	0	1	0	0	0	0	1	1	0	0	1	0	1	?	0	0	1	1	1	1
8	202R	3	0	0	0	0	1	0	0	0	0	1	1	0	0	1	0	1	?	0	0	1	1	1	1
			r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	?	r	r	r	r	r	r
9	204	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	?	0	0	1	0	1	1
10	204R	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	?	0	0	1	0	1	1
			r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	?	r	r	r	r	r	r
11	207	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	?	0	0	0	0	1	0
12	207R	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	?	0	0	0	0	0	0
			r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	?	r	r	r	r	?	r
13	208	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	?	0	0	1	0	1	?
14	208R	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	?	0	0	1	0	0	?
			r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	?	r	r	r	?	?	?
15	292	4	1	1	0	0	0	0	0	0	0	1	1	0	0	0	0	0	?	0	0	1	0	1	0
16	292R	4	1	1	0	0	0	0	0	0	0	1	1	0	0	0	0	0	?	0	0	1	0	1	0
			r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	?	r	r	r	r	r	r
17	386	5	0	0	?	?	?	0	0	0	0	?	0	?	0	?	0	0	1	0	0	1	0	1	0
18	386R	5	0	0	1	1	1	0	0	0	0	1	0	1	0	1	0	0	1	0	0	?	0	1	0
			r	r	?	?	?	r	r	r	r	?	r	?	r	?	r	r	r	r	r	?	r	r	r
19	482	6	0	0	1	1	0	0	1	1	0	1	1	0	1	0	0	0	?	1	?	1	1	1	1
20	482R	6	0	0	1	1	0	0	1	1	0	1	1	0	1	0	0	0	?	1	?	1	1	1	1
			r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	?	r	?	r	r	r	r
21	488	6	0	0	0	0	0	0	0	0	0	1	0	1	0	0	1	1	?	0	?	?	0	1	?
22	488R	6	0	0	1	0	0	0	0	0	0	1	0	1	0	0	1	1	?	0	?	?	1	0	?
			r	r	n	r	r	r	r	r	r	r	r	r	r	r	r	r	?	r	?	?	n	n	?
23	497	6	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	?	0	?	1	1	0	1
24	497R	6	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	?	0	?	1	1	0	1
			r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	?	r	?	r	r	r	r
sum r			11	11	10	10	11	11	11	11	12	9	12	11	12	11	12	12	1	12	7	10	10	9	9
%			100	100	91	100	100	100	100	100	100	90	100	100	100	100	100	100	100	100	100	100	91	90	100

Appendix 4 Fragment and allelic frequencies in each population

Pop.	Total	Abb	Bak	Beb	Del	Did	Fin	Gur	Gin	Gon	Har
locus	Relative frequency of fragments in each population										
82	0.773	0.857	0.944	0.913	0.692	0.870	0.783	0.850	0.833	0.818	0.789
84	0.345	0.238	0.333	0.609	0.192	0.174	0.522	0.600	0.333	0.227	0.211
86	0.954	0.952	0.944	1.000	0.885	1.000	0.957	1.000	0.889	0.909	0.947
89	0.960	1.000	0.944	1.000	0.923	1.000	0.957	1.000	0.944	1.000	0.789
93	0.987	1.000	1.000	1.000	0.923	1.000	1.000	1.000	1.000	1.000	1.000
96	0.990	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
99	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
101	0.317	0.571	0.056	0.217	0.154	0.391	0.391	0.150	0.333	0.727	0.263
102	0.479	0.714	0.667	0.826	0.269	0.478	0.348	0.400	0.889	0.182	0.632
104	0.994	1.000	0.944	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
107	0.086	0.000	0.111	0.000	0.077	0.043	0.087	0.050	0.222	0.136	0.000
108	0.252	0.333	0.278	0.304	0.077	0.261	0.217	0.250	0.389	0.182	0.263
111	0.994	1.000	1.000	1.000	0.962	1.000	1.000	1.000	1.000	1.000	1.000
113	0.655	0.810	0.889	0.870	0.615	0.609	0.478	0.850	0.778	0.455	0.632
116	0.193	0.238	0.111	0.174	0.077	0.087	0.217	0.250	0.333	0.136	0.316
118	0.748	0.810	0.556	0.739	0.577	0.826	0.217	0.900	0.944	0.955	0.842
119	0.697	0.714	0.778	0.739	0.731	0.739	0.870	0.600	0.722	0.773	0.737
128	0.759	0.524	0.833	0.913	0.769	0.739	0.773	0.850	0.833	0.636	0.778
130	0.288	0.619	0.222	0.348	0.077	0.087	0.348	0.150	0.056	0.136	0.316
132	0.796	0.857	0.778	0.870	0.769	0.522	0.783	0.950	0.722	0.864	0.947
134	0.153	0.143	0.333	0.391	0.038	0.087	0.174	0.200	0.056	0.227	0.105
136	0.853	0.952	1.000	0.957	0.731	0.913	0.739	1.000	1.000	1.000	0.947
137	0.672	0.571	0.667	0.652	0.423	0.870	0.522	0.750	0.944	0.773	0.579
139	0.473	0.714	0.389	0.522	0.654	0.609	0.348	0.550	0.556	0.500	0.895
142	0.990	1.000	1.000	0.957	0.962	1.000	1.000	1.000	1.000	1.000	1.000
144	0.994	0.952	1.000	0.957	1.000	1.000	1.000	1.000	1.000	1.000	1.000
147	0.242	0.238	0.222	0.304	0.000	0.435	0.261	0.150	0.056	0.364	0.158
149	0.078	0.048	0.000	0.043	0.038	0.130	0.087	0.000	0.000	0.091	0.105
150	0.340	0.429	0.167	0.435	0.192	0.652	0.348	0.400	0.278	0.182	0.368
151	0.511	0.571	0.556	0.565	0.846	0.435	0.609	0.300	0.389	0.409	0.579
154	0.766	0.857	0.778	0.696	0.885	0.727	0.870	0.800	0.667	0.864	0.684
159	0.416	0.333	0.278	0.348	0.462	0.652	0.478	0.250	0.111	0.455	0.421
160	0.384	0.714	0.556	0.348	0.231	0.174	0.348	0.300	0.611	0.545	0.474
162	0.866	0.952	0.889	0.870	0.923	0.870	0.652	0.950	0.833	0.955	0.842
163	0.912	0.905	0.833	0.955	1.000	1.000	0.913	0.950	0.833	1.000	0.947
166	0.124	0.048	0.056	0.087	0.077	0.043	0.348	0.100	0.111	0.227	0.105
170	0.700	0.762	0.833	0.870	0.385	0.739	0.522	0.850	0.611	0.818	0.789
174	0.960	0.905	1.000	1.000	0.962	1.000	0.870	1.000	0.889	0.955	1.000
177	0.767	0.952	0.778	0.870	0.500	0.783	0.652	0.900	0.889	0.818	0.789
178	0.880	1.000	0.944	0.957	0.692	0.957	0.435	0.750	0.778	1.000	0.947
179	0.122	0.048	0.056	0.000	0.000	0.130	0.261	0.200	0.167	0.227	0.211
182	0.194	0.286	0.111	0.043	0.000	0.174	0.045	0.150	0.278	0.364	0.263
187	0.168	0.095	0.000	0.087	0.269	0.130	0.304	0.250	0.000	0.318	0.158
189	0.765	0.857	0.889	0.783	0.462	0.913	0.609	0.850	0.833	0.864	0.737
192	0.210	0.095	0.167	0.217	0.038	0.304	0.087	0.250	0.278	0.273	0.368

Pop.	Hir	Jim	Kem	Mek	Met	Sha	She	But	Tep	Won	Yay	Zeg
locus	Relative frequency of fragments in each population											
82	0.545	0.818	0.762	0.471	0.769	0.565	0.909	0.870	0.636	0.952	0.421	0.880
84	0.136	0.500	0.667	0.059	0.346	0.348	0.273	0.478	0.318	0.333	0.474	0.200
86	0.864	1.000	1.000	0.824	0.885	0.957	0.955	1.000	1.000	1.000	1.000	1.000
89	0.909	0.955	1.000	0.941	0.923	1.000	1.000	0.826	1.000	1.000	1.000	1.000
93	1.000	1.000	1.000	0.882	0.962	1.000	1.000	1.000	1.000	1.000	0.947	1.000
96	1.000	1.000	0.952	0.941	1.000	1.000	1.000	1.000	1.000	1.000	0.842	1.000
99	1.000	1.000	1.000	0.941	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
101	0.500	0.227	0.238	0.706	0.385	0.304	0.318	0.130	0.364	0.095	0.105	0.360
102	0.227	0.500	0.333	0.235	0.423	0.261	0.727	0.478	0.318	0.714	0.842	0.280
104	1.000	1.000	1.000	0.941	1.000	0.957	1.000	1.000	1.000	1.000	1.000	1.000
107	0.045	0.045	0.000	0.000	0.192	0.000	0.227	0.043	0.045	0.095	0.316	0.160
108	0.227	0.364	0.429	0.176	0.308	0.217	0.318	0.391	0.227	0.095	0.105	0.160
111	1.000	1.000	1.000	0.941	0.962	1.000	1.000	1.000	1.000	1.000	1.000	1.000
113	0.591	0.773	0.762	0.353	0.615	0.522	0.818	0.522	0.545	0.667	0.895	0.480
116	0.318	0.227	0.190	0.412	0.346	0.391	0.045	0.087	0.091	0.143	0.000	0.120
118	0.682	0.545	0.619	0.941	0.615	0.957	0.909	0.739	0.591	0.952	0.842	0.840
119	0.727	0.545	0.857	0.647	0.731	0.826	0.636	0.739	0.364	0.762	0.158	0.840
128	0.409	0.727	0.762	0.176	0.808	0.913	0.818	0.826	0.818	0.905	0.737	1.000
130	0.091	0.182	0.714	0.647	0.231	0.217	0.500	0.130	0.136	0.286	0.421	0.520
132	0.909	0.818	0.857	0.941	0.692	0.913	0.818	0.696	0.909	0.857	0.789	0.400
134	0.000	0.136	0.048	0.059	0.308	0.087	0.455	0.000	0.045	0.143	0.316	0.040
136	0.864	0.864	0.952	0.941	0.808	0.826	0.727	0.826	0.682	0.762	0.421	0.920
137	0.727	0.636	0.476	0.824	0.500	0.783	0.818	0.652	0.545	0.619	0.737	0.840
139	0.455	0.273	0.333	0.294	0.385	0.304	0.591	0.609	0.227	0.429	0.158	0.560
142	1.000	0.955	1.000	0.941	1.000	1.000	0.955	1.000	1.000	1.000	1.000	1.000
144	1.000	0.955	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
147	0.136	0.318	0.143	0.000	0.192	0.130	0.500	0.435	0.455	0.429	0.211	0.120
149	0.045	0.000	0.238	0.000	0.077	0.043	0.045	0.130	0.091	0.048	0.316	0.120
150	0.136	0.273	0.524	0.235	0.308	0.391	0.136	0.304	0.182	0.333	0.842	0.400
151	0.409	0.364	0.762	0.294	0.500	0.652	0.727	0.478	0.273	0.619	0.053	0.640
154	0.773	0.591	0.667	0.647	0.846	0.696	0.818	0.957	0.818	0.762	0.526	0.800
159	0.227	0.318	0.619	0.471	0.269	0.783	0.227	0.565	0.091	0.381	0.737	0.600
160	0.409	0.500	0.143	0.412	0.346	0.130	0.591	0.304	0.409	0.524	0.105	0.400
162	0.773	0.864	0.810	0.882	0.923	0.913	0.909	0.783	0.864	0.810	0.895	0.880
163	0.682	0.818	1.000	0.941	0.923	0.696	0.955	0.913	0.955	0.857	0.947	1.000
166	0.091	0.045	0.143	0.118	0.231	0.043	0.273	0.000	0.136	0.000	0.053	0.320
170	0.455	0.364	0.952	0.765	0.615	0.652	0.864	0.696	0.500	0.762	0.895	0.840
174	0.955	1.000	1.000	0.824	0.923	0.957	1.000	1.000	1.000	1.000	1.000	0.880
177	0.773	0.500	0.714	0.941	0.846	0.739	0.818	0.609	0.591	0.857	1.000	0.720
178	0.955	0.909	1.000	0.882	0.846	0.870	1.000	0.913	0.818	0.857	0.895	1.000
179	0.227	0.091	0.095	0.000	0.154	0.043	0.227	0.043	0.136	0.095	0.053	0.200
182	0.500	0.136	0.190	0.529	0.231	0.130	0.136	0.087	0.091	0.143	0.105	0.360
187	0.091	0.136	0.429	0.000	0.192	0.174	0.182	0.087	0.364	0.190	0.105	0.040
189	0.636	0.636	0.667	1.000	0.654	0.870	0.864	0.783	0.500	0.762	0.842	0.960
192	0.273	0.182	0.095	0.059	0.077	0.435	0.273	0.348	0.318	0.143	0.158	0.200

Pop.	Total	Abb	Bak	Beb	Del	Did	Fin	Gur	Gin	Gon	Har
locus					Relative frequency of fragments in each population						
194	0.215	0.333	0.167	0.304	0.077	0.261	0.136	0.250	0.278	0.318	0.105
195	0.535	0.857	0.500	0.696	0.154	0.348	0.409	0.450	0.611	0.591	0.632
198	0.200	0.286	0.333	0.348	0.115	0.087	0.130	0.350	0.222	0.227	0.316
202	0.642	0.905	0.667	0.696	0.346	0.609	0.409	0.700	0.722	0.727	0.632
204	0.193	0.190	0.278	0.174	0.077	0.217	0.130	0.350	0.167	0.136	0.211
207	0.964	1.000	0.944	1.000	1.000	0.913	1.000	1.000	0.944	1.000	0.947
208	0.580	0.381	0.500	0.391	0.692	0.870	0.478	0.900	0.111	0.545	0.632
210	0.786	1.000	0.778	0.739	0.923	0.652	0.783	0.700	0.778	0.727	0.947
214	0.326	0.571	0.333	0.391	0.115	0.130	0.217	0.400	0.389	0.409	0.316
221	0.363	0.524	0.222	0.174	0.115	0.478	0.261	0.350	0.444	0.591	0.368
223	0.130	0.095	0.111	0.043	0.038	0.261	0.043	0.250	0.167	0.091	0.211
224	0.412	0.524	0.444	0.609	0.192	0.435	0.261	0.200	0.333	0.364	0.368
225	0.450	0.810	0.333	0.478	0.077	0.739	0.391	0.350	0.722	0.545	0.526
227	0.355	0.429	0.444	0.565	0.115	0.478	0.261	0.400	0.222	0.136	0.368
229	0.080	0.000	0.056	0.261	0.038	0.087	0.043	0.150	0.056	0.000	0.000
232	0.214	0.143	0.167	0.217	0.000	0.348	0.435	0.450	0.222	0.273	0.158
233	0.183	0.238	0.167	0.174	0.077	0.130	0.043	0.300	0.278	0.318	0.158
235	0.500	0.619	0.500	0.652	0.231	0.696	0.391	0.500	0.222	0.409	0.368
238	0.271	0.333	0.222	0.348	0.038	0.478	0.304	0.150	0.444	0.409	0.316
242	0.088	0.000	0.000	0.000	0.038	0.348	0.000	0.150	0.000	0.000	0.105
246	0.424	0.476	0.444	0.522	0.115	0.739	0.478	0.450	0.222	0.136	0.474
249	0.092	0.095	0.000	0.261	0.038	0.130	0.087	0.150	0.056	0.045	0.053
250	0.162	0.095	0.167	0.435	0.038	0.217	0.000	0.300	0.056	0.091	0.000
252	0.145	0.095	0.167	0.174	0.038	0.087	0.043	0.350	0.222	0.182	0.105
255	0.086	0.095	0.056	0.217	0.038	0.130	0.000	0.200	0.000	0.045	0.000
262	0.172	0.000	0.111	0.087	0.192	0.174	0.304	0.200	0.111	0.136	0.158
265	0.615	0.714	0.667	0.304	0.846	0.522	0.783	0.350	0.444	0.773	0.421
271	0.185	0.095	0.167	0.130	0.115	0.261	0.130	0.500	0.278	0.227	0.053
275	0.924	1.000	0.889	0.957	1.000	0.870	0.957	0.900	0.833	0.955	0.842
279	0.042	0.000	0.000	0.000	0.077	0.000	0.000	0.000	0.111	0.182	0.000
293	0.996	1.000	1.000	1.000	0.962	1.000	1.000	1.000	1.000	1.000	1.000
309	0.017	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
312	0.162	0.381	0.278	0.174	0.000	0.304	0.000	0.000	0.222	0.045	0.263
315	0.032	0.000	0.000	0.043	0.000	0.000	0.000	0.150	0.000	0.136	0.000
320	0.061	0.000	0.118	0.000	0.000	0.130	0.000	0.000	0.111	0.045	0.000
322	0.941	0.952	0.824	1.000	0.923	0.870	1.000	0.950	0.833	0.955	0.947
326	0.085	0.100	0.000	0.087	0.000	0.087	0.130	0.050	0.056	0.682	0.053
330	0.101	0.105	0.200	0.174	0.043	0.043	0.043	0.263	0.067	0.190	0.056
332	0.108	0.316	0.067	0.000	0.000	0.174	0.000	0.000	0.200	0.286	0.222
335	0.719	0.684	0.733	0.826	0.870	0.652	0.696	0.632	1.000	0.286	0.778
339	0.649	0.789	0.643	0.870	0.500	0.565	0.636	0.842	0.867	0.524	0.625
354	0.245	0.333	0.154	0.409	0.067	0.217	0.045	0.313	0.154	0.095	0.375
371	0.090	0.000	0.000	0.095	0.000	0.043	0.150	0.067	0.154	0.050	0.071
446	0.601	0.706	0.600	0.750	0.211	0.571	0.286	0.643	0.750	0.421	0.750
462	0.459	0.571	0.455	0.750	0.158	0.333	0.333	0.571	0.600	0.389	0.818

Pop.	Hir	Jim	Kem	Mek	Met	Sha	She	But	Tep	Won	Yay	Zeg
locus	Relative frequency of fragments in each population											
194	0.182	0.318	0.190	0.176	0.192	0.174	0.182	0.217	0.364	0.190	0.158	0.160
195	0.591	0.545	0.857	0.529	0.462	0.478	0.500	0.522	0.455	0.476	0.684	0.560
198	0.136	0.136	0.190	0.000	0.462	0.000	0.227	0.043	0.318	0.190	0.211	0.080
202	0.591	0.545	0.714	1.000	0.615	0.565	0.818	0.391	0.636	0.476	0.737	0.800
204	0.136	0.091	0.571	0.000	0.192	0.217	0.364	0.043	0.182	0.238	0.263	0.080
207	1.000	1.000	0.810	1.000	0.923	0.957	1.000	1.000	1.000	1.000	1.000	0.800
208	0.636	0.455	0.762	0.118	0.654	0.609	0.773	0.652	0.455	0.667	0.368	0.840
210	0.727	0.545	0.905	0.588	0.808	0.783	0.773	0.870	0.727	0.762	0.789	0.920
214	0.364	0.364	0.190	0.294	0.462	0.174	0.500	0.043	0.455	0.429	0.474	0.240
221	0.545	0.500	0.286	0.647	0.385	0.304	0.273	0.261	0.318	0.190	0.316	0.520
223	0.136	0.182	0.333	0.059	0.231	0.087	0.045	0.087	0.091	0.095	0.053	0.160
224	0.273	0.364	0.476	0.235	0.462	0.217	0.591	0.435	0.455	0.571	0.632	0.600
225	0.682	0.318	0.476	0.706	0.385	0.391	0.318	0.261	0.318	0.429	0.316	0.480
227	0.273	0.409	0.238	0.059	0.769	0.130	0.727	0.217	0.682	0.238	0.211	0.320
229	0.045	0.091	0.143	0.000	0.038	0.087	0.227	0.000	0.136	0.048	0.263	0.000
232	0.045	0.318	0.238	0.176	0.308	0.087	0.227	0.174	0.227	0.143	0.211	0.160
233	0.227	0.273	0.143	0.235	0.192	0.174	0.227	0.087	0.182	0.190	0.158	0.120
235	0.273	0.318	0.762	0.118	0.692	0.478	0.727	0.435	0.500	0.619	0.632	0.720
238	0.136	0.227	0.190	0.118	0.231	0.174	0.273	0.217	0.136	0.429	0.526	0.320
242	0.000	0.091	0.238	0.000	0.115	0.087	0.091	0.000	0.136	0.143	0.158	0.200
246	0.227	0.227	0.667	0.059	0.538	0.435	0.682	0.304	0.364	0.524	0.526	0.640
249	0.000	0.045	0.048	0.000	0.154	0.043	0.227	0.043	0.182	0.095	0.053	0.160
250	0.000	0.045	0.000	0.118	0.308	0.000	0.545	0.130	0.273	0.333	0.158	0.200
252	0.182	0.227	0.143	0.176	0.115	0.043	0.227	0.043	0.273	0.143	0.105	0.120
255	0.000	0.091	0.095	0.000	0.154	0.130	0.045	0.000	0.045	0.048	0.158	0.280
262	0.227	0.318	0.286	0.000	0.115	0.130	0.227	0.130	0.364	0.238	0.105	0.120
265	0.545	0.636	0.857	1.000	0.423	0.522	0.773	0.261	0.429	0.762	0.579	0.920
271	0.045	0.500	0.143	0.059	0.231	0.000	0.318	0.217	0.182	0.143	0.105	0.160
275	0.955	0.909	1.000	0.882	0.923	0.913	0.909	1.000	0.773	0.857	0.947	1.000
279	0.182	0.045	0.048	0.000	0.038	0.000	0.000	0.174	0.000	0.000	0.053	0.000
293	1.000	0.955	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
309	0.000	0.000	0.000	0.059	0.077	0.000	0.000	0.000	0.000	0.190	0.000	0.040
312	0.364	0.045	0.048	0.412	0.154	0.130	0.000	0.130	0.045	0.095	0.105	0.440
315	0.045	0.000	0.000	0.059	0.000	0.000	0.045	0.000	0.045	0.143	0.000	0.040
320	0.045	0.273	0.143	0.000	0.077	0.000	0.091	0.130	0.000	0.048	0.158	0.000
322	0.955	0.955	1.000	1.000	0.840	1.000	0.955	1.000	1.000	0.952	0.895	0.880
326	0.045	0.091	0.000	0.118	0.038	0.000	0.000	0.087	0.045	0.048	0.000	0.120
330	0.053	0.158	0.105	0.067	0.040	0.000	0.143	0.000	0.190	0.150	0.105	0.080
332	0.263	0.053	0.000	0.467	0.000	0.095	0.000	0.000	0.000	0.1.00000	0.053	0.240
335	0.842	0.579	0.895	0.667	0.680	0.857	0.667	0.636	0.667	0.850	0.842	0.600
339	0.833	0.526	0.684	0.733	0.880	0.476	0.667	0.286	0.714	0.158	0.778	0.680
354	0.111	0.105	0.263	0.143	0.360	0.286	0.316	0.143	0.190	0.333	0.556	0.320
371	0.000	0.000	0.056	0.071	0.000	0.263	0.158	0.190	0.200	0.111	0.222	0.042
446	0.692	0.471	0.632	0.643	0.762	0.588	0.625	0.526	0.722	0.588	0.778	0.682
462	0.583	0.286	0.737	0.667	0.476	0.643	0.500	0.158	0.389	0.389	0.625	0.227

Appendix 5: List of Loci and that showed significant differentiation ($p < 0.001$) among populations and the percent proportion of bands present among individuals analysed in each population

Pop/locus	**82**	**84**	**101**	**102**	**107**	**113**	**116**	**118**	**119**	**128**	**130**	**132**	**134**	**136**	**147**	**150**	**151**	**159**	**160**	**163**	**170**	**177**	**178**
Abbay	86	24	57	71	0	81	24	81	71	52	62	86	14	95	24	43	57	33	71	90	76	95	100
Finoteselam	77	50	36	36	5	50	23	18	86	77	36	77	18	73	27	36	59	50	32	91	50	64	45
Gondar	82	23	73	18	14	45	14	95	77	64	14	86	23	100	36	18	41	45	55	100	82	82	100
Kemisse	76	67	24	33	0	76	19	62	86	76	71	86	5	95	14	52	76	62	14	100	95	71	100
Mekelle	47	6	71	24	0	35	41	94	65	18	65	94	6	94	0	24	29	47	41	94	76	94	88
Zeghie	88	20	36	28	16	48	12	84	84	100	52	40	4	92	12	40	64	60	40	100	84	72	100
Bako	94	33	6	67	11	89	11	56	78	83	22	78	33	100	22	17	56	28	56	83	83	78	94
Bebeka	91	61	22	83	0	87	17	74	74	91	35	87	39	96	30	43	57	35	35	91	87	87	96
Didessa	87	17	39	48	4	61	9	83	74	74	9	52	9	91	43	65	43	65	17	100	74	78	96
Guraferda	85	60	15	40	5	85	25	90	60	85	15	95	20	100	15	40	30	25	30	95	85	90	75
Jimma	82	50	23	50	5	77	23	55	55	73	18	82	14	86	32	27	36	32	50	82	36	50	91
Meti	77	35	38	42	19	62	35	62	73	81	23	69	31	81	19	31	50	27	35	92	62	85	85
Sheko	91	27	32	73	23	82	5	91	64	82	50	82	45	73	50	14	73	23	59	95	86	82	100
Tepi	64	32	36	32	5	55	9	59	36	82	14	91	5	68	45	18	27	9	41	95	50	59	82
Yayu	42	47	11	84	32	89	0	84	16	74	42	79	32	42	21	84	5	74	11	95	89	100	89
Butajira	87	48	13	48	4	52	9	74	74	83	13	70	0	83	43	30	48	57	30	91	70	61	91
Dolo-Mana	69	19	15	27	8	62	8	58	73	77	8	77	4	73	0	19	85	46	23	100	38	50	69
Ginnr	83	33	33	89	22	78	33	94	72	83	6	72	6	100	6	28	39	11	61	83	61	89	78
Harar	79	21	26	63	0	63	32	84	74	74	32	95	11	95	16	37	58	42	47	95	79	79	95
Hirna	55	14	50	23	5	59	32	68	73	41	9	91	0	86	14	14	41	23	41	68	45	77	95
Shashemene	57	35	30	26	0	52	39	96	83	91	22	91	9	83	13	39	65	78	13	70	65	74	87
Wondogenet	95	33	10	71	10	67	14	95	76	90	29	86	14	76	43	33	62	38	52	86	76	86	86
Pop/locus	**182**	**187**	**189**	**195**	**198**	**202**	**204**	**207**	**208**	**221**	**225**	**227**	**235**	**242**	**246**	**250**	**265**	**271**	**309**	**312**	**332**	**335**	**462**
Abbay	29	10	86	86	29	90	19	100	38	52	81	43	62	0	48	10	71	10	0	38	29	62	38
Finoteselam	5	32	59	36	14	41	14	100	45	27	41	27	41	0	50	0	77	14	0	0	0	68	27
Gondar	36	32	86	59	23	73	14	100	55	59	55	14	41	0	14	9	77	23	0	5	27	27	32
Kemisse	19	43	67	86	19	71	57	81	76	29	48	24	76	24	67	0	86	14	0	5	0	81	67
Mekelle	53	0	100	53	0	100	0	100	12	65	71	6	12	0	6	12	100	6	6	41	41	59	12
Zeghie	36	4	96	56	8	80	8	80	84	52	48	32	72	20	64	20	92	16	4	44	24	60	20
Bako	11	0	89	50	33	67	28	94	50	22	33	44	50	0	44	17	67	17	0	28	6	61	28
Bebeka	4	9	78	70	35	70	17	100	39	17	48	57	65	0	52	43	30	13	0	17	0	83	65
Didessa	17	13	91	35	9	61	22	91	87	48	74	48	70	35	74	22	52	26	0	30	17	65	30
Guraferda	15	25	85	45	35	70	35	100	90	35	35	40	50	15	45	30	35	50	0	0	0	60	40
Jimma	14	14	64	55	14	55	9	100	45	50	32	41	32	9	23	5	64	50	0	5	5	50	18
Meti	23	19	65	46	46	62	19	92	65	38	38	77	69	12	54	31	42	23	8	15	0	65	38
Sheko	14	18	86	50	23	82	36	100	77	27	32	73	73	9	68	55	77	32	0	0	0	64	36
Tepi	9	36	50	45	32	64	18	100	45	32	32	68	50	14	36	27	41	18	0	5	0	64	32
Yayu	11	11	84	68	21	74	26	100	37	32	32	21	63	16	53	16	58	11	0	11	5	84	53
Butajira	9	9	78	52	4	39	4	100	65	26	26	22	43	0	30	13	26	22	0	13	0	61	13
Dolo-Mana	0	27	46	15	12	35	8	100	69	12	8	12	23	4	12	4	85	12	0	0	0	77	12
Ginnr	28	0	83	61	22	72	17	94	11	44	72	22	22	0	22	6	44	28	0	22	17	83	33
Harar	26	16	74	63	32	63	21	95	63	37	53	37	37	11	47	0	42	5	0	26	21	74	47
Hirna	50	9	64	59	14	59	14	100	64	55	68	27	27	0	23	0	55	5	0	36	23	73	32
Shashemene	13	17	87	48	0	57	22	96	61	30	39	13	48	9	43	0	52	0	0	13	9	78	39
Wondogenet	14	19	76	48	19	48	24	100	67	19	43	24	62	14	52	33	76	14	19	10	10	81	33

Appendix 6: Eigen Values by Axis and their respective percentage

Axis No.	EigenValue	Percent	Axis No.	EigenValue	Percent
1	110.602	29.3	26	3.640	1.0
2	41.499	11.0	27	3.462	0.9
3	26.939	7.1	28	3.223	0.9
4	18.634	4.9	29	2.905	0.8
5	18.085	4.8	30	2.756	0.7
6	14.883	3.9	31	2.560	0.7
7	14.553	3.8	32	2.439	0.6
8	13.180	3.5	33	2.191	0.6
9	12.134	3.2	34	1.974	0.5
10	10.304	2.7	35	1.900	0.5
11	10.197	2.7	36	1.585	0.4
12	9.631	2.5	37	1.387	0.4
13	8.275	2.2	38	1.300	0.3
14	8.247	2.2	39	1.194	0.3
15	7.539	2.0	40	1.127	0.3
16	6.902	1.8	41	0.991	0.3
17	6.522	1.7	42	0.784	0.2
18	6.032	1.6	43	0.572	0.2
19	5.664	1.5	44	0.476	0.1
20	5.438	1.4	45	0.415	0.1
21	5.186	1.4	46	0.340	0.1
22	5.036	1.3	47	0.264	0.1
23	4.534	1.2	48	0.148	0.0
24	4.131	1.1	49	0.048	0.0
25	3.903	1.0	50	0.028	0.0
			Sum	378.051	100.0

Appendix 7 cpSSR data of the 22 populations of *C. africana*

No.	Population	Sample	ccmp2	ccmp10	Haplotype
1	Abbay	134	188	106	C
2	Abbay	135	188	106	C
3	Abbay	136	188	106	C
4	Abbay	380	187	106	B
5	Abbay	381	187	106	B
6	Abbay	382	187	106	B
7	Abbay	383	187	106	B
8	Abbay	512	187	106	B
9	Bako	002	188	106	C
10	Bako	009	187	106	B
11	Bako	011	187	106	B
12	Bako	012	187	106	B
13	Bako	036	188	106	C
14	Bako	308	187	106	B
15	Bako	309	187	106	B
16	Bako	310	187	106	B
17	Bebeka	010	187	106	B
18	Bebeka	017	187	106	B
19	Bebeka	019	187	106	B
20	Bebeka	020	187	106	B
21	Bebeka	023	187	106	B
22	Bebeka	024	187	106	B
23	Bebeka	030	187	106	B
24	Bebeka	037	187	106	B
25	Butajira	047	187	106	B
26	Butajira	048	187	106	B
27	Butajira	394	187	106	B
28	Butajira	395	187	106	B
29	Butajira	396	187	106	B
30	Butajira	397	187	106	B
31	Butajira	398	187	106	B
32	Butajira	399	187	106	B
33	Butajira	798	187	106	B
34	Didessa	148	187	106	B

No.	Population	Sample	ccmp2	ccmp10	Haplotype
102	Hirna	126	187	106	B
103	Hirna	127	187	106	B
104	Hirna	129	187	106	B
105	Hirna	130	187	106	B
106	Hirna	193	187	106	B
107	Hirna	194	187	106	B
108	Hirna	195	187	106	B
109	Hirna	197	187	106	B
110	Hirna	269	187	106	B
111	Hirna	270	187	106	B
112	Hirna	271	187	106	B
113	Hirna	272	187	105	A
114	Jimma	039	187	106	B
115	Jimma	040	187	106	B
116	Jimma	041	187	106	B
117	Jimma	260	187	106	B
118	Jimma	261	187	106	B
119	Jimma	262	187	106	B
120	Jimma	263	187	106	B
121	Jimma	264	187	105	A
122	Jimma	265	187	106	B
123	Jimma	266	187	106	B
124	Kemmise	162	187	105	A
125	Kemmise	163	187	105	A
126	Kemmise	164	187	106	B
127	Kemmise	165	187	106	B
128	Kemmise	166	187	106	B
129	Kemmise	167	187	106	B
130	Kemmise	168	187	106	B
131	Kemmise	169	187	106	B
132	Kemmise	440	187	106	B
133	Mekelle	097	187	106	B
134	Mekelle	098	187	106	B
135	Mekelle	099	187	106	B

No.	Population	Sample	ccmp2	ccmp10	Haplotype
35	Didessa	150	187	105	A
36	Didessa	155	187	105	A
37	Didessa	372	187	106	B
38	Didessa	373	187	106	B
39	Didessa	374	187	106	B
40	Didessa	375	187	105	A
41	Didessa	376	187	106	B
42	Didessa	379	187	106	B
43	Didessa	476	187	106	B
44	Dolo-Mana	138	187	105	A
45	Dolo-Mana	139	187	105	A
46	Dolo-Mana	142	187	105	A
47	Dolo-Mana	145	187	105	A
48	Dolo-Mana	146	187	106	B
49	Dolo-Mana	147	187	106	B
50	Dolo-Mana	152	187	105	A
51	Dolo-Mana	800	187	106	B
52	Finoteselam	178	187	105	A
53	Finoteselam	179	187	105	A
54	Finoteselam	180	187	106	B
55	Finoteselam	181	187	106	B
56	Finoteselam	186	187	106	B
57	Finoteselam	187	187	106	B
58	Finoteselam	188	187	106	B
59	Finoteselam	229	187	106	B
60	Finoteselam	230	187	106	B
61	Ginnr	088	187	106	B
62	Ginnr	094	187	106	B
63	Ginnr	095	187	106	B
64	Ginnr	096	187	105	A
65	Ginnr	157	187	106	B
66	Ginnr	158	187	105	A
67	Ginnr	160	187	106	B
68	Ginnr	161	187	105	A

No.	Population	Sample	ccmp2	ccmp10	Haplotype
136	Mekelle	100	187	106	B
137	Mekelle	101	187	106	B
138	Mekelle	102	187	106	B
139	Mekelle	103	187	106	B
140	Mekelle	104	187	106	B
141	Mekelle	412	187	106	B
142	Meti	025	187	106	B
143	Meti	027	187	106	B
144	Meti	033	187	106	B
145	Meti	069	187	106	B
146	Meti	083	187	106	B
147	Meti	086	187	106	B
148	Meti	121	187	106	B
149	Meti	124	187	106	B
150	Meti	125	187	106	B
151	Meti	424	187	106	B
152	Meti	494	187	106	B
153	Shashemene	115	187	105	A
154	Shashemene	117	187	106	B
155	Shashemene	118	187	105	A
156	Shashemene	122	187	105	A
157	Shashemene	123	187	106	B
158	Shashemene	442	187	105	A
159	Shashemene	443	187	106	B
160	Shashemene	444	187	106	B
161	Sheko	182	187	105	A
162	Sheko	183	187	106	B
163	Sheko	184	187	105	A
164	Sheko	189	187	106	B
165	Sheko	190	187	105	A
166	Sheko	191	187	106	B
167	Sheko	289	187	106	B
168	Sheko	290	187	106	B
169	Sheko	532	187	106	B

No.	Population	Sample	ccmp2	ccmp10	Haplotype
69	Ginnr	522	187	106	B
70	Ginnr	523	187	106	B
71	Gondar	170	188	106	C
72	Gondar	171	187	106	B
73	Gondar	172	187	106	B
74	Gondar	173	187	106	B
75	Gondar	174	188	106	C
76	Gondar	175	188	106	C
77	Gondar	176	187	106	B
78	Gondar	177	187	106	B
79	Gondar	483	188	106	C
80	Gondar	484	188	106	C
81	Gondar	485	187	106	B
82	Gondar	489	188	106	C
83	Guraferda	003	187	106	B
84	Guraferda	004	187	106	B
85	Guraferda	005	187	106	B
86	Guraferda	006	187	106	B
87	Guraferda	013	187	106	B
88	Guraferda	014	187	106	B
89	Guraferda	015	187	106	B
90	Guraferda	016	187	106	B
91	Harar	087	187	106	B
92	Harar	089	187	106	B
93	Harar	090	187	106	B
94	Harar	149	187	106	B
95	Harar	151	187	106	B
96	Harar	281	187	106	B
97	Harar	282	187	106	B
98	Harar	283	187	106	B
99	Harar	545	187	106	B
100	Harar	546	187	106	B
101	Harar	797	187	106	B

No.	Population	Sample	ccmp2	ccmp10	Haplotype
170	Tepi	071	187	106	B
171	Tepi	074	187	106	B
172	Tepi	077	187	106	B
173	Tepi	078	187	106	B
174	Tepi	079	187	106	B
175	Tepi	081	187	106	B
176	Tepi	082	187	106	B
177	Tepi	222	187	106	B
178	Wondogenet	105	187	106	B
179	Wondogenet	106	187	105	A
180	Wondogenet	107	187	106	B
181	Wondogenet	337	187	106	B
182	Wondogenet	338	187	106	B
183	Wondogenet	339	187	106	B
184	Wondogenet	340	187	106	B
185	Wondogenet	341	187	105	A
186	Yayu	054	187	106	B
187	Yayu	055	187	106	B
188	Yayu	056	187	106	B
189	Yayu	057	187	106	B
190	Yayu	058	187	106	B
191	Yayu	059	187	106	B
192	Yayu	060	187	106	B
193	Yayu	795	187	106	B
194	Zeghie	110	187	106	B
195	Zeghie	111	187	106	B
196	Zeghie	112	187	106	B
197	Zeghie	113	188	106	C
198	Zeghie	114	187	106	B
199	Zeghie	116	187	106	B
200	Zeghie	410	187	106	B
201	Zeghie	411	187	106	B
202	Zeghie	794	188	106	C

Appendix 8 Tree Form Scoring Sheet

Forking (F) height (m)	Forking score
<3	1
3-6	2
6-9	3
9-12	4
12-15	5
≥15	6
Bole shape	**Bole shape (S) score**
Slanting and twisting	1
large bends and some waves	2
small bends	3
Fairly straight	4
Straight but not clear bole	5
Straight clear bole	6
Forking shape (F*S) interaction	**Tree Form**
<6	1 (poor)
6-12	2
12-18	3
18-24	4
24-30	5
≥30	6 (Excellent)

Appendix 9 Nei's unbiased genetic distance among the 22 populations of *Cordia africana* at cpSSRs

Bak	0.009																				
Beb	0.064	0.022																			
Del	0.406	0.320	0.231																		
Did	0.143	0.085	0.037	0.053																	
Fin	0.110	0.057	0.017	0.090	0.010																
Gur	0.143	0.085	0.038	0.053	0.014	0.010															
Gin	0.007	0.025	0.137	0.532	0.235	0.196	0.235														
Gon	0.064	0.022	0.000	0.231	0.038	0.017	0.038	0.137													
Har	0.064	0.022	0.000	0.231	0.038	0.017	0.038	0.137	0.000												
Hir	0.075	0.029	0.001	0.174	0.016	0.002	0.016	0.152	0.001	0.001											
Jim	0.078	0.031	0.002	0.163	0.012	0.001	0.012	0.156	0.002	0.002	0.004										
Kem	0.110	0.057	0.017	0.090	0.010	0.012	0.010	0.196	0.017	0.017	0.002	0.001									
Mek	0.064	0.022	0.000	0.231	0.038	0.017	0.038	0.137	0.000	0.000	0.001	0.002	0.017								
Met	0.064	0.022	0.000	0.231	0.038	0.017	0.038	0.137	0.000	0.000	0.001	0.002	0.017	0.000							
Sha	0.276	0.203	0.133	0.011	0.008	0.032	0.008	0.387	0.133	0.133	0.091	0.083	0.032	0.133	0.133						
She	0.159	0.099	0.048	0.034	0.015	0.007	0.015	0.254	0.048	0.048	0.024	0.019	0.007	0.048	0.048	0.002					
But	0.064	0.022	0.000	0.231	0.038	0.017	0.038	0.137	0.000	0.000	0.001	0.002	0.017	0.000	0.000	0.133	0.048				
Tep	0.064	0.022	0.000	0.231	0.038	0.017	0.038	0.137	0.000	0.000	0.001	0.002	0.017	0.000	0.000	0.133	0.048	0.000			
Won	0.119	0.065	0.022	0.075	0.013	0.014	0.013	0.207	0.022	0.022	0.005	0.002	0.014	0.022	0.022	0.021	0.012	0.022	0.022		
Yay	0.064	0.022	0.000	0.231	0.038	0.017	0.038	0.137	0.000	0.000	0.001	0.002	0.017	0.000	0.000	0.133	0.048	0.000	0.000	0.022	
Zeg	0.003	0.014	0.017	0.306	0.077	0.049	0.077	0.036	0.017	0.017	0.023	0.025	0.049	0.017	0.017	0.192	0.090	0.017	0.017	0.057	0.017
	Abb	Bak	Beb	Del	Did	Fin	Gur	Gin	Gon	Har	Hir	Jim	Kem	Mek	Met	Sha	She	But	Tep	Won	Yay

CURRICULUM VITAE

Personal details

Name: Abayneh Derero Dimenso

Sex: Male

Date and place of birth: 28.01.1971, Addis Ababa

Nationality: Ethiopian

Marital Status: Married

Education

2004-to date: PhD student at the Georg-August-Universität Göttingen

1998: MSc in Forestry from the Swedish University of Agricultural Sciences (SLU) (Thesis: Natural regeneration in a broad leaved Afromontane rain forest, southwest Ethiopia)

1993: BSc in Forestry from Haramaya University (the then Alemaya University of Agriculture)

1985-1989: Nefas-Silk Comprehensive Secondary School, Addis Ababa

1978-1985: Freheiwot Elementary School, Addis Ababa

Employment

1999-2003: Associate Researcher (to date) and National Coordinator for Tree Seeds and Tree Improvement Research Program, in the Ethiopian Institute of Agricultural Research (EIAR)

1994-1996: Forestry expert at Kaffa Agricultural Bureau in the Southern Peoples Regional State