

Maintenance of genetic diversity in *Cordia africana* Lam., a declining forest tree species in Ethiopia

Abayneh Derero · Oliver Gailing · Reiner Finkeldey

Received: 20 May 2010 / Accepted: 20 May 2010 / Published online: 9 June 2010
© The Author(s) 2010. This article is published with open access at Springerlink.com

Abstract We assessed genetic variation in a total of 22 populations of the tree species *Cordia africana* Lam. in Ethiopia and analyzed the country-wide impact of population history, forest disturbance and alteration of land use on extant intraspecific diversity. Amplified Fragment Length Polymorphisms (AFLPs) and chloroplast microsatellite markers were investigated. The analyses of the AFLP data revealed high diversity in all investigated populations: the percentage of polymorphic loci (PPL) ranged from 62.2% to 92.2% and Nei's gene diversity from 0.220 to 0.320 within the populations. The mean PPL and the mean diversity within populations were 85.7% and 0.287, respectively. The analysis of molecular variance revealed a moderate level of differentiation ($\Phi_{ST}=0.07$, $p<0.001$) among the populations. The Mantel test proved a significant but low correlation ($r=0.31$, $p<0.001$) between the geographic distance and the genetic differentiation matrices. Only three different cpDNA haplotypes were observed; no more than two haplotypes were found in any population. The

dominant haplotype with an overall frequency of 81% was observed in all populations. The level of differentiation among the populations was comparatively low at chloroplast DNA ($G_{ST}=0.18$, R_{ST} , $N_{ST}=0.22$). The observed patterns and levels of genetic variation within and among the populations indicate that efficient gene flow via pollen and seed is likely to be the main factor contributing to the maintenance of genetic diversity in natural and disturbed conditions.

Keywords AFLP · *Cordia africana* · CpSSR · Haplotype · Genetic variation · Conservation

Introduction

Human impacts restrict the current distribution of species and population sizes and modify the demography of many tropical forest plants. A rapid decline of the distribution area of a species and decreasing population sizes are frequently assumed to negatively affect forest genetic resources and result in irreversible losses of intraspecific diversity (Namkoong et al. 1996).

However, the genetic consequences of declining population sizes and increasing fragmentation are difficult to predict (Lowe et al. 2005) since trees are long living, predominantly outcrossing organisms with usually high fecundity and efficient means of gene dispersal by pollen and seed. This combination of life history characteristics shaped patterns of genetic diversity under natural conditions in the past resulting in high diversity within populations and moderate to low differentiation among populations (Hamrick and Godt 1996). The same factors may also be responsible for the maintenance of genetic diversity in case of declining populations unless extremely low numbers of plants are surviving (Hamrick 2004). Thus, genetic diversity of tropical forest plants may be less vulnerable to human impact than

Communicated by O. Savolainen

A. Derero (✉)
Ethiopian Institute of Agricultural
Research/Forestry Research Center,
P. O. Box 30708, Addis Ababa, Ethiopia
e-mail: abaynehdd@yahoo.com

A. Derero · O. Gailing · R. Finkeldey
Georg-August-Universität Göttingen,
Section Forest Genetics and Forest Tree,
Breeding, Büsgenweg 2,
37077 Göttingen, Germany

O. Gailing
School of Forest Resources and Environmental Science,
Michigan Technological University,
Houghton, MI 49931, USA

often assumed and efficient conservation measures may be implemented even for species with severely reduced distribution areas and numbers of surviving plants (Finkeldey and Hattmer 2007 and literature cited therein).

Numerous events during the history of populations shaped their extant genetic structures. Thus, it is mandatory to consider natural processes including the spatial distribution of naturally regenerated trees and populations and the occurrence of populations in different climatic regions or ecosystems even for rapidly declining species in order to better understand human impacts on their genetic resources. Furthermore, different human activities such as selective harvesting, conversion of natural forests to plantations or fragmentation are expected to affect genetic structures differently (Namkoong et al. 1996).

Human impacts on genetic diversity are unlikely to affect different parts of the plant genome in the same way (Hedrick 2001). For example, variation at genes with a direct impact on economically important traits may differently respond to selective removal of superior phenotypes than variation patterns at ‘neutral’ genes (Kohn et al. 2006). Thus, experimental studies aimed to investigate human impact on intraspecific diversity patterns need to consider differences concerning the mode of inheritance, average variation levels and functional importance of the investigated marker systems (Garant and Kruuk 2005).

Cordia africana Lam. is a tree species native to tropical Africa and tropical Arabia. 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). Generally, the species grows in areas with altitudes between 550 and 2,600 m a.s.l. and annual rainfall of 700 to 2,000 mm (Friis 1992). It is an early colonizer in forest re-growth and is often found along forest margins; it often regenerates in clearings and forest gaps (Derero et al. 2003; Fichtl and Admasu 1994). Flowers are complete and pollinated predominantly by bees. The fruits are edible, and seed dispersal is mainly carried out by mammals and birds.

C. africana is a fast-growing and highly valued timber tree in Ethiopia used for high-quality furniture, doors, windows, cabinet making, drums, beehives, joinery and interior construction, mortars, panelling and veneering (ICRAF 1998). The contribution of scattered trees of *C. africana* to improve soil fertility and its importance as a (coffee) shade tree in traditional agroforestry systems has been documented (Teketay and Tegineh 1991; Yadessa et al. 2001).

The current distribution, the habitats and the populations of *C. africana* are severely affected by deforestation, fragmentation and selective logging. The northern part of the country has been extremely deforested; in the north, the species is currently only represented by scattered trees on farmlands, in and around homesteads, church compounds and

graveyards and a few spots where a relatively continuous forest exists.

The general objective of the study was to determine the genetic variation in populations of *C. africana* over a wide geographical distance and to investigate genetic variation of populations representing various levels of forest fragmentation in Ethiopia. The life history traits of *C. africana* and the rapid decline of the species in Ethiopia make it a suitable model to assess the large-scale consequences of over-exploitation and forest destruction on the genetic resources of a valuable, common tropical tree species. We expected that (1) human impact including fragmentation does not reduce genetic variation at neutral markers within extant populations and that (2) genetic differentiation among populations mainly reflects the current spatial distribution of the species, but not different land use systems, ecological conditions or disturbance regimes. We used AFLPs and cpSSRs to investigate the genetic patterns in both the nuclear and chloroplast genomes.

Materials and methods

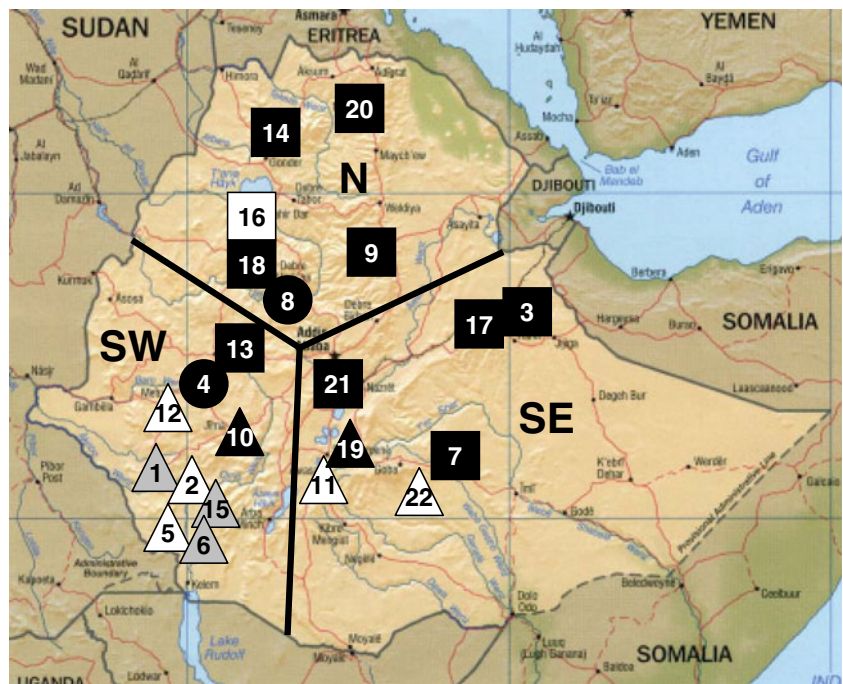
Sampling and plant material collection

A total of 22 different populations were sampled from all main growing regions of *C. africana* in Ethiopia (Fig. 1). Plant material (leaves) was 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. Up to 32 plants were sampled at each location; the number of investigated samples varied between populations and used marker systems (see below).

The sampled populations grow in an altitudinal range from 950 m to 2,150 m a.s.l., and from 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 and Fig. 1). The populations were grouped according to the following criteria:

- Grouping into three geographical units following the physiographic units of the country, and referred to as the north (N), south-west (SW) and south-east (SE). The north was represented by populations entirely from the northern highlands, which is a highly denuded area. The SW was represented by both highland and lowland populations including one population from the western part of the Central Plateau. It has the vast majority of the remnant high forests and deforestation is less advanced than in the northern part. The SE (populations from the SE highlands and two populations from the Rift Valley) is ‘intermediate’ with regard to levels of deforestation since it is more

Fig. 1 Approximate location of investigated *Cordia africana* populations. Numbers as in Table 1. Symbol colour: *black*, scattered population; *white*, natural population; *grey*, shade trees in coffee plantation. Symbol shape: *triangle*, moist evergreen montane forest ecosystem (MEMF); *square*, dry evergreen montane forest and grassland complex (DEMFG); *circle*, *Combretum–Terminalia* woodland ecosystem (CTW)



deforested than the SW, but represents a much better forest condition compared to the north.

- Grouping according to population types: 13 scattered populations (remnant trees, regenerations and planted trees from wildings on traditional agroforestry systems,

in and around homesteads, in church compounds and in graveyards) were distinguished from six continuous (undisturbed or disturbed) forests and three coffee plantations, i.e. natural forests converted to coffee plantations about three decades ago. The occurrence of

Table 1 Investigated *Cordia africana* populations, listed in decreasing order of expected heterozygosity within populations (H)

| No | Population | Altitude (m) | Seed zone ^a | Number ^b | PPL | H | SE (H) |
|----|-------------|--------------|------------------------|---------------------|------|-------|--------|
| 1 | Meti | 1,150–1,350 | 23 | 26 | 92.2 | 0.320 | 0.016 |
| 2 | Sheko | 1,150–1,400 | 23 | 22 | 88.9 | 0.320 | 0.016 |
| 3 | Harar | 1,750–2,050 | 21 | 19 | 86.7 | 0.306 | 0.017 |
| 4 | Didessa | 1,300–1,650 | 11 | 23 | 92.2 | 0.304 | 0.016 |
| 5 | Guraferda | 1,000–1,150 | 23 | 20 | 91.1 | 0.303 | 0.015 |
| 6 | Bebeka | 1,050–1,250 | 23 | 23 | 87.8 | 0.301 | 0.017 |
| 7 | Ginnr | 2,000 | 18 | 18 | 84.4 | 0.300 | 0.017 |
| 8 | Abbay | 1,600–1,900 | 11 | 21 | 86.7 | 0.299 | 0.018 |
| 9 | Kemisse | 1,500 | 15 | 21 | 87.8 | 0.297 | 0.017 |
| 10 | Jimma | 1,500–1,800 | 23 | 22 | 87.8 | 0.296 | 0.015 |
| 11 | Wondogenet | 1,800–2,000 | 24 | 21 | 92.2 | 0.296 | 0.016 |
| 12 | Yayu | 1,250–1,550 | 23 | 19 | 87.8 | 0.295 | 0.017 |
| 13 | Bako | 1,500–1,700 | 20 | 18 | 83.3 | 0.294 | 0.017 |
| 14 | Gondar | 1,850–2,000 | 19 | 22 | 91.1 | 0.294 | 0.015 |
| 15 | Tepi | 1,000–1,100 | 23 | 22 | 88.9 | 0.290 | 0.015 |
| 16 | Zeghie | 1,800–1,900 | 22 | 25 | 91.1 | 0.290 | 0.016 |
| 17 | Hirna | 1,750–2,000 | 21 | 22 | 82.2 | 0.284 | 0.017 |
| 18 | Finoteselam | 1,750–1,950 | 20 | 22 | 82.2 | 0.277 | 0.017 |
| 19 | Shashemene | 1,950 | 24 | 23 | 83.3 | 0.271 | 0.018 |
| 20 | Mekelle | 2,000–2,100 | 15 | 17 | 72.2 | 0.268 | 0.019 |
| 21 | Butajira | 1,850–1,900 | 20 | 23 | 82.2 | 0.259 | 0.017 |
| 22 | Dolo-Mana | 1,500–1,600 | 24 | 26 | 62.2 | 0.220 | 0.018 |

No population number as in Fig. 1, PPL percentage of polymorphic AFLP loci, H expected heterozygosity, SE standard error of H

^a Seed zone according to Aalbæk (1993)

^b Sample size for AFLP analyses

trees in scattered populations and in particular artificial stand establishment possibly affect the reproduction effective population size (N_e) and may result in decreased levels of genetic variation.

- Grouping according to three natural ecosystems (IBC 2005): *Combretum–Terminalia* woodland ecosystems (CTW; two populations), dry evergreen montane forest and grassland complex (DEMF; ten populations), and moist evergreen montane forest ecosystems (MEMF; ten populations).
- Grouping according to tree seed zones. Tree seed zones were delineated for Ethiopia according to Aalbæk (1993) as a basis to regulate the use of forest reproductive material for all important planted forest species in the country. Populations belong to nine different seed zones.
- Grouping according to altitudinal range: 15 highland populations (with altitude equal or above 1,500 m a.s.l.), five lowland populations (with altitude below 1,500 m a.s.l) and two intermediate populations (extending from the lowland to the highland) were distinguished. The occurrence of trees at different altitudes potentially affects migration patterns of this zoogamous and zoochorous tree species.

DNA extraction and molecular analyses

Total genomic DNA was extracted from dried leaves (taking about 20 mg or 1 cm² dried leaves) using the DNeasy 96 Plant Kit (Qiagen, Hilden, Germany) following the manufacturer's instructions. The amount and quality of the DNA was checked on 1% agarose gels.

To run AFLPs, samples from different populations were distributed randomly on the PCR plates. Genome-wide variation patterns were analysed following the protocol of Vos et al. (1995) with minor modifications. The genomic DNA was digested with two different restriction enzymes, a six-cutter (EcoRI; 5'-G↓AATTC-3') and a four-cutter (MseI; 5'-T↓TAA-3'). The first PCR (preamplification) amplified the restriction-ligation fragments with the primer combination E01/M03, and the second PCR amplified the preamplification products with the primer combination E41/M74, a fluorescent-dye-labelled selective forward primer and a reverse primer, respectively (nomenclature of primers according to <http://www.keygene.com/keygene/pdf/PRIMERCO.pdf>). All samples with a large number of ambiguous fragments in particular size ranges or low reproducibility were excluded from further analyses. Sample sizes varied from 17 to 26 per population with a mean of 21.6 for AFLPs (Table 1).

Ten fluorescence-labelled consensus chloroplast microsatellite primers (cpSSRs; ccmp1 to ccmp10; Weising and Gardner 1998), were tested on eight individuals to amplify

DNA in *C. africana*. Primers with amplification products in the expected range were tested on 60 samples (three samples per population) to detect polymorphism. Finally, only the two polymorphic loci were used to analyze eight to twelve plants from all 22 *C. africana* populations. The mean sample size was 9.2 plants per population for cpSSR analyses.

The genotyping for both AFLP and cpSSR was done on the automated capillary sequencer ABI Prism 3100® Genetic Analyser with polymer 3100 POP-6™ (Applied Biosystems, New York, USA) using the Genescan © version 3.7 (Applied Biosystems) and Genotyper © 3.7 NT software of Applied Biosystems. The repeatability of the observed patterns for both AFLPs and cpSSRs was checked for 12 samples. The scoring for AFLPs was carried out manually as “1” for fragment present, “0” for fragment absent and “?” for dubious data such as unusual weak peaks for specific fragments. 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. A total of 90 highly reproducible AFLP loci were scored.

AFLP data analyses

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 (Zhitovovsky 1999), was used for the computation of allelic frequencies from the observed frequencies of fragments at each marker locus in each population assuming Hardy–Weinberg (HW) genotypic frequencies. Based on the estimates of allele 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 and expected heterozygosity or Nei's gene diversity (Nei 1973) were computed. Differentiation among populations was investigated by an analysis of molecular variance (AMOVA) using the software Arlequin 3.1 (Excoffier et al. 2005). Groupings of populations were conducted as explained above. 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. The covariance components (σ_i^2 's) were used to compute fixation indices (Slatkin 1991). These fixation indices computed by AMOVAs are denoted as Φ_{ST} .

Spatial structure of AFLP data were investigated by calculating the regression of $F_{ST}/(1-F_{ST})$ for population

pairs on the natural logarithm of corresponding geographic distances as suggested by Rousset (1997). The correlation between genetic and geographic distances was tested for significance by a Mantel test (Mantel 1967) based on the two matrices of pairwise distances ($F_{ST}/(1-F_{ST})$ and natural logarithm of geographic distance). Calculations were separately made for all investigated populations and for populations of a particular region (Southeast, Southwest or North).

CpSSR data analysis

The variants identified at the polymorphic cpSSR loci were combined to haplotypes. Pie charts showing the frequency of haplotypes in each population were plotted on a map to show their geographic distribution pattern. Genetic diversity within populations (H_S), total diversity (H_T), and population differentiation values (G_{ST} , R_{ST} , N_{ST}) were computed using the Permut-cpSSR programme (Pons and Petit 1996).

Results

AFLP variation within populations

The mean percentage of PPL ranged from 62.2% to 92.2% with a mean of 85.7% (Table 1). The mean within-population diversity (H) based on the 22 populations was 0.287. The population gene diversity (H_j) ranged from 0.22 to 0.32. The three highest diversities were recorded in two continuous populations and on a farmland, and the lowest value in a disturbed forest.

Levels of mean genetic diversity parameters within groups were rather similar if populations were grouped according to different criteria (Table 2). The populations from the southwest, the coffee plantations and the *Combretum-Terminalia* woodland ecosystems exhibited both the highest mean PPLs and diversities whereas the former was the highest in the intermediate altitudes and seed zones 19 and 22 and the latter in the lowlands and seed zone 23.

Population differentiation at AFLPs

The AMOVA revealed moderate, but significant variation among the populations ($\Phi_{ST}=0.072$, $p<0.001$). The analyses of the populations according to altitude and geographical units revealed low but significant differentiation among these groups (Table 3). However, there was no significant differentiation among population types, tree seed zones and natural ecosystems.

A highly significant spatial structure was identified by the Mantel test based on the matrices of pairwise genetic distances ($F_{ST}/(1-F_{ST})$) and the natural logarithm of pairwise geographic distances (Fig. 2). However, no significant

correlation was found if populations from only a single region were included indicating spatial structure only on a large, country-wide but not on a regional scale.

Variation at cpSSRs

Out of the ten consensus chloroplast microsatellite primers, seven amplified DNA in *C. africana* (ccmp2, ccmp3, ccmp4, ccmp6, ccmp7, ccmp8 and ccmp10). However, only two of the primers (ccmp2 and ccmp10) were polymorphic. A total of three haplotypes were characterized from the two polymorphic cpSSR loci (ccmp2 and ccmp10) (Table 4).

The majority (81%) of the individuals belong to haplotype B, which was detected in all the populations analysed. Eight of the populations were fixed for haplotype B. Haplotype C was limited to a few (18%) populations in the NW mountain massif (three populations belonging to the northern unit, one population belonging to the SW unit). Haplotype A had a wider distribution (41% of the populations) (Fig. 3). 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.

Discussion

Diversity within populations

The assessment of genetic diversity within populations of *C. africana* was based on the observation of AFLP variation. The use of AFLPs allows to monitor numerous loci simultaneously and is a straightforward approach for species if no sequence-based genomic information is available (Kremer et al. 2005), even though diversity estimates may be biased if based on dominant markers (Szmidt et al. 1996). We used the Bayesian method proposed by Zhivotovsky (1999) to estimate allele frequencies at the putatively dominant marker loci and assumed HW proportions within populations (inbreeding coefficient $F=0$). No data from codominant markers are available confirming that populations of *C. africana* are in or close to Hardy–Weinberg equilibrium (HWE) at most loci. However, *C. africana* is a long-living tropical tree with efficient means of gene dispersal both via pollen and seed. These life history characteristics suggest that the species is predominantly or exclusively outcrossing and that populations do not strongly deviate from HWE (Hamrick and Godt 1996).

C. africana exhibits a relatively high mean diversity within populations ($H=0.287$), which is characteristic for many widely distributed, abundant tree species (Hamrick et al. 1992). The mean diversity estimates of ten neotropical, outcrossing tree species using AFLPs under the assumption

Table 2 Average genetic diversities of *C. africana* populations grouped according to different criteria

| No | Population category | | No of populations | PPL | | | H | | |
|----|---------------------|--------------|-------------------|------|------|------|-------|-------|-------|
| | | | | Min | Max | Mean | Min | Max | Mean |
| 1 | Geographical units | North | 6 | 72.2 | 91.1 | 85.2 | 0.268 | 0.299 | 0.288 |
| | | Southeast | 7 | 62.2 | 92.2 | 81.9 | 0.220 | 0.306 | 0.277 |
| | | Southwest | 9 | 83.3 | 92.2 | 88.9 | 0.290 | 0.320 | 0.303 |
| 2 | Types | Scattered | 13 | 72.2 | 92.2 | 84.8 | 0.259 | 0.306 | 0.288 |
| | | Continuous | 6 | 62.2 | 92.2 | 85.6 | 0.220 | 0.320 | 0.287 |
| | | Coffee | 3 | 87.8 | 92.2 | 89.6 | 0.290 | 0.320 | 0.304 |
| 3 | Natural ecosystems | CTW | 2 | 86.7 | 92.2 | 89.5 | 0.299 | 0.304 | 0.302 |
| | | DEMF | 10 | 72.2 | 91.1 | 84.3 | 0.259 | 0.300 | 0.287 |
| | | MEMF | 10 | 62.2 | 92.2 | 86.2 | 0.220 | 0.320 | 0.291 |
| 4 | Seed zones | SZ-11 | 2 | 86.7 | 92.2 | 89.5 | 0.299 | 0.304 | 0.302 |
| | | SZ-15 | 2 | 72.2 | 87.8 | 80.0 | 0.268 | 0.297 | 0.283 |
| | | SZ-18 | 1 | 84.4 | 84.4 | 84.4 | 0.300 | 0.300 | 0.300 |
| | | SZ-19 | 1 | 91.1 | 91.1 | 91.1 | 0.294 | 0.294 | 0.294 |
| | | SZ-20 | 3 | 82.2 | 83.3 | 82.6 | 0.259 | 0.294 | 0.277 |
| | | SZ-21 | 2 | 82.2 | 86.7 | 84.5 | 0.284 | 0.306 | 0.295 |
| | | SZ-22 | 1 | 91.1 | 91.1 | 91.1 | 0.290 | 0.290 | 0.290 |
| | | SZ-23 | 7 | 87.8 | 92.2 | 89.3 | 0.290 | 0.320 | 0.304 |
| 5 | Altitude | SZ-24 | 3 | 62.2 | 92.2 | 79.2 | 0.220 | 0.296 | 0.262 |
| | | Highland | 15 | 62.2 | 92.2 | 83.7 | 0.220 | 0.306 | 0.283 |
| | | Intermediate | 2 | 87.8 | 92.2 | 90.0 | 0.295 | 0.304 | 0.300 |
| | | Lowland | 5 | 87.8 | 92.2 | 89.8 | 0.290 | 0.320 | 0.307 |

Abbreviations as in Table 1

of HWE ($F=0$) was 0.23 (Kremer et al. 2005); values for individual species ranged from 0.14 to 0.32. The analysis of the same set of species under the assumption of full selfing ($F=1$) resulted in a similar mean diversity of about 0.24; values for individual species ranged from 0.17 to 0.38 (Kremer et al. 2005). The overall mean value of the within population gene diversity summarized from 38 studies of outcrossing species was 0.27 at dominant RAPD markers (Nybohm 2004).

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. No strong and consistent differences with regard to the levels of genetic variation (PPL; H_j) were found between scattered and continuous populations. The average diversity within intensively managed populations in coffee plantations is even slightly higher in comparison to continuous forests. Likewise, no consistent tendency for decreased levels of genetic diversity in the heavily denuded North of Ethiopia was found in comparison to other parts of the country.

Table 3 Genetic differentiation among 22 populations of *C. africana* based on AMOVA

| Source of variation | <i>df</i> | SSD | MSD | Variance component | Percentage of variation | <i>p</i> Value |
|---|-----------|----------|-------|--------------------|-------------------------|----------------|
| Among populations | 21 | 610.74 | 29.08 | 0.85 | 7.3 | <0.001 |
| Within populations | 453 | 4,895.83 | 10.81 | 10.81 | 92.7 | <0.001 |
| Among altitudinal groups | 2 | 104.06 | 52.09 | 0.22 | 1.9 | <0.01 |
| Among populations within altitudinal groups | 19 | 506.67 | 26.58 | 0.74 | 6.3 | <0.001 |
| Within populations | 453 | 4,895.83 | 10.81 | 10.81 | 91.9 | <0.001 |
| Among geographic units | 2 | 113.24 | 56.62 | 0.19 | 1.7 | <0.001 |
| Among populations within geographic units | 19 | 497.49 | 26.18 | 0.71 | 6.1 | <0.001 |
| Within populations | 453 | 4,895.83 | 10.81 | 10.81 | 92.3 | <0.001 |

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

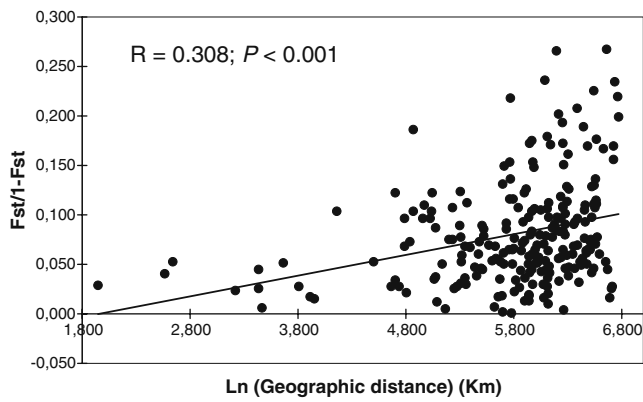


Fig. 2 Regression of pairwise genetic differentiation ($F_{ST}/1-F_{ST}$) at AFLPs on geographic distance (ln distance in km) for 22 populations of *Cordia africana*

Thus, our results support the hypothesis that human impact did not yet severely diminish genetic diversity in intensively altered populations.

This main conclusion is based on the observation of a dominant, anonymous marker system and a rough assessment of genetic diversity. The loss of rare alleles at more variable gene markers and genes under selection cannot be ruled out. The percentage of PPL is a more sensitive variation parameter with regard to the loss of rare alleles than the ‘expected heterozygosity’ (H_j). We observed an average PPL in scattered populations which is only 1% lower than the corresponding value in continuous forests. PPL in coffee plantations is even on average 4% higher than in continuous forests. Thus, there is no indication for losses of rare ‘neutral’ variation at least at the dominant loci observed. Our results do not allow conclusions on genetic variation in adaptive traits as neutral and adaptive variation levels can be poorly correlated (Hedrick 2001). Currently, no sequence data are available for *C. africana* or closely related species, and this is the main hindrance for an assessment of adaptive variation patterns.

Differentiation among populations

C. africana is a previously common, but now rapidly declining tree species in Ethiopia. The current distribution suggests restricted or possibly even completely absent extant

Table 4 Haplotype identification in *Cordia africana* based on chloroplast microsatellites and total frequency of observed individuals with particular haplotypes

| Haplotype | ccmp2 (bp) | ccmp10 (bp) | % proportion |
|-----------|------------|-------------|--------------|
| A | 187 | 105 | 13 |
| B | 187 | 106 | 81 |
| C | 188 | 106 | 6 |

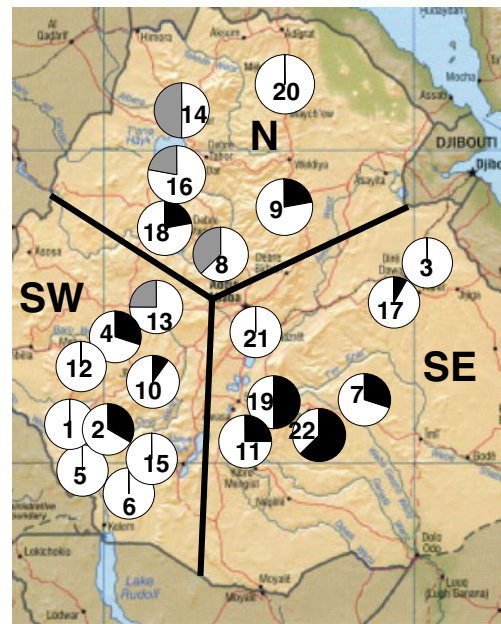


Fig. 3 Spatial distribution of chloroplast haplotypes in *Cordia africana*. Black, haplotype A; white, haplotype B; grey, haplotype C. Haplotype definitions according to Table 4

gene flow among widely isolated patches. Genetic differentiation among populations at neutral gene loci is mainly influenced by gene flow where strong differentiation arises through isolation. However, the current distribution is a poor predictor of past gene flow for long-lived but rapidly declining tropical tree species. Spatially isolated remnants might have been parts of much larger, interconnected populations even a single generation ago.

We used Φ_{ST} to quantify differentiation, which measures like G_{ST} and related estimators, population differentiation relative to the total diversity of a hypothetical population comprised of all investigated individuals. The use of these measures to assess differentiation among populations is arguable in particular if hypervariable markers are investigated (Jost 2008). However, calculation of Φ_{ST} allows interesting comparisons to earlier studies, and AFLPs are only moderately variable markers due to their dominant behaviour with no more than two alleles.

We found a high proportion of the total diversity within populations. This is a commonly made observation in forest trees and reflects their efficient means of gene dispersal in combination with long lifespan (Austerlitz et al. 2000). Approximately 11% of the total diversity at isozyme loci was found to reside among populations of tropical forest trees (Hamrick et al. 1992; Loveless 1992).

Differentiation among populations of *C. africana* is higher at cpDNA haplotypes ($G_{ST}=0.182$) in comparisons to AFLPs ($\Phi_{ST}=0.073$). Since G_{ST} measures differentiation relative to the total diversity, the higher value of G_{ST} at cpDNA is partially due to the low total diversity observed

at chloroplast haplotypes. Differentiation at cpDNA markers is much higher in comparison to anonymous AFLPs for most forest trees, since cpDNA is typically maternally inherited in angiosperms and differentiation reflects only seed, but not pollen dispersal. The mean differentiation among populations at maternally inherited markers (G_{ST}) was 0.637 in a survey covering 124 species (Petit et al. 2005).

In addition to presumably efficient pollen dispersal by insects, endozoochorous seed dispersal further promotes gene exchange among populations of *C. africana* in Ethiopia. The fruits of *C. africana* are indehiscent and are consumed by a wide range of mammals and birds, some of which may migrate and travel long distances before defecating the seeds. The ability of the species to grow in mountain areas and the Rift Valley, in the lowlands as well as in river gorges potentially contributes to efficient seed dispersal among the populations of *C. africana*. We interpret the comparatively low differentiation at cpDNA haplotypes as an indication of extensive seed dispersal and limited impact of population history on extant genetic structures.

Differentiation among populations of tropical trees at isozyme loci is much lower for animal dispersed species ($G_{ST}=0.050$) in comparison to species with abiotic means of seed dispersal ($G_{ST}=0.138$) (Loveless 1992). Results presented here support the claim that endozoochorous seed dispersal increases the genetic connectivity among spatially isolated patches, reducing the degree of differentiation among populations or subpopulations.

Even though differentiation among populations is quite low, we observed a significant correlation between spatial and genetic distances taking all investigated populations into account. Thus, gene flow does not completely prevent the development and maintenance of any spatial genetic structure. This conclusion is also supported by the restricted spatial occurrence of the rarest cpDNA haplotype (haplotype C).

Apart from the weak, but significant differentiation among the three geographic regions we found a significant differentiation among population groups only if altitudinal groups were distinguished. Thus, the hypothesis that genetic variation primarily reflects the distribution of the species, but not human impact or ecological conditions, is also supported by our data.

Based on the overall analysis of AFLP polymorphisms, we expect that most of the loci investigated are selectively neutral, and stronger differentiation at loci under selection cannot be ruled out.

Conclusions and outlook

The extant genetic structures of *C. africana* in Ethiopia possibly reflect the much wider distribution of the species in the country in the past. Studies on ongoing pollen flow and seed migration are needed to assess the vulnerability of the remaining populations.

Our results indicate that patterns of natural diversity and differentiation are not yet severely affected although *C. africana* is rapidly declining. Measures to conserve the genetic resources of the species in Ethiopia are urgent and, as shown by the genetic diversity patterns, still feasible. Since scattered populations and those under intense management in coffee plantations are not less variable than continuous populations, we conclude that conservation of genetic resources of *C. africana* in Ethiopia must not be confined to a few natural populations in continuous forests. The maintenance of highly variable populations is also feasible in agroforestry systems including intensively used coffee plantations. This view is supported by the weak, but significant spatial structure calling for a broad coverage of conservation measures in all parts of the country including regions without large continuous forests.

Likewise, forest reproductive material can be harvested in managed populations without the risk of losing genetic variation, if harvesting operations follow appropriate guidelines concerning the number and distribution of seed trees (Finkeldey and Hattemer 2007). This is particularly important in those areas largely denuded from natural forests in the northern and eastern part of Ethiopia.

Acknowledgements We are indebted to Outi Savolainen and two anonymous reviewers for their constructive criticism and suggestions to improve the manuscript. The authors are thankful to Olga Artes, Marius Ekué and Hans H. Hattemer for support in the lab, and during data analyses and drafting the first version of the manuscript. The first author is grateful to the German Academic Exchange Service (DAAD) for a scholarship and financial support for the field work.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

- Aalbæk A (1993) Tree seed zones for Ethiopia. Forestry Research Center/National Tree Seed Project, Addis Ababa, p 120
- Austerlitz F, Mariette S, Machon N, Gouyon PH, Godelle B (2000) Effects of colonization processes on genetic diversity: differences between annual plants and tree species. *Genetics* 154:1309–1321
- Derero A, Bekele T, Näslund BÅ (2003) Population structure and regeneration of woody species in a broadleaved Afromontane rain forest, South-west Ethiopia. *Eth J Nat Res* 5(2):255–280
- Excoffier L, Laval G, Schneider S (2005) Arlequin ver. 3.1: an integrated software package for population genetics data analysis. *Evol Bioinf Online* 1:47–50
- Fichtl R, Admasu A (1994) Honeybee Flora of Ethiopia. Magraf Verlag, Weikersheim, p 510
- Finkeldey R, Hattemer HH (2007) Tropical forest genetics. Springer, Berlin, p 315
- Friis I (1992) Forests and Forest Trees of Northeast Tropical Africa. Kew Bulletin Additional Series XV. HMSO, London. pp 396

- Garant D, Kruuk LEB (2005) How to use molecular marker data to measure evolutionary parameters in wild populations. *Mol Ecol* 14:1843–1859
- Hamrick JL (2004) Response of forest trees to global environmental changes. *Forest Ecol Manage* 197:323–335
- Hamrick JL, Godt MJW, Sherman-Broyles SL (1992) Factors influencing levels of genetic diversity in woody plant species. *New For* 6:95–124
- Hamrick JL, Godt MJW (1996) Effects of life history traits on genetic diversity in plant species. *Phil Trans Royal Soc London, Ser B* 351:1291–1298
- Hedrick PW (2001) Conservation genetics: where are we now? *TREE* 16:629–636
- IBC (2005) National Biodiversity Strategy and Action Plan. Institute of Biodiversity Conservation, Addis Ababa, p 115
- ICRAF (1998) Agroforestry Data Base: *Cordia africana* Lam. ICRAF
- Jost L (2008) G(ST) and its relatives do not measure differentiation. *Mol Ecol* 17:4015–4026
- Kohn MH, Murphy WJ, Ostrander EA, Wayne RK (2006) Genomics and conservation genetics. *TREE* 21:629–637
- Kremer A, Caron H, Cavers S, Colpaert N, Gheysen G, Gribel R, Lemes M, Lowe AJ, Margis R, Navarro C, Salgueiro F (2005) Monitoring genetic diversity in tropical trees with multilocus dominant markers. *Heredity* 95:274–280
- Loveless MD (1992) Isozyme variation in tropical trees: patterns of genetic organization. *New For* 6:67–94
- Lowe AJ, Boshier D, Ward M, Bacles CFE, 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, Milligan BG (1994) Analysis of population structure with RAPD markers. *Mol Ecol* 3:91–99
- Mantel NA (1967) The detection of disease clustering and a generalized regression approach. *Cancer Res* 27:209–220
- Namkoong G, Boyle T, Gregorius H-R, Joly H, Savolainen O, Wickneswari R, Young A (1996) Testing criteria and indicators for assessing the sustainability of forest management: genetic criteria and indicators. Center for International Forestry Research, Bogor
- Nei M (1973) Analysis of gene diversity in subdivided populations. *Proc Natl Acad Sci USA* 70:3321–3323
- Nybom H (2004) Comparison of different nuclear DNA markers for estimating intraspecific genetic diversity in plants. *Mol Ecol* 13:1143–1155
- Petit RJ, Duminil J, Fineschi S, Hampe A, Salvini D, Vendramin GG (2005) Comparative organization of chloroplast, mitochondrial and nuclear diversity in plant populations. *Mol Ecol* 14:689–701
- Pons O, Petit RJ (1996) Measuring and testing genetic differentiation with ordered and unordered alleles. *Genetics* 144:1237–1245
- Rousset F (1997) Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. *Genetics* 145:1219–1228
- Slatkin M (1991) Inbreeding coefficients and coalescence times. *Genet Res* 58:167–175
- Szmidt AE, Wang XR, Lu MZ (1996) Empirical assessment of allozyme and RAPD variation in *Pinus sylvestris* (L) using haploid tissue analysis. *Heredity* 76:412–420
- Teketay D, Tegineh A (1991) Shade trees of coffee in Hararghe, Eastern Ethiopia. *Int Tree Crop J* 7:17–27
- Vekemans X (2002) AFLP-SURV version 1.0. Distributed by the author. Laboratoire de Genetique et Ecologie Vegetale, Universite Libre de Bruxelles, Belgium
- Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, Zabeau M (1995) AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Res* 23:4407–4414
- Weising K, Gardner RC (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
- Yadessa A, Itanna F, Olsson M (2001) Contribution of indigenous trees to soil properties: the case of scattered trees of *Cordia africana* Lam. in croplands of western Oromia. *Eth J Nat Res* 3(2):245–270
- Zhivotovsky LA (1999) Estimating population structure in diploids with multilocus dominant DNA markers. *Mol Ecol* 8:907–913