CONSERVATION GENETICS OF THE FRANKINCENSE TREE

Addisalem Ayele Bekele

Thesis committee

Promotor

Prof. Dr. F.J.J.M. Bongers

Personal chair at the Forest Ecology and Forest Management Group

Wageningen University

Co-promotors

Dr. M.J.M. Smulders

Business Unit manager, Wageningen UR Plant Breeding

Wageningen University and Research centre

Dr. K. Tesfaye Geletu

Assistant professor, Institute of Biotechnology

Addis Ababa University

Other members

Prof. Dr. N.P.R. Anten, Wageningen University

Prof. Dr. R.G.A. Boot, Tropenbos International and Utrecht University

Prof. Dr. O. Hardy, Université Libre de Bruxelles, Belgium

Prof. Dr. B. Zwaan, Wageningen University

This research was conducted under the auspices of the Graduate School of Production Ecology and Resource Conservation (PE&RC).

CONSERVATION GENETICS OF THE FRANKINCENSE TREE

Addisalem Ayele Bekele

Thesis

Submitted in fulfilment of the requirements for the degree of Doctor

at Wageningen University

By the authority of the Rector Magnificus

Prof. Dr. A.P.J. Mol,

In the presence of the

Thesis Committee appointed by the Academic Board

to be defended in public

on Monday 23 May 2016

at 4 p.m. in the Aula.

Addisalem Ayele Bekele

Conservation genetics of the frankincense tree

158 pages

PhD thesis, Wageningen University, Wageningen, NL (2016)

With references, with summaries in Dutch and English

ISBN 978-94-6257-686-5

Contents **Pages**

Chapter 1	Introduction	7
Chapter 2	Genomic sequencing and microsatellite marker development for *Boswellia papyrifera,* an economically important but threatened tree native to dry tropical forests	23
Chapter 3	Genetic diversity and differentiation of frankincense tree (*Boswellia papyrifera* (Del.) Hochst) and strategies for its conservation	47
Chapter 4	Fine-scale spatial genetic structure in the frankincense tree *Boswellia papyrifera* and implications for conservation	71
Chapter 5	General Discussions	94
	References	117
	Summary	133
	Samenvatting	139
	Acknowledgments	145
	Short Biography	151
	List of Publications	153
	PE&RC Training and Education Statement	154
	The FRAME project	156

CHAPTER 1

INTRODUCTION

Forests and woodland resources

Forests provide many products and ecological services to society and thereby contribute to the local as well as national economy of the country where the forest are located. The future value of forest resources is determined by the ways humans manage these resources in the ever changing environment in today's world. In recent times the survival of a large number of species is threatened due to deforestation, habitat degradation and fragmentation (FAO, 2014). The consequences of these threats would mean a loss of the goods and services that can be obtained from the resource and these would be translated into reduced economic opportunities for present and future generations.

In Ethiopia, while information on forest and woodland vegetation cover is limited, most reports indicate that Ethiopia's natural forests have been under a severe pressure from human impacts since at least a century ago (EFAP, 1994; Reusing, 2000; FAO, 2003). Forest resources have been depleted driven by increased human demands mainly for agricultural lands, settlement areas, fuelwood, construction wood and generating income from timber sales. Reusing (2000) reported severe deforestation and degradation on 27.4 % of Ethiopian forests. The pressure has been relatively less in the southern and southwestern regions and these are the areas in which the extensive and relatively undisturbed forests exist (Reusing, 2000; FAO, 2003; WBISPP, 2004). Deforestation is highly pronounced in the central, Northern and Northeastern parts of the country. These regions are left with fragmented forest patches and sparsely distributed old trees on the landscapes as signature of forests in the past. Tree density has been decreasing in almost all forested areas due to selective logging. Natural regeneration is scarce in a considerable portion of forests (Reusing, 2000; Abiyu et al., 2010, Alemu et al. (2015).

More than 65% of Ethiopia's land mass is located in dry land areas and they are associated with tropical dry forest and woodland vegetation (NCSS, 1993). Woodlands of Combretum-Terminalia and Acacia-Commiphora are the two

dominant dry land vegetation types (Eshete et al., 2011; Lemenih and Kassa, 2011) in the country. These vegetation possess many economically important tree species including *B. papyrifera* which is the source of economically important commodity known as frankincense (Vollesen, 1989; Gebrehiwot et al., 2003; Lemenih et al., 2007)**.** Woodlands and lowland dry forests have been threatened in similar manner since 1960s (Lemenih et al., 2011; Alemu et al., 2015).

Recognizing the rapid deforestation and the urgent need of forest resource conservation, Ethiopian government established state-owned Forest Priority Areas (FPAs), National Parks, Game Reserves, Sanctuaries and Controlled Hunting Areas in 1970's to conserve forest resources (Reusing, 2000; Amente, 2006). However, this conservation approach which precluded local people from accessing the resource was not effective enough to protect the forests. The annual rate of deforestation increased from 1 % during the period 1990-2000 to 1.1 % between 2000-2005 (FAO, 2005). The failure of conventional approaches (State owned) led to an attempt to initiate a participatory forest management (PFM) that involved the community living in and around the forests in the 1990s. The government in collaboration with NGO's (FARM Africa, German Technical Cooperation (GTZ) and SOS Sahel) initiated and implemented PFM in various forests in the country (e.g. Chilimo, Bonga and Bale mountain national forest priority areas). These pilot phase PFM projects have become effective in delivering benefits to the local community and improving conservation of the forests (Kubsa et al., 2003; Tesfaye et al., 2010; Amente, 2006, FARM Africa, 2015).

Conservation genetics and molecular markers

The diverse species in the forest and the genetic variability contained within them are the sources of the enormous range of goods and services we obtain from trees and forests such as timber, construction wood, gum and resins, soil conservation and carbon sequestration. Knowledge of the level of genetic diversity of the species and how this is spatially distributed among populations

is crucial to identify populations of conservation priority and sampling representative genetic diversity for conservation. Moreover, understanding the ways in which human disturbances influence the genetic diversity pattern are important components of formulating effective management that ensure the preservation and sustainable use of the resources for human needs (Moran, 2002; Edwards et al., 2011). Conservation genetics deals with the understanding of genetic diversity pattern of the species, genetic factors that affect extinction risk and genetic management regimes required to minimise these risks. It basically focuses on scientifically sound conservation and management of genetic resourcesultimately aiming at ensuring the continuity of the benefits they provide (Frankham et al., 2002). This involves the use of genetic markers (Ouborg et al., 2006; Moran, 2002; Edwards et al., 2011). Many modern genetic markers, such as microsatellites or simple sequence repeats (SSRs) and amplified fragment-length polymorphisms (AFLPs), among other applications, are routinely used for the assessment of populations genetic variation pattern, estimating seed and pollen dispersal distance, parentage analysis, and kinship studies (Booy et al., 2000; González-Martínez et al., 2006). Compared with other types of molecular markers, SSRs have many advantages including high abundance, high polymorphism, co-dominant inheritance, and reproducibility (Chase et al., 1996).

Genetic diversity pattern in tree species

Genetic diversity pattern of plants, among other factors, is governed by genetic and demographic processes such as population size, genetic drift, breeding system and gene dispersal of the species (Booy et al., 2000), and their interaction with environment (Lawton-Rauh, 2008). The genetic structure of trees differs from other organisms in several key respects. Much of the genetic diversity (> 90 % of the total diversity) of most tree species is found within populations rather than among populations (Hamrick, 2004). Trees species are widespread, long-lived, mostly outcrossing, have large populations, and able to disperse genes over long distances. These characteristics, coupled with native

environments that are often variable, have enabled forest tree species to maintain high levels of genetic diversity and evolve into some of the most genetically diverse organisms in existence (Hamrick, 1992; Hamrick and Godt, 1996). Widespread species may have historically consisted of large and continuous populations that were less susceptible to the random loss of genetic variation, thus such species are expected to have higher genetic diversity (Hamrick et al., 2004). Long life span maintains high level of genetic diversity because it ensures representation of many cohorts (overlapping generations) within a population.

Long distance gene dispersal leads to higher intraspecific genetic diversity. Because this increases, the probability of new mutations to be spread and incorporated into additional populations increases as well. Among populations, extensive pollen and seed dispersal may homogenize the genetic variation between populations and thus lead to a decrease of their genetic differentiation. In contrast, restricted gene flow to shorter distances, increases differentiation between populations as a function of the spatial distance between them (Savolainen et al., 2007).

Genetic consequences of deforestation and population fragmentation

Anthropogenic impacts such as habitat loss due to deforestation, fragmentation, and population degradation (regeneration and density) may cause the interruption of the genetic and ecological processes and subsequently lead to a loss of genetic diversity and change in spatial distribution pattern of the genetic variation. Human induced forest fragmentation is considered to be one of the primary causes of tropical biodiversity loss (Bogaert et al., 2011; FAO, 2014). Continuous forests in most continents have been replaced by highly fragmented landscapes during the last 1000 years (Hamrick, 2004). Fragmentation produces small and isolated populations, which are expected to be more vulnerable to extinction (local) due to demographic, genetic and environmental stochastic events. Theoretically, reductions of gene flow, increased random genetic drift,

and inbreeding may lead to decreased genetic diversity and increased among population genetic differentiation (Frankham et al., 2002). The most immediate effect of fragmentation is rapid loss of genetic variation, specifically, a lower proportion of polymorphic loci and a reduction in the number of alleles per loci because a significant portion of individuals can be lost from the original population (Young et al., 1996; Aguilar et al., 2008). This effect depends on the effective population size within fragments and genetic structure pattern of the original populations previous to fragmentation (Aguilar et al., 2008). If there was a high level of genetic exchange within the original population, the fragments may still contain similar genetic diversity pattern, high level of genetic diversity and low level of differentiation. In the long-term, the spatial isolation of populations limits the pollen and seed dispersal among remnant populations causing a loss of heterozygosity (Sebbenn et al., 2011; Kamm et al., 2010) particularly if the isolation persists over several successive generations without being reversed. This would mean that relatively recent processes of fragmentation may not cause a loss of genetic diversity through genetic drift, since the effect of genetic drift is influenced directly by the number of generations in which populations size remain small (Young et al., 1996).

The genetic bottleneck combined with the limited gene flow may cause elevated genetic drift and inbreeding which subsequently result in a low level of genetic diversity (White et al., 2001; Kamm et al., 2010).In agreement with these predictions, some empirical studies of gene flow verified reduced pollen and seed immigration into isolated fragments increased differentiation among populations. This effect is observed, for example, in an insect pollinated tree species *Copaifera langsdorffii* (Sebbenn et al., 2011). However, other studies demonstrated contrasting findings for instance, in tropical tree species*Swietenia humilis* (White et al., 2001), *Cordia africana* (Derero et al., 2010), and *Sorbus domestica* (Kamm et al., 2010). This implies that the theoretical assumption about the impacts of population fragmentation on the genetic diversity and differentiation do not necessarily apply to all species and that not all

fragmentation episodes necessarily result in genetic erosion of plant populations.From a conservation perspective, understanding how fragmentation affects gene flow among remnant populations in such fragmented landscapes is important to manage genetically isolated populations by formulating management strategies that enhance the connectedness of populations through gene flow, and therefore, the level of genetic variation (Richards, 2000). For instance, in a meta-population that originates from fragmentation of a population with formerly widespread occurrence, restoration of gene flow as management intervention can reduce the detrimental effect of inbreeding in small subpopulations (Booy et al., 2000).

Fine-scale genetic structuring (FSGS) and gene dispersal within population

Tree species have a large proportion of their genetic diversity within their populations (Hamrick et al., 1992) and this large amount of genetic variation is locally sub-structured in space within the populations as a result of different processes. Life forms (e.g woody or herbaceous species), density of the population, the breeding systems of the species and gene dispersal are important determinants of intra-population (fine-scale) spatial genetic structure (FSGS), a decrease of pairwise kinship with increasing pairwise distances (Vekemans and Hardy, 2004). Therefore, investigating the pattern of FSGS and how it varies between generations provides insight into the intensity and spatial scale of gene dispersal and other fundamental aspects of species' biology, such as mating system (Segelbacher et al., 2010).

Dry forest and woodland resources and frankincense

B. papyrifera is one of the dominant species in Combretum-Terminalia dry forest/woodland vegetation (Lemenih and Kassa, 2011). The species produces a commercial resin known as frankincense which is internationally traded product for its uses as ingredients in cosmetic, detergent, food flavor and perfumes productions (Lemenih and Teketay, 2003). Local people obtain financial benefit from sales of frankincense and employment opportunities in field processing of the products (Eshete et al., 2005; Woldeamanuel, 2011). In

addition to the frankincense production, *B. papyrifera* is a source of good quality honey (Gebrehiwot et al., 2003). Different parts of *B. papyrifera:* the leaves, roots, bark and the frankincense itself, are used traditionally as medicines against various diseases among which are leprosy, bronchitis, bronchial asthma, typhoid and viral hepatitis (Lemenih and Teketay, 2003). Researches on the medicinal use of frankincense from *Boswellia* species revealed its potential for modern pharmaceutical uses. It was discovered that frankincense contains chemical constituents that have the ability to target cancer cells of different tissue origins. For example, frankincense oil derived from *Boswellia carteri* kills bladder tumor cell while the frankincense from *Boswellia sacra* does the same with breast cancer cells and pancreatic cancer cell lines (Lin, 2013; Ni et al., 2012). A recent study identified cancer-killing properties of frankincense inlate-stage ovarian cancer (Evans, 2013). It was also discovered that frankincense boosts the body immune system (Khajuria et al., 2008). These examples indicate the enormous pharmaceutical use of frankincense in the future including its potential in preventing and treating cancer. Compared to other *Boswellia* species *(B. serrata and B. carterii)*, frankincense from *B. papyrifera* contains high quantities of incensole acetate (Paul et al., 2012), a resin constituent that affects the emotional state of the mammalian central nervous system (Moussaieff et al., 2008). *B. papyrifera* is therefore one of the most important wild tree genetic resources of Ethiopia with current and potential economic importance and societal values.

Ecology and life history characteristics of *B. papyrifea*

This study focuses on *B. papyrifera*, a tropical woodland species which is under increasing threat from various anthropogenic factors. *B. papyrifera* is a widespread species, its distribution stretching from Ethiopia in the Horn of Africa to Cameroon and Nigeria to the west (Vollesen, 1989; Zhang et al., 2013) growing in Sudanian and Sahelian climatic regions (Figure 1-1).

Source: http://www.ars-grin.gov/cgi-bin/npgs/html/taxon.pl?310554

Fig. 1-1: Geographical distribution of *B. papyrifera* across the world

B. papyrifera is a slow-growing tree species which attains 20 cm DBH at the age of 73 years (Tolera et al., 2013), hence its life span is >100 years. It is propagated from seeds and root sprouts (Lemenih and Kassa, 2011). Propagation through branch cutting has also been proved successful (Haile et al., 2011). The breeding system, pollination and seed dispersal mode of the species are not well known. *B. papyrifera* is a monosious species (separate male and female flowers on the same tree). Insect pollination (entomophily) is believed to be one of the species pollination mechanism as it was observed that honeybees frequently visit its flowers (Fitchl and Admasu, 1994; Gebrehiwot et al. 2003). Seeds were small (~2mm), light, papery. These are seed characteristics adapted for wind dispersal (anemophily). The same seed characteristics were also observed in *B. ovalifoliolata.* The different life stages of the species are presented in Fig 1-2.

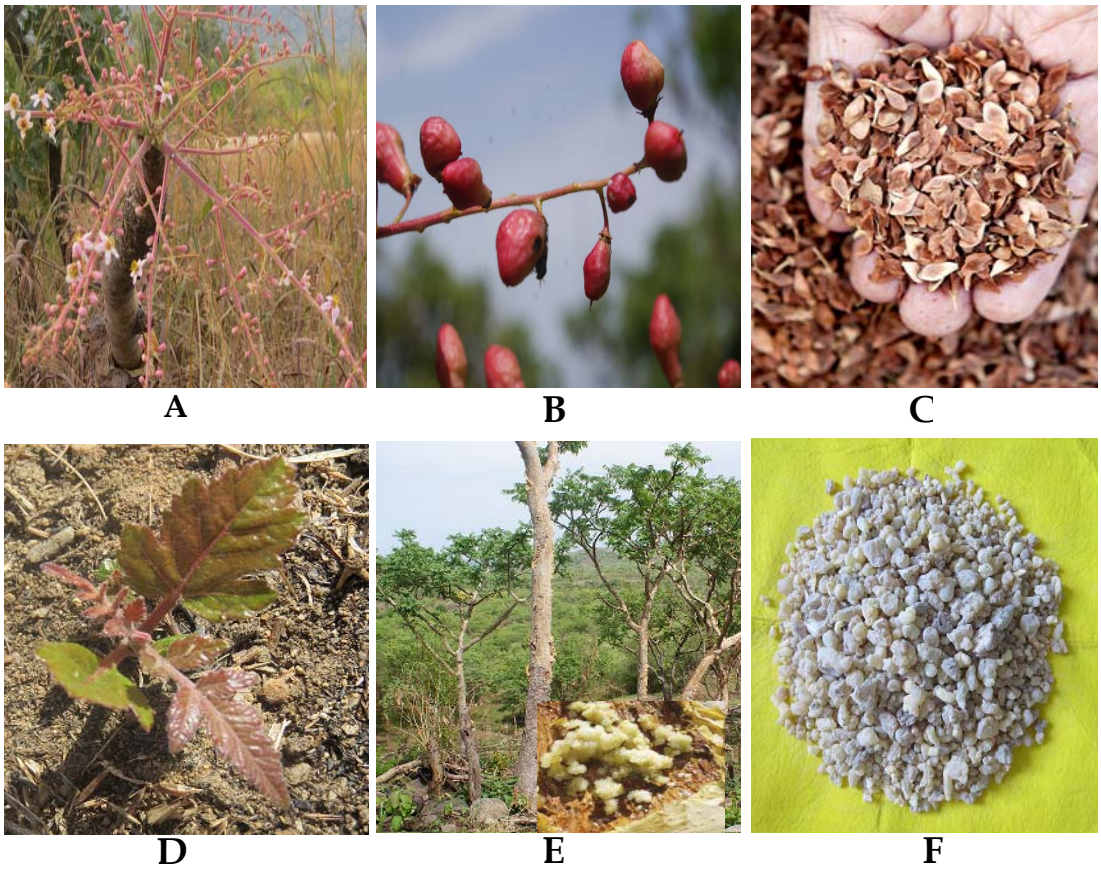


Fig. 1-2: Pictures showing the Flower (A), Fruits (B), Seeds (C), Seedling (D), mature trees producing frankincense and the dried resin Frankincense (F).

Conservation status of *B. papyrifera* and the genetic threats

Dry-lands contain highly resilient species adapted to the seasonal pattern of rainfall and recurrent drought that prevail in these ecosystems. However, powerful forces of resource exploitation and vegetation degradation overpower the resilience of dry-land ecosystems and constitute potentially serious threats to the genetic resources **(**Janzen, 1988; Mortimore et al., 2009). Woodlands in Ethiopia have been facing several challenges and the potential range of forest communities with *B. papyrifera* is greatly reduced and degraded since late 1960's (Gebrehiwot, 2003; Lemenih et al., 2007). Alemu et al. (2015) detected 1.56 to 1.68% annual rate of woodland conversion during 1985 to 2010 in the North western low lands. The forests with *B. papyrifera* have suffered degradation largely from anthropogenic habitat destruction because of: the frequent occurrences of fires, threats from open livestock grazing, conversion of forests

to agricultural lands (e.g large-scale sesame production in North and North Western Ethiopia) and over exploitation of forests mainly for wood products (Abiyu et al., 2010; Eshete et al., 2011; Groenendijk et al., 2012; Tolera et al., 2013). *B. papyrifera* populations have been decreased tremendously and fragmented to smaller patches (Abiyu et al., 2010), regeneration is blocked for the last 50 years in some of the populations (Tolera et al., 2013), and adult trees which are still in reproductive and productive ages have been dying in some localities for unidentified reasons (Abiyu et al., 2010; Groenendijk et al., 2012). Projections based on the current rate of adult mortality and regeneration bottlenecks predicted a 90 % loss of currently existing *B. papyrifera* trees in 50 years time. This loss will be associated with 50 % loss of frankincense production in 15 years time (Groenendijk et al., 2012). In conclusion, *B. papyrifera* populations have been threatened by fragmentation, decline in population size, adult mortality, and lack of regeneration due to the human-driven factors. As depicited in Fig. 1-3, the threats or combinations of them might have impacted the genetic processes: impaired gene flow, elevated inbreeding and genetic drift (the blue coloured box in Fig. 3-1). Subsequently, these might have lead to a loss of genetic diversity, fine scale genetic structuring and divergense of populations which may also lead to loss of genetic variation if fragmentation followed by further population bottlenecks.

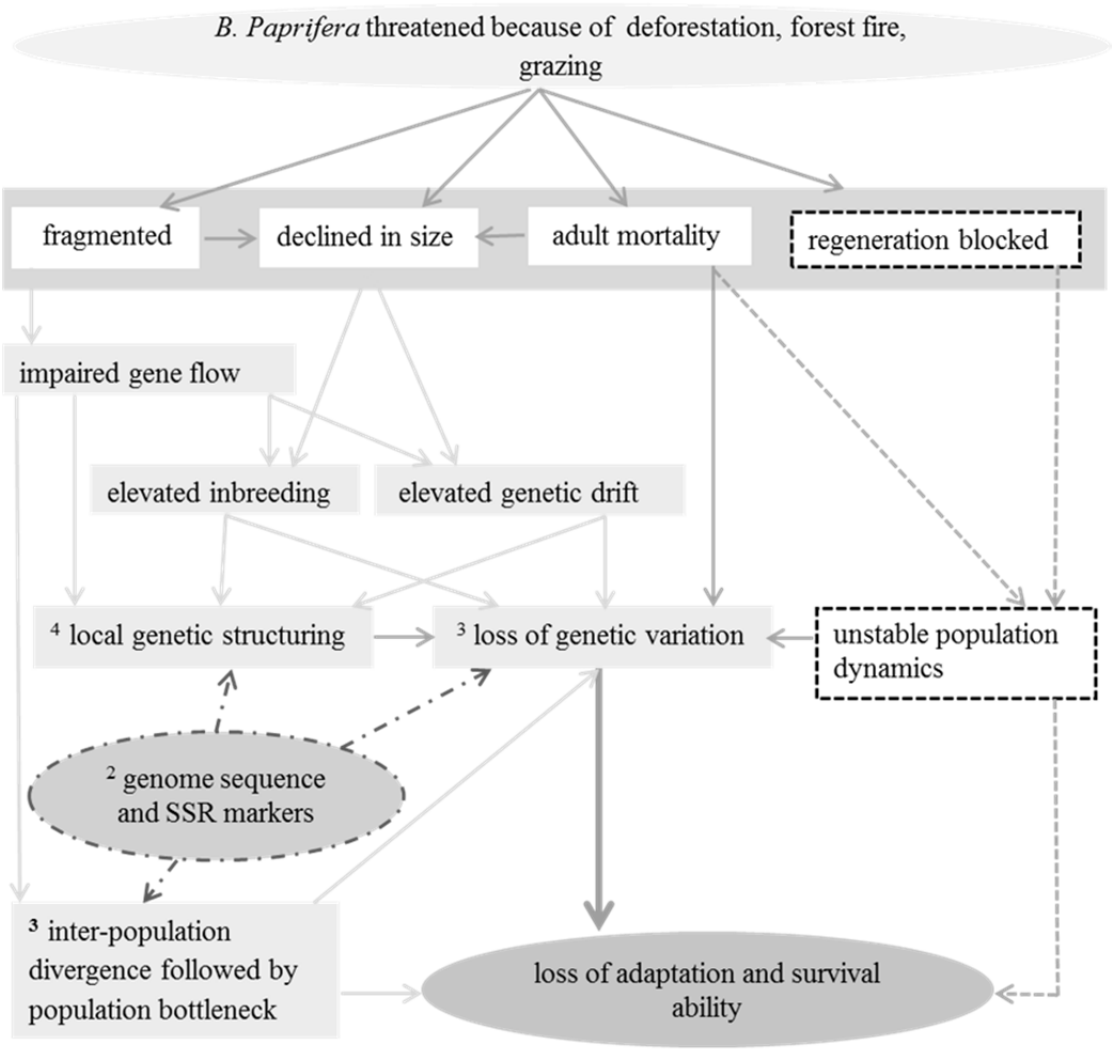


Fig. 1-3: A flowchart of the probable genetic consequences of anthropogenic threats to the genetic processes and genetic diversity of *B. papyrifera.* The Boxes numbered 2-4 were the focus of this study.

Research objectives

Conserving forest genetic resources is vital, as it constitutes a unique and irreplaceable resource for the future, including sustainable economic growth and progress and environmental adaptation. This study generally deals with the conservation genetics of *B. papyrifera*, which is a one of the dominant dry forest or woodland species of enormous economical use, and ecological and societal value. Despite its diverse use and notable economic importance and the ongoing threats, *B. paprifera* is under no conservation measures. The frankincense is harvested from the existing natural forests. *B. papyrifera* forest,

however, is depleting. The genetic tools needed to characterize the genetic diversity pattern are hardly available for tree species in Ethiopia, particularly those of the dryland woodland tree species including *B. papyrifera*. This study specifically aimed at 1) sequencing the genome and developing microsatellite markers 2) characterizing genetic diversity within and among *B. papyrifera* populations in Ethiopia 3) characterizing fine-scale spatial genetic structure and gene dispersal within *B. papyrifera* populations 4) assessing the change in level of genetic diversity and fine-scale genetic structure across adult and seedling generations.

Thesis Outline

This thesis comprised of the following chapters.

Chapter 1 provides the background information on the forest and woodland resources, the existing threats and the expected genetic consequences on *B. papyrifera* within the Ethiopian forest history context. The utility of molecular genetic markers in conservation of forest genetic resources are indicated.

In Chapter 2 part of the genomic DNA of *B. papyrifera* was sequenced applying Illumina paired-end sequencing and we assembled 444927 contigs (i.e. longer stretch of sequences). Genome size of the species was determined. Sets of 46 popymorphic microsatellite markers are developed from the Illumina reads. In addition several contigs which represent part of genes of the terpene synthesis pathways are identified.

In Chapter 3 of this dissertationI studied the genetic diversity and differentiation of 12 populations sampled across Ethiopia applying a subset (10) of the 46 markers developed in chapter 2. The genetic analysis in this study generated valuable results that increased our understanding about the level and spatial patterns of the genetic diversity of *B. papyrifera* in Ethiopia and provided information useful for identifying genetic clusters and populations for

conservation *in situ*, in natural forests and fragmented landscapes and for *ex situ* conservation.

Chapter 4 assessed the fine-scale genetic structure (FSGS) and a change in level of genetic diversity across generations within two *B. papyrifera* populations (Guba-Arenja and Kurmuk) considering two cohorts (adults and saplings) in each of the populations. These analyses showed adult B. papyrifera trees (multi-generation) are genetically related up to a distance of 62 m. No loss of genetic diversity was detected from adult to sapling cohorts. The outcomes from this investigation have led to recommendation of the distances to be considered for sampling unrelated plant material (seeds or vegetative part) for *ex situ* conservation as well as for plantation programs of other purposes.

Chapter 5 presents the general discussion where the results described in chapter 2-4 are synthesized and connected. The chapter discusses the implications of the findings for conservation and management of the genetic resource of this species within broader scope encompassing all relevant findings from previous studies. Research problems to be addressed in line with the conservation of the species and sustainable production of frankincense are outlined.

Research methodology

Description of the studied populations

In Ethiopia *B. papyrifera is* distributed over a wide geographical range in the Western, North western, Northern, North western and central parts of the country (Fig. 1-4), in Amhara, Tigray, Benishangul-Gumz, Oromia and Afar region.

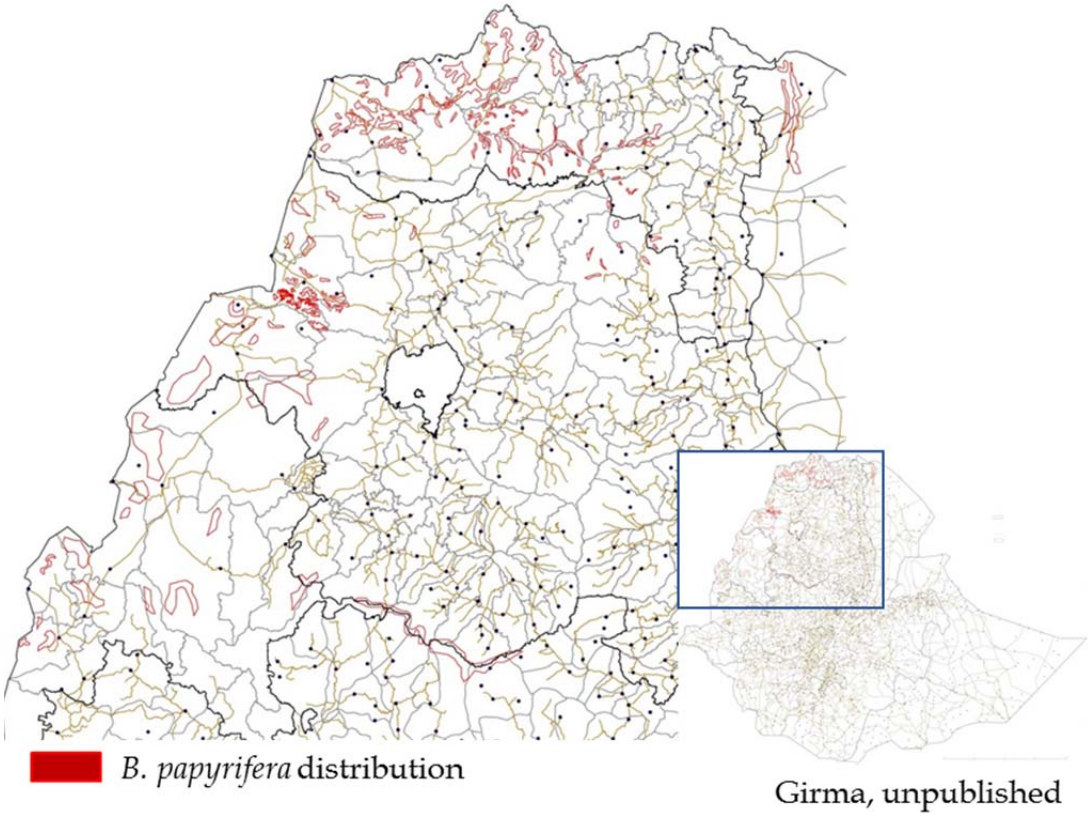


Fig. 1-4: Geographical distribution of *B. papyrifera* in Ethiopia

The populations vary in terms of the level of disturbances. The North and North eastern populations are severly degraded while the Western populations are relatively intact and anthropogenic interference is relatively recent. The populations ranged from small (< 10 ha), degraded (< 50 stem per ha) without saplings, to extensive (> 50 ha), dense (> 400 stems per ha) and stocked with seedlings and saplings (examples in Fig.1-5). The study was designed taking into account the different disturbance regime to obtain representative population samples for investigation.

For analysing the FSGS and change in genetic diversity and structuring between cohorts, groups with age gaps are required. For these purposes two

populations with abundant regeneration/saplings were considered from the Western part of Ethiopia.

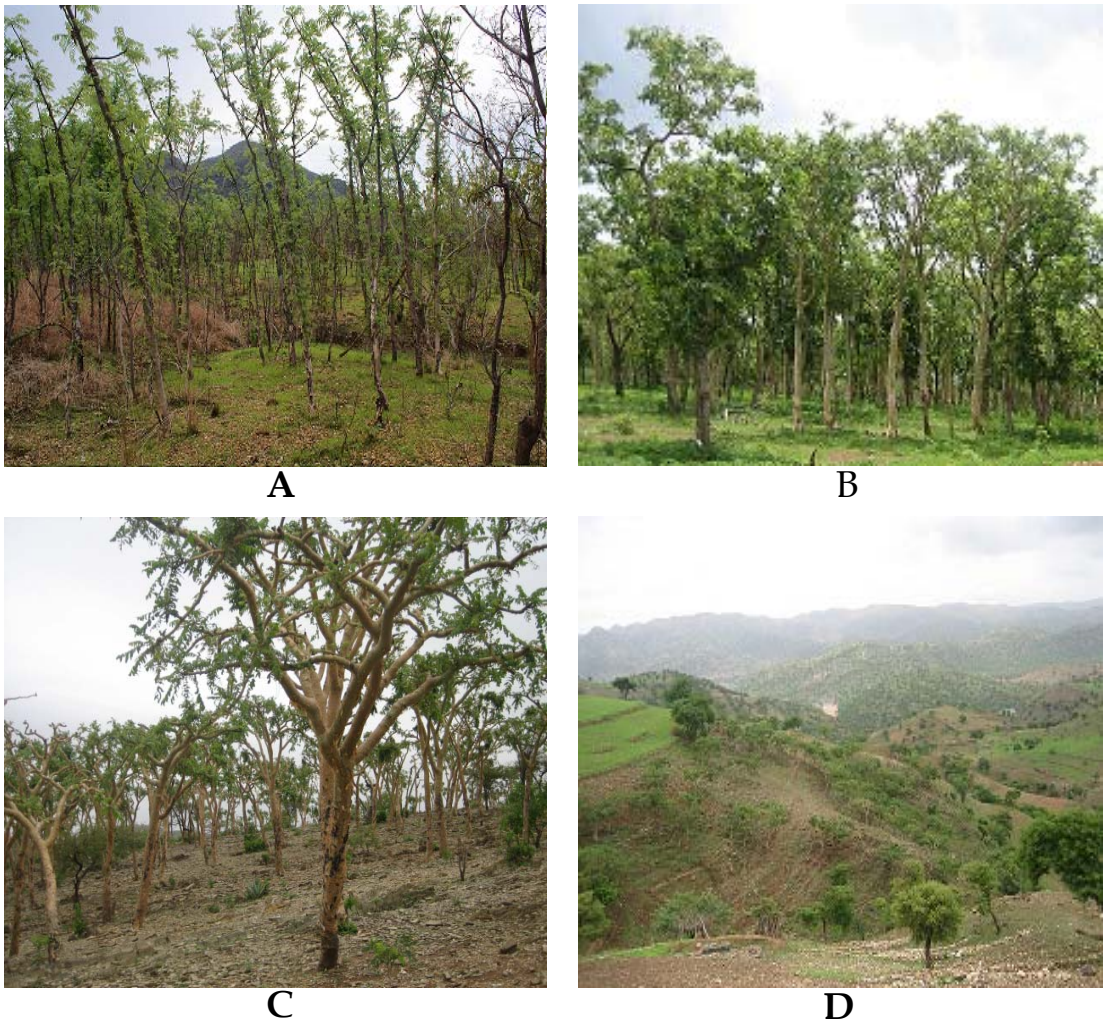


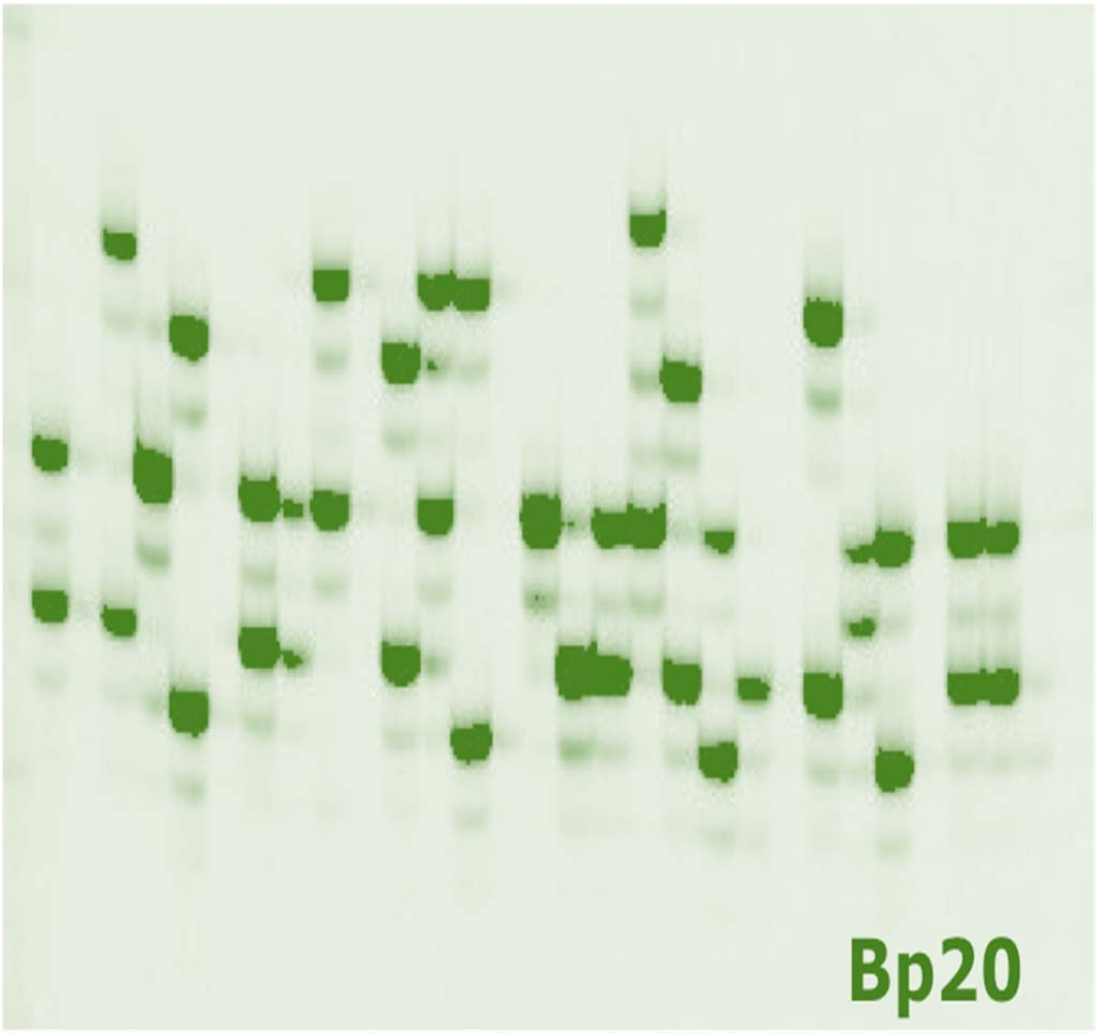
Fig. 1-5: *B. papyrifera* populations under different disturbance levels: dense young populations stocked with juvenile individuals and seedlings (A), dense population lacking juvenile individuals (B), highly degraded old populations (C), fragmented population patches and scattered trees on farm landscapes (D).

CHAPTER 2

GENOMIC SEQUENCING AND MICROSATELLITE (SSR) MARKER DEVELOPMENT FOR FRANKINCENSE TREE Boswellia papyrifera

A.B. Addisalem, G.D. Esselink, F. Bongers, M.J.M. Smulders

Published in Aob Plants 7: plu086 (2015).



Abstract

Microsatellite (or SSR) markers are highly informative DNA markers often used in conservation genetic research. Next generation sequencing enables efficient development of large numbers of SSR markers at lower costs, also in under-studied species. *Boswellia papyrifera* is an economically important tree species for frankincense production, an aromatic resinous gum exudate from the bark of the tree. It grows in dry tropical forests in Africa and is threatened by a lack of rejuvenation. For efficient conservation of source populations, seed orchards and germplasm collections the genetic structure needs to be assessed. Sequences were generated using Illumina paired-end sequencing of genomic DNA. The genome size was estimated at 705 Mb per haploid genome. The reads contained 1 microsatellite repeat per 5.7 kb. Based on a subset of these repeats 46 polymorphic SSR markers were developed which amplified 2-12 alleles in 10 genotypes. This set included 30 trinucleotide repeat markers, 4 tetranucleotide repeat markers, 6 pentanucleotide markers and 6 hexanucleotide repeat markers. Several markers were cross-transferable to *B. pirrotae* and *B. popoviana*. In addition, retrotransposons were identified, the reads were assembled and several contigs were identified with similarity to genes of the terpene and terpenoid backbone synthesis pathways, which form the major constituents of the bark resin.

Keywords: SSR; terpene biosynthesis; terpenoid; resin; conservation genetics; tropical dry forest.

Introduction

Many tree species are harvested from wild populations. For the implementation of conservation programs, it is, amongst others, essential to understand the distribution of their genetic diversity. Conservation genetics studies the dynamics of genetic variation of a species over space and time using genetic markers (Karp et al., 1997; Burczyk et al., 2006; González-Martinez et al., 2006; Frankham, 2010; Nybom et al., 2014). Microsatellite or Simple sequence repeat (SSR) markers have been widely applied in quantifying the level of genetic variation, analyzing the spatial organization of the genetic variability, describing the demography and history of populations, parentage analyses and the direct estimates of gene flow in forest trees (e.g., Smulders et al., 2008; Primmer, 2009; Allan and Max, 2010). These repeats are abundant in the genome, very polymorphic and multi-allelic (thus highly informative), have co-dominant inheritance (allowing a direct measurement of heterozygosity), and markers based on them are frequently transferable across related species (Chase et al., 1996; Smulders et al., 1997; Brondani et al., 1998; Pastorelli et al., 2003; Selkoe and Toone, 2006; Allan and Max, 2010; Fan et al., 2013). Development of SSR markers has traditionally been a long, labour intensive, and economically costly process, usually yielding small numbers of useable markers (e.g., Selkoe and Toone, 2006; Zalapa et al., 2012; Vukosavljev et al., 2014). Recently, however, Next Generation Sequencing (NGS) technologies have simplified generating large amounts of sequences at affordable cost, thus facilitating thedevelopment of large numbers of molecular markers, including SSRs and SNPs (Ekblom and Galindo, 2011; Edwards et al., 2011; Castoe et al., 2012; Smulders et al., 2012; Lande et al., 2013; Vukosavljev et al., 2014), as well as chloroplast sequences for phylogeographical studies (Van der Merwe et al., 2014). The development of these markers has thus become feasible also for species for which no prior sequence information exists (Smulders et al., 2012) including for understudied but economically important crops (Zalapa et al., 2012).

Marker development can be based on very short length sequences from genomic DNA sequences or cDNA (RNA-seq). Both sets of reads will do, but there is a difference with regard to further data mining. RNA-seq data can be *de novo* assembled into a (partial) transcriptome (Yang and Smith, 2013) with caveats, partly depending on the assembler used (Shahin et al., 2012). A common denominator appears to be that multiple assemblers need to be compared (Nakasugi et al., 2014), but the final result can be compared to the transcriptome of other species. It is not straightforward (Vicedomini et al., 2013) to assess the quality of a *de novo* assembly of short reads of genomic DNA from a heterozygous species for which no prior sequence information is available, especially if the species is heterozygous, the genome is large and contains many repeats. As it is generally easier to extract DNA from dry material of wild species collected in the field (on silica gel) than to try to extract good quality RNA from fresh samples or samples specifically prepared for RNA extraction, we can expect many studies that generate a large number of shotgun sequences of short length for marker development. What additional information can reliably be extracted from a single library of short reads of genomic DNA is an open question.

B. papyrifera is currently the number one Frankincense-producing tree species in the world (Coppen, 2005). Frankincense is an aromatic resinous gum exudate from the bark of the tree. Frankincense has great economic value in the world market as ingredient in pharmaceuticals, cosmetics and as church incense (Groom, 1981; Tucker, 1986; Lemenih and Teketay, 2003). In Ethiopia besides its value in national economy, it has a significant contribution in the local livelihoods providing up to one-third of annual household income, especially in the Northern regions of the country (Lemenih et al., 2003, 2007; Woldeamanuel, 2011).

The population size of *B. papyrifera* is declining in Ethiopia (Abiyu et al., 2010; Groenendijk et al., 2012; Tolera et al., 2013), Eritrea (Ogbazghi et al., 2006) and

Sudan (Abtew et al., 2012). Little or no successful tree regeneration is found in most of its natural habitat and mortality of adult trees is increasing, resulting in a deteriorating species conservation status. Despite its economic importance and its poor conservation status presently very few conservation efforts exist and none is supported by genetic information as no genetic markers have been developed for the species.

In the present study, we applied Illumina paired-end sequencing technology to sequence genomic DNA of *B. papyrifera* with the goal of identifying microsatellite repeats and developing SSR markers.This work produced 46 polymorphic SSR markers validated on 10 genotypes. Amplification was also tested on samples of two closely related species, *B. pirrotae* and *B. popovina.* The reads were also assembled into the first genomic resource for this species, and we present a couple of structural and functional analyses on them.

Material and Methods

Plant material

Boswellia papyrifera is one of six Boswellia species that grow in various parts of Ethiopia. The *B. papyrifera* genotype used for Illumina paired-end sequencing was collected in a natural population at Kafa Humera Wuhdet (14.05265N latitude; 37.13078E longitude) in North-West Ethiopia. Young leaves were collected from growing shoot tips of the plant and preserved in silica gel while in the field and during transportation to lab for DNA extraction. A genomic DNA library for Illumina paired-end sequencing was prepared from 4 microgram of DNA following the PCR-based gel-free illumina TruSeq DNA sample prep protocol and sequenced as 2 x 100 nt paired-end reads on an Illumina HiSeq at Greenomics, Wageningen UR, Wageningen, the Netherlands.

Plant material for SSR marker development

For testing of the SSR loci a set of 12 genotypes was used. Ten of the genotypes represented populations of *B. papyrifera* collected from 10 different regions of Ethiopia. The remaining two genotypes, *B. pirrotae* and *B. popovina,* were included for testing the cross-transferability of the markers to closely related species. The *B. pirrotae* sample was from the North-western part of Ethiopia. *B. popoviana*is endemic to Socotra Island, Yemen, and the dried leaf sample was obtained through the Edinburgh Royal Botanical Garden, UK.

DNA extraction

Total DNA was isolated from silica dried young leaves following the cetyl trimethylammonium bromide (CTAB) protocol of Fulton et al. (1995). As large amount of phenolic compounds was expected because of the resin content in the leaves, the protocol was modified by the addition of 2% pvp-40 in the extraction buffer and 1% mercaptoethanol in the microprep buffer of Fulton et al. (1995), added immediately before use. The extraction was followed by purification steps using DNeasy (Qiagen, Venlo, The Netherlands) according to Smulders et al. (2010). DNA yield and quality were visually assessed on a 1% Agarose gel.

Sequence filtering

The raw reads were error-corrected using musket (Liu et al., 2012). This error-corrected set was used for the repeat assembly. Prinseq-lite 0.20.04 (Schmieder and Edwards, 2012) was used for quality control and filtering of reads (minimum read length of 50 nt, minimum average base quality of 25, maximum ambiguous nt (N) of 1) after which the data were used for SSR mining. After low complexity trimming (minimum DUST score of 7 for removal of low complexity reads, removal of duplicate reads, also with Prinseq-lite), paired-end reads with overlapping sequences were connected using COPE (Liu et al., 2012) in the full mode. Reads were filtered for chloroplast sequences by mapping the reads against the closest chloroplast genome available, which is

the one of *Citrus sinisensis*, using bowtie2 (Langmaed and Salzberg 2012, settings -D 20 -R 3 -N 1 -L 20 -i S,1,0.50 –a).

Repeat analysis

Reads from the highly repeated fraction of the genome were extracted and assembled using RepARK (REPetitive motif detection by Assembly of Repetitive K-mers; Koch et al., 2014). The motifs present in the repetitive contigs were counted and analysed by blastn (e-value 1e-5) against Repbase v19.08 (database of repetitive DNA elements, Jurka et al., 2005).

Assembly and annotation

A *de novo* draft assembly was created from the filtered reads using Soapdenovo 2.21 (Li et al., 2012, settings -K 41 -M 3 -d 4). The gaps emerging during the scaffolding process by SOAPdenovo were closed using GapCloser (vs 1.12). The contigs larger than 1000bp of the draft assembly were analysed and functionally annotated using Blast2GO (Coneza et al., 2005).

SSR mining and design of primers

Five million of the filtered but not assembled reads were analyzed with PAL_FINDER 0.02.03 (Castoe et al., 2012) to identify SSRs using slightly adjusted criteria: at least six contiguous repeat units for dinucleotide repeats, four for tri- and tetranucleotide repeats, and three for penta- and hexanucleotide repeats (Castoe et al., 2012) used six units for trunicleotide repeats). Following Castoe et al. (2012) the reads with multiple SSR loci were considered a 'compound' repeat if the SSRs had a different repeat motif, but a 'broken' repeat if the SSRs had the same motif. Reverse-complement repeat motifs (e.g., TG and CA) and translated or shifted motifs (e.g., TGG, GTG, and GGT) were grouped together, so that there were a total of four unique dinucleotide repeats, 10 unique trinucleotide repeats, and so on.

A subset of over 70,000 SSR-containing reads was used to further screen potentially amplifiable SSR loci (PALs): loci for which PCR primers could be designed. Primer designing followed the default parameters specified in Primer3 (Rozen and Skaletsky, 2000). The reads were then screened for differences in lengths of those sequences that contained these primers (as in Vukosavljev et al., 2014). At these loci the sequenced plant may be heterozygous, thus indicating that the locus is polymorphic. These formed the group of potentially polymorphic loci.

SSR loci amplification and analyses of polymorphism

PCR were performed in a total volume of 10 µl reaction mix containing 4 µl 2ng/ul DNA, 5 µl MP mix from Qiagen kit,0.8 µl *(2 µM)* universal fluorescent labeled primer and 0.2 µl mix of the forward and reverse primers. The fluorescent labelling method described in Schuelke (2000) was adapted to label the primers for analyses of the PCR products with a laser detection system. For this the forward primers were labelled with a universal M13 sequence (AACAGGTATGACCATGA) at the 5′ end while the reverse primers were tailed with GTTT at their 5′ end according to Brownstein et al. (1996) to reduce stutter bands (both tailing sequences are not shown in the sequences in Table 1). A thermal cycling profile was set at 15 minutes of initial denaturation at 95 ^{o}C, followed by 30 cycles of 30 second denaturation at 94 ^{o}C, 45 second annealing at 56 ^{o}C and 45 second extension at 72 ^{o}C. This was followed by additional 8 cycles with 53 ^{o}C annealing temperature to facilitate the annealing of the fluorescent dye-labeled M13 primer, and a final extension step of 10 minutes at 72 ^{o}C. After amplification 10 µl water was added. Fluorescently labelled amplicons were resolved on a 4200 or 4300 Licor DNA analyzer.

Results

Next generation sequencing

Genomic DNA of one *B. papyrifera* individual was sequenced in order to obtain a library to mine for microsatellite repeats. One lane on an Illumina HiSeq produced 143,458,368 raw reads. Based on k-mer counts the estimated genome size of *B. papyrifera* was 705 Mb, sequenced at 36x. After error correction and filtering reads for short sequences, sequences with ambiguities (Ns) and low complexity, and excluding redundant sequences, 120,479,203 (84%) paired-end reads and 10,851,777 single-end reads remained.

SSR identification

A search of SSRs in a subset of 5 million Illumina paired-end reads identified 170,832 reads (3.4%) containing SSR repeats. In these reads a total of 175,607 repeat loci (dinucleotide through hexanucleotide repeats) were identified, which corresponds to 1 SSR locus per 5.7kb. Fig. 2-1 shows the frequency of the top-20 repeat motifs. These include all dinucleotide motif repeats (of at least 6 repeat units long), of which AC and AT repeats were the most abundant. Of the trinucleotide repeats (of at least 4 repeat units) AAT and AAC were the most frequent, followed by TTC. Excluding the dinucleotide repeats, the remaining 70,415 SSR loci were screened for the presence of sufficient forward and reverse flanking sequences suitable to design primers. This yielded 29,886 (42%) Potentially Amplifiable Loci (PALs). Further filtering of these PALs by applying the most stringent criteria aimed at selecting single copy loci yielded 4,071 potentially amplifiable SSR loci**.**

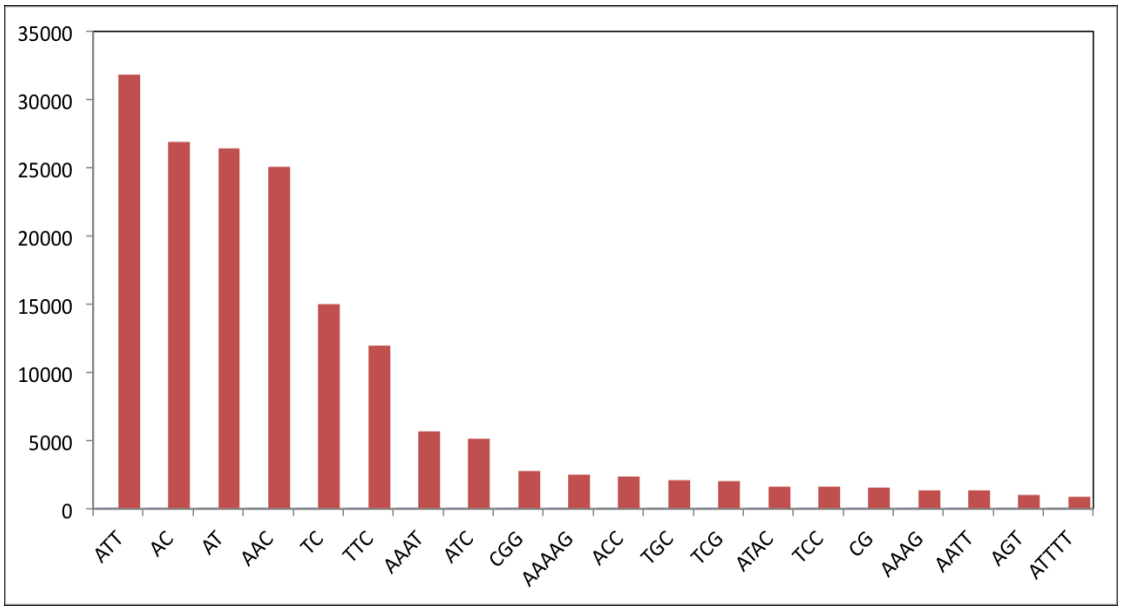


Fig. 2-1: The 20 most frequent simple sequence repeat motifs obtained, sorted according to frequency.

Polymorphism and amplification of SSR loci

A total of 136 SSR loci (117 randomly picked and 19 loci predicted to be potentially polymorphic as they appeared to have two different alleles in the sequence reads) were tested for amplification and degree of polymorphism in 10 randomly chosen individuals from different populations. Of the 117 randomly picked loci 82 primer pairs amplified a high quality PCR product, of which 37 (45%) were polymorphic with a banding pattern that could be scored clearly (Table 2-1). Of the 19 primer pairs predicted to be polymorphic, 13 amplified bands of which 9 loci (69%) were polymorphic, indicating a significantly higher rate of polymorphism *(Chi-squared test, $p < 0.005$)* compared to randomly picked loci. The final set of 46 markers included 30 trinucleotide repeat markers, 4 tetranucleotide repeats, 6 pentanucleotides and 6 hexanucleotide repeats. The number of alleles across the polymorphic loci varied between 2 and 12 with an average value of 4.8 alleles in 10 genotypes. Several of the polymorphic markers with 10-12 alleles were TTC repeats. The heterozygosity per locus ranged widely from 0.10 to 0.89 (average 0.43)**.**As the test set was small the number of alleles is probably an underestimation.

As depicted in Table 2-1, most of the SSRs successfully amplified in *B. pirrotae* and in *B. popoviana* (not all markers were tested in the latter species as the amount of DNA was limited). Amplification, even in the same size range as the alleles in *B. papyrifera*, is not proof that the marker is polymorphic, but heterozygosity (two different alleles in the expected range) is. Based on that criterion at least 19 of the 46 markers are polymorphic in *P. pirrotae* and at least 8 of 33 tested are polymorphic in *B. popoviana*.

Table 2-1: Forty six polymorphic microsatellite markers developed for *Boswellia papyrifera* and their cross-transferability to *Boswellia pirrotae* and *Boswellia popoviana.*

Name	Primer sequence (5'----> 3')	Repeat motif	A[1]	Alleles size range (bp)	Quality (Smulders et al.,1997)	Ho[2] based on 10 *B. papyrifera* genotypes	Other Boswellia species[3]
Bp01	F:TTGTTAAGGCTTTTCTCCTC R:GTTGCTTATCTTTGGCTGAG	(AAG)6	4	119-134	2	0.34	Br=het Bv=hom
Bp02	F:TGAGAAGTTTACCCTTTATGTTT R:TCTCTGCCTCTTCTTCTTATTT	(ATT)13	7	195-219	2	0.78	Br=hom Bv=hom
Bp03	F:ATGGGGAAAGGTTAAAGATC R:CTGCACAACACAAGTTAAGC	(ATC)6	3	123-129	1	0.1	Br=het Bv=het
Bp04	F:TATCAACACTTTTGTTTTGC R:CAATTCGAGTCTCCTCAAC	(TTC)8	2	182-197	3	0.2	Br=het Bv=het
Bp05	F:GGAGCAGGTACCTTGTATGT R:AACAGATCTCTTGGTTTGATT	(AAC)7	5	232-250	1	0.8	Br=hom Bv=hom
Bp06	F: GATCTCCACTTGATCAGGAC R:ACATGGAAAATTGAAAGCAC	(TTC)9	8	263-297	1	0.5	Br=het Bv=het
Bp07	F:GAAACTTTGTGGGTGTTTGT R:TCATCCTCTGACATATCCATT	(ATT)8	3	284-293	1	0.34	Br=hom Bv=hom
Bpo8	F:TTTTCTGTGTTTTGTACGCA R:GCATGCAAGAAATAGGAGAG	(ATT)6	3	207-213	2	0.11	Br=no ampl Bv=no ampl

Table 2-1 continued

Name	Primer sequence (5'----> 3')	Repeat motif	A[1]	Alleles size range (bp)	Quality (Smulders et al., 1997)	Ho[2] based on 10 *B. papyrifera* genotypes	Other Boswellia species[3]
Bp09	F:TTGATCAATTATTTCGGACA R:AAAATGCAAGTCCTTTGTAA	(ATT)11	7	292-331	1	0.78	Br=no ampl Bv=het
Bp10	F:CTTTGGCAGATTCAAATAGG R:GACACAAGAAAATTGAGGGA	(TTC)6	4	197-213	1	0.11	Br=het Bv=het
Bp11	F:AGAGAATTCCCTAAGGAGAGA R:TCTACAATAGCCCAGCAACT	(TTC)9	6	284-307	1	0.78	Br=hom Bv=het
Bp12	F:ACCCATGATAAAGAGTTCCA R:GAGAACGCCGTTTGAGTT	(ATT)10	7	238-302	2	0.56	Br=het Bv=no ampl
Bp13	F:ATAATTTCCCACCAGGAGAT R:CAACGAACTACAAGTATTGAATG	(ATT)7	3	227-239	1	0.22	Br=hom Bv=hom
Bp14	F:GGCAATTATTTGATCGCTAC R:ATGACATTCATTCGTAACCC	(ATT)15	8	198-253	1	0.44	Br=het Bv=hom
Bp15	F:TATATGCCTTGCTAAGCGTT R:AAACTCCGAGCTGACTACAC	(ATC)10	7	301-337	1	0.78	Br=het Bv=hom
Bp16	F:AAAACTTTGTTTCCTCTCCA R:TCAGAAGGAAGCACTTCAAC	(TCC)11	2	218-221	1	0.33	Br=hom Bv=hom
Bp17	F:AGCAATATTTCCAAAGGACA R:CTGCCCAATAACATAGTTCC	(TTC)11	6	200-215	1	0.4	Br=no ampl Bv=hom

Table 2-1 continued

Name	Primer sequence (5'----> 3')	Repeat motif	A[1]	Alleles size range (bp)	Quality (Smulders et al.,1997)	Ho[2] based on 10 *B. papyrifera* genotypes	Other Boswellia species[3]
Bp18	F:TTATCTTGTAGTGGGATGGG R:GAGAACTGGTAATCACATGAAA	(TTC)12	6	221-262	2	0.67	Br=hom Bv=no ampl
Bp19	F:GTGCCAGAATTCAGGTATGT R:GGTTGTGAGTCCACCATTAT	(TTC)13	5	287-321	2	0.1	Br=het Bv=hom
Bp20	F:TGCTTTATGACTTTGTTGAGA R:GAACCATCATGCAATTAGTTT	(TTC)15	10	227-266	2	0.5	Br=het Bv=hom
Bp21	F:CAGAGTTAATAATATAAGTAGCAGCA R:CTATGTTCATACTTAGAAAAGTTGG	(TTC)16	12	117-299	1	0.6	Br=hom Bv=hom
Bp22	F:TAAAACCATTTTCAGCAAGG R:AGAACCAGACCTTCAAATCA	(TTC)17	11	237-307	1	0.7	Br=hom Bv=het
Bp23	F:GCGAATTTGCTCTGTAATTCR:TAAGA CCCCAAGAAATTGAA	(TTC)20	11	224-266	2	0.8	Br=het Bv=hom
Bp24	F:TATTTGTCAACAGATTGGGG R:CAGTCTAAGTCCACAAACTCC	(CGGGG)3	2	241-251	1	0	Br=hom Bv=hom
Bp25	F:ATCATCATCAGGTGAAGACC R:ATGTCGTTTTCGACTTTCG	(TCTCGC)3	4	261-279	1	0.22	Br=hom Bv=hom
Bp26	F:AAATCATGTTTGGCTAATGG R:TGCAAATGCAAATTAATGG	(TGCC)6	3	235-247	1	0.34	Br=hom Bv=hom

Table 2-1 continued

Name	Primer sequence (5'----> 3')	Repeat motif	A[1]	Alleles size range (bp)	Quality (Smulders et al.,1997)	Ho[2] based on 10 *B. papyrifera* genotypes	Other Boswellia species[3]
Bp27	F:CTCTAGATGCATAGGGATGG R:AAATATAATCCTAAACCTTGCG	(TCCGGG)3	2	240-246	1	0.25	Br=no ampl Bv=no ampl
Bp28	F:CAAATCCTTGTGATTTCTCC R:AAGTAGCCATAAATAATCATAGGG	(AAGAG)3	4	262-272	1	0.14	Br=het Bv=hom
Bp29	F:ATTTCACAAATCACTTTCGC R:TTAACAAGTAACGCTAACGC	(TC)10(AGCG)5	6	249-264	1	0.43	Br=hom Bv=het
Bp30	F:ATATGCTAGAGACTTGGCCC R:TTTTCAATGCTTGGATGC	(TTGGGC)3	3	200-212	1	0.34	Br=hom Bv=hom
Bp31	F:CAGAACAAAAGTGACAGTTAGC R:GAGGCAAAGAGACTTGACC	(AGAGC)4	4	277-307	2	0.75	Br=hom Bv=no ampl
Bp32	F:TCATAACTTCCAAAATTGAGC R:TTTCTATCTTTGGATCAATGC	(TCTG)4	3	144-156	1	0.11	Br=hom Bv=no ampl
Bp33	F:CGTCTACCTCCTTCTCTTCC R:GTACTAAACCCTCCGTTCG	(TCTCC)3	2	171-181	3	0.33	Br=het Bv=no ampl
Bp34	F:AGAGAACATCCCAAGAATCC R:AGGATGGAGAGCCCTAGC	(ATGGAG)4	4	183-193	1	0.56	Br=het

Table 2-1 continued

Name	Primer sequence (5'----> 3')	Repeat motif	A[1]	Alleles size range (bp)	Quality (Smulders et al.,1997)	Ho[2] based on 10 *B. papyrifera* genotypes	Other Boswellia species[3]
Bp35	F:GGCTCCTCGCTAACCGACC R:CTCCCAGTCGAGATCGAGCC	(TTGGCG)4	2	224-230	1	0.1	Br=hom
Bp36	F:GGTATAAAGAGAAAGGGATAGAG G R:CACAATTTACTGGCAATGG	(TGTGC)3	4	211-226	2	0.89	Br=hom
Bp37	F:ATCTCGCATTCCTACATCC R:ACGACCTCTTCATCTAACCC	(ATGC)5	2	277-283	1	0.11	Br=hom
Bp38	F:GTTGAGAATGAGAAGAACGG R:CATCAACTTCCTCAAATTCC	(ATC)7,(8)	5	243-273	1	0.22	Br=het
Bp39	F:TCATGGAATAAGAAACCAAA R:TCTTAACATTTCGTCTGCTG	(ATC)8, (9)	8	247-298	2	0.6	Br=het
Bp40	F:AAACAAATATACGTGGCACA R:TCCAAGTGAACATCCAAAAT	(ATT)8,(14)	3	240-255	2	0.3	Br=hom
Bp41	F:TGGGTTTAAAGTATTCTAAAAGG R:CATTAGAAGAGGCAAAATGG	(ATT)8,(9)	4	230-252	2	0.22	Br=hom
Bp42	F:TTATAAGCAGAGCAAATTATAGC R:CTAATTTCGCAATTTAAGGC	(ATT)10,(11)	6	228-264	2	0.4	Br=hom

Table 2-1 continued

Name	Primer sequence (5'----> 3')	Repeat motif	A[1]	Alleles size range (bp)	Quality (Smulders et al.,1997)	Ho[2] based on 10 *B. papyrifera* genotypes	Other Boswellia species[3]
Bp43	F:CCAAGCCTATACACTTCTTCA R:GATGAATTGGGCTTAGATTG	(TTC)6,(8)	6	272-293	3	0.89	Br=het
Bp44	F:CCATATGGGGATATAGGTCA R:TTGGCCAAGAAGAAACTTAG	(ATT)6,(7)	4	226-235	2	0.25	Br=het (out of range)
Bp45	F:AACAGTTGGTTTAACAACGC R:CTTAAAAGGGAACTGGAAGG	(AACAAG)3,(4)	3	281-293	1	0.67	Br=het
Bp46	F:ATATTCAATTTATCTGTGTGACG R:TTTGATTTCAAAGGAAAACG	(ATATT)3, (4)	2	256-271	2	0.75	Br=hom

[1] A= number of alleles in 10 *B. papyrifera* genotypes.

[2] Ho = observed heterozygosity (a tentative figure, as the 10 individuals are from 10 different populations.

[3] Amplification was also tested in one individual of *B. pirrotae* (Br) and one of *B. popoviana* (Bv) except where no Bv is indicated. Hom=homozygous and Het=heterozygous, always with products in the same size range as the alleles in *B. papyrifera*, except where noted that they were out of range. No ampl = no amplification.

Sequence assembly and annotation

The Illumina reads are the first genomic resource generated in the genus Boswellia. The repeat fraction was assembled based on K-mer frequency. This produced 49,576 contigs of repeats that were present at least 50X (median length 139 bp, mean length 224 bp, N50 238 bp, maximum length 21,153 bp, total sum = 5,74 Mbp). Next, 1533 contigs had blastn hits with RepBase, mostly with Copia (639 hits) and Gypsy (523) retrotransposons, alongside EnSpm (114), hAT (72), Satellite (29), TY (23), Harbinger (16), YPrime (14), Helitron (12) and SCTRANSP (3). Intermixed with these elements were hits to the ribosomal RNAs (LSU 56 hits, SSU 41) and also to Caulimoviridae viruses (11).

Using all data in a *de novo* assembly with Soapdenovo, 444,927 contigs were obtained with a median of 375 bp, a mean contig length of 690 bp, a N50 of 1085 bp, and a maximum contig size of 19236 bp (total sum = 307 Mb genomic DNA sequence). The contigs larger than 1000 bp were blasted against Genbank, and 65,467 were annotated with GO terms (Fig. 2-2; note that these are overlapping classes).

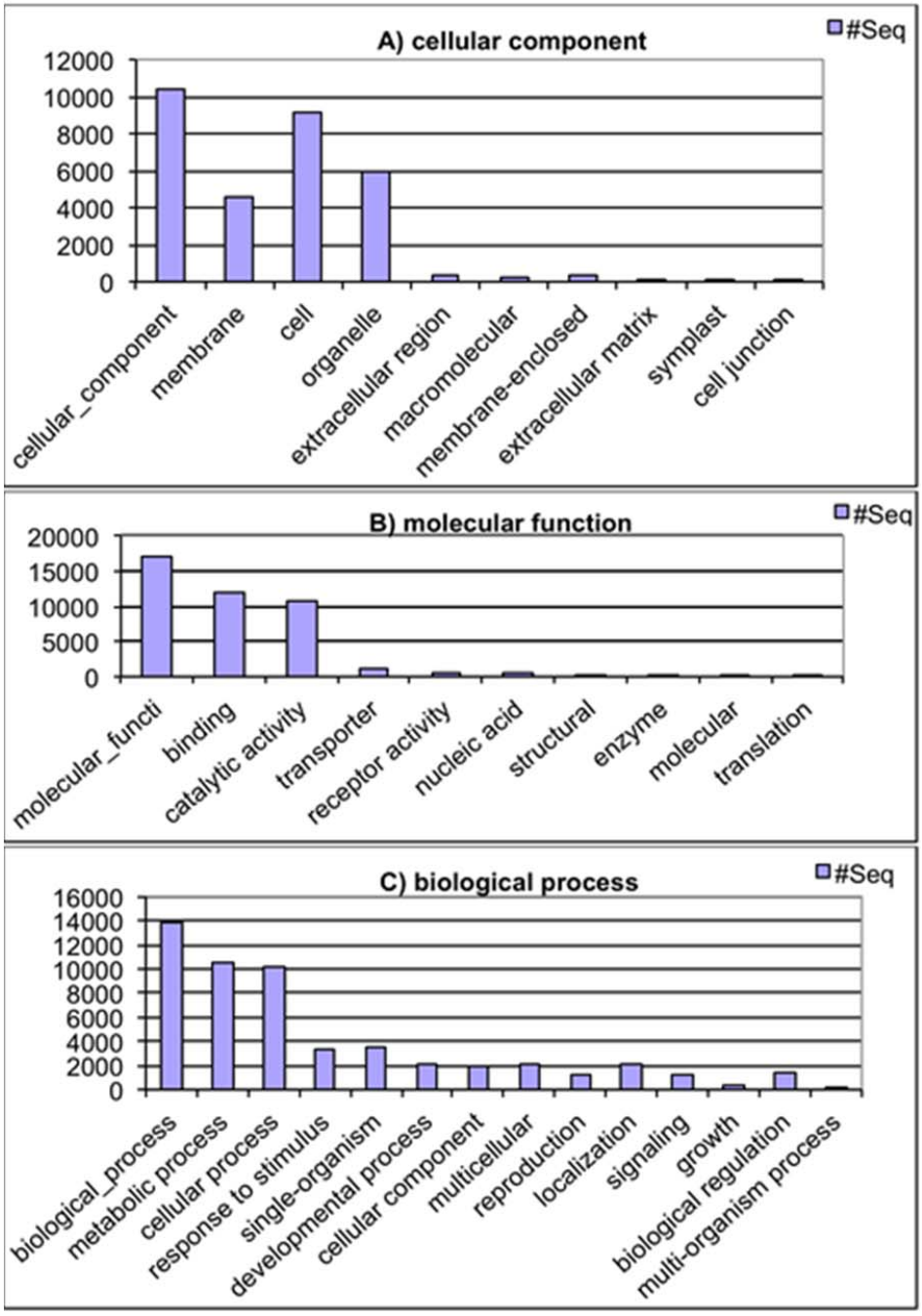


Fig. 2-2: Representation of ontology assignments of the *B. papyrifera* contigs. **A,** the 31,086 GO terms of cellular components, **B**, the 42,423 GO terms of molecular function and **C**, the 54,256 GO terms of biological processes. Note that these are overlapping classes.

Terpene biosynthesis genes

Asseffa et al. (2012) conducted a biophysical and chemical study on resins of *Boswellia* species with special emphasis on *B. papyrifera*. Using the list of identified components, we found eight contigs which represent part of genes of the terpene synthesis pathways, namely pinene synthase, limonene synthase (2x), isoprene synthase (4x) and gamma-terpinene synthase. We also searched for the enzymes that are involved in terpenoid backbone biosynthesis (according to the KEGG pathway database). Table 2-2 lists the enzymes of the mevalonate and non-metavolate (MEP/DOXP) pathways, the two pathways of synthesis of terpenoid building blocks in plants,that were found among the annotation results. Two of the key enzymes of the MEP pathway, 2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase (EC 2.7.7.60) and 4-(cytidine 5'-diphospho)-2-C-methyl-D-erythritol kinase (EC: 2.7.1.148), were not recognised in the set of scaffolds, but reciprocal tBlastx (at 1e-5) against these enzymes identified in Arabidopsis did reveal hits with respectively 3 and 2 contigs.

Table 2-2: MEP/DOXP and mevalonate pathway genes found among the contigs of *B. papyrifera.*

MEP/DOXP pathway	Name	EC nr
DXS	1-deoxy-D-xylulose 5-phosphate synthase	EC 2.2.1.7
DXR	1-deoxy-D-xylulose 5-phosphate reductoisomerase	EC 1.1.1.267
MDS	2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase	EC 4.6.1.12
HDS	4-hydroxy-3-methylbut-2-enyl diphosphate synthase	EC 1.17.7.1
IDI	isopentenyl diphosphate isomerase	EC 5.3.3.2
GPPS	geranyl-diphosphate synthase	EC 2.5.1.1
GGPPS	Geranylgeranyl diphosphate synthase	EC 2.5.1.29
CPS	copalyl diphosphate synthase	EC 5.5.1.12
KS	kaurene synthase	EC 4.2.3.19
mevalonate pathway		
AACT	acetyl-CoA C-acetyltransferase	EC 2.3.1.9
HMGS	hydroxymethylglutaryl-CoA synthase	EC 2.3.3.10
HMGR	hydroxymethylglutaryl-CoA reductase	EC 1.1.1.34
MK	mevalonate kinase	EC 2.7.1.36
PMK	5-phosphomevalonate kinase	EC 2.7.4.2
MDC	Mevalonate-5-pyrophosphate decarboxylase	EC 4.1.1.33
IDI	isopentenyl diphosphate isomerase	EC 5.3.3.2
FPPS	Farnesyl diphosphate synthase	EC 2.5.1.10

Discussion

We have developed the first set of 46 SSR markers for *B. papyrifera*. The markers amplified between 2 and 12 alleles in individuals from 10 different populations across Ethiopia. We based the marker development on DNA sequences from one individual. Most of the markers tested were chosen randomly, but the subset for which we assessed, from the sequence reads, that they probably had two alleles in this individual, gave a significantly higher success rate compared to the randomly chosen ones. This assessment is a technically easy screening step that would improve the efficiency of marker development in an outbreeding species, even if only sequences from one individual have been generated, as is often the case. It is probably not as efficient as a strategy that generates transcriptome sequences from multiple individuals with the specific aim of testing only those loci on gel for which polymorphisms in repeat length exist among the reads obtained from these individuals (Vukosavljev et al., 2014).

The SSRs were developed based on a set of Illumina paired-end DNA sequence reads from young leaves of a single individual of *B. papyrifera*. The distribution of these reads indicated a genome size of 705 Mb. This is close to the estimate of 682 Mb for *B. serrata*, the only Boswellia species listed in the Kew Gardens C-value database.

Mobile elements that are present in multiple copies in the genome were analysed based on sequence homology in K-mers that occurred at high frequency (Koch et al., 2014). We have identified a series of retrotransposons, the most common being Copia and Gypsy elements. As these elements are present in large numbers our Illumina reads probably were a sufficiently good source to determine presence and relative frequency of various elements.

We also assembled all reads of our paired-end short read library and obtained 307 Mb of unique sequences. The quality of this assembly is difficult to assess without independent other sources such as libraries of different insert sizes, and we therefore did not compare the results of various assemblers (as e.g. Shahin et al., 2012 did) or merged assemblies (Vicedomini et al., 2013). Our resource was searched for genes that are expected to be involved in production of the major compounds of the resin, which in *B. papyrifera* concerns diterpenes, triterpenes and nortriterpenes (Basar, 2005; Asseffa et al., 2012; Bekana et al., 2014). The contigs of our assembly gave significant hits for most genes of the core terpene and terpenoid pathways. We have not done an in depth analysis of the sequences in these contigs, as extracting the complete Boswellia homologs of these genes would need more bioinformatics steps and independent validation, e.g. by PCR and Sanger sequencing. However, the results indicate that for many genes of interest at least partial sequence information is present.

Conclusion

Based on Illumina paired-end sequences we have developed a set of polymorphic SSR markers for *B. papyrifera* and two sister species, which will be useful for studying genetic diversity within and differentiation between *Boswellia* populations. We also generated the first genomics resource in Boswellia.

Accession Numbers

Accession number in ENA/Genbank for the set of DNA sequences on which the SSR markers were developed: ERS403283.

Acknowledgements

We thank Alan Forrest, the Edinburgh Botanical Garden, for providing the *B. popoviana* sample. Koen Pelgrom and Doret Wouters are thanked for helping in the

lab, Robert van Loo for help in analysing the genes involved in secondary component synthesis.

CHAPTER 3

GENETIC DIVERSITY AND DIFFERENTIATION OF FRANKINCENSE TREE (Boswellia papyrifera (Del.) Hochst) ACROSS ETHIOPIA AND IMPLICATION FOR CONSERVATION

A.B. Addisalem, F. Bongers, T. Kassahun, M.J.M. Smulders

Published in Forest Ecology and Management 360, 253–260 (2016).

Abstract

Boswellia papyrifera is used to produce frankincense, a bark resin that has been a commodity of domestic and international trade since ancient times. It is harvested from natural forests. The tropical dry forest (Terminalia-Combretum) woodland ecosystems in which *B. papyrifera* is one of the dominant species, are facing anthropogenic threats. In Ethiopia *B. papyrifera* populations have decreased tremendously to smaller and isolated remnant patches, and many forests in the North-western and North-easternparts of Ethiopia completely lack recruitment of saplings. This regeneration bottleneck, in combination with adult mortality, threatens the persistence of the species. Devising an effective strategy to conserve wild genetic resources needs information on the genetic diversity and the pattern of genetic differentiation across the species area. In the present study we analysed adult trees sampled in twelve populations across the growing area of the species in Ethiopia for genetic diversity and spatial genetic differentiation using 10 polymorphic microsatellite loci. The mean level of observed and expected heterozygosity were 0.669 and 0.681 respectively, and these levels were similar for trees from larger populations and those from degraded populations. A moderate level of among populations genetic differentiation (F_{ST} = 0.084) was detected. Genetic distance between populations was correlated with geographic distance (r = 0.663, $p<0.05$). STRUCTURE analysis distinguished four distinct genetic clusters corresponding to regions with different environmental conditions. In the Western populations we detected recruitment of many seedlings and saplings, which is a significant novel finding as most of the other populations are completely devoid of saplings. We conclude that currently a high level of genetic variation is still maintainedin *B. papyrifera* adult trees across the species' range in Ethiopia including the highly degraded remnant *B. papyrifera* population patches scattered on farm and pasture lands. An effective conservation strategy for the species has to take into account the geographic distribution of source populations.

Keywords: Conservation genetics; frankincense; microsatellite marker; genetic diversity; genetic differentiation

Introduction

Widespread anthropogenic disturbance has resulted in the fragmentation and conversion of forest and woodland ecosystems into pasture and agricultural land (Lindenmayer and Fischer, 2006). Terminalia-Combretum woodlands in which *Boswellia papyrifera* is one of the dominant species (Vollesen, 1989) are no exception. These woodlands are threatened by various human-driven factors in a large part of its geographical distribution (Ogbazghi et al., 2006; Abtew et al., 2012; Abiyu et al., 2010; Eshete et al., 2011). As a result *B. papyrifera* populations have decreased tremendously to ever smaller and isolated remnant patches. In addition, a severe regeneration bottleneck is threatening the persistence of this species as many populations are devoid of seedlings and small recruiting individuals. This situation has been described in Eritrea (Ogbazghi et al., 2006), Sudan (Abtew et al., 2012) and North-western Ethiopia (Abiyu et al., 2010; Eshete et al., 2011). *B. papyrifera*is used for the production of frankincense, the dried bark resin that has been a commodity of domestic and international trade since ancient times (Groom, 1981; Gebrehiwot et al., 2003). In modern days, *B. papyrifera* is an important economic resource because of the growing demand of frankincense for diverse industrial uses, including food flavours, cosmetic ingredients (e.g. lotions, soaps, and ointment formulation), and perfumes (Tucker, 1986; Coppen, 2005). Frankincense is exclusively harvested from trees in natural forests.

In Ethiopia *B. papyrifera* is commonly distributed in the dry lowlands in the Western, Northern, North-western, and North-eastern part of the country including the river basins of the Blue Nile, Jamma, Tekeze, and Zarema. The only population in the central part of the country is located along the Jamma River Gorge, which is a tributary to the Blue Nile river. Most of the populations lack rejuvenation. Tolera et al. (2013) indicated that no recruitment has occurred

for the last 50 years in the North-western (Metema) populations. Some of the populations in the North-east are so severely degraded that only few scattered remnant trees are left on farmlands and inaccessible hill tops (Addisalem, personal observation, 2011). A high adult mortality is leading to ever smaller populations (Abiyu et al., 2010; Groenendijk et al., 2012). The ongoing threats were predicted to lead to a 90% decline in number of trees in 50 years and this loss might lead to a 50% associated loss of frankincense yield in 15 years (Groenendijk et al., 2012). Groenendijk et al. (2012) and Lemenih et al. (2014) underscored the need of implementing intensive management to sustainably conserve *B. papyrifera* and the incense production from the species.

Devising an effective strategy to conserve wild plant genetic resources requires basic information on the genetic diversity, differentiation among populations, and understanding the underlying processes that play role in determining the level and distribution of the genetic variation (Booy et al., 2000).Large populations of naturally outbreeding species usually have extensive genetic diversity. In long-lived outcrossing plants such as trees, the majority of the genetic diversity occurs within populations (Hamrick et al, 1992; Frankham, 1995; Hamrick and Godt, 1996). Pairs of populations farther away from each other will be genetically differentiated because of limited gene flow explained as movement of individuals, their seeds, or pollen over longer distances (Wright, 1943; Schaal et al., 1998). Geographic distance alone, however, may not explain all of the genetic differentiation because patterns of variation may also be impacted by environmental variables that are important for growth or survival of the species (Ledig and Fryer, 1972; Duminil et al., 2013). In addition, anthropogenic interference may affect the genetic variation and structure of a population (Frankham, 1995; Arenset al., 2007). Changes in population demography as a result of a rapid decline in population size may result in substantial loss of genetic diversity and may lead to increased differentiation across populations as a result of genetic drift if it continues for more than one or two generations (Frankham, 1995).

In this study we determined the genetic diversity and population structure of Frankincense tree *(B. papyrifera*) populations across the species range in Ethiopia. We hypothesized that (i) the genetic differentiation among populations is low since *B. papyrifera* is a long-lived and widespread species and (ii) populations farther away from each other are genetically more distant than populations close to each other due to the limited gene dispersal caused by spatial distance. We used the resulting information to identify to-be-conserved population representative of the genetic diversity of the species.

Material and methods

Study species

B. papyrifera is one of the twenty species in the genus *Boswellia* (*Burseraceae* family; Vollesen, 1989). The speciesis adeciduous monoecious treewith sweet scented flowers that are frequently visited by honeybees for pollen and nectar (Fitchl and Admasu, 1994). Fruits are about 2 cm long, usually containing three tapered seeds (Vollesen, 1989). Next to the stem being the source of frankincense, the flowers are the source of honey, the leaves provide animal fodder and various parts of the tree are used as traditional medicines.

Field sampling

Adult trees were sampled in twelve populations across the species range in Ethiopia (Supplementary kmz file). Population were sampled from natural forests, or fragmented and degraded landscapes in areas where natural forests were not available. The geographical location and altitude of the populations were recorded using a handheld Garmin Dakota 20 GPS. Characteristics of the sampled populations (altitude, temperature, rainfall and soil) are listed in Table 3-1.

Table 3-1. Ecological description of the sampled regions according to Eshete et al. (2011), Teshome (2013), Alemu et al. (2015) and own data/observations. NA = data not available; KTD: Kola-Temben-Dambar; AB: Abergelle; TRG: Tekeze-River-Gorge; AF-BIR: Afar- Bir Hale; B-KUR: Benishanguil-Kurmuk; B-SHK: Benishanguil-Sherkole; MAW: Metema-Awlaba; MQ: Metema-Quara; JRG: Jamma-River-Gorge; KHB: Kafa-Humera-Baeker; KHW: Kafa-Humera-Waldaba; TA-SJ: Tach-Armachew-Sanja. The geographic coordinates are in decimal form.

Pop.	Latitude	Longitude	Altitude (m)	Rainfall (mm)	Maximum Temp. (°C)	Population Size (ha)	Rejuvenation	Density (stems/ha)	Soil type
KTD	13.812969	38.995745	1771	700	NA	<10	Absent	<50	NA
AB	13.457950	38.858440	1622	849	29.3	> 50	Absent	200-400	clayey, shallow, not fertile
TRG	13.744338	38.207686	1015	857	NA	20-50	Absent	50-100	NA
AF-BIR	13.892760	40.092710	597	301	45	<10	few seedlings	<50	NA
B-KUR	10.563689	34.303058	680	960	39	> 50	abundant seedlings and saplings	>400	clayey to sandy-loam, red, deep, moist, fertile

Table 3-1 continued

B-SHK	10.660723	34.692611	879	1245-1350	42.0	> 50	abundant seedlings and saplings	100-200	clayey to sandy-loam, red, deep, moist, fertile
MAW	12.690783	36.313999	882	800-900	35.7	> 50	absent	200-400	clayey, deep, fertile
MQ	12.566644	36.067952	597	800-900	35.7	> 50	absent	200-400	NA
JRG	10.065459	38.467988	1120	880	NA	10-20	absent	50-100	NA
KHB	13.959036	36.858947	811	400-650	41.7	>50	few seedlings present	200-400	Combisols, luvisols, vertisols
KHW	14.055500	37.129460	985	400-650	41.7	20-50	few seedlings present	200-400	Combisols, luvisols, vertisols
TA-SJ	13.093871	37.268437	953	750	NA	20-50	absent	100-200	NA

A total of 344 genotypes were sampled (25-43 individuals per population except for one population (AF-BIR) where sampling was possible only for 8 individuals because of inaccessibility of the area). As mechanisms of pollination and seed dispersal are not well known, a minimum of 25 mdistance was kept between sampled individuals to reduce the risk of sampling genetically related individuals. Young leaves were sampled from growing shoots of the plants and dried and stored on silica gel until DNA extraction**.**

DNA extraction, PCR amplification and genotyping

Total DNA was isolated from dried young leaves following the cetyl trimethylammonium bromide (CTAB) protocol of Fulton et al. (1995) modified with 2% pvp-40 in the extraction buffer and 1% mercaptoethanol in the microprep buffer, added immediately before use, and followed by purification using DNeasy (Qiagen, Venlo, The Netherlands) according to Smulders et al. (2010). DNA yield and quality were visually assessed on a 1% agarose gel. Ten polymorphic microsatellite markers: Bp02, Bp11, Bp17, Bp18, Bp20, Bp21, Bp22, Bp23, Bp29 and Bp39 (Addisalem et al. 2015) were used for genotyping.

PCR amplification was in a 10 µl reaction mix containing 4 µl 2ng/µl DNA, 5 µl MP mix from Qiagen kit, 0.8 µl(2 µM) universal fluorescently labeled primer and 0.2 µl mix of the forward and reverse primers. For fluorescently labelling the forward primers were extended with a universal M13 sequence (AACAGGTATGACCATGA) at the 5' end (Schuelke et al., 2000) while the reverse primers were tailed with GTTT at their 5' end according to Brownstein et al. (1996) to reduce stutter bands. The PCR cycling profile consisted of 15 min denaturation at 95 ^{o}C, followed by 30 cycles of 30 sec denaturation at 94 ^{o}C, 45 sec annealing at 56 ^{o}C and 45 sec extension at 72 ^{o}C. This was followed by 8 cycles at 53 ^{o}C annealing temperature to facilitate the annealing of the fluorescent dye-labeled M13 primer, and a final extension step of 10 min at 72 ^{o}C. After amplification 10 µl water was added for dilution and increase the

volume of the product. Fluorescently labelled amplicons were resolved on a 4200 or 4300 Licor DNA analyzer. Bands were scored manually.

Data analyses

Based on the microsatellite data, the total number of alleles, effective number of alleles, mean values of expected and observed heterozygosity, and inbreeding coefficient were estimated and mean values over all loci and populations computed using GenAlEx 6.5 (Peakall and Smouse 2006, 2012). The genetic differentiation based on allelic frequency (F_{ST}), AMOVA, and pair-wise Nei's unbiased genetic distance were also analysed using this program. The correlation of genetic distance and geographic distances of populations were analysed using Isolation by distance, web service (IBDWS; Jensen et al., 2005) using pairwise genetic distances calculated using GenAlEx 6.5 and geographical distances calculated with Geographic Distance Matrix Generator 1.2.3 (Ersts, [internet]). The coefficient of determination was calculated in order to determine the proportion of the genetic distance attributable to geographic distance.

A Principal Component Analysis (PCA) was performed to explore the genetic relationship of the populations using GenAlEx 6.5. Individual-based population assignment was used to identify differentiated genetic clusters using the Bayesian assignment clustering method implemented in STRUCTURE 2.3.4 (Pritchard et al., 2000). The optimal number of genetic clusters (K) was determined using a 10000 cycle burn-in period and 100000 Monte-Carlo Markov Chains (MCMC), using the admixture model (which assumes individuals may have mixed ancestry) and assuming correlated allele frequencies among subpopulations. Simulations were repeated 20 times for K = 1 to K = 10. The K that best explained the data(most likely number of clusters) was inferred following Evanno et al. (2005). STRUCTURE analysis calculated membership coefficient of individuals (individuals Q-matrix) for each of the defined genetic clusters and the proportion of ancestry of each population in each of the clusters

(population Q-matrix) by averaging the membership coefficient of all the individuals of a population. Subsequently, populations were assigned to a cluster based on the highest average proportion of ancestry, no matter what this value wasfor the rest of the clusters (Porras-Hurtado et al., 2013).

Results

Sampling across Ethiopia

Most populations visited for sampling were degraded, lacking rejuvenation and young individuals. The most depauperate populations were AF-BIR and KTD in the North eastern region, where few trees were scattered across hilltops and farmland. In contrast, we discovered populations in the West that were well stocked with seedlings and young recruiting individuals (saplings). This discovery is new for Ethiopia, as several authors (Eshete et al., 2011; Groenendijk et al., 2012; Tolera et al., 2013) reported the complete absence of recruitment for at least five decades in the North western and North easternparts of Ethiopia. The West is less affected by agricultural activities and the woodlands are relatively intact.

Genetic diversity detected with the microsatellite markers

Ten of the 12 populations were genotyped with the 10 SSR loci. Populations AF-BIR and JRG were genotyped with only 5 loci because of failure in amplification and weak signals in the other loci, probably due to poor DNA quality as only mature leaves could be sampled. An average of 16.6 different alleles were scored across the 10 microsatellite loci used, ranging from 10 alleles at Bp02 and Bp18 to 28 alleles at Bp22 (Table 3-2). As expected, this was much higher than the number of alleles detected when developing the markers (6-12 alleles per marker, average 8.3), as these were tested on a small set of samples only (Addisalem et al. 2015). The observed heterozygosity (H_O; average 0.669) was similar to the expected heterozygosity (H_E; average 0.681) across the markers, and hence the inbreeding coefficient (F_{IS}) was very low (average 0.01, range -

0.034 to 0.077). Only marker Bp11 had a lower level of heterozygosity (0.31), even though it amplified 11 different alleles.

Table 3-2: Performance of the microsatellite markers on over 300 samples taken across Ethiopia. Total number of alleles scored per locus (A_L), mean values of observed (H_O) and expected heterozygosity (H_E), and inbreeding coefficient (F_{IS}) per locus across 10 populations (excluding AF-BIR and JRG) of *B. papyrifera* in Ethiopia.

Locus name	A_L	H_O	H_E	F_{ST}	F_{IS}
Bp02	10	0.713	0.735	0.061	0.031
Bp11	11	0.311	0.310	0.061	-0.002
Bp17	11	0.586	0.730	0.096	0.197
Bp18	10	0.526	0.527	0.121	0.002
Bp20	12	0.806	0.791	0.077	-0.019
Bp21	24	0.829	0.815	0.084	-0.017
Bp22	28	0.795	0.742	0.065	0.049
Bp23	24	0.820	0.790	0.093	-0.038
Bp29	14	0.622	0.554	0.100	-0.123
Bp39	22	0.773	0.811	0.081	0.047
Mean	16.6	0.669	0.681	0.084	0.013

Intra-population genetic diversity

The heterozygosity for each population varied between 0.622 to 0.727 for H_O and 0.661 to 0.725 for H_E (Table 3-3). The heterozygosity of KTD, which is a population of scattered remnant trees on farmlands, was comparable (H_O = 0.701 and H_E = 0.661) to that of the large populations. The mean number of alleles (N_A) and effective number of alleles (N_E) per populations were 7.78 (range 3.6 to 8.9) and 4.02 (range 2.5 to 4.4), respectively (Table 3-3), and also N_E was similar across populations.

Table 3-3: Allele frequencies and heterozygosity, averaged across 10 loci, for 12 populations of *B. papyrifera* in Ethiopia. N: number of individuals genotyped per population; N_A: total number of alleles; N_E: effective number of alleles; H_O: observed heterozygosity; H_E: expected heterozygosity; *: genotyped over 5 loci namely Bp02, Bp17, Bp18, Bp21 and Bp39. For population abbreviations see Table 3-1.

Population	N	N_A	N_E	H_O	H_E	F_{IS}
TRG	30	6.9	4.2	0.679	0.692	0.024
AB	30	6.7	3.9	0.635	0.676	0.049
KTD	30	5.9	3.3	0.701	0.661	0.061
AF-BIR*	8	3.6	2.5	0.810	0.590	-0.436
B-KUR	28	8.6	4.1	0.669	0.666	-0.018
B-SHK	43	8.7	4.3	0.637	0.671	0.028
MQ	30	8.1	3.8	0.669	0.668	-0.004
MAW	30	8.7	4.4	0.727	0.725	-0.015
JRG*	25	5.4	3.1	0.660	0.640	-0.081
TA-SJ	30	7.7	3.9	0.698	0.683	-0.034
KHB	30	7.6	3.8	0.622	0.669	0.077
KHW	30	8.9	4.4	0.651	0.693	0.057
Mean		7.78	4.02	0.669	0.680	0.010

Inter-population differentiation and genetic relationship

The average F_{ST} across all 12 populations, based on 5 markers, was 0.0109 ($P<0.001$; Table 3). When the two populations (JRG and AF-BIR) that were genotyped with only 5 markers, were excluded from the calculation, the $F_{ST\ for}$ 10 populations based on all 10 markers was slightly lower (0.084; $P<0.001$). Nei's unbiased pair-wise genetic distances ranged from 0.017 to 0.411

(Table 3-4). Generally, a large genetic distance was observed between the Western (B-SHK and B-KUR) and North eastern region (KTD, AB, and TRG) populations, which are geographically most distant. B-SHK and AB were genetically the most distant from each other (0.411) followed by B-SHK and TRG (0.386), and B-SHK and KTD (0.364).

Genetic distance was significantly correlated ($r = 0.663$, $p < 0.001$) with geographic distance (Fig. 3-1). About 44% ($r^2 = 0.44$) of the genetic distance could be due to the geographic distance among the populations.

Table 3-4: Nei's unbiased pairwise genetic distance (below diagonal) and geographic distance (in km, above diagonal) among sampled *B. papyrifera* populations in Ethiopia. For population abbreviations see Table 3-1.

	TRG	**AB**	**KTD**	**B-KUR**	**B-SHK**	**MQ**	**MAW**	**TA-SJ**	**KHB**	**KHW**
TRG	0	77	86	553	514	266	236	125	148	122
AB	0.163	0	42	591	550	319	289	177	223	198
KTD	0.178	0.145	0	626	585	346	316	203	231	203
B-KUR	0.183	0.328	0.258	0	44	295	323	429	469	496
B-SHK	0.386	0.411	0.364	0.207	0	260	287	390	436	462
MQ	0.294	0.338	0.287	0.259	0.241	0	30	143	177	202
MAW	0.194	0.242	0.214	0.236	0.248	0.066	0	113	153	176
TA-SJ	0.206	0.209	0.149	0.175	0.243	0.152	0.153	0	106	108
KHB	0.179	0.163	0.094	0.231	0.261	0.210	0.127	0.076	0	31
KHW	0.196	0.183	0.140	0.230	0.227	0.189	0.109	0.083	0.017	0

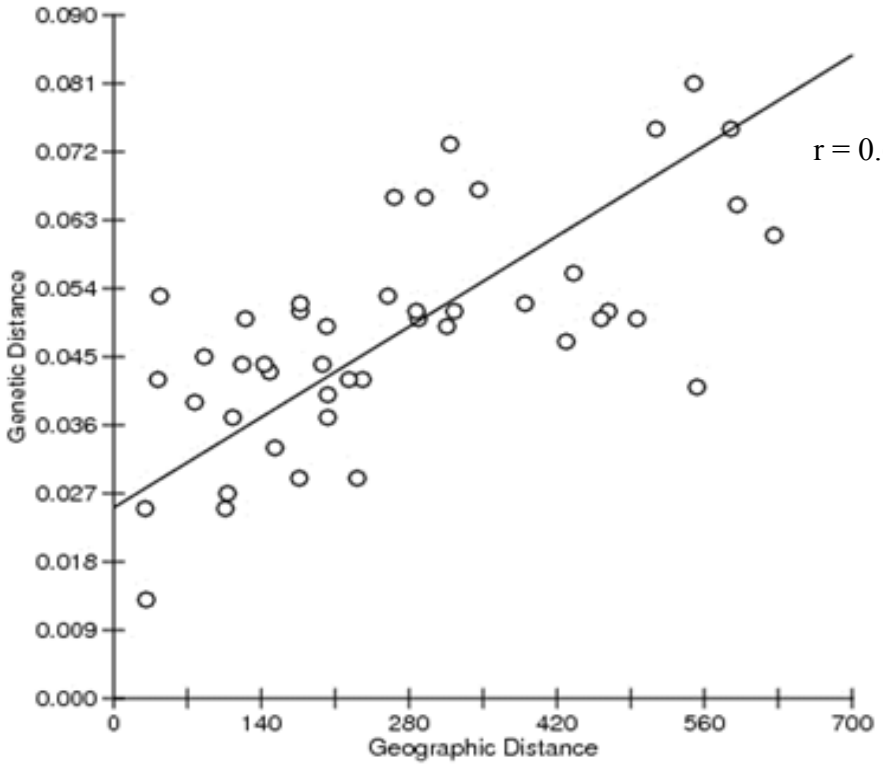


Fig. 3-1: Correlation of genetic distance (F_{ST}) with geographic distance (Km) for all possible pairs of the 10 surveyed populations of *B. papyrifera* in Ethiopia.

Population structure

Consistent with the low F_{ST} and the fact that most diversity is present among sampled trees within populations, a PCA plot revealed somewhat differentiated populations, with a large overlap of individuals from different populations (Fig.3-2).

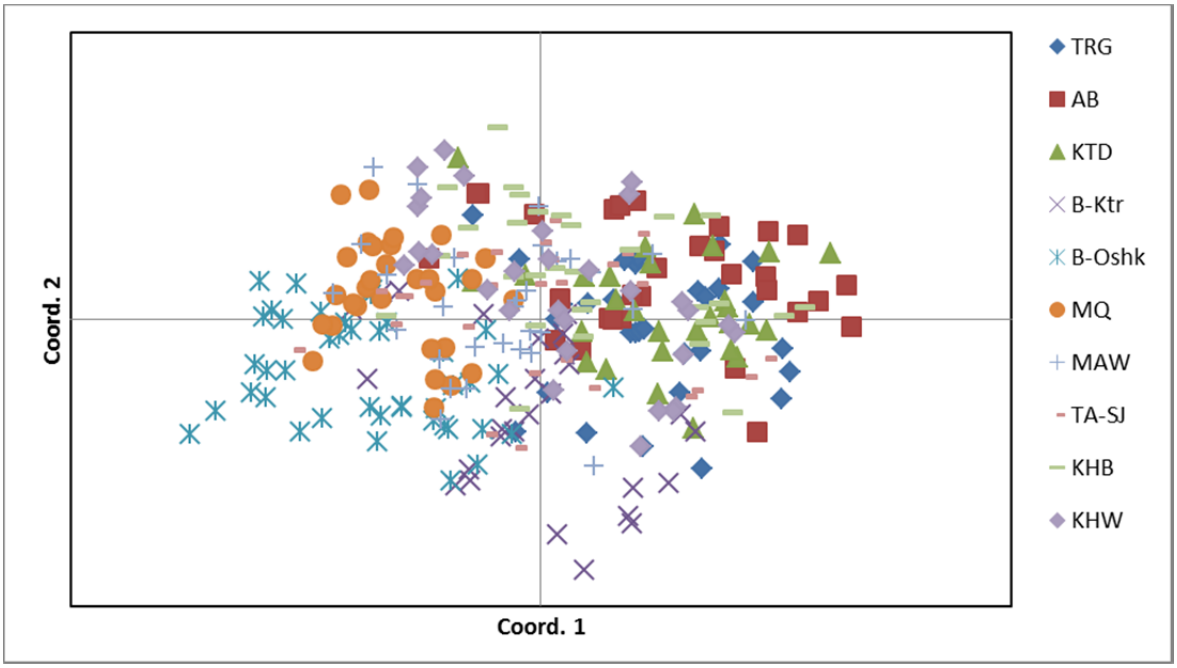


Fig. 3-2: Principal Component Analysis (PCA) of the 311 *B. papyrifera* individuals from 10 populations based on genotyping with 10 microsatellite loci.For population abbreviations see Table 3-1.

With regard to the analysis of population structure, the most likely number of clusters (K) was 4 (not shown). The STRUCTURE analysis identified two clusters containing 3 populations and two clusters each with two populations (Table 4-5). Two of the four clusters were clearly distinct while the other two showed individuals with mixed posterior probability of cluster membership. One of the populations (TA-SJ) was assigned to the cluster with low (45.9%) posterior probability (Table 5). These four genetic clusters corresponded to different geographical regions but also to different environmental conditions. The clusters are hereafter referred to as North Eastern (NE), Western (W), North Western (NW) and Northern (N) clusters based on the geographical locations of the accessions (Fig. 3-3). The analyses placed TRG, AB, and KTD within one cluster (NE cluster). The maximum geographic distance among these population was 85 km. B-KUR and B-SHKwere grouped together as one cluster (W cluster) and these populations were 44 km apart. MQ and MAW corresponded to a single cluster (NW cluster).MQ and MAW were only 30 km

apart. TA-SJ and the populations from the Northern part of the country (KHB, KHW) were assigned together (N cluster). Although TA-SJ is located more than 100 km from KHB and KHW, it was assigned to this cluster but with a posterior probability below 0.5 (0.459), and the probability for the next cluster, NW, was 0.339.

Table 3-5: Proportion of ancestry (calculated as average membership coefficient of all the individuals of a population) in each of the four *B. papyrifera* genetic clusters. Grey colour represent the four genetic clusters.

	Populations	Average proportion of ancestry in each of the four clusters				Genetic clusters in which populations assigned
1	TRG	0.776	0.026	0.158	0.041	NE
2	AB	0.823	0.015	0.100	0.062	
3	KTD	0.852	0.018	0.077	0.052	
4	B-KUR	0.134	0.590	0.222	0.054	W
5	B-SHK	0.026	0.875	0.060	0.038	
6	MQ	0.043	0.047	0.821	0.090	NW
7	MAW	0.057	0.072	0.617	0.253	
8	TA-SJ	0.142	0.060	0.339	0.459	N
9	KHB	0.125	0.018	0.066	0.790	
10	KHW	0.089	0.043	0.104	0.764	

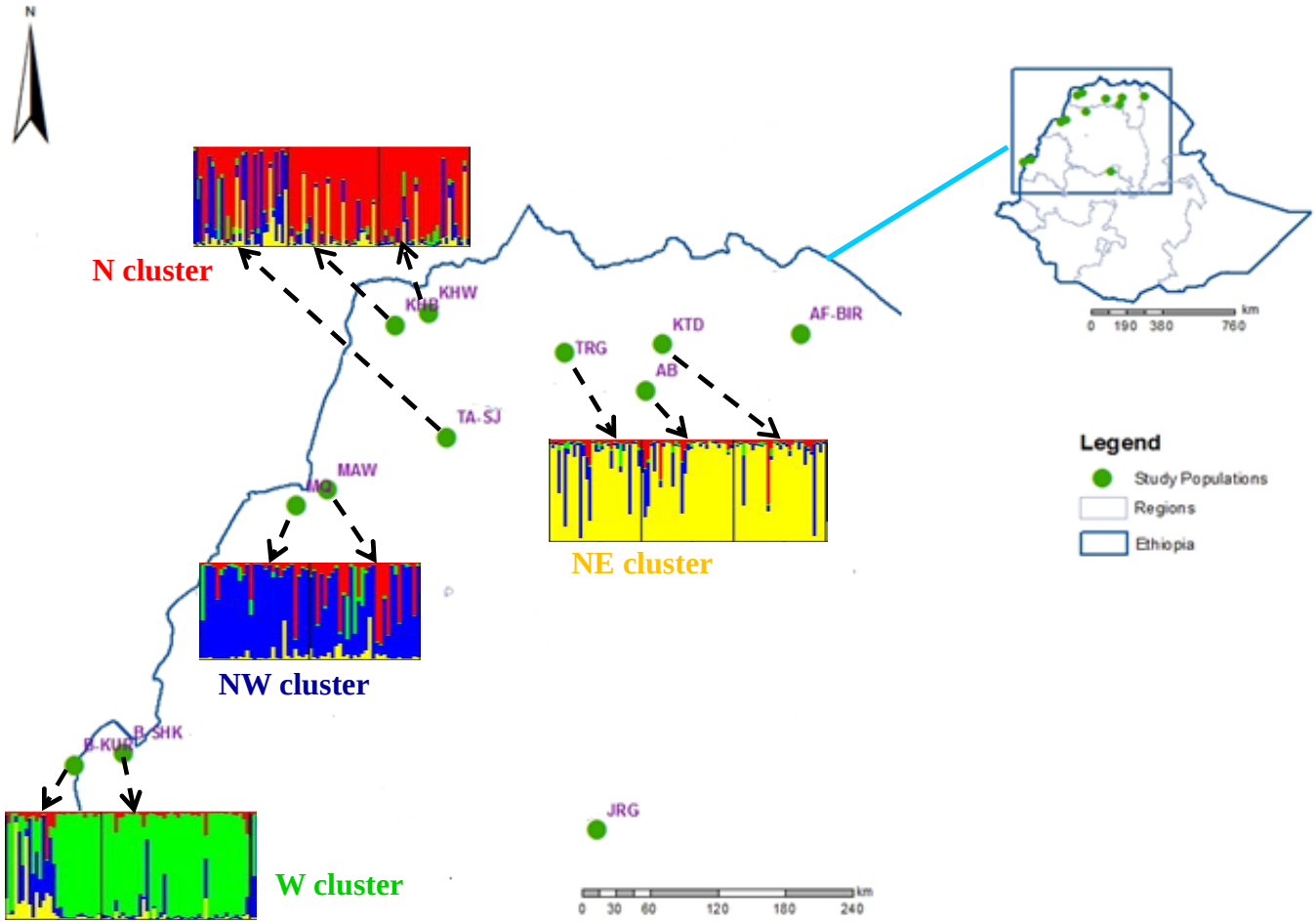


Fig. 3-3: Geographic distribution of the four genetic clusters of *B papyrifera* in Ethiopia. Clusters are identified based on STRUCTURE analysis. Each arrow indicates one population in a cluster. The cluster is subdivided into populations with the black solid line. AF-BIR and JRG were not included in the analysis. For population abbreviations see Table 3-1.

Discussion

Intra-population genetic diversity of *B. papyrifera* in Ethiopia

The present study examined the genetic diversity and pattern of population differentiation in the frankincense tree *B. papyrifera* in Ethiopia. The populations of *B. papyrifera* have declined tremendously in size, are fragmented and degraded, and many populations in the North and Northwest of Ethiopia have no recruitment of young saplings at all (Eshete et al., 2011; Tolera et al., 2013).

A small population size and decreased connectivity (gene flow) may erode the populations' genetic variability and drive strong spatial genetic structuring (Schaal et al., 1998; Kramer, 2008). In this study, however, we detected ample genetic variation ($N_A = 8$) and no inbreeding ($H_O = 0.669$, $H_E = 0.681$, $F_{IS} = 0.01$) in the adult trees of the populations we sampled from various parts of Ethiopia, including those from the most degraded populations. The heterozygosity was comparable to estimates obtained for other tropical tree species for example *Acacia Senegal* (mean $H_O = 0.709$, Omondi et al., 2010), one of the Terminalia-Combretum woodland tree species co-occurring with *B. papyrifera* in Ethiopia. One of the *B. papyrifera* populations (KTD) consisted of only remnant trees scattered on farmlands, yet it had the second highest heterozygote in the study ($H_O = 0.701$). In *B. papyrifera* the intraspecific genetic diversity apparently is still present in the mature trees. This is probably due to the long life span of individual trees (Smulders et al., 2008) in combination with the fact that the ongoing forest fragmentation and degradation processes in Ethiopia are relatively recent events. We do not know whether *B. papyrifera* tolerates selfing, but the main threat to the populations is one of recruitment (Eshete et al., 2011; Tolera et al., 2013).

Forests in Ethiopia were still unaffected by man in 1900 (EFAP, 1994; Reusing, 2000). Lemenih et al. (2007) related the degradation of these dry tropical forest ecosystems to continuous human settlement, which was intensified since the late

1960s. Tolera et al. (2013) studied tree rings and estimated that the mature trees in existing *B. papyrifera*forests established themselves from seeds between 1903 and 1955, but not later. Their results implied an absence of regeneration during the last 60 years. Based on our population genetic data the *B. papyrifera* trees that we sampled still represent the diversity in their parents, which grew in the large and continuous forests that existed in the 19th century. A second positive note is that we discovered populations in the Western region (Benishangul-Gumuz) with a high density of seedlings and saplings, indicating that natural regeneration still takes place in these populations. This region has fewer agricultural activities and the forests are relatively undisturbed. Traditionally this low and hot region did not see cattle herding because of the occurrence of Trypanosoma. This situation is currently changing but the livestock population in the region of Assosa is still very low, and particularly in Boswellia woodlands no cattle occur, only a few goats, donkeys and camels (Addisalem, personal observation). Possibly the deeper soil and the higher amount of rainfall also contribute to good microsites for seedling development. This region has a much higher amount of rainfall (as much as 1350 mm annual precipitation in B-SHK) than the North of Ethiopia (Table 3-1; Eshete et al., 2011; Teshome, 2013; Alemu et al., 2015).

Population differentiation

The average F_{ST} across all the 12 populations, based on data from 5 markers, was 0.0109 ($P<0.001$) while it was a bit lower (0.084) when the two populations (AF-BIR and JRG) were excluded from the calculation. The higher F_{ST} among all 12 populations may be caused by geographical distance, isolation of the populations and/or adaptation to the environment. JRG is geographically the most distant while AF-BIR is isolated from other populations by mountains. The latter is also located in very harsh climatic conditions, namely a very hot (> 45 ^{o}C) area with rainfall as low as 300 mm. These populations would be interesting for a follow-up study. Sampling of young leaves, as we did here, turned out to be problematic as

rainfall was irregular and unpredictable. A possible alternative would be to sample bark tissue for DNA extraction (e.g., Novaes et al., 2009), but this needs to be tested for *B. papyrifera* as the exudate in the bark may hamper the extraction of good quality DNA.

The F_{ST} obtained was comparable to the genetic differentiation observed in the out-crossing tree species *Swietenia macrophylla* (F_{ST} = 0.097, Lemes et al., 2003) and in *Populus euphratica* (F_{ST} = 0.093, Wang et al., 2011), a dioecious, out-crossing and wind-pollinated desert tree species. The sampled *B. papyrifera* populations were at least 30 km apart (MAW and MQ in the North West). *B. papyrifera* is a species characterized by bright and fragrant flowers (Fitchl and Admasu, 1994). The flowers are frequently visited by honeybees for pollen and nectar (Fitchl and Admasu, 1994), hence honeybees are considered one of the pollinators of the species. Insect pollination (entomophily) was confirmed in two closely related species, *Boswellia serrata* (Sunnichan et al., 2005) and *B. ovalifoliolata* (Raju et al., 2012), while wind pollination was ruled out in *B. serrata*. The seed dispersal mechanism of the species is not well known. In *B. ovalifoliolata,* a closely related species, the seeds disseminate up to a distance of 400 m (Raju et al., 2012) indicating limited seed dispersal. For *B. papyrifera* such a seed dispersal distance may not be very effective to genetically connect populations. Hence, the isolation-by-distance pattern of genetic differentiation across populations likely is the consequence of pollen dispersal by insects.

In its growing regions the species is also distributed along river basin networks, with occasional chances for long distance seed dispersal along the traversing rivers. Likewise, *Populus nigra*, a riparian species in Europe, had a F_{ST} of 0.081 (based on microsatellites) among river systems (Smulders et al., 2008). However, the contribution of water to seed dispersal was lower than anticipated in many

riparian species. We do not know whether water is an effective seed dispersal agent for *B. papyrifera.*

There was a clear pattern of isolation-by-distance (IBD), which explains 44% of the variation among populations. The findings support our prediction that *B. papyrifera*populations are spatially structured as the result of their geographical isolation. However, genetic differentiation may also be caused by genetic drift and natural selection (Ledig and Fryer, 1972; Slatkin, 1987; Duminil et al., 2013). The individual-based STRUCTURE analysis resolved the 10 populations into four genetically differentiated clusters that correlated with the geographical locations of the sampled populations. The four clusters, North eastern (NE), Western (W), North western (NW) and Northern (N), also correspond to environmentally different conditions in terms of precipitation and soil conditions. The NE populations grow in regions characterized by higher altitude (above 1000 m a. s. l.), modest levels of rainfall (mean annual precipitation 850 mm), cooler temperatures (average maximum temperature 29.3 ^{o}C) and less fertile, clayey and shallow (15 cm deep) soil (Eshete et al., 2011). The W populations are located at lower altitude (below 900 m a. s. l) in hot (maximum temperature of 39.42 ^{o}C) but moist areas that receive up to 1350 mmannual precipitation. The soil is deep red, and clayey to sandy-loam in texture (Teshome, 2013). The NW populations are located at lowland altitudes (below 900 m a. s. l.), which are also hot (around 36 ^{o}C) and receive modest precipitation (800-900 mm) annual precipitation. The soil is deeper (27 cm), more fertile and clayey in texture (Eshete et al., 2011). Finally, the N populations are located in areas that are dry (less than 650 mm precipitation) and hot (maximum temperature 41 ^{o}C) (Alemu et al., 2015). Hence, the population genetic differentiation detected in *B. papyrifera* may differentiate populations that are also adapted to local environmental conditions, caused by precipitation, temperature, altitude and soil conditions or combinations of these factors. This

needs to be confirmed with further studies such as common garden experiments or provenance studies.

Conservation of populations *in situ* and *ex situ*

High levels of genetic variation are still presentin the adult trees in *B. papyrifera* populations across the species' range in Ethiopia. Yet, the demographic threats should not be discounted because most of the extant populations of *B. papyrifera* are now degraded in terms of their size, spatially fragmented, and they lack recruitment of new generations of trees (Eshete et al., 2011; Groenendijk et al., 2012; Tolera et al., 2013).

Sampling for conservation and *in situ* management of genetic resources should ensure the majority of extant variation is preserved. Conserving among-population diversity should focus on reserving the most genetically distinctive populations whilst conserving within-population diversity should preserve large core populations that will not lose diversity due to genetic drift (Namkoong, 1988). In *B. papyrifera*, therefore, selection of populations for conservation should take into account the geographical distance among the populations, but also the existence of distinct clusters, (NE, W, NW and N) and the fact that they grow in different environmental conditions. Within each cluster, populations of conservation priority have to be selected taking into account the level of genetic diversity, the regeneration status of the populations and the level of threat populations are experiencing. Delineating large reserves with more than one population within each cluster would ensure a sample of a gene pool that captures the diversity and uniqueness existing in the standing populations of the species.

Our findings suggest that the conservation value of degraded adult *B. papyrifera* populations under human influenced landscapes should not be overlooked, because such populations (e.g. KTD) still maintain genetic diversity comparable to

the genetic diversity of the continuous forests that once existed. However, the *in situ* or *circa situm* (conservation in human influenced landscapes)management of such depleted populations needs active enrichment planting, not only within the current populations to replace dying old trees, but also in areas between the current populations, to improve and facilitate the genetic connectedness of individuals and populations, and thereby to ensure that the genetic diversity is maintained to the future.

Although only scarce genetic data were obtained so far, the population at Afar (AF-BIR), because of its geographical isolation, the depletion of the populations, and its ability to grow in the harsh environmental conditions (Table 3-1) may deserve special attention of conservation measures. This population is marginal in terms of microclimate (located in the hottest place on earth, the Danakil depression in Ethiopia, which barely supports vegetation). The effects of climate change in the region, namely a rise in temperature and declining and unpredictable rainfall (Joanne et al., 2005; NMA, 2007; EPA, 2010), on the population dynamics (e.g. regeneration, population expansion and shift in ecological range) of the species is not known. However, it is expected that such trends of climate change may result in expansion of drought in arid areas such as Afar-Bir Hale, suggesting increasing risk of population degradation and subsequent risk of local extinction. In view of this, it is highly advisable to preserve this population *ex situ* in laboratory and field gene banks, additionally to more common *in situ* conservation. The same may account for the quite isolated JRG population.

The close correlation of the genetic clustering to very diverse environmental conditions suggests the need of establishing provenance trial plantations in each of the four sites, both to preserve all available local and regional genetic variation *ex situ,* and to use these as source populations for future breeding programs. Urgent

action is needed if people keep valuing frankincense as an important commodity, both for local use as for the international market.

Acknowledgements

This research is supported by Netherlands Fellowship Program (Nuffic) and the Dutch-Ethiopian FRAME programme 'FRAnkincense, Myrrh and arabic gum: sustainable use of dry woodland resources in Ethiopia', funded by the Netherlands Foundation for the Advancement of Scientific Research in the Tropics (NWO-WOTRO, grant W01.65.220.00). We thank the district Agricultural offices of all the study regions for allowing us to undertake the research in the area. We are also grateful to their staff members for their technical support during the fieldwork and sampling. We thank the Ethiopian Biodiversity Institute (EBI) for permitting us to transfer the plant material needed for the research to the Netherlands

CHAPTER 4

FINE-SCALE GENETIC STRUCTURE IN THE FRANKINCENSE TREE Boswellia papyrifera AND IMPLICATIONS FOR CONSERVATION

A.B. Addisalem, J. Duminil, D. Wouters, F. Bongers, M.J.M. Smulders

Accepted (with minor revisions) for publication in
Tree Genetics & Genomes (TGG)

Abstract

The fine-scale genetic structure and how it varies between generations depends on the spatial scale of gene dispersal and other fundamental aspects of species' biology, such as the mating system. Such knowledge is crucial for the design of genetic conservation strategies. This is particularly relevant for species that are increasingly fragmented such as *Boswellia papyrifera*. This species occurs in dry tropical forests from Ethiopia, Eritrea and Sudan and is an important source of frankincense, a highly valued aromatic resin obtained from the bark of the tree. This study assessed the change in genetic diversity and fine-scale spatial genetic structure (FSGS) between two cohorts (adults and seedlings) from two populations (Guba-Arenja and Kurmuk) in the west of Ethiopia, and inferred intra-population gene dispersal in the species, using microsatellite markers. The mean expected heterozygosity per cohort ranged between 0.664 and 0.724 with very low mean F (0.016). The spatial analyses based on kinship coefficient(F*ij*) revealed a significant positive genetic correlation up to a distance of 62 m. Spatial genetic structure was relatively weak (*Sp* = 0.0023–0.0137) indicating gene dispersal is extensive within the populations. Historical gene dispersal distances of 103and 124 m were estimated from FSGS patterns for Guba-Arenja and Kurmuk populations, respectively. The high heterozygosity, the low fixation index andthe low *Sp* values found in this study are consistent with outcrossing as the (predominant) mating system in *B. papyrifera.* We suggest that seed collection for *ex situ* conservation and reforestation programmes of *B. papyrifera* should use trees separated by distances of 100 m but at least over 60 mto reduce genetic relatedness among seeds from different trees.

Keywords: Fine-scale genetic structure, gene dispersal, spatial autocorrelation, gene flow, conservation

Introduction

The spatial genetic structure (SGS) of plant populations is determined by various processes including gene flow and local selection at different life-history stages. In plant species, gene dispersal is mediated by pollen and seed and along with other factors it influences how genetic diversity is structured within and between populations. Seed movement allows migration and range expansion, while the spatial scale of gene flow via both seed and pollen has important implications for the maintenance of genetic diversity (Peakall et al., 2003; Moran and Clark, 2011). A limited dispersal will enhance population differentiation, and in the long term it may hinder the ability of a species to colonize new sites or to shift the range in response to environmental changes (Moran and Clark, 2012).

At a fine spatial scale, the genetic similarity is in general higher among neighbouring than among more distant individuals. A local genetic structure can develop due to limited gene flow, generally when seed-mediated gene flow is low even in a large continuous population (Vekemans and Hardy, 2004; Segelbacher et al., 2010; Moran and Clark, 2012). The structure is reinforced in subsequent generations due to bi-parental inbreeding and genetic drift if pollen movement and seed dispersal are limited (King et al., 2012; Moran and Clark, 2012). On the other hand, extensive dispersal and immigration would weaken FSGS over generations (Peakall and Smouse, 2008). The strength of FSGS is expected to increase in low density, as opposed to high density populations (Vekemans and Hardy, 2004). It is significantly related to the breeding system and is higher in selfing species than mixed mating and outcrossing species (Vekemans and Hardy, 2004).

Knowledge of the scale of gene dispersal within the populations of the species is crucial information for the design of a conservation strategy that will be effective for the maintenance of genetic diversity of the species. This is particularly relevant for species that are increasingly fragmented, such as

Boswellia papyrifera. A comparison of the genetic structure of trees from different cohorts (across generations) identifies the temporal changes in genetic structure and (genetic) diversity (Chibicki and Burczyk, 2010b). The genetic diversity and FSGS among adults reflect the historical population diversity and structure, whereas the effects of the threats will be detectable in the juvenile life stages (seedlings/saplings). Therefore, comparing the FSGS of trees across different generations can also shed light on the different demographic and genetic processes in the formation and maintenance of the genetic structure.

B. papyrifera is an important source of frankincense, a valuable non-timber forest product (NTFP) due to its use as raw material for industries such as perfumeries and pharmaceuticals. Frankincense is one of the export commodities of Ethiopia, Sudan and Eritrea (Coppen, 2005; Lemenih, 2005; Ogbazghi et al., 2006; Abtew et al., 2012).

In many areas in Ethiopia, the natural stands of *B. papyrifera* now consist of fragmented patches of forest because of land-use changes, grazing, fire and adult mortality. Populations in Northern (Abergele) and Northwestern (Metema) Ethiopia are suffering from lack of saplings and small trees (Abiyu et al., 2010; Eshete et al., 2011; Girma et al., 2013; Tolera et al., 2013). In these *Boswellia* populations, non-permanent seedlings (seedlings of which the shoots dieback during the dry season and reappear in the rainy season) are abundant but do not grow up to the sapling stage mainly because of the impacts from frequent forest fire and intensive grazing in the area (Groenendijk et al., 2012;Tolera et al., 2013). Recruitment failure is also occurring in most of the other populations across Ethiopia. A similar situation was reported from Eritrea (Ogbazghi et al., 2006, Rijkers et al., 2006) and Sudan (Adam and El Tayeb, 2008; Abtew et al., 2012). These findings outline that *B. papyrifera* populations are threatened across its whole distribution area. *Boswellia sacra* in Oman (Farah, 2008), and both *B. serrata* (Sunnichan et al., 2005) and *B. ovalifoliolata* (Raju et al., 2012) populations in India also show regeneration bottlenecks. During our fieldwork in Western Ethiopia we discovered populations that did have a

young cohort of seedlings and saplings next to the mature, flowering trees (Addisalem et al., 2016). These unique, naturally regenerating populations enabled us to study the FSGS of the species, for seedling and adult cohorts. This study aimed at assessing: (i) differences in genetic diversity, if any, between the parent and progeny cohorts, (ii) the FSGS within two cohorts in each of two populations, and (iii) the historical gene dispersal patterns at local scale as inferred from FSGS patterns. The results are discussed in view of an urgent need to define strategies of conservation for this emblematic species.

Materials and Methods

Study site and population description:

B. papyrifera (Burseraceae) is a monoecious tree species with sweet-scented flowers, frequently visited by honeybees for pollen and nectar (Fitchl and Admasu, 1994), indicating honeybees as one of its pollinators. Two closely related species, *B. serrata* and *B. ovalifoliolata*, are also insect pollinated (Sunnichan et al., 2005; Raju et al., 2012). Fruits of *B. papyrifera* are about 2 cm long, usually containing three tapered seeds (Vollesen, 1989). The seed dispersal mode of *B. papyrifera* is not well known. Seeds of *B. Ovalifoliolata* disseminate up to a distance of 400 m (Raju et al., 2012).

The study was conducted in the Western part of Ethiopia, which is the only place, as far as we know, in the species' distribution range in Ethiopia where *B. papyrifera* regeneration and stand recruitment amply occurs. After a preliminary survey of the distribution of the species within the region, two populations, Guba-Arenja located at 11.05N, 35.16E and Kurmuk 10.56858N, 34.297E (Fig. 4-1), which are both continuous populations, were selected for this study. The two populations were treated as 'blind' spatial replications for the FSGS analysis of the species. All four life cycles/size classes described in Groenendijk et al. (2012) were present in the structure of these two populations: seedlings (0-2.0 cm Root Collar Diameter (RCD)), saplings (RCD >2 cm and < 6 cm Diameter at Breast Height (DBH)), juvenile trees (individuals with DBH 6-10 cm) and

adult trees (>10 cm DBH). The relative abundance of the different life cycles/size classes, however, varied between the two populations.

In Guba-Arenja, the sampled area was a strip running along a slope from the top of a hill (highest altitude = 764 m) downhill (lowest altitude = 720 m), where it was embedded with a seasonal river-course. It contained a patchily distributed population of *B. papyrifera* intermittent with grasslands. Recently a highway had been constructed across the slope. This population was characterized by patches of abundant seedlings and saplings, and a few small and big-sized adult trees. Unlike Groenendijk et al.'s (2012) observations in Metema population, we observed that several small *B. papyrifera* individuals as tall as only 1.5 m were already flowering in this population.

The Kurmuk population was located in a relatively flat area with altitude 662–680 m surrounded by a hilly area in the west and northwest directions. The *B. papyrifera*population in this area was characterized by presence of abundant seedlings, saplings, small trees, and a few scattered big adult trees. Spatially, individuals were uniformly distributed.

A comparison of the genetic structure of trees from different cohorts (across generations) identifies the temporal changes in genetic structure and (genetic) diversity (Chibicki and Burczyk, 2010b). In this study the FSGS pattern of two life stages, adult and seedling cohorts, were investigated within each population. For the purpose of this study, we considered all mature trees presumably capable of producing pollen and seeds, (i.e., trees that were greater than ~10 cm DBH) as "adults" and they were all considered as potential parents. Our "seedling" cohort consisted of small plants up to a height of ~ 50 cm. These "seedlings" are the youngest plants present in the population but they may be up to several years old already (Birhane et al., unpublished data). Their height distinguishes them from larger sapling and juveniles (not-yet

reproducing plants), and allowed us to get the largest contrast in plant age and developmental stage possible within the populations.

Sampling individuals and plant material

We sampled 106 adult trees and 105 seedlings with young leaves from Guba-Arenja (22-hectare plot) and 80 adult trees and 60 seedlings from Kurmuk (36-hectare plot). The difference in number of samples was due to the limitation of trees bearing young leaves during the sample collection. Trees and seedlings that did not bear young leaves were not sampled. All samples were georeferenced. Samples of young leaves were immediately stored in silica gel, which was regenerated regularly until the samples were dry. The adult (considering trees >10 cm DBH) density in Guba-Arenja was 58 individuals per hectare, whereas it was 87 trees per hectare in Kurmuk (Teshome, 2013; Teshome et al., in prep.). The geographical distribution of sampled adult and seedling individuals presented in Fig.4-1.

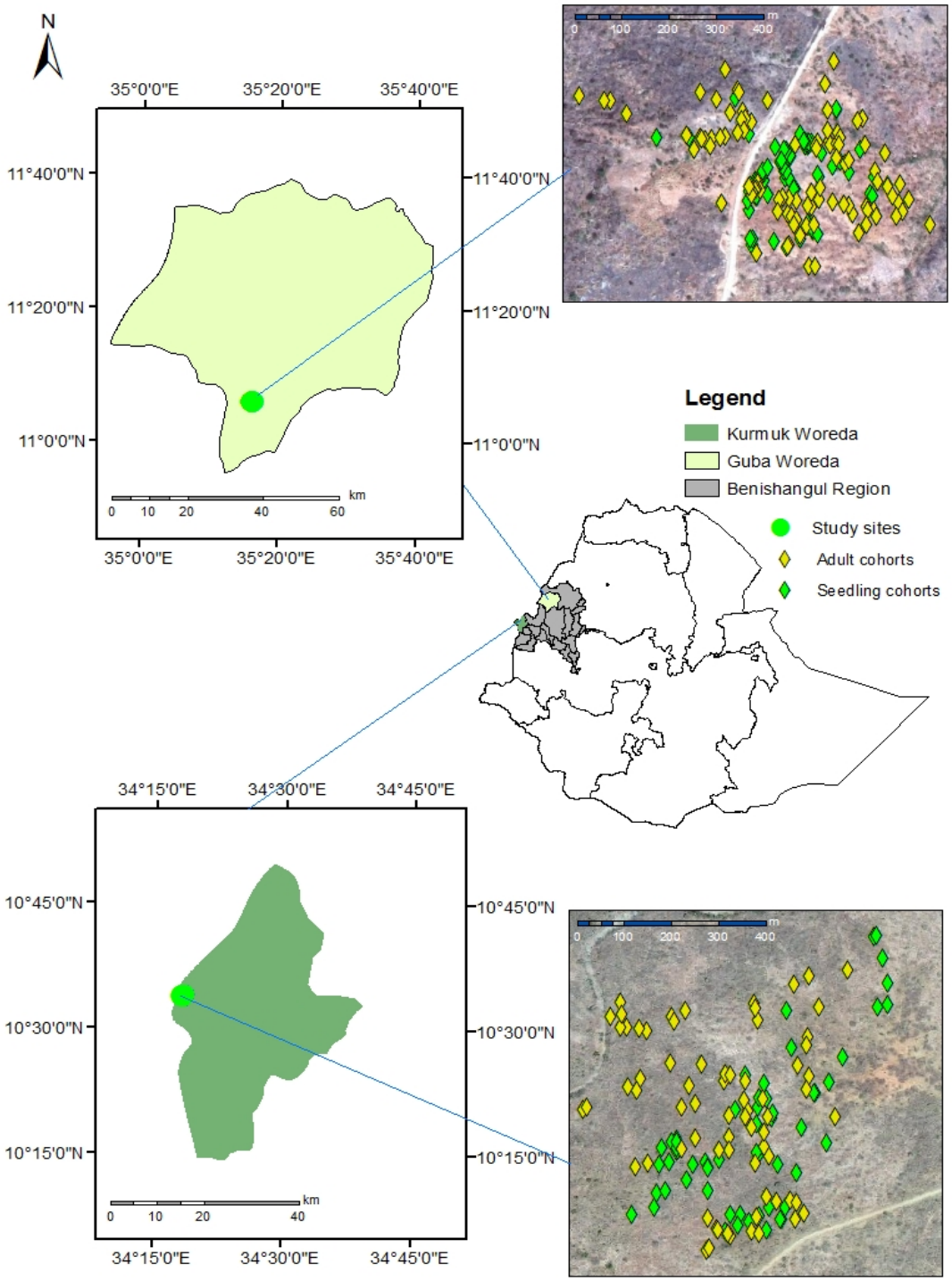


Fig. 4-1: Study sites and geographical distribution of sampled adult (yellow) and seedling (green) individuals in the study sites in Guba (upper panel) and Kurmuk (lower panel) districts, Western Ethiopia. The two sites are 108 km apart.

Laboratory analyses of the samples

DNA extraction

Total DNA was isolated from dried young leaves following the CTAB protocol. As large amounts of phenolic compounds were expected because of the resin content in the leaves, the protocol was modified by the addition of 2% pvp-40 in the extraction buffer and 1% mercaptoethanol in the microprep buffer following the methods described in Fulton et al. (1995). Purification was performed using DNeasy (Qiagen, Venlo, the Netherlands) procedures following Smulders et al. (2010). DNA yield and quality were visually assessed on 1% agarose gels.

PCR and genotyping

The samples were genotyped with six polymorphic microsatellite loci, namely Bp17, Bp20, Bp21, Bp22, Bp23 and Bp39 (Table 1) developed previously for this species (Addisalem et al. 2015). The loci were selected based on their level of polymorphism in terms of the total number of alleles (10-28 alleles per locus) in a previous population structure study (Addisalem et al., 2016) that included 12 populations across Ethiopia. Descriptions of the primers, the PCRs reaction mix, labeling of the primers and the thermal cycling profile of the PCRs are detailed in Addisalem et al. (2015). Forward primers were labelled with IRD700 or IRD800 fluorescent dye. After amplification the mixture was diluted with 10 µl water, and 5-6 µl of fluorescently labelled PCR products was resolved on a 4200 or 4300 Licor DNA analyzer. The gel was reloaded once or twice, as long as the resolution was sufficient, otherwise they were (re) analyzed on a new gel. The bands were scored manually as a codominant marker. The length of the alleles was estimated using Sequamark 10 bp DNA step ladder.

Data analyses

Genetic diversity analysis

Genetic diversity was estimated separately for the two populations and for the two cohorts (adults and seedlings). The total number of alleles (N_A), the effective number of alleles (N_E), the observed (H_O) and expected heterozygosity

(H_E) and the fixation index (F) were calculated for each population and each cohort using GenAlEx 6.5 (Peakall and Smouse 2006, 2012). Genetic differentiation between the populations from the two sites (seedling and adult cohorts combined) and between cohorts within each population (adults versus seedlings) was analysed based on F_{ST} estimates.

Fine-scale Spatial Genetic Structure

FSGS can be described by computing pairwise relatedness coefficients between individuals in the sample and analysing their relationship with the spatial distance separating the individuals (Vekemans and Hardy, 2004). In this study we characterized the FSGS using the kinship relatedness coefficient (F_{ij}) (Hardy & Vekemans, 1999) and the autocorrelation coefficient (r) (Peakall et al., 2003). Analyses were based upon pair-wise genetic and geographic distance matrices calculated using data from six microsatellite markers and field GPS records of sampled plants. Variable distance class sizes (20, 40, 80, 160, 300, 500 and 850 m) were specified so that the first distance interval would calculate relatedness coefficientsbased on all pairwise comparisons within a distance of 0–20 m, the second analysis for 20–40 m, and so on until the total spatial distance was covered. We used the same distance classes in both populations and both cohorts to be able to compare the results between the two sites and between the two cohorts within a population.

The FSGS pattern based on the kinship or co-ancestry coefficient (F_{ij}), i.e. the probability that a random gene from individual *i* is identical to a random gene from individual *j* (Vekemans and Hardy 2004), was analysed using SPAGeDi 1.4b (Hardy and Vekemans 2002). The slope of regression ($\hat{b_F}$) (kinship coefficient regressed over logarithmic geographic distance) and the historical gene dispersalparameter (σ) were estimatedusing the same programme. The significance of the regression slope $\hat{b_F}$ was tested against the null hypothesis (the overall absence of FSGS, $\hat{b_F}$= 0) by comparing the observed values with those obtained after 1000 random permutations of individuals among positions.

The extent of the FSGS was quantified using the *Sp* statistic (*Sp*), calculated as - $\hat{b_F}$ / (1- $\hat{F_1}$) described in Vekemans and Hardy (2004), where $\hat{b_F}$is the regression slopeand $_{F1}$is the mean F_{ij}between individuals belonging to the first distance class containing adequate pairs of individuals to precisely estimate the $F_{ij.}$The neighbourhood size and the gene dispersal distance σ_g were jointly estimated for each population relying on SGS patterns following the iterative procedure described in Hardy et al. (2006). The method posits that F_{ij} is expected to decay linearly with the ln(distance) at a rate inversely proportional to the product $D_E\sigma_g^2$ for a distance range between σ_g and ca. 20 σ_g, where D_E is the effective density of reproductive individuals and σ_g^2 is the axial variance of gene dispersal distance between two generations. The joint estimate is sensible to the quality of D_E (Hardy et al., 2006). Accordingly, different values of D_E were tested, considering ½, ¼ and 1/10 of the adult densities to account for the lifetime variation in reproductive success among adult trees (Hardy et al., 2006).

The pattern of FSGS is affected by the mating system, gene dispersal as well as ecological factors; hence different situations can be expected to yield different autocorrelation patterns. We applied the nonparametric heterogeneity test described by Smouse et al. (2008), implemented in GenAlEx 6.5, to test the statistical significance of the spatial autocorrelation patterns (analogues to the pattern based on kinship coefficient) observed between the two populations and the two cohorts within each of the populations. The method for calculating autocorrelation coefficient (r) is described in Peakall and Smouse (2006, 2012). The statistical significance of r was determined using 999 permutations, randomizing genotypes among distance classes. Within a specific distance class, spatial genetic structure was considered significant when the observed r-value fell outside of the 95% CI and when the error bar did not intersect r = 0 (Peakall and Smouse 2006, 2012). Single-class (t^2) and multi-class test statistic (ω) were computed with 999 permutations.

Results

Genetic diversity

Across the six microsatellite loci and the 330 genotypes (adults and seedlings in both populations) 98 alleles were detected in total with an average of 16.3 alleles per locus. The number of alleles per locus (A_L) ranged from 5 (at locus Bp39) to 30 (at locus Bp22). The common alleles in these populations were those found in our previous studies that included 12 populations from across Ethiopia (Addisalem et al., 2016), but more alleles were detected for some of the loci (e.g. 30 alleles at Bp22) and fewer for other loci (e.g. 5 alleles at Bp39). The mean H_O and H_E across all cohorts and across the two populations were 0.689 and 0.693 respectively (Table 4-1). MeanH_O and $H_E$were higher (H_O = 0.819and H_E = 0.825) when the locus with the fewest number of alleles (Bp39) was excluded from the estimation. The intra-population fixation index (F) ranged from -0.03 in Kurmuk Adults to 0.07 in Kurmuk seedlings (Table 4-2) with an average of 0.016 for all populations. Examination of genetic differentiation between the two study sites, Guba-Arenja and Kurmuk (adult and seedling cohorts in each population combined), resulted in low genetic differentiation (F_{ST} = 0.038). The pair-wise F_{ST} between adult and seedling cohorts were 0.010 in Guba-Arenja and 0.004 in Kurmuk population.

Table 4-1: The six microsatellite loci (from Addisalem et al. 2015) used for the FSGS analysis and gene dispersal study for two populations of *Boswellia papyrifera* in western Ethiopia. N, number of individuals screened; A_L, total number of alleles scored per locus; H_O, observed heterozygosity; H_E, expected heterozygosity; F_{IS}, fixation index; SE, standard error.

Loci	N	A_L	H_O (SE)	H_E (SE)	F_{ST}	F_{IS}
Bp17	77	9	0.706	0.708	0.075	0.003
Bp20	73	12	0.808	0.762	0.043	-0.062
Bp21	81	19	0.861	0.855	0.030	-0.006
Bp22	80	30	0.818	0.889	0.021	0.080

Bp23	81	23	0.870	0.843	0.035	-0.031
Bp39	79	5	0.073	0.103	0.028	0.285
Mean (SE)	76(4)	16.3	0.689 (0.025)	0.693 (0.029)	0.038 (0.008)	0.045 (0.052)

Genetic diversity between generations/cohorts

The number of alleles was the same in both cohorts within each of the populations (Table 4-2). In Guba-Arenja, the total number (N_A) was 14 and the effective number (N_E) was 7 in both adult (GA) and seedling (GS) cohorts. In Kurmuk slightly fewer alleles (N_A = 10 and N_E = 4) were detected in both cohorts. The heterozygosity was 0.707 and 0.724 respectively in adult (GA) and seedling (GS) cohorts in Guba-Arenja population. Similarly, the heterozygosity slightly increased from 0.664 to 0.677 from adult (KA) to seedling (KS) cohorts in Kurmuk population (Table 4-2).

Table 4-2: Genetic diversity parameters of *B. papyrifera* in Guba-Arenja and Kurmuk populations, West Ethiopia. N, the number of screened individuals; N_A, total number of alleles; N_E, effective number of alleles; H_O, observed heterozygosity; H_E, expected heterozygosity; F, Fixation index; numbers in parenthesis, standard error.

Cohort	N	N_A	N_E	H_O	H_E	F
Guba-Arenja adults	92 (1)	14(4)	7(2)	0.701(0.139)	0.707(0.140)	0.01(0.01)
Guba-Arenja seedlings	91(1)	14(3)	7(2)	0.701(0.136)	0.724(0.139)	0.02(0.029)
Kurmuk adults	65(3)	10(2)	4(1)	0.678(0.110)	0.664(0.110)	-0.03(0.053)
Kurkum seedlings	55(1)	10(2)	4(1)	0.675(0.121)	0.677(0.098)	0.07(0.105)

Fine-scale Spatial Genetic Structure (FSGS)

The kinship coefficient (F_{ij}) and the autocorrelation coefficient (r) showed a similar FSGS pattern. Here, we will first present the SPAGeDi results on kinship and the extent of SGS (*Sp*) analysis in detail, and then report the results from GenAlEx on the cross-population heterogeneity test.

Kinship analysis and extent of FSGS

An overall positive signal of local spatial genetic structure was detected in the two *B. papyrifera* populations except in the Kurmuk seedling cohort. A significant genetic relatedness (F_{ij}) was detected up to 62m, 61m and 31m respectively in Guba-Arenja adult (GA), Guba-Arenja seedling (GS) and Kurmuk adult (KA) cohorts (Fig. 4-2A, Fig. 4-2B and Fig. 4-2C). In Kurmuk seedling (KS) $F_{ij}$was not significant over the whole distance range (Fig. 4-2D). We also analysed adult and seedling cohorts combined in each population (GC and KC) for comparison at landscape level. The results showed significant positive SGS up to a distance of 62 m in GC (Fig. 4-2E) and no significant relatedness in KC (Fig. 4-2F).

The kinship coefficient over the shortest distance (F_1) was 0.0355 and 0.0291 respectively in GA and GS. In Kurmuk the estimates were 0.0776 for KA and 0.0050 for KS (Table 3). The regression slope ($\hat{b_F}$) of the mean kinship coefficients on the logarithm of spatial distance were negative and significant in GA, GS and KA all with whereas it was not significant in KS. The statistic for the extent of the spatial genetic structure (*Sp*) was 0.0124 ± 0.0058(SE) in GA, 0.0137± 0.0131(SE) in GS, 0.0109 ± 0.0038 (SE), in KA and 0.0023 ± 0.0036 (SE) in KS. Historical gene dispersal parameters (σ) estimated based on adults assuming effective densities (D_E) equal to D/2, D/4 and *D*/10 were respectively 46 ± 11.5, 64 ±14 and 103 ± (NA) m for Guba-Arenja and 65 ± 25, 81 ± 25, 124 ± (NA) m for Kurmuk populations (Table 4-3).

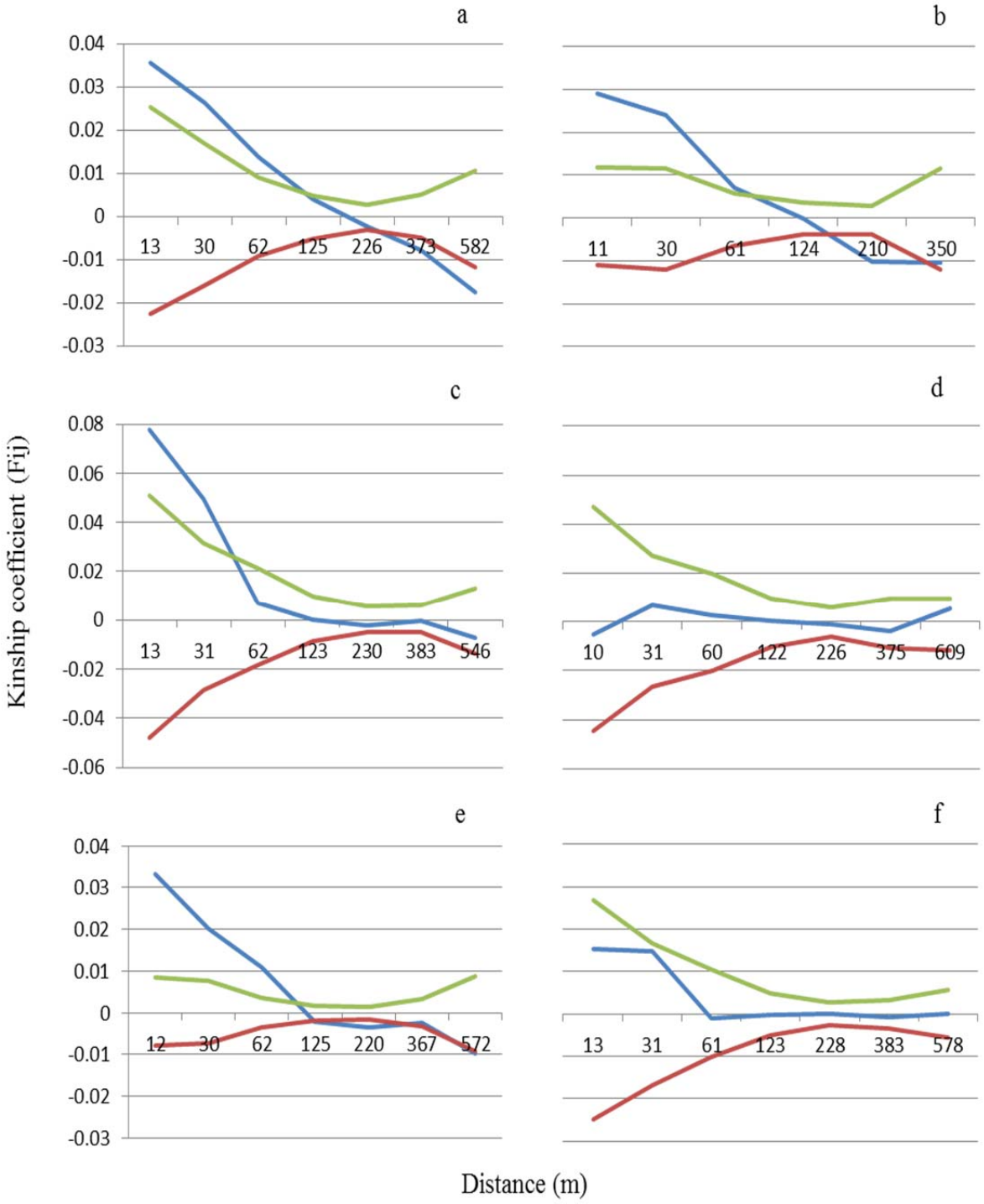


Figure 4-2: The kinship coefficient (*Fij*) plotted against geographical distances. (A) Guba-Arenja adult cohorts; (B) Guba-Arenja seedling cohorts; (C) Kurmuk adult cohorts; (D) Kurmuk seedling cohorts; (E) Guba-Arenja seedling and adult cohorts combined; (F) Kurmuk seedling and adult cohorts combined.

Table 4-3: FSGS intensity (*Sp*), estimation of the neighbour size (Nb) and the gene dispersal distance σ_g with their respective 95% confidence intervals for the Kurmuk and the Guba-Arenja populations using three estimates of effective densities (DE = D/2, D/4 and D/10). Standard error (SE) in parenthesis; $\hat{b}_{F,\ regression}$ slope$\hat{F}_{1,}$ the mean kinship coefficient (F_{ij})between individuals in the first distance class.

			SGS Parameters				**Gene Dispersal parameters**		
Population	**Cohorts**	**N**	$\hat{b}_F$	$\hat{F}_1$	***Sp***	**Adult trees per km²**	D_E	**Nb (± SE)**	**σ_g (m) (± SE)**
Guba-Arenja	GA	97	0.0119	0.0355	0.0124 (0.0058)	2900 1450 580	D/2 D/4 D/10	80 (40) 75 (35) 77 (NA)	46 (11.5) 64 (14) 103 (NA)
	GS	94	0.0133	0.0291	0.0137 (0.0131)				
Kurmuk	KA	71	0.0100	0.0776	0.0109 (0.0038)	4350 2175 870	D/2 D/4 D/10	234 (212) 181(126) 169 (NA)	65 (25) 81 (26) 124 (NA)
	KS	58	0.0023	0.0050	0.0023 (0.0036)				

Heterogeneity analyses of the FSGS

The spatial autocorrelationanalysis revealed a similar pattern of autocorrelation to the kinship analysis over comparable distances in each of the cohorts and populations. The test for cross-population heterogeneity based on the single-class test (t^2) and multi-class criterion (ω) showed that the pattern of the genetic structure did not significantly differ between the two populations (Guba-Arenja and Kurmuk) and between the two cohorts (adult Vs seedlings) within each of the populations (Table 4-4). The test between populations considering only adult cohorts also did not show significant difference.

Table 4-4: Heterogeneity tests of FSGS between populations and cohorts. t^2: single-class and (ω) multi-class test criterion. GA: Guba-Arenja Adults, GS: Guba-Arenja Seedlings, KA: Kurmuk Adults, KS: Kurmuk Seedlings, GC: Guba-Arenja adult and seedling cohorts Combined and KC: Kurmuk adult and seedling cohorts Combined; DC: distance classes

	DC(m)	0-40	40-80	80-150	150-300	300-500	500-850	ω	p
GA vs GS	t^2	0.162	0.073	2.140	3.589	0.351	0.588	13.52	0.335
	p	0.692	0.809	0.144	0.058	0.556	0.446		
KA vs KS	t^2	14.375	0.797	0.148	0.000	0.684	0.684	20.14	0.063
	p	0.001	0.357	0.695	0.982	0.423	0.410		
GC vs KC	t^2	0.281	2.12	0.047	0.906	0.145	1.216	10.42	0.587
	p	0.599	0.153	0.823	0.400	0.692	0.262		
GA vs KA	t^2	6.851	0.091	0.595	0.059	0.138	0.456	14.24	0.304
	p	0.007	0.766	0.472	0.825	0.737	0.525		

Discussion

Genetic diversity between generations/cohorts

A high level of genetic diversity was detected in two population of *B. papyrifera* in Western Ethiopia, both in terms of number of alleles and heterozygosity. The level of genetic diversity was similar in adult and seedling cohorts in both Guba-Arenja and Kurmuk populations. The high mean number of alleles detected per cohorts (10-14) suggests a large effective population size. The level of heterozygosity detected in these two populations (H_O = 0.689, H_E = 0.693) is comparable to results observed in other tropical tree species when measured with microsatellite markers. For example, H_E = 0.67 and H_O = 0.71 in *Acacia senegal* (Omondi et al., 2010), H_E = 0.78 and H_O = 0.75 in *Swietenia macrophylla* (Lemes et al., 2003) and H_E = 0.58–0.68 and H_O = 0.59–0.77 in *Parashorea stellata* (Tiep et al., submitted). *Acacia senegal* is a tree species occurring with *B. papyrifera* in the same habitat (Terminalia-Combretum woodlands) while *S. macrophylla* and *P. stellata*are effectively outbreeding long-lived tree species occurring in natural tropical forests. We found low F_{ST} between the two populations (Guba-Arenja Versus Kurmuk, F_{ST} = 0.031), which is expected as the populations were geographically close. The two cohorts (adult versus seedling cohorts) in the two populations were not significantly differentiated.

The mean number of alleles (N_A) and effective number of alleles (N_E) per population were slightly higher in Guba-Arenja (N_A = 14, N_E = 7) than in Kurmuk (N_A = 10, N_E = 4). The number of alleles were consistently the same for adult and seedling cohorts while the heterozygosity was slightly higher in seedling cohorts of both populations indicating that no loss of genetic diversity nor selection for heterozygosity was detected across generations and that the diversity in adult cohorts is adequately represented in their progeny (Table 4-2). The high heterozygosity, the low fixation index andthe low *Sp* values may indicate extensive intra-population gene dispersal. *Sp* in predominantly outcrossing or self-

incompatible species is lower (average 0.0126) than the *Sp* (average 0.1431) in predominantly selfing species, with intermediate *Sp* (average 0.0372) in mixed mating species (Hardy et al., 2006). The *Sp* values for *B. papyrifera* (0.0023 – 0.0137) are close to the out-crossing species average (0.0126). The mating system of *B. papyrifera* has not been studied. However, two taxonomically close species are out-crossing (*B. serrata* - Sunnichan et al., 2004; *B. ovalifoliolata* - Raju et al., 2012). These speciesalso showed a pre-zygotic self-incompatibility mechanism. Our results strongly suggest that outcrossing is the (predominant) mating system in *B. papyrifera* as well, and it may have a similar self-incompatibility system. In out-crossing species pollen dispersal contributes to gene dispersal (Vekeman and Hardy, 2004). Therefore, the extensive within population gene dispersal detected in *B. papyrifera* may also be a result of the contribution of extensive pollen dispersal.

Fine-scale spatial genetic structure (FSGS)

An overall positive signal of local spatial genetic structure was detected in the two *B. papyrifera* populationsexcept in Kurmuk seedling cohorts. It is common to find a high relatedness at small distances. From the kinship analyses, a significant genetic structure was detected up to a distance of 62 m in the two *B. papyrifera* populations. In Kurmuk, however, no significant spatial genetic structure was detected for the seedlings over the whole distance range. Limited gene flow reinforces FSGS across generations (King et al., 2012; Moran and Clark 2012). In contrast to this, however, there is the possibility that FSGS weakens over generations due to extensive dispersal and immigration (Smouse et al. 2008). In *B. papyrifera* the FSGS did not significantly increase from adult to seedling populations. The absence of significant SGS in the seedling cohorts of Kurmuk (KS) may suggest that seed dispersal was efficient these last years (seedling cohort) in this population, whereas it was not before (adult cohort). This may be due to an influx of seeds from the

forest on a nearby slope, which has become more degraded (decreased tree density) and more open from time to time. Rainwater may cause surface soil runoff from the steep exposed area, taking the seeds with it. Indeed, on the slope very few seedlings were found, and the flat sampled area had deeper, moist and fertile soil which is a conducive condition for germination of seeds and establishment of seedlings and saplings.

The intensity of FSGS in *B. paprifera* (*Sp* = 0.0023 – 0.0137) was similar to that of two other tropical tree species, *Dalbergia nigra* (*Sp* = 0.017, Buzatti et al., 2012) a wind-dispersed species and *Protium spruceanum* (mean *Sp* = 0.008, Vieiraet al., 2010) a bird-dispersed tree species. Both of them are insect-pollinated. The latter is in the same family as *B. paprifera (Burseraceae*).

In plant species FSGS is related to life-form, population density and the mating system of the species. Vekemans and Hardy (2004), in their assessment based on 17 tree species, found a mean *Sp* of 0.0102. Our *B. papyrifera Sp*, particularly in Guba-Arenja (0.0124) was only slightly higher. The difference in *Sp* estimates of the two *B. papyrifera* population were not in line with the expectation that *Sp* are significantly higher in low-density compared to high-density populations as shown for example in Hardy et al. (2006). In our study the Guba-Arenja adult population was less dense (58 individuals/hectare) than that in Kurmuk (87 individuals per hectare) whereas it had *Sp* values (0.0124) similar to Kurmuk (0.0109).

Gene dispersal distance was relatively shorter in the less dense population in contrast to what was expected. The gene dispersal estimation considering D_E = D/10, were 103 m in Guba-Arenja but 124 m in the denser population in Kurmuk. In studies in other species (e.g *Chamaecrista fasciculata,* Fenster, 1991; *Heliconia acuminata,* Cortes et al., 2013) higher gene dispersal distance was found in low-density populations. The absence of the consistency of the gene dispersal distance

with population density in different populations may be due to the differences of the sites in other ecological factors that affect gene dispersal such as for example diversity of pollinators, seed dispersers and microsite conditions affecting germination and seedling survival. In mixed stands, high tree density can limit dispersal by serving as a physical barrier to the movements of propagules (Curtu et al., 2015).

The heterogeneity test for autocorrelation showed that neither the spatial genetic structure across populations (Guba-Arenja (GC) versus Kurmuk (KC)) nor the SGS between cohorts (GA versus GS; KA versus KS) was significantly different. The absence of significance difference in FSGS pattern might be due to immigration of pollen or seed into the seedling generation. Long life span ensures the representation of many cohorts within a population. This enables such species to maintain diverse individuals across cohorts, ensuring high genetic diversity within populations (King et al., 2012; Moran and Clark, (2012). As a long-lived, widespread species with extensive gene dispersal, *B. papyrifera* may maintain its genetic diversity across generations in continuous natural forest landscapes, but it may be at risk of losing the genetic diversity in the future in fragmented landscapes, in those cases where the inter-fragment distance is beyond the scale of the gene dispersal.

Conservation implications

This study provides the first assessment of the FSGS and gene dispersal patterns in *B. papyrifera*. Such information can be valuable for management objectives that seek conservation of the genetic resources of the species. The high heterozygosity in the adult and seedling cohorts, the low fixation index/inbreeding coefficients and the low *Sp* values suggest that FSGS in *B. papyrifera* is relatively weak and gene dispersal is extensive within the populations. The FSGS in one of the seedling cohort populations might be the consequence of seed deposit from sources

originating outside of the sampled area. The current FSGS analysis results provide a guide for sampling of individuals and seed collection within conservation units for ex-situ conservation and reforestation programmes. Our findings suggest that within the wider conservation units/provenances the collection of seeds for *ex situ* conservation and reforestation programmes of *B. papyrifera* should use trees separated by distances of preferably 100 m but at least over 60 m to reduce genetic relatedness among seeds from different trees. In the case of populations located in foothills where water-driven seed transport is probable, and knowing that small-scale runoff occurred, the use of smaller inter-tree distances may be sufficient, for instance a distance of around 30 m.

The *B. papyrifera* populations in Ethiopia are increasingly fragmented, and scattered remnant trees and populations dominate the growing region of the species. The species has maintained high level of genetic diversity in the adult trees in all populations studied (Addisalem et al., 2016). Most of these populations do not have younger plants, however, preventing future regeneration. Although the findings of this study are important for managing and conserving the extant trees and populations in the fragmented landscapes, this study only provides baseline information on the spatial structuring and dispersal of genes in the species. Facilitating seed survival and dispersal, and the survival of emerging seedlings are crucial for maintaining the populations viable and ensure their sustainability. The fact that the Western populations of the present study do show high genetic diversity in combination with seedling recruitment is encouraging in that sense: if regeneration is made possible in other areas in Ethiopia, the genetic diversity may be conserved in all the current populations. Conservation measures thus should focus on protection of the current populations and on effective and long-lasting regeneration of those populations.

Acknowledgement

This study was financially supported by Netherlands Fellowship program (NUFFIC), the Netherlands. The field research was partly supported by Abdou-Salam Ouédraogo Fellowship through Bioversity International and the CGIAR Research Programme on Forests, Trees and Agroforestry, Rome, Italy.

CHAPTER 5

GENERAL DISCUSSION

Introduction

Dry tropical forests and woodland vegetation are characterized by economically and ecologically beneficial plant species. However, due to the ever-growing dependence of humans on natural resources, these forested areas are increasingly overexploited and converted to other land uses. This occurs continuously and increasingly, despite that this is at the expense of the economic benefit, and of the wealth of ecological services this vegetation provides, now and potentially in the future (Chidumayo and Marunda, 2010). Tropical woodlands, mainly Combretum-Terminalia and Acacia-Commiphora, possess a variety of such valuable tree species.*B. papyrifera* is one of these economically important dry forest/woodland tree species (Vollesen, 1989; Gebrehiwot, 2003; Lemenih et al., 2007). The species provides frankincense, a commodity traded internationally because of its value as ingredients in pharmaceuticals, food and cosmetic industries (Tucker, 1986). Local people benefit from the sales of frankincense, honey production and employment opportunities in the field production and processing of frankincense (Eshete et al., 2005; Woldeamanuel, 2011). Different parts of *B. papyrifera* are used traditionally as medicines against various diseases (Lemenih and Teketay, 2003; Zhang et al., 2013). In modern medicine, it was discovered that frankincense from *Boswellia* species contains chemical constituents that have the ability to target cancer cells of different tissue origins (Lin, 2013; Ni et al., 2012; Evans, 2013) and boost immune systems (Khajuria et al., 2008). These findings have shed light on the potential of frankincense for pharmaceutical use. *B. papyrifera* is, therefore, one of the most important wild tree genetic resources of Ethiopia with current and potential economic importance.

In Ethiopia, however, vegetations which are habitats for *B. papyrifera* and other woodland species, are threatened by various human-driven factors since the late 1960's (Lemenih, 2007; Alemu et al., 2015). Because of these threats *B. papyrifera* populations have decreased tremendously to ever smaller and fragmented patches.

Regeneration is scarce or absent in most of the populations (Abiyu et al., 2010, Eshete et al., 2011; Groenendijk et al., 2012; Tolera et al., 2013; Alemu et al., 2015). This causes a risk of losing genetic diversity and highlights the need of urgent conservation measures to be taken to preserve this irreplaceable genetic resource and ensure the sustainable production of frankincense.

The present study dealt with the conservation genetics of *B papyrifera*, that is, the genomic resources (sequences and SSR markers) were developed and applied in genetic analysis of thepopulations at local and regional geographical scales in Ethiopia. The results increased our understanding about the level of genetic diversity within the populations, patterns of differentiation among populations, the genetic structure at smaller geographical scale (FSGS), the genetic diversity across generations and the processes predominantly playing a role in influencing the patterns. This population-genetic information can assist in devising conservation measures to preserve a considerable portion of the existing genetic variation of the species (Moran, 2002; Frankham, 2010; Edwards et al., 2011). Fig. 5-1 summarizes the threats to *B. papyrifera* (Box number 1-4), the focus of the study (Box number 5-8), the application of the outcome (Box number 9-10) in sampling populations for *in situ* and *ex situ* conservation and management of the existing genetic diversity.

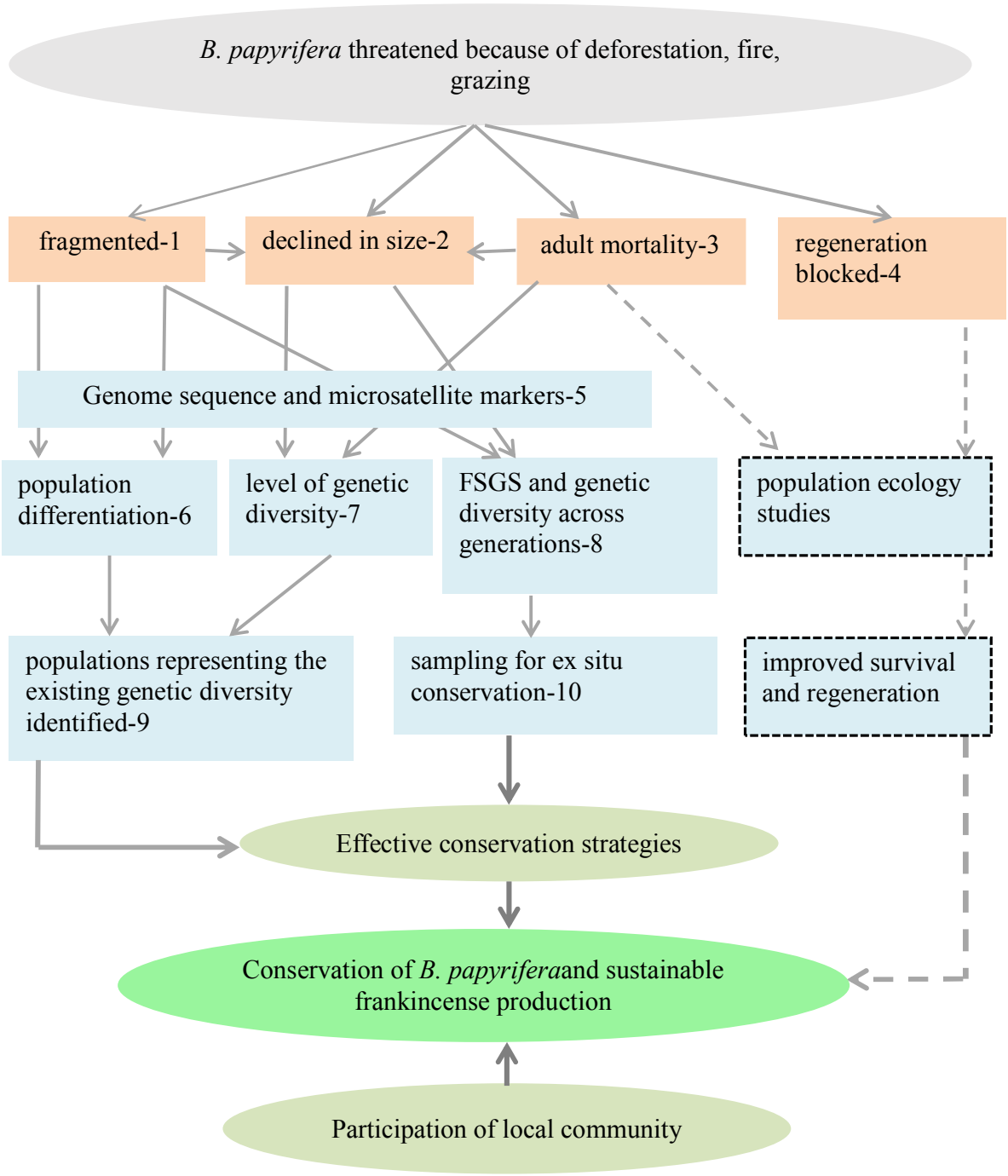


Fig. 5-1: Schematic presentation of the conservation genetics of *B. papyrifera*. The boxes in dashed line and connected with dashed arrows show the role of

population ecology studies in preserving the species and sustainable frankincense production.

In this General Discussion I will outline the main results (Fig.5-1, box number 5-8) of my thesis work and discuss them in a context of (a) the utility of the genetic tools (genetic markers) for conservation of tropical woodland trees, (b) the potential of the markers for application in related species, (c) the preservation of the populations representing the existing genetic diversity *in situ* in the natural forests and the challenges from the regeneration bottlenecks, (d) the conservation values of remnant *B. papyrifera* trees and population patches in human influenced landscapes, (e) conservations of marginal and isolated populations,(f) the implication of the local genetic structure (FSGS) in sampling of material for *ex situ* conservation and other plantation programmes and (g) the challenges in conserving the existing genetic diversity and what should be studied in the future to generate more information and to effectively conserve the species and the woodland resources.

Genome sequencing and microsatellite markers for *B. papyrifera*

Molecular markers provide a means of characterizing genetic variation in natural populations and thus to produce estimates of genetic parameters that are of considerable importance for conservation and management of genetic resources. In this study, Illumina paired-end sequencing was applied and the first genomic resource (sequence data and SSR markers) were generated for *B. papyrifera*. A set of 46 polymorphic SSR markers, most of which cross-amplified in two sister species, *B. pirrotae* and *B. popoviana*, were developed from the short Illumina reads (Chapter 2). A subset of ten SSR markers were applied for characterizing the genetic diversity and differentiation among populations of *B. papyrifera* sampled across Ethiopia. The markers were used to generate genetic information that allowed us to distinguish four conservation units for the species (Chapter 3). Furthermore, the local genetic structuring (FSGS) and gene dispersal within populations of the

species were also determined for two populations in the West of Ethiopia where regeneration was still taking place (Chapter 4). Based on these data we determined the distance to be kept between trees to obtain genetically unrelated genotypes when sampling for plant material collection for *ex situ* conservation and plantations of any purpose.

From the short reads 444 927 contigs (i.e. longer stretch of sequences) were assembled and several contigs showed similarity to genes of the terpene and terpenoid backbone synthesis pathways, which form the major constituents of the bark resin.

The cross-amplification of the SSRs is an indication of their potential use for genetic studies of the remaining *Boswellia* species. With further exploration, this data provides opportunity to develop more SSR markers for further studies of this and other closely related species. The sequence data generated forms the start of a valuable genomic resource for various applications. The data serves as a basis for generating more information needed for conservation and management of the species. It has enormous potential for assembling the chloroplast genome using advanced bio-informatics approaches and applying in other genetic studies relevant to conservation, for instance studying the phylogenetics and historical processes responsible for the contemporary distribution of populations (phylogeography) of the species.

Molecular markers have been effective for generating of information of conservation relevance and the recent advances in DNA sequencing technology have been boosting a rapid development of molecular markers. However, application in developing countries is hindered by the unavailability of the technologies and the lack of knowledge and skills to apply the technologies. In Ethiopia genetic diversity has been assessed for few tree species, namely *Cordia*

africana (Derero, 2007), *Hagenia abyssinica* (Bekele, 2008*), Juniperus procera* (Sertse et al., 2011), *Prunus africana* (Mihretie, 2015), *Lobelia giberroa* (Kebede et al., *2007*) and *B. papyrifera* (this study) and all of the studies were conducted abroad (Europe). This would imply the need of promoting the new techniques and technologies, building the sectors involved in forest genetic resources in terms of material and financial resources and equipping the technical staff with the necessary skills and knowledge.

Genetic diversity pattern and the breeding system of B. *papyrifera*

In this study the assessment of genetic diversity showed that the populations studied across Ethiopia (CHAPTER 3) have maintained a high level of genetic diversity. Most tree species display high levels of genetic variation because they often combine characteristics that ensure the preservation of genetic diversity, such as long live span, overlapping generations, large continuous populations, wide geographical range, outcrossing breeding systems, and relatively long distance gene dispersal (Hamrick, 2004; Moran, 2002). The findings of this study are in line with these expectations. *B. papyrifera* is a slow-growing and geographically widespread tree species (Figure 1 in CHAPTER 1). The studied populations are confined to Ethiopia but are distributed over a wide geographical range in the Western, North western, Northern, North western and central parts of the country. We attributed the high genetic diversity to the long life span and wide geographical distribution of this species.

High within-population genetic variation is a typical genetic structure for outcrossing species because an outcrossing mating system promotes gene dispersal and therefore heterozygosity (White et al., 2001). In *B. papyrifera* the mean level of inbreeding detected across the populations was very small (mean F = 0.013, NS) (CHAPTER 3). Considering the fact that the level of genetic diversity is relatively

high and inbreeding is low we concluded that outcrossing is a predominant breeding system in *B. papyrifera*.

Population fragmentation effects on genetic diversity of *B. papyrifera*

Previous studies in Ethiopia (e.g. Abiyu et al., 2010; Eshete et al., 2011; Groenendijk et al. 2012; Tolera et al., 2013) indicated that *B. papyrifera* populations are threatened in terms of the population size, fragmentation of populations, lack of recruiting individuals and high mortality of aged trees. Ultimately, threatened species will end up with a low level of genetic diversity. This idea was empirically supported by many studies (e.g. Farwig et al., 2008, Sertse et al., 2011; Yineger et al., 2014). The findings of this study differ from this theoretical expectation. *B. papyrifera* populations are not yet genetically threatened despite the ongoing threatening factors including population fragmentation (CHAPTER 3). Population fragmentation often leads to decreased population size and increased population isolation. These effects theoretically lead to reduced level of genetic diversity and increased genetic differentiation among populations due to reduced gene flow, increased level of genetic drift and mating among closely related individuals (Aguilar et al., 2008). There is increasing evidence, however, that plant species respond in diverse ways to fragmentation leading to different effects than expected. This study revealed that small, fragmented and sparsely distributed *B. papyrifera* populations (KTD, AF-Bir and JRG) have retained comparable levels of genetic diversity to larger and more continuous populations. This may be due to the following two reasons: 1) fragments were large enough to prevent loss of alleles except may be a few rare alleles; 2) the time that has passed since the onset of forest fragmentation (about 5 decades) is short in comparison to the life span of adult trees, which can be over 100 years of age, and hence insufficient to cause inbreeding and subsequent loss of heterozygosity.

Intra-population spatial genetic structure (FSGS) and genetic diversity across generations

The study that considered two populations from western Ethiopia (CHAPTER 4) showed a weak genetic structure, comparable with other outcrossing tree species. The comparison of the FSGS across generations revealed weaker genetic relationship in juvenile generations compared to adults. Local spatial clustering of closely related individuals arises from genetic and demographic factorsand the complex interaction of such factors can have unexpected results (King et al., 2012; Moran and Clark, 2012; Curtu et al., 2015). The intensity of SGS is strongly influenced by species dispersal ability, i.e., FSGS is expected to increase in species with shorter seed and pollendispersal distances (Cavers et al., 2005). The findings of chapter 4 thus suggest that gene dispersal is extensive in continuous *B. papyrifera* populations. The weaker FSGS in seedling cohorts may be due to seed immigration from outside of the sampled area (CHAPTER 4). Particularly in Kurmuk this may have been promoted by forest stands which have become more open and more degraded (in terms of density) due to human disturbances. The study demonstrated genetic relatedness of individuals up to a maximum distance of 62 m. This means that sampling of trees for seed and other plant material collection for *ex situ* conservation and plantation programmes should consider this as a minimum distance to be kept between trees to collect unrelated genotypes. The genetic diversity estimates (number of alleles and heterozygosity) did not differ between adult and seedling cohorts, consistent with the long life span and extensive gene dispersal.

Predominantly outcrossing species such as *B. papyrifera* are not adapted to high levels of inbreeding and may therefore suffer from increased bi-parental inbreeding if this would occur as a result of limited gene flow because of severe forest fragmentation. It was shown that fragmentation affects outcrossing plant species more than self–compatible species by reducing their genetic diversity

(Honnay and Jacquemyn, 2007; Ng et al., 2009). This implies that in out-crossing species the gene flow among fragmented populations and scattered remnant trees has to be maintained for the populations to remain genetically connected and to maintain the level of genetic diversity over generations. Currently, we have not observed this in *B. papyrifera*, but if the loss of habitat and reduction of number of trees continues, especially in the North of Ethiopia where it is already very severe (Abiyu et al, 2010; Eshete et al., 2011), then it would compromise the potential of regeneration.

Genetic differentiation among *B. papyrifera* populations

The genetic differentiation in *B. papyrifera* (F_{ST} = 0.084) was comparable to the genetic differentiation observed in other outcrossing tree species, such as *Swietenia macrophylla* (F_{ST} = 0.097, Lemes et al., 2003) and in the wind-pollinated desert tree species *Populus euphratica* (F_{ST} = 0.093, Wang et al., 2011). Honeybees are considered as one of the pollinators of *B. papyrifera*. The seed dispersal syndrome of the species is not well known. In a sister species (*B. ovalifoliolata,* Raju et al., 2012), however, seed characteristics are considered to be adapted to anemochory (small size, light weight, papery and winged) and seed dispersal was detected up to 400 m. *B. papyrifera* seeds have similar characteristics. Assuming the same dispersal distance for *B. papyrifera*, seed dispersal might not be very effective to genetically connect geographically far apart populations whereas it indicates the possibility of extensive within-population dispersal. This suggests that the relatively low level of among-population differentiation we detected for this species may be due to long-distance pollen dispersal that may connect populations. There is theoretical and empirical support that pollen dispersal is often over much larger distances than seed dispersal in tree species, for instance in oak and beech (Petit et al.,2005; Chen et al., 2008). However, for tree species with wind-dispersed seeds and limited movement of pollinators, seed dispersal distance is expected to exceed pollen dispersal distance. In *B. papyrifera*the seed

and pollen dispersal mechanisms are not well known. The hypothesis that pollen dispersal is over much larger distances than seed dispersal needs further investigation.

Tree species with larger, more continuously distributed populations generally have higher genetic diversity at the population level and less diversity among populations than those that are originated from a limited number of individuals and/or that have been isolated for considerable number of generations (Hamrick, 2004). Our findings show that the majority of the genetic diversity in *B. papyrifera* is present within populations and the level of differentiation among populations is low. High levels of within-population diversity and low among-population differentiation have also been reported for *Hagenia abyssinica* (Bekele, 2009), *Cordia africana* (Derero, 2011)and *Juniperus procera* (Sertse et al., 2011), which are tree species investigated in Ethiopia.For our study we sampled adult trees, which means that the genetic diversity pattern represents the populations that existed in the past (19th century) in large and continuous populations, so forest fragmentation must be (much) more recent than 100 years.

Correlation of genetic distance with the geographic distances revealed a clear isolation by distance pattern with 44% of the variation accounted for by geographic distance. This strongly suggests that geographical distance between populations is one of the factors driving differentiation and it has to be considered in sampling representative genetic diversity for conservation of the species.

The genetic structure analysis which involved assigning of individual genotypes to a genetic cluster detected four distinct genetic clusters, each containing 2-4 populations. The structuring assigned geographically closer populations to the same genetic cluster (CHAPTER 3). The four clusters corresponded to different ecological conditions in terms of temperature, rainfall and soil conditions. This

pattern might be due to the influence of the selection pressure, leading to local adaptation. As SSRs are selectively neutral markers (Holderegger et al., 2006), the existence of local adaptation cannot de derived from our data. However, it is wise to consider these ecological factors during conservation planning, in addition to the spatial distances between populations, and for the moment assume at least four different clusters to be conserved.

Conservation of *B. papyrifera* in natural habitat and the challenges from regeneration bottlenecks

In view of the increasing pressure on forest resources and the associated risk of losing the diverse benefits they provide to human wellbeing, conservation of biodiversity and genetic resources is an important responsibility. Limited funding is a challenge for genetic resource conservation and management. Hence, selecting a set of sites among several available ones is one of the approaches to systematically plan and efficiently implement conservation measures (FAO, 2014).

Effective in situ conservation demands acquiring genetic diversity representing a gene pool of the species in question to represent the level of the existing genetic diversity while taking into account the genetic differentiation among populations. Conserving within-population diversity should target preserving large populations that will not lose diversity due to drift (Namkoong, 1988). We determined 4 genetic clusters based on the STRUCTURE analysis results presented in CHAPTER 3. The genetic clustering of the populations strongly suggests the need of establishing at least one in situ conservation unit in each of the four genetic clusters to represent the existing genetic diversity. Accordingly, we recommend a minimum of four conservation sites for *B. papyrifera* which represent the Western, North-western, Northern and North genetic clusters. Each of the conservation sites should include one large core population or more than one population that will not lose genetic diversity in the subsequent generations due to genetic drift and

inbreeding, which are the causes for loss of genetic diversity in small populations. The FSGS study demonstrated relatedness of individuals over a maximum distance of 62 m (CHAPTER 4). On the basis of this finding, the sampling for seed and other plant material collection for *ex situ* conservation and plantation for restoration should ensure that over 60 m (preferably 100 m) between trees distance is maintained to acquire nonrelated/less related genotypes.

As discussed in CHAPTER 3, the high genetic variation within populations and the limited differentiation among populations suggests larger and continuously distributed B. populations in the past. It is assumed that the forest cover of Ethiopia was much larger and continuous in the not so distant past (EFAP, 1994; Reusing, 2000). The genetic differentiation pattern shown in this study and those of the previous studies on other Ethiopian tree species (*Hagenia abyssinica* (Bekele, 2009), *Cordia africana* (Derero, 2011) and *Juniperus procera* (Sertse et al., 2011)) supports the notion that *B. papyrifera* populations were larger and continuously distributed in the past. In particular, the genetic data indicate that the trees we sampled must, themselves, have been part of such continuous forests, or their parents must have been. The findings of this study particularly imply that populations of other forest tree species co-existing with *B. papyrifera* may have been also larger and continuous. Hence, the conservation strategies developed for *B. papyrifera*and sites recommended herein might also serve to preserve other species in Combretum-Terminalia and Acacia-Commiphora woodlands. This would be particularly effective for *Boswellia* species growing together with it in the same geographical range (*e.g. B. pirottae*).

Reversing the loss of forests, ensuring persistence of species and transfer of the genetic diversity into the future generation largely depends on the occurrence of regeneration. While this study demonstrated that *B. papyrifera* is not genetically threatened (yet), maintaining genetic viability of the species for coming

generations is questionable due to the regeneration bottleneck which is prevalent in the majority of *B. papyrifera* populations. Previous studies (e.g. Eshete et al., 2011; Groenendijk et al., 2012; Tolera et al., 2013) identified forest fire, grazing and trampling by domestic animals as causes for the regeneration bottleneck of *B. papyrifera*. From a genetics perspective, inbreeding depression also can prevent plant growth at young life stages such as seedlings and saplings. This was observed, for instance, in Douglas-fir, Ponderosa Pine and Noble Fir (Sorensen and Miles, 1982). Inbreeding may be expected in degraded and fragmented small populations like *B. papyrifera*. Our genetic data, however, showed that *B. papyrifera* still maintains its genetic variation, so inbreeding depression is not a revenant factor for the regeneration bottleneck.

Seed dispersal operating together with micro-site suitability for regeneration and growth of established seedlings influence the species range expansion and colonization of new sites (Chybicki and Burczyk, 2010; Chen et al., 2008). The regeneration status of *B. papyrifra* hugely varies from population to population. The Western region (Benishangul-Gumuz) has a high density of seedlings and saplings (observations during the field survey for sampling). Surprisingly, in several places in this region *B. papyrifera* stands are even mainly composed of saplings and seedlings, with only scattered very few or no adult trees in the stand. This means that the seeds may have come from nearby populations and established in different locations. This peculiar situation, needs to be further investigated. We need to understand the seed/propagule sources, the microsite conditions favouring the regeneration and whether the low livestock population pressure in the area, due to the presence of Trepanosoma, contributes to the rejuvenation and better persistence of the species. This information can then be used to apply to the various parts of the distribution range where this species is under threat and therewith the future use of its important resin frankincense.

The common mechanisms of regeneration in dry forests and woodlands include current-year seedlings as well as seedling re-sprouts that occur after dieback of vegetative shoots taking advantage of the extensive root systems and the larger carbohydrate reserves in the tap root (Kennard et al., 2002; Birhane et al., 2012; Chirwa et al., 2015). Fire is one of the few dry forest and woodland determinants that can affect regeneration differently, depending on its timing, intensity and frequency, as intense and frequent fires retard seedling recruitment into saplings. The growth ability after injury from fire varies with the age or size of a plant and also with the type and severity of injury. That is, seedlings that grew and build enough root reserves between fires survive the next fire (Chidumayo and Marunda, 2010). Fire is one of the major causes for blocked regeneration in *B. papyrifera* (Eshete et al., 2011; Groenendijk et al., 2012; Tolera et al., 2013). However, there is a huge variation in response of the seedlings to fire occurrence, as in some places (e. g. Western Ethiopia) fire seems to enhance re-sprouting (Fig.5-2) rather than to damage the seedlings (personal observation) which may be due to differences in fire intensity. This clearly suggests that determining the number of years needed to achieve high enough growth and development of seedlings and root sprouts to be able to resist the damage from the next fire is an important aspect to be considered in future studies.

Fig. 5-2: seedlings re-sprouting shortly (4-6 weeks) after fire (Benishangul-Gumuz, in Guba-Arenja population).

The potential of fragmented landscapes/populations for conservation of *B. papyrifera* genetic resource and management strategies

Under subsistence farming some tree species and fragmented forest patches remain while the rest is converted from forests to other land uses. Such human-modified landscapes can harbour valuable genetic resources important for in-situ conservation, while at the same time they can be used for collecting genetically diverse and locally adapted seeds for effective restoration of degraded forests (Harvey et al., 2006; Trimble et al., 2014), given that the management approaches ensure maintenance of genetically viable populations across the landscapes. Our study showed that the fragmented Ethiopian landscapes dominated by human activities, particularly farm and grazing lands, still contain *B. papyrifera* populations with the same level of genetic diversity as landscapes with relatively continuous populations, but the lack of regeneration indicates that it is not a sustainable situation. Therefore, in complement with the conservation in the natural forests and woodlands, active conservation and management of these

valuable genetic resources existing in remnant and fragmented *B. papyrifera* populations (e.g. Fig. 5-3) is highly recommended.

Fig. 5-3: Scattered *B. papyrifera* trees in human-dominated landscapes (settlement areas and farmlands).

For this, the fragmented populations and scattered remnant trees have to be genetically connected for the level of genetic diversity to be maintained over generations. This may be done by promoting gene flow between the remnant trees and fragmented populations through reducing the isolation distance, for instance by building corridors between the patches, and through increasing overall population density, both within the fragmented populations and across the whole landscape. Ensuring natural regeneration in fragmented landscapes is unlikely largely due to soil disturbance and uprooting of seedlings during agricultural practices (e.g. ploughing, weeding) and other factors as browsing and trampling by grazing animals. Fencing cattle out of large areas destined for forest regeneration may not work. Management intervention, therefore, needs a different approach than we apply in natural forests (protection of the forest from human interference). One system involves planting of nursery-raised seedlings and subsequent protection of out-planted seedlings at the planting spot. Birhane et al. (2012) showed that mycorrhizal inoculation of *B. papyrifera* seedlingsimproves the

trees' ability to survive in the harsh climate. The finding implies that applying such management would improve the field establishment of out-planted seedlings, which is one of the bottlenecks for establishing *B. papyrifera* plantations.

Another potential means of regeneration of *B. papyrifera* is vegetative propagation (Eshete et al., 2011; Lemenih and Kassa, 2011). Propagation through cuttings has been tested and it has become successful (Fig.5-4) in Northern (Kaffa-Humera) (Haile et al., 2011) and North eastern Ethiopia (around KTD). Such a propagation method appears more appropriate in the human-dominated landscapes as compared to planting of nursery-raised seedlings since the latter are more susceptible to damage due to up-rooting and other disturbances imposed by agricultural activities.

Fig. 5-4: *B papyrifera* plantation trial established from cuttings (picture taken during a dry season).

Conservations of marginal and isolated populations

Where situations may lead to the risk of local extinction, *in situ* conservation should be complemented with *ex situ* conservation of the genetic resources. This study revealed that isolated and degraded populations (e.g. JRG and Af-Bir) still contain genetic diversity comparable to large and continuous populations (CHAPTER 3). Despite its substantial genetic diversity, the AF-Bir population is

marginal at present in terms of a microclimate that barely supports vegetation suggesting local extinction risk for the species. However, *B. papyrifera* trees are existing in the landscape regardless of the harsh ecological conditions. This might be due to local adaptation of the populations to the environmental conditions. Obviously, this highlighted the need of *in situ* conservation of these populations which might have already developed adaptation ability to such harsh ecological conditions. However, as these populations are highly threatened and *in situ* conservation alone cannot guarantee protection from further threatening factors and local extinction; we highly recommend complementing the *in situ* conservation with *ex situ* conservation for this population. Similarly, *ex situ* conservation strategy is needed for JRG population because this population is isolated from other populations and severely degraded as evidenced by very low tree density and absence of regeneration. The *ex situ* approach could be preservation of seeds and other plant material in cryopreservation in gene banks depending on the longevity of seed and promoting production plantation of representative provenances, either from seeds or as cuttings.

Conclusions and outlook

In this thesis, genetic methods were utilized to characterize the genetic diversity pattern of frankincense tree, *B. papyrifera,* an economically important but ecologically threatened tree native to tropical woodland and dry forests. The findings indicated that the species is not genetically threatened despite the ongoing ecological and demographic threats from anthropogenic factors.

The study raised a number of conservation related issues for future studies. The following questions stand out as particularly important to be investigated.

Provenance test for climate change and local adaptation

Habitat heterogeneity (elevation, slope, aspect, moisture etc.) may have preserved adaptive genetic variation that, when recombined and exposed to selection in newly colonized habitats, has given rise to the local adaptation of tree populations (Hamrick, 2004). Our study identified four distinct populations, which corresponded to different ecological conditions in terms of mainly temperature, rainfall and soil conditions. This pattern might have been developed due to the influence of local selection pressure operating coupled with geographical distances among the populations. The existence of local adaptation should be investigated by provenance trials, which are reciprocal common garden tests where tree populations collected from different regions are tested for survival, growth, phenology, productivity, and other adaptive traits to study the environment effects (Krutovsky et al., 2012). Within the provenances populations may have special adaptation due to the different microsite conditions, therefore the experiment should consider various sub-populations existing within the provenance for the test.

Climate change puts many species at risk of extinction because of its effect on habitat suitability, the rate depending on the species potential to expand their geographical range and restore their population under the new environmental regime (Chen et al., 2008; Hamrick, 2004). Climate change in Ethiopia was predicted to lead to a rise in temperature and precipitation. The effects of these climate changes on population dynamics relevant to conservation and restoration of the species, such as regeneration, potential for shift in ecological range, remain largely unknown. The data obtained from provenance tests are useful to assess the adaptive range of trees and can be used in defining suitable habitat for the species that will be at risks of the predicted climate changes (Krutovsky et al., 2012).

The provenance trial established for the species may serve four purposes: improving traits suitable for adaptation to varying environmental conditions including climate changes, determining suitable habitat ranges under predicted climate changes, (ex situ) conservation of the species and improving traits associated with frankincense production.

Phylogeography and phylogenetic study of *Boswellia* species

The distinct genetic clusters in geographically separate regions emphasize the need of understanding the evolutionary history of the species to identify more specific evolutionary lineages and diversity hotspots deserving priority for conservation. The establishment of such independent historicalevolutionary lineages is influenced by demographic processes and ancestral relationships especially in long-lived plant species such as trees (Schaal et al., 1998; Savolainen et al., 2007). The understanding of the historical divergence (Debout et al., 2011) and geographic pattern of tree populations has been greatly enhanced by phylogeographic studies (Schaal et al., 1998). *B. papyrifera* is a widespread species hence, phylogeographic study is crucial to identify evolutionary lineages and diversity centers for broad scale conservation of the species. Chloroplast genome evolves slowly, hence differences in chloroplast sequences can persist over long time (Palmer, 1987). With respect to this, the sequence data we have generated has enormous potential for phylogenetic reconstruction of species in the genus (*Boswellia)* and phylogeographic study of *B. papyrifera*.

Social perspectives

Deforestation and forest fragmentation in Ethiopia is widespread and has resulted in fragmented populations. This is the dominant situation in much of the forests and woodlands of the country and it is expected to get worse in the near future as human population expansion and the associated livelihood dependence on forests are increasingly competing for forest lands (Reusing, 2000; FAO, 2005, 2014).

Forest management and conservation initiatives should consider participation of local people because their livelihood depends on forests. A state-community joint forest management approach, namely participatory forest management (PFM), that involves the community living in and around the forests has become effective forest management approach in the country. Notable example of such practices are the pilot-phase PFM initiated in mid-1990's in Chilimo and Bonga forests which later also expanded to Bale mountain national forest priority areas (the Waldayaa Jiraatoota Bosonaa (WAJIB) means forest dwellers association approach). These pilot forest conservation approaches have proved a positive progress in reversing deforestation, improving regenerations of forests and improving the livelihood of participating group (Kubsa et al., 2003; Tesfaye et al., 2010; Amente, 2006; FARM Africa, 2008; FARM Africa, 2015).

PFM as an approach of forest management is being recognized in Ethiopia at federal level and in various states. It is refocusing its principle on making forests pay specially from Non-Timber-Forest Products (NTFP) and recently from sales of carbon stocks (involving the REDD+ scheme: Reducing Emissions from Degradation and Deforestation) and payment for ecosystem services (FARM Africa, 2015). This approach is under trial in conservation and management of different forest resources (e.g. Kafa Biosphere reserve based on forest coffee, spices and honey production). In woodland resources, in which *B. papyrifera* is one of the dominant NTFP source, one example of these trials is the forest management utilisation cooperative (FMUC) project, which has been implemented in Famatsere Kebele, Benishangul Gumuz (FARM Africa, 2015). The project that focuses on raising income from Frankincense sales has demonstrated a positive impact on household income resilience and pointed to the possibility of successfully implementing PFM (FARM Africa, 2015). This approach would be particularly effective for conservation of forests genetic resources in landscapes dominated by human activities which contains a considerable *B. papyrifera* forest patches and the

associated genetic resources integrated in different off-forest land uses, such as agroforestry systems. Approaches considering the potential range of products and services derived from the forests would produce more benefits, hence, making forest conservation more attractive to the community. Boswellia forests are sources of a range of NTFP such as honey production and low land bamboo (used for furniture) in addition to frankincense (Gebrehiwot et al., 2003; FARM Africa, 2015). Exploring the potential of the forests and producing multiple products would be more beneficial and fit the demand of the people from the different community/societal settings. This may enable to expand and effectively conserve the forests and the genetic resources there in.

REFERENCES

Abiyu, A., Bongers, F., Eshete, A., Gebrehiwot, K., Kindu, M., Limenih, M., Moges, Y., Ogbazghi, W. Sterk, F.J. 2010. Incense woodlands in Ethiopia and Eritrea: regeneration problems and restoration possibilities. In Bongers, F. and Tenningkeit, T. (eds.), degraded forest in eastern Africa: Management and restoration, Earthscan, London, pp. 123–132.

Abtew, A.A., Pretsch, J., Mohamoud, T.E., Adam, Y.O. 2012. Population status of Boswellia papyrifera (Del.) Hochst in the dry woodlands of Nuba Mountains, South Kordofan State, Sudan. Agriculture & Forestry 54(8), 41–50.

Adam, A.A., El Tayeb, A.M. 2008. A comparative study of natural regeneration on B. papyrifera and other tree species in Jebel Marra Darfur, Sudan. Research Journal of Agriculture and Biological Sciences 4(1), 94-102.

Addisalem, A.B., Bongers, F., Kassahun, T, Smulders, M. J.M. 2016. Genetic diversity and differentiation of the frankincense tree (Boswellia papyrifera (Del.) Hochst) across Ethiopia and implications for its conservation. Forest Ecology and Management 360, 253–260. doi: 10.1016/j.foreco.2015.10.038.

Addisalem, A.B, Esselink, G.D., Bongers, F., Smulders, M.J.M. 2015. Genomic sequencing and microsatellite marker development for *Boswellia papyrifera*, an economically important but threatened tree native to dry tropical forests. Aob Plants 7: plu086. doi: 10.1093/aobpla/plu086

Aguilar, R., Quesada, M., Ashworth, L., Herrerlasdiego, Y., Lobo, J. 2008. Genetic consequences of habitat fragmentation in plant populations: susceptible signals in plant traits and methodological approaches. Molecular Ecology 24, 5177 – 88.

Alemu, B., Garedew, E., Eshetu, Z., Kassa, H., 2015. Land Use and Land Cover Changes and Associated Driving Forces in North Western Lowlands of Ethiopia. International Research Journal of Agricultural Science and Soil Science 5, 28-44. doi 10.14303/irjas.2014.063.

Allan, G.J., Max, T.L. 2010. Molecular genetic techniques and markers for ecological research. Nature Education Knowledge 3(10):2

Amente, G. 2006. Integrated and participatory forest management in the Bale Mountains of Ethiopia. International symposium towards sustainable livelihood and ecosystems in mountainous regions. 7-9 March 2006, Chiang Mai, Thailand.

Arens, P., van der Sluis, T., van 't Westende, W.P.C., Vosman, B., Vos, C.C., Smulders, M. J.M., 2007. Genetic population differentiation and connectivity among fragmented Moor frog (*Rana arvalis*) populations in the Netherlands. Landscape Ecology 22, 1489-1500. doi 10.1007/s10980-007-9132-4

Assefa M, Dekebo H, Kassa H, Habtu A, Fitwi G, Redi-Abshiro M. 2012. Biophysical and chemical investigations of frankincense of *Boswellia papyrifera* from North and Northwestern Ethiopia. *Journal of Chemical and Pharmaceutical Research* 4:1074-1089.

Basar, S. 2005. Phytochemical investigations on *Boswellia* species. Comparative Studies on the Essential Oils, Pyrolysates and Boswellic Acids of *Boswellia carterii* Birdw., *Boswellia serrata* Roxb., *Boswellia frereana* Birdw., *Boswellia neglecta* S. Moore and *Boswellia rivae* Engl. PhD thesis, Universität Hamburg, Germany.

Bekana, D., Kebede, T., Assefa, M., Kassa, H. 2014. Comparative phytochemical analyses of resins of Boswellia species (*B. papyrifera* (Del.) Hochst., *B. neglecta* S. Moore, and *B. rivae* Engl.) from Northwestern, Southern, and Southeastern Ethiopia. *ISRN Analytical Chemistry* 2014: 374678. http://dx.doi.org/10.1155/2014/374678

Bekele, T. 2008. Colonization history, phylogeography and conservation genetics of the gravely endangered tree species *Hagenia abyssinica* (Bruce) J.F. Gmel from Ethiopia. PhD thesis, Goettingen University, Germany.

Bekele, T., Gailing, O., Umer, M., Finkeldey, R. 2009. Chloroplast DNA haplotype diversity and postglacial recolonization of *Hagenia abyssinica* (Bruce) J.F. Gmel. in Ethiopia. Plant Syst Evol. 280, 175 –185.

Birhane, E., Sterck, F. J., Fetene, M., Bongers, F., Kuyper, T.W. 2012. Arbuscular mycorrhizal fungi enhance photosynthesis, water use efficiency, and growth of frankincense seedlings under pulsed water availability conditions. Oecologia 169, 895–904.

Booy, G., Hendriks, R.J.J., Smulders, M.J.M., van Groenendael, J.M., Vosman, B. 2000. Genetic diversity and the survival of populations. Plant Biology 2, 379-395. doi 10.1055/s-2000-5958

Brondani, R.P.V., Brondani, C., Tarchini, R., Grattapaglia, D. 1998. Development, characterization and mapping of microsatellite markers in *Eucalyptus grandis* and *E. urophylla*. Theoretical and Applied Genetics 97, 816-827.

Bogaert, J., Barima, Y. S.S., Mongo, L.I.W., Bamba, I., Mama, A., Toyi, M. Lafortezza, R. 2011. Forest Fragmentation: Causes, Ecological Impacts and Implications for Landscape Management. In Li, C., Lafortezza, R., Chen, J. Landscape Ecology in Forest Management and Conservation.: Challenges and Solutions for Global Change, Higher Education Press, Beijing, pp. 273-396.

Brownstein, M.J., Carpten, J.D., Smith, J.R. 1996. Modulation of Non-Templated Nucleotide Addition by Taq DNA Polymerase: Primer Modifications that Facilitate Genotyping. BioTechniques 20, 1004-1010.

Burczyk, .J, Adams, W.T., Birkes, D.S., Chybicki, I.J. 2006. Using Genetic Markers to Directly Estimate Gene Flow and Reproductive Success Parameters in Plants Based on Naturally Regenerated Seedlings. *Genetics* 173, 363-372.

Burcham, J. Frankincense Fit for a King (One, Anyway). 2011. New York times, http://www.nytimes.com/2011/12/08/garden/replicating-the-slightly-plantable-gifts-of-the-magi-in-the-garden.html#

Buzatti, R.S.O., Ribeiro, R.A., de Lemos Filho J.P., Lovato, M.B. 2012. Fine-scale spatial genetic structure of *Dalbergia nigra* (Fabaceae), a threatened and endemic tree of the Brazilian Atlantic Forest. Genetics and Molecular Biology 35(4), 838-846.)

Castoe, T.A., Poole, A.W., de Koning, A.P.J., Jones, K.L., Tomback, D. 2012. Rapid Microsatellite identification from Illumina paired-end genomic sequencing in two birds and a snake. *PLoS ONE*7: e30953. doi: 10.1371/journal.pone.0030953.

Cavers, S., Navarro, C. Lowe, A.J. 2005. Targeting genetic resource conservation in widespread species: a case study of *Cedrela odorata* L. Forest Ecology and Management 197, 285–294.

Chase, M., Kessel, R., Bawa, K. 1996. Microsatellite markers for population and conservation genetics of tropical trees. American Journal of Botany 83(1),51-57.

Chen, X-Y., Fan, X-X., Hu, X-S. 2008. Roles of seed and pollen dispersal in natural regeneration of *Castanopsis fargesii* (Fagaceae): Implications for forest management. Forest Ecology and Management 256, 1143–1150.

Chidumayo, E., Marunda, C. 2010. Dry forests and woodlands in Sub-Saharan Africa: context and challenges. in Emmanuel N. Chidumayo, E and Gumbo, D.J. (eds.). The dry forests and woodlands of Africa managing for products and services. Earthscan, London, pp. 1–10.

Chirwa, P.W., Larwanou, M., Syampungani, S. AbalolaA, F.D. 2015. Management and restoration practices in degraded landscapes of Southern Africa and requirements for up-scaling. International Forestry Review 17, (S3) 2015 31.

Chybicki, I.J., Burczyk, J. 2010b. Realized gene flow within mixed stands of *Quercus robur* L. and *Q. petraea* (Matt.) L. revealed at the stage of naturally established seedling. Molecular Ecology 19, 2137–2151.

Conesa, A., Götz, S., García-Gómez, J.M., Terol, J., Talón, M., Robles, M. 2005. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics 21,3674-6.

Coppen, J.J.W. 2005. Overview of international trades and markets. In: Chikamai, B. and, Casadei, E. eds. Production and marketing of gum resins: frankincense, myrrh and opoponax. Network for Natural Gums and Resins in Africa (NGARA), Publication Series No. 5, Nairobi, Kenya: NGARA, KEFRI, 5–34.

Cortes, M.C., Uriarte, M., Lemes, M.R., Gribel, R., Kress, W.J., Smouse, P.E., Bruna, E.M. 2013. Low plant density enhances gene dispersal in the Amazonian understory herb *Heliconia acuminata*. Molecular Ecology 22, 5716–5729.

Curtu, A.L., Craciunesc, I., Enescu, C.M., Vidalis, A., Sofletea. N. 2015. Fine-scale spatial genetic structure in a multi-oak-species (*Quercus* spp.) forest. Journal of Biogeosciences and Forestry 8, 324-332.

Debout,G.D.G., Doucet, J-L. Hardy, O.J. 2011. Population history and gene dispersal inferred from spatial genetic structure of a Central African timber tree, *Distemonanthus benthamianus* (Caesalpinioideae). Heredity 106, 88-99.

Derero, A. 2007.Genetic variation in *Cordia afrcana* Lam. In Ethiopia. PhD thesis, GoettingenUniversity, Germany.

Derero, A. Gailing, O. Finkeldey, R. 2010. Maintenance of genetic diversity in *Cordia africana* Lam., a declining forest tree species in Ethiopia. Tree Genetics &Genomes.http://www.springerlink.com/content/754412q52l2rr91x/fulltext.pdf

Duminil, J., Brown, R.P., Ewédjè, E.B.K., Mardulyn, P., Doucet, J., Hardy, O.J., 2013. Large-scale pattern of genetic differentiation within African rainforest trees: insights on the roles of ecological gradients and past climate changes on the evolution of *Erythrophleum* spp (Fabaceae). BMC Evolutionary Biology 13,195. doi: 10.1186/1471-2148-13-195

Edwards, C.E., Parchman, T.L., Weekley, C. 2011. Assembly, gene annotation and marker development using 454 floral transcriptome sequences in *Ziziphus celata* (Rhamnaceae), a Highly Endangered, Florida Endemic Plant. DNA Research 19, 1–9.

EFAP, 1994. Ethiopian Forestry Action Program. Final Report, Ministry of Natural Resources Development and Environmental Protection, Addis Ababa, Ethiopia.

Ekblom, R., Galindo, J. 2011. Applications of next generation sequencing in molecular ecology of non-model organisms. Heredity 107,1-15. doi:10.1038/hdy.2010.152.

EPA. 2010.Afar national regional state programme of plan on Adaptation to climate change. Environmental Protection Authority of the Federal Democratic Republic of Ethiopia, Semera, Ethiopia

Ersts, P.J. [Internet]. Geographic distance matrix generator (version 1.2.3). American Museum of Natural History, Center for Biodiversity and Conservation. http://biodiversityinformatics.amnh.org/open_source/gdmg. Accessed on 2014-8-6.

Eshete, A., Sterck, F.J., Bongers, F. 2011. Diversity and production of Ethiopian dry woodlands explained by climate - and soil - stress gradients. Forest Ecology and Management 261(9), 1499-1509.

Eshete, A., Teketay, D. Hulten, N.2005. The socioeconomic status of *B. papyrifera* (Del.) Hochst. in Northern Ethiopia: The case of northern Gonder Zone. Forests, Trees and livelihood 15, 55-74.

Evanno, G., Regnaut, S., Goudet, J., 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Molecular Ecology 14, 2611–2620. doi: 10.1111/j.1365-294X.2005. 02553.x

Evans, M. 2013. The cancer-killing properties of frankincense in ovarian cancer. http://soundcloud.com/university-of-leicester/mark-evans)

Fan, L., Zhang, M.Y., Liu, Q.Z., Li, L.T., Song, Y., Wang, L.F., Zhang, S.L., Wu, J. 2013. Transferability of Newly Developed Pear SSR Markers to Other Rosaceae Species. Plant Molecular Biology Reporter 31,1271–1282. doi: 10.1007/s11105-013-0586-z

FAO, 1984. Land use, Production regions and farming systems: Assistance to land use planning. Technical report No. 3, Rome, Italy.

FAO, 2003. Role of Planted Forests and Trees Outside Forests in Sustainable Forest Management in the Republic of Ethiopia. Planted Forests and Trees Working Papers, Working Paper 29. Forest Resources Development Service, Forest Resources Division. FAO, Rome Italy.

FAO, 2010. Global Forest Resources Assessment 2010 Main report. FAO Forestry paper no. 163. Rome, Italy.

FAO, 2014. The state of world's forest genetic resources. Rome Italy.

FAO. 2005. Global forest resources assessment 2005, Rome, Italy.

Farah, M.H. 2008. Non-timber forest products (NTFP) extraction in arid environments: Landuse change, frankincense production and the sustainability of *Boswellia sacra* in Dhofaar (Oman). PhD Dissertation, the University of Arizona, United States of America.

FARM Africa and SOS Sahel. 2008. Community profile and settlement dynamics in four woredas of Oromiya National Regional State: Dallo Mena, Harana Buluq, Goba and Nansabo. Addis Ababa, Ethiopia.

FARM Africa, 2015. Making forest conservation benefit local communities: participatory forest management in Ethiopia.http://www.farmafrica.org/downloads/participatory-forest-management-in-ethiopia.pdf.

Farwig, N., Braun, C., Böhning-Gaese, K. 2008. Human disturbance reduces genetic diversity of an endangered tropical tree, *Prunus africana* (Rosaceae). *Conservation Genetics* 9, 317–326.

Fenster, C.B. 1991. Gene flow in *Chamaecrista fasciculata*(Leguminosae). I. Gene dispersal. Evolution, 45, 398–409.

Fichtl, R., Admasu, A. 1994. Honeybee flora of Ethiopia. Margraf Verlag, Weikershem pp 510.

Frankham, R., Ballou, J.D., Briscoe, D.A. 2010. Introduction to conservation genetics (2nd edition). Cambridge University Press, United Kingdom. pp 644.

Frankham, R., 1995. Effective population size/adult population size ratios in wildlife: a review. Genetics Research (Camb.) 66, 95–107.

Frankham, R., Ballou, J.D., Briscoe, D.A. 2002. Introduction to conservation genetics (2nd edition). Cambridge University Press, New York. Pp 617.

Fulton, T.M., Chunwangse, J., Tanksley, S.D., 1995. Microprep protocol for extraction of DNA from tomato and herbaceous plants. Plant Molecular Biology Reporter 13, 207–209.

Gebrehiwot, K., Muys, B., Haile, M., Mitloehner, R., 2003. Introducing *Boswellia papyrifera* (Del.) Hochst and its non-timber forest product, frankincense. International Forestry Review 5, 348-353.

Girma, A., Skidmorea, A.K., de Biea, C.A.J.M., Bongers, F., Schlerf, M. 2013. Photosynthetic bark: Use of chlorophyll absorption continuum index to estimate Boswellia papyrifera bark chlorophyll content. International Journal of Applied Earth Observation and Geoinformation 23, 71–80

González-Martínez, S.C., Krutovsky, K.V., Neale, D.B. 2006. Forest-tree population genomics and adaptive evolution. *New Phytologist* 170, 227–238.

Groenendijk, .P, Eshete, W., Sterk, F., Zuidman, P.and Bongers F. 2012. Limitations to sustainable Frankincense production: blocked regeneration, high adult mortality, and declining population. Journal of Applied Ecology 49, 164–173.

Groom, N., 1981. Frankincense and Myrrh: a study of the Arabian incense trade. Longman. 285 pp.

Haile, G., Gebrehiwot, K., Lemenih, M. Bongers, F. 2011. Time of collection and cutting sizes affect vegetative propagation of *Boswellia papyrifera*(Del.) Hochst through leafless branch cuttings. Journal of Arid Environments 75, 873-877

Hamrick, J.L., Godt, M.J.W. 1996. Effects of life history traits on genetic diversity in plant species. Philos. Trans. R. Soc. Lond. B. 351, 1291–1298. dOI: 10.1098/rstb.1996.0112

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

Hamrick, J. L., Milton, J. B., Linhart, Y. B. 1979. Levels of Genetic Variation in Trees: Influence of life history characteristics. Proceedings of the Symposium on Isozymes of North American Forest Trees and Forest Insects. July 27, 1979, Berkeley, California.

Hamrick, J.L. 2004. Response of forest trees to global environmental changes. Forest Ecology and Management 197, 323–335.

Hardy, O.J., Maggia, L., Bandou, E., Breyne, P., Caron, H., Chevallier, M.E., Doligez, A., Dutech, C., Kremer, A., Latouche-Halle, C., Troispoux, V., Veron, V., Degen, B. 2006. Fine-scale genetic structure and gene dispersal influences in 10 Neotropical tree species. Molecular Ecology 15, 559–571.

Hardy, O.J., Vekemans, X. 2002. SPAGeDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. Molecular Ecology Notes 2: 618-620.

Harvey, C.A., Medina, A., Sa´nchez, D.M., Vi´lchez, S., Herna´ndez, B., Sanez, J.C., Maes, J.M., Casanoves, F., Sinclair, F.L. 2006. Patterns of animal diversity in

different forms of tree cover in agricultural landscapes. Ecological Applications 16(5), 1986–1999.

Holderegger, R., Kamm, U., Gugerli, F. 2006.Adaptive vs. neutral genetic diversity: implications for landscape genetics. Landscape Ecology 21,797–807.

Honnay, O., Jacquemyn, H. 2007. Susceptibility of common and Rare Plant Species to the genetic Consequences of Habitat fragmentation. Conservation Biology 21(3), 823–831. DOI: 10.1111/j.1523-1739.2006. 00646.x

Janzen, D.H. 1988. Tropical dry forests: the most endangered major tropical ecosystem. In Wilson, E.O. and Peter F.M. (eds.). Biodiversity, National Academy Press, Wshington, DC, pp.130-137.

Jensen, J.L., Bohonak, A.J., Kelley, S.T., 2005. Isolation by distance, web service. BMC Genetics 6, 13. doi: 10.1186/1471-2156-6-13

Joanne, P., Asnake, A., Kasaye, H., 2005. Livelihoods/Emergency assessment in Afar region. Oxfam International. Addis Ababa, Ethiopia. pp 44.

Jurka, J., Kapitonov, V.V., Pavlicek, A., Klonowski, P., Kohany, O., Walichiewicz, J. 2005. Repbase Update, a database of eukaryotic repetitive elements. *Cytogenetic and Genome Research* 110,462-467

Kamm, U., Gugerli, F., Rotach, P., Edwards, P., Holderegger, R. 2010. Open areas in a landscape enhance pollen-mediated gene flow of a tree species: evidence from northern Switzerland. Landscape Ecol 25:903–911.

Karp, A., Kresovich, S., Bhat, K.V., Ayad, W.G., Hodgkin, T. 1997. Molecular tools in plant genetic resources conservation: A guide to the technologies; In: IPGRI Technical Bulletin No. 2. International Plant Genetic Resources Institute, Rome Italy.

Kebede, M., Ehrich, D., Taberlet, P., Nemomissa, S., Brochmann, C. 2007. Phylogeography and conservation genetics of a giant lobelia *Lobelia giberroa* in Ethiopian and Tropical East African mountains**.** Molecular Ecology 16,1233–1243**.**

Kennarda, D.K., Gouldb, K., Putza, F.E., Fredericksenc, T.S., Morales, F. 2002. Effect of disturbance intensity on regeneration mechanisms in a tropical dry forest. Forest Ecology and Management 162, 197–208.

Khajuria, A.,Gupta, A.,Suden, P., Singh, S., Malik, F., Singh, J. 2008. Immunomodulatory activity of biopolymeric fraction BOS 2000 from Boswellia serrata. Phytother Res 22(3), 340-8.

King, R., Zalucki, J.M. 2012. Potential Inbreeding in a Small Population of a Mass Flowering Species, *Xanthorrhoea johnsonii* (Xanthorrhoeaceae): Is Your Mother My Father? American Journal of Plant Sciences 3, 303–312. doi: 10.4236/ajps.2012.33036.

Koch, P., Platzer, M., Downie, B.R. 2014. RepARK--de novo creation of repeat libraries from whole-genome NGS reads. *Nucleic Acids Research* 42(9):e80. doi: 10.1093/nar/gku210.

Kramer, A.T., Ison, J.L., Ashley, M.V., Howe, H.F., 2008. The Paradox of Forest Fragmentation Genetics. Conservation Biology 22(4), 878–885.

Krutovsky, K.V., Burczyk, J., Chybicki, I., Finkeldey, R., Pyhäjärvi, T., Robledo-Arnuncio, J. J. 2012. Gene flow, spatial structure, local adaptation, and assisted migration in trees. In Schnell, R.J. andPriyadarshan, P.M. (Eds.). Genomics of Tree Crops, Springer, New York, pp. 71-116.

Kubsa, A., Mariame, A., Amante, G., Lipp, H–J., Tadesse, T. 2003. Wajib: an alternative forest conservation approach for Ethiopia's forests. XII World Forestry Congress., Quebec, Canada.

Kubsa, A., Tadesse, T. 2003. Granting exclusive user rights to the forest dwellers in the state owned forest: the WAJIB approach in Ethiopia. Proceedings of second international workshop on participatory forestry in Africa. February 18-22, Arusha, Tanzania.

Lance, S.L., Love, C.N., Nunziata, S.O., O'Bryhim, J.R., Scott, D.E., Wesley Flynn, R.W., Jones, K.L. 2013. 32 species validation of a new Illumina paired-end approach for the development of microsatellites. *PLoS ONE* 8(11): e81853. doi:10.1371/journal. pone.0081853

Langmead, B., Salzberg, S.L. 2012. Fast gapped-read alignment with Bowtie 2. *Nature Methods* 9:357–359. doi:10.1038/nmeth.1923

Lawton-Rauh, A. 2008. Demographic processes shaping genetic variation. Current opinion in plant biology. 11(2), 103-109.

Ledig, F. T., Fryer, J. H., 1972. A pocket of variability in *Pinus rigida* Mill. Evolution 26, 259-266.

Lemenih, M., Abebe, T., Mats, O. 2003. Gum and resin resources from some *Acacia*, *Boswellia*, and *Commiphora* species and their economic contributions in Liban, South-East Ethiopia. *Journal of Arid Environments* 55, 465–482.

Lemenih, M., Kassa, H. 2011. Management guide for sustainable production of frankincense. CIFOR, Bogor, Indonesia

Lemenih, M., Arts, B., Wiersum, K.F., Bongers, F. 2014. Modelling the future of *Boswellia papyrifera* population and its frankincense production. Journal of Arid Environments 105, 33-40. DOI: 10.1016/j.jaridenv.2014.02.006

Lemenih, M., Feleke, S., Tadesse, W. 2007. Constraints to smallholders production of frankincense in Metema district, North-western Ethiopia. Journal of Arid Environments 71, 393–403.

Lemenih, M., Teketay, D. 2003. Frankincense and myrrh resources of Ethiopia: medicinal and industrial uses. SINET Ethiopian Journal of Science26,161-172.

Lemes, M.R., Gribel, R., Proctor, J., Grattapaglia, D., 2003. Population genetic structure of mahogany (*Swietenia macrophylla* King, Meliaceae) across the Brazilian Amazon, based on variation at microsatellite loci: implications for conservation. Molecular Ecology 12, 2875–2883. doi: 10.1046/j.1365-294X.2003. 01950.x

Li, R., Zhu, H., Ruan, J., Qian, W., Fang, X., Shi, Z., Li, Y., Li, S., Shan, G., Kristiansen, K. Li, S., Yang. H., Wang, J., Wang, J. 2010. De novo assembly of human genomes with massively parallel short read sequencing. *Genome Research* 20,265–272.

Lin, H.K., Suhail, M.M., Fung, K.M., Woolley, C.L., Young, D.G. 2013. Extraction of biologically active compounds by hydrodistillation of Boswellia species gum resins for anti-cancer therapy. OA Alternative Medicine. 1(1):4

Lindenmayer, D.B., Fischer, J. 2006. Habitat Fragmentation and Landscape Change: An Ecological and Conservation Synthesis. Island press, New York, USA.

Liu, B., Yuan, J., Yiu, S.M., Li, Z., Xie, Y., Chen, Y., Shi, Y., Zhang, H., Li, Y., Lam, T.W., Luo, R. 2012. COPE: an accurate k-mer-based pair-end reads connection tool to facilitate genome assembly. *Bioinformatics* 28:2870–2874.

Liu, Y., Schröder, J., Schmidt, B. 2012. Musket: a multistage k-mer spectrum-based error corrector for Illumina sequence data. *Bioinformatics* 29:308-315. doi: 10.1093/bioinformatics/bts690

Mihretie, Z., Schueler, S., Konrad, H., Bekele, E., Geburek, T. 2015. Patterns of genetic diversity of *Prunus africana* in Ethiopia: hot spot but not point of origin for range-wide diversity. Tree Genetics & Genomes ,11(6):1-13. DOI: 10.1007/s11295-015-0945-z

Moran, E.V., Clark, J.S. 2012. Between-Site Differences in the Scale of Dispersal and Gene Flow in Red Oak. PLoS ONE 7(5), e36492. doi:10.1371/ journal. pone.0036492

Moran, P. 2002. Current conservation genetics: building an ecological approach to the synthesis of molecular and quantitative genetic methods. Ecology of Freshwater Fish 11, 30–55.

Moran, E.V., Clark, J.S. 2011. Estimating seed and pollen movement in a monoecious plant: a hierarchical Bayesian approach integrating genetic and ecological data. Molecular Ecology 20, 1248–1262.

Mortimore, M., Anderson, S., Cotula, L., Davies, J., Faccer, K., Hesse, C., Morton, J., Nyangena, W., Skinner, J. and Wolfangel C. 2009. Dryland Opportunities: a new paradigm for people, ecosystems and development, IUCN, Gland, Switzerland

Moussaieff, A., Shein, N. A., Tsenter, J., Grigoriadis, S., Simeonidou, C., Alexandrovich, A.G., Trembovler, V., Ben-Neriah, Y., Schmitz, M. L., Fiebich B. L., Munoz, E., Mechoulam, R., Shohami, E. *2008.* Incensole acetate: a novel

neuroprotective agent isolated from *Boswellia carterii. Journal of Cerebral Blood Flow & Metabolism*28, 1341–1352.

Nakasugi, K., Crowhurst, R., Bally, J., Waterhouse, P. 2014. Combining transcriptome assemblies from multiple de novo assemblers in the allo-tetraploid plant *Nicotiana benthamiana*. *PLoS One* 9(3): e91776. doi: 10.1371/journal.pone.0091776.

Namkoong, G. 1988. Sampling for germplasm collections. HortScience 23, 79-81.

National Conservation Strategy Secretariat (NCSS). 1993. National conservation strategy Vol. 1: National policy on the resources base, its utilization and planning for sustainability. Addis Ababa, Ethiopia. 131p.

Ng, C., Lee, S., Ng, K., Muhammad, N., Ratnam, W. 2009. Mating system and seed variation of Acacia hybrid (A. mangium ×A. auriculiformis). Journal of Genetics 88, 25–31.

Ni, X., Suhail, M.M., Yang, Q.,, Cao, A., Fung, K-M., Postier, R.G., Woolley, C., Young, G., Zhang, J., Lin, H-K. 2012. Frankincense essential oil prepared from hydrodistillation of *Boswellia sacra* gum resins induces human pancreatic cancer cell death in cultures and in a xenograft murine model. BMC Complement Altern Med. 12, 253. doi: 10.1186/1472-6882-12-253.

NMA. 2007. Climate change national adaptation programme of action (NAPA) of Ethiopia. The Federal Democratic Republic of Ethiopia, Ministry of Water Resources, Addis Ababa, Ethiopia.

Novaes, R.M.L., Rodrigues, J.G., Lovato, M.B., 2009. An efficient protocol for tissue sampling and DNA isolation from the stem bark of Leguminosae trees. Genetics and Molecular Research 8 (1): 86-96. http://www.funpecrp.com.br/gmr/year2009/vol8-1/pdf/gmr542.pdf

Nybom, H., Weising, K., Rotter, B. 2014. DNA fingerprinting in botany: past, present, future. *Investigative Genetics* 5:1. doi:10.1186/2041-2223-5-1

Ogbazghi, W., Rijkers, T., Wessel, M., Bongers, F., 2006. Distribution of the frankincense tree Boswellia papyrifera in Eritrea: the role of environment and land use. Journal of Biogeography 33, 524–535. DOI: 10.1111/j.1365-2699.2005.01407.x

Omondi, S.F., Eliud, K., Dangasuk, O.G., Chikamai, B., Odee, D.W., Cavers, S., Khasa, D.P. 2010. Genetic Diversity and Population Structure of *Acacia senegal* (L) Willd. in Kenya. Tropical Plant Biology 3, 59–70. DOI 10.1007/s12042-009-9037-2

Ouborg, N.J., Angeloni, F. 2010. An essay on the necessity and feasibility of conservation genomics. Conservation Genetics 11:643–653.

Palmer, J.D. 1987. Chloroplast DNA Evolution and Biosystematic Uses of Chloroplast DNA Variation Author(s): Source: The American Naturalist 130, Supplement: Plant Molecular Evolution, pp. S6-S29.

Pastorelli, R., Smulders, M.J.M., Van 't Westende, W.P.C., Vosman, B., Giannini, R., Vettori, C., Vendramin, G.G. 2003. Characterisation of microsatellite markers in *Fagus sylvatica* L. and *Fagus orientalis* Lipsky. *Molecular Ecology Notes* 3, 76-78.

Paul, M., Brüning, G., Bergmann, J., Jauch, J. 2012. A thin-layer chromatography method for the identification of three different olibanum resins (Boswellia serrata, Boswellia papyrifera and Boswellia carterii, respectively, Boswellia sacra).Phytochem Anal. 23(2), 184-9.

Peakall, R., Ruibal, M., Lindenmayer, D.B. 2003. Spatial autocorrelation analysis offers new insights into gene flow in the Australian Bush Rat, *Rattus fuscipes*. Evolution57(5), 1182–1195.

Peakall, R., Smouse, P.E. 2006. GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Molecular Ecology Notes 6, 288–295.

Peakall, R., Smouse, P.E. 2012. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research - an update. Bioinformatics 28, 2537–2539. DOI: 10.1093/bioinformatics/bts460

Petit, R.J.A, Mousadik, E., Pons, O. 1998. Identifying populations for conservation on the basis of genetic markers. Conservation Biology12, 844–855.

Porras-Hurtado, L., Ruiz, Y., Santos, C., Phillips, C., Carracedo, Á., Lareu, M.V. 2013. An overview of STRUCTURE: applications, parameter settings, and supporting software. Frontiers in Genetics 4, 98. DOI: 10.3389/fgene.2013.00098

Primmer, C.R. 2009. From Conservation Genetics to Conservation Genomics. *Annals of the New York Academy of Sciences* 1162,357–368.

Pritchard, J.K., Stephens, M., Donnelly, P. 2000. Inference of Population Structure Using Multilocus Genotype Data. Genetics 155, 945-959.

Raju, A.J.S., Lakshmi, P.V.K., Ramana, V., Chandra, P.H. 2012. Entomophily, ornithophily and anemochory in the self-incompatible *Boswellia ovalifoliolata* Bal. & Henry (Burseraceae), an endemic and endangered medicinally important tree species. Journal of Threatened Taxa 4(7), 2673–2684.

Reusing, M. 2000. Change detection of natural high forest in Ethiopia using remote sensing and GIS techniques. International Archives of Photogrammetry and Remote Sensing XXXIII, Part B7. Amsterdam

Richards, C.M. 2000. Inbreeding Depression and Genetic Rescue in a Plant Metapopulation. The American naturalist 155 (3), 383-394.

Rozen, S., Skaletsky, H.J. 2000. Primer3 on the WWW for general users and for biologist programmers. In: Krawetz S, Misener S (eds.) Bioinformatics methods and protocols: Methods in molecular biology. Totowa NJ, Humana Press, 365-386.

Savolainen, O., Pyhajarvi, T., Knurr, T. 2007. Gene flow and local adaptation in trees. Annu. Rev. Ecol. Evol. Syst. 38, 595-619.

Schaal, B.A., Hayworth, D.A, Olsen, K.M., Rauscher, J.T., Smith, W.A. 1998. Phylogeographic studies in plants: problems and prospects. Molecular Ecology 7, 465-474.

Schmieder, R., Edwards, R. 2011. Quality control and preprocessing of metagenomic datasets. *Bioinformatics* 27, 863-864.

Schroeder, J.W., Tran, H.T., Dick, C.W. 2014. Fine scale spatial genetic structure in *Pouteria reticulata* (Engl.) Eyma (Sapotaceae), a dioecious, vertebrate dispersed tropical rain forest tree species. Global Ecology and Conservation. Global Ecology and Conservation 1, 43–49.

Schuelke, M., 2000. An economic method for the fluorescent labelling of PCR fragments: A poor man's approach to genotyping for research and high-throughput diagnostics. Nature Biotechnology 18, 233–234.

Sebbenn, A.M., Carvalho, A.C.M., Freitas, M.L.M., Moraes, S.M.B., Gaino, A.P.S., da Silva, C.J.M., Jolivet, C., Moraes, M.L.T. 2011. Low levels of realized seed and pollen gene flow and strong spatial genetic structure in a small, isolated and fragmented population of the tropical tree *Copaifera langsdorffii* Desf. Heredity106, 134–145.

Segelbacher, G., Cushman, S.A., Epperson, B.K., Fortin, M-J., Francois, O., Hardy. O.J., Holderegger, R., Taberlet, P., Waits, L.P., Manel, S. 2010. Applications of landscape genetics in conservation biology: concepts and challenges. Conservation Genetics 11, 375–385.

Selkoe, K. A., Toone, R.J. 2006. Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. *Ecology Letters* 9: 615–629. doi: 10.1111/j.1461-0248.2006.00889x.

Sertse, D., Gailing, O., Eliades, N-G.,Finkeldey, R. 2011. Anthropogenic and natural causes influencing population genetic structure of *Juniperus procera*Hochst. ex Endl. in the Ethiopian highlands. Genet Resour Crop Evol 58, 849–859 DOI 10.1007/s10722-010-9623-z

Shahin, A., van Gurp, T., Peters, S.A., Visser, R.G.F., van Tuyl, J.M., Arens, P. 2012. SNP markers retrieval for a non-model species: a practical approach. *BMC Research Notes* 5:79. doi:10.1186/1756-0500-5-79.

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

Smouse P, Peakall R, Gonzales E. 2008. A heterogeneity test for fine-scale genetic structure. Molecular Ecology 17: 3389–3400. doi: 10.1111/j.1365-294X.2008.03839.

Smulders, M.J.M., Bredemeijer, G., Rus-Kortekaas, W., Arens, P., Vosman, B. 1997. Use of short microsatellites from database sequences to generate

polymorphisms among *Lycopersicon esculentum* cultivars and accessions of other *Lycopersicon* species. Theoretical and Applied Genetics 94: 264–272.

Smulders, M.J.M., Cottrell, J.E., Lefèvre, F., van der Schoot, J., Arens, P., Vosman, B., Tabbener, H.E., Grassi, F., Fossati. T., Castiglione, S., Krystufek, V., Fluch, S., Burg, K., Vornam, B., Pohl. A., Gebhardt, K., Alba, N., Agúndez, D., Maestro, C., Notivol, E., Volosyanchuk, R., Pospíšková, M., Bordács, S., Bovenschen, J., van Dam, B.C., Koelewijn, H-P., Halfmaerten, D., Ivens, B., van Slycken. J, Vanden Broeck, A., Storme, V., Boerjan, W. 2008. Structure of the genetic diversity in Black poplar (*Populus nigra* L.) populations across European river systems: consequences for conservation and restoration. *Forest Ecology and Management* 255, 1388-1399.

Smulders, M.J.M., Esselink, G.D., Everaert, I., De Riek, J., Vosman, B. 2010. Characterisation of sugar beet (*Beta vulgaris* L. ssp. vulgaris) varieties using microsatellite markers. *BMC Genetics* 11:41. doi:10.1186/1471-2156-11-41.

Smulders, M.J.M, Van der Schoot, J., Arens, P., Vosman, B. 2001. Trinucleotide repeat microsatellite markers for black poplar (*Populus nigra* L.). *Molecular Ecology Notes* 1, 188-190.

Smulders, M.J.M, Vukosavljev, M., Shahin, A., van de Weg, W. E., Arens, P. 2012. High throughput marker development and application in horticultural crops. *Acta Horticulturae (ISHS)* 961:547-551. http://www.actahort.org/books/961/961_72.htm

Sorensen, F.C., Miles, R.S. 1982. Inbreeding depression in height, height growth, and survival of Douglas-fir, Ponderosa pine, and Noble fir to 10 Years of Age Forest Sci., 28(2), 283-292.

Sunnichan, V.G., Morhanram, H.Y., Shivanna, K.R., 2005. Reproductive biology of *Boswellia serrata*, the source of salaiguggul, an important gum-resin. Botanical Journal of the Linnean Society 147, 73–82.

Tesfaye, Y., Roos, A., Campbell, B.M., Bohlin, F. (2010). Forest incomes and poverty alleviation under participatory forest management in the Bale Highlands, Southern Ethiopia. *International Forestry Review,* 12(1), 66-77.

Teshome et al. in prep Population structure and density of *Boswellia papyrifera* in the dry woodlands of Benishangul Gumuz Regional State, Ethiopia**.**

Teshome, M. 2013. Structure and composition of woody plants in *Boswellia* dominated woodland of Western Ethiopia. MSc Thesis.

Tiep, H.V., Quang, T.H., De Groot, G.A., Boot, R.G.A., Zuidema, P.A. In prep. Forest fragmentation causes genetic erosion and differentiation in spite of high genetic exchange in remnant populations of a threatened tropical tree species in Vietnam

Tolera, M., Sass-Klaassen, U., Eshete, A., Bongers, F., Sterck, F.J., 2013. Frankincense tree recruitment failed over the past half century. Forest Ecology and Management 30, 65–72. doi: 10.1016/j.foreco.2013.04.036

Trimble, M. J., Aarde, R. Supporting conservation with biodiversity research in sub-Saharan Africa's human-modified landscapes. Biodivers Conserv, doi10.1007/s10531-014-0716-4

Tucker, A.O., 1986. Frankincense and Myrrh. Economic Botany 40, 425-433. doi 10.1007/BF02859654

Tuskan, G.A., Gunter, L.E., Yang, Z.K., Yin, T., Sewell, M.M., DiFazio, S.P. 2004. Characterization of microsatellites revealed by genomic sequencing of *Populus trichocarpa*. *Canadian Journal of Forestry Research* 34, 85–93

Van der Merwe, M., McPherson, H., Siow, J., Rossetto, M. 2014. Next-Gen phylogeography of rainforest trees: exploring landscape-level cpDNA variation from whole-genome sequencing. *Molecular Ecology Resources* 14, 199–208.

Vekemans, X., Hardy, O.J. 2004. New insights from fine-scale spatial genetic structure analyses in plant populations. Molecular Ecology 13, 921-935.

Vicedomini, R., Vezzi, F., Scalabrin, S., Arvestad, L., Policriti, A. 2014. GAM-NGS: genomic assemblies merger for next generation sequencing. *BMC Bioinformatics* 14(Suppl 7): S6 doi:10.1186/1471-2105-14-S7-S6.

Vieira, F.A., Fajardo, C.G., de Souza, A.M., de Carvalho, D. 2010. Landscape-Level and Fine-Scale Genetic Structure of the Neotropical Tree *Protium spruceanum* (*Burseraceae*). International Journal of Forestry, doi:10.1155/2010/120979

Vollesen, K., 1989. Burseraceae. In: Hedberg, I. and Edwards, S. eds. Flora of Ethiopia, Volume 3. National Herbarium, Addis Ababa University, Addis Ababa and Uppsala University, Uppsala. pp. 442-478.

Vukosavljev, M., Esselink, G.D., Van 't Westende, W.P.C., Cox, P., Visser, R.G.F., Arens, P., Smulders, M.J.M. 2014. Efficient development of highly polymorphic microsatellite markers based on polymorphic repeats in transcriptome sequences of multiple individuals.*Molecular Ecology Resources* (in press). doi:10.1111/1755-0998.12289

Wang, J., Li, Z., Guo, Q., Ren, G., Wu, Y. 2011. Genetic variation within and between populations of a desert poplar (*Populus euphratica*) revealed by SSR markers. Annals of Forest Science 68, 1143–1149. DOI 10.1007/s13595-011-0119-6

Wang, Y., Qin, Y., Du, Z., Yan, G., 2012. Genetic diversity and differentiation of the endangered tree *Elaeagnus mollis* Diels (*Elaeagnus* L.) as revealed by Simple Sequence Repeat (SSR) Markers. Biochemical Systematics and Ecology 40, 25–33. doi: 10.1016/j.bse.2011.09.009

White, F. 1983. The vegetation of Africa, a descriptive memoir to accompany the UNESCO/AETFAT/UNSO Vegetation Map of Africa (3 Plates, Northwestern Africa, Northeastern Africa, and Southern Africa, 1: 5,000,000). UNESCO, Paris.

White, G.M., Boshier, D.H. Powell, W. 2001. Increased pollen flow counteracts fragmentation in a tropical dry forest: An example from *Swietenia humilis* Zuccarini. Ecology 99 (4), 2038-2042.

Woldeamanuel, T. 2011. Dryland resources, livelihoods and institutions: Diversity and dynamics in use and management of gum and resin trees in Ethiopia. PhD thesis, Wageningen UR, The Netherlands.

Woody Biomass Inventory and Strategic Planning Project (WBISPP). 2004. Forest Resources of Ethiopia. Addis Ababa, Ethiopia.

Wright, S., 1943. Isolation by distance. Genetics 28(2), 114–138.

Yang, Y., Smith, S.A. 2013. Optimizing *de novo* assembly of short-read RNA-seq data for phylogenomics. *BMC Genomics* 14:328. doi:10.1186/1471-2164-14-328

Yineger, H., Schmidt D. J., Hughes, J.M. 2014. Genetic structuring of remnant forest patches in an endangered medicinal tree in North-western Ethiopia. BMC Genetics 15:31, http://www.biomedcentral.com/1471-2156/15/31

Young, A., Boyle, T., Brown, T. 1996. The population genetic consequences of habitat fragmentation for plants. Trends in Ecology and Evolution 11, (10), 413–418.

Zalapa, J.E., Cuevas, H., Zhu, H., Steffan, S., Senalik, D., Zeldin, E., McCown, B., Harbut, R., Simon, P. 2012. Using next-generation sequencing approaches to isolate simple sequence repeat (SSR) loci in the plant sciences. *American Journal of Botany* 99(2), 193-208. doi: 10.3732/ajb.1100394.

Zhang, Y., Ning, Z., Lu, C., Zhao, S., Wang, J., Liu, B. Xu, X., Liu, Y. 2013. Triterpenoid resinous metabolites from the genus Boswellia: pharmacological activities and potential species-identifying properties. Chemistry Central Journal 7:153, http://journal.chemistrycentral.com/content/7/1/153

SUMMARY

Summary

Forests provide many products and ecological services and thereby contribute to both local and national economies. Dry tropical forests and woodland vegetations are characterized by economically and ecologically beneficial plant species. *B. papyrifera* is a main tree species of the extensive Combretum-Terminalia dry forests/woodlands in Africa and it is one of the most important wild tree genetic resources of Ethiopia with current and potential economic importance, and with large ecological and societal values. The species produces a commercial frankincense which is internationally traded because of its value as ingredients in cosmetic, detergent, food flavor and perfumes productions, and because of its extensive use as incense during religious and cultural ceremonies in many parts of the world. Frankincense is increasingly valued for its potential in modern pharmaceutical uses such as boosting the body immune system and specially its ability to target cancer cells of different tissue origins (e.g. bladder tumor cells, breast cancer cells, ovarian cancer).

Dryland species are adapted to the seasonal pattern of rainfall and recurrent drought that prevail in these ecosystems. Disturbances and overexploitation of natural resources overpower the resilience the ecosystems and constitute potentially serious threats to the genetic resources (Janzen, 1988; Mortimore et al., 2009). Hence, the future value of forest resources is determined by the ways humans manage these resources in the ever changing environment in today's world. The forests with *B. papyrifera* has been increasingly overexploited at the expense of the economic benefit, and the wealth of ecological services it provides, now and potentially in the future. Populations have declined in size, are increasingly fragmented, have a regeneration which is blocked for the last 50 years, and have adult productive trees which are dying fast. Projections showed a 90 % loss in the coming 50 years of existing *B. papyrifera* trees in current populations and this is associated with a 50 % loss of frankincense production in 15 years time.

The threats have impacts on population genetic processes such as genetic drift, inbreeding, and gene flow. The rise in genetic drift, elevated level of inbreeding and limited gene flow subsequently causes a loss of genetic diversity, increase in genetic relatedness of the individuals within populations (Fine-scale spatial genetic structure (FSGS) and an increase in divergence among populations. Understanding the genetic diversity pattern is vital in determining the conservation strategy and sampling of the genetic resources.

This study addressed the conservation genetics of *B. papyrifera:* we developed the genetic tools and applied those to characterize the genetic diversity pattern of the species. The study considered 12 *B. papyrifera*populations in Ethiopia which vary in terms of disturbance regimes, population size, occurrence of regeneration and variation in stem density. We assessed the generational change in genetic diversity and structuring and used two cohort groups (adults and seedlings) in two populations from Western Ethiopia. Both populations were unique in having abundant small seedlings and saplings.

This study specifically aimed at 1) sequencing the genome and developing microsatellite markers 2) characterizing genetic diversity within and among *B. papyrifera* populations in Ethiopia 3) characterizing fine-scale spatial genetic structure and gene dispersal within *B. papyrifera* populations 4) assessing within-population genetic structure and the change in level of genetic diversity across adult and seedling generations.

We applied Illumina paired-end next generation sequencing and sequenced part of the *B. papyrifera* genome. Based on the short reads we developed 46 sets of microsatellite markers of which most are cross-transferable to two other Boswellia tree species: *B. popoviana and B. pirrotae.* The findings suggested the potential use of

the SSR markers for further genetic studies of *B. papyrifera* and the genetic characterization of closely related species.

We characterized the genetic diversity of the 12 populations and how it is partitioned among populations across Ethiopia. Most populations visited for sampling were degraded, lacking rejuvenation and young individuals. In the North eastern region some were depauperated (AF-BIR and KTD), with only few trees scattered across hilltops and farmland. The regeneration status of *B. papyrifra* hugely varies from population to population. In contrast to expectations we discovered populations in the Western region (Benishangul-Gumuz) with a high density of seedlings and saplings, indicating that natural regeneration still takes place. We assumed this to be due to less impact from fewer agricultural activities, low cattle population and a more suitable ecological conditions for regeneration in terms of soil and rainfall amount. We recommended detailed studies to be carried out in this area to understand the reasons for the abundant regeneration. This information can then be used to apply to the various parts of the distribution range where this species is under threat and therewith the future use of its important resin frankincense.

Despite the threats the populations are experiencing, we detected ample genetic variation in the adult trees of the populations we sampled from various parts of Ethiopia, including those from the most degraded populations. This diversity is likely due to the widespread distribution of *B. papyrifera*, the long life span of the speciesand the fact that the ongoing forest fragmentation and degradation processes in Ethiopia are relatively recent events. This means that the *B. papyrifera* trees we sampled still represent the diversity present in the formerly large and continuous forests that existed in the 19th century. The cross-generation diversity analysis suggests that no loss of genetic diversity nor selection for heterozygosity occurred across generations and that the diversity in adult cohorts is adequately represented in their progeny. The high level of genetic diversity and the low level

of inbreeding and extensive within population gene dispersal suggests that out-crossing is the predominant breeding system in *B. papyrifera.*

Low levels of population differentiation and isolation-by-distance patterns were detected. Based on STRUCTURE analysis, populations were grouped into four genetic clusters: the North eastern (NE), Western (W), North western (NW) and Northern (N) part of Ethiopia. The clusters corresponded to environmentally different conditions in terms of temperature, rainfall and soil conditions. We recommend that geographical distance among populations is the main factor to be considered in sampling for conservation. Additionally, however, also the ecological difference across populations should play a role during conservation planning.

Based on the genetic clustering of the populations we recommend a minimum of four conservation sites for *B. papyrifera,* representing the Western, North-western, Northern and North genetic clusters. Each of the conservation sites should include a population large enough not to lose genetic diversity in the subsequent generations due to genetic drift and inbreeding, the causes for loss of genetic diversity in small populations. The conservation units recommended herein might also serve to conserve other species in Combretum-Terminalia woodlands specially *Boswellia* species growing together within the same geographical range (*e.g. B. pirottae*).

This dissertation also assessed the FSGS of two populations in western Ethiopia to evaluate the extent of gene dispersal within populations and the change in genetic diversity across generations. We detected a low FSGS value in both populations and found that individuals are significantly related up a distance of 62 m. Based on these findings we recommend that seed collection for *ex situ* conservation and

plantation programmes should come from trees at least 60 m, but preferably 100 m apart, to be able to represent unrelated genotypes.

Outcrossing species like *B. papyrifera* may not be adapted to high levels of inbreeding and may therefore suffer more from increased bi-parental inbreeding caused by limited gene flow because of fragmentation. We highly recommend that conservation enables continued gene flow among fragmented populations and scattered remnant trees to ensure that the existing level of genetic diversity will continue to the future generations.

The study raises a number of conservation related issues that need further study. The existence of regeneration bottlenecks in most existing populations is an urgent prevailing problem that needs to be solved to ensure the continuity of the existing genetic diversity, species survival and sustainable production of frankincense. In the face of predicted climate change, understanding the species potential range expansion and determining the suitable habitat range are important aspects to be evaluated. The recent State-community joint forest management initiatives (PFM) involving the communities living in and around the forests in the actual use and management of the forests has become a promising forest management scenario in the country. Promoting this approach in different forests in the country including the dry forests and woodlands would enable to achieve conservation goals while meeting the livelihood demand of the local community for forest products.

SAMENVATTING

Bossen leveren producten en ecologische diensten en dragen daarmee bij aan lokale en nationale economieën. Droge tropische bossen bevatten economisch en ecologisch interessante plantensoorten. *Boswellia papyrifera* is een van de belangrijkste boomsoorten van het Combretum-Terminalia droge tropische bostype in Afrika. De soort is één van de belangrijkste genetische bronnen van Ethiopië, met huidig en potentieel economische belang en met grote ecologische waarden. *B. papyrifera* wordt gebruikt voor de productie van een hars (frankincense of wierook), dat internationaal wordt verhandeld omdat het een ingrediënt is in cosmetica, wasmiddelen, voedsel en parfums, en wereldwijd gebruikt als wierook tijdens religieuze en culturele ceremonies. Wierook wordt ook steeds meer gewaardeerd om zijn potentieel gebruik in farmaceutische toepassingen zoals het stimuleren van het menselijk immuunsysteem en speciaal het vermogen om zich te richten tegen kankercellen van diverse oorsprong (inclusief blaastumorcellen, borstkankercellen, en eierstokkanker).

Soorten uit het droge tropische bos zijn aangepast aan de afwisseling van regenseizoenen met droge periodes. Verstoringen en overexploitatie van natuurlijke hulpbronnen tasten de veerkracht van ecosystemen aan en vormen potentieel ernstige bedreigingen voor genetische bronnen. Om deze reden wordt de toekomstige waarde van het bosbestand bepaald door de manier waarop mensen deze middelen beheren. De bossen met *B. papyrifera* worden in toenemende mate overgeëxploiteerd, wat ten koste gaat van de economische voordelen en de rijkdom van de ecologische diensten die het levert, nu en in de toekomst. De populaties van deze boomsoort zijn afgenomen in omvang, zijn in toenemende mate gefragmenteerd, en sinds 50 jaar is regeneratie nagenoeg afwezig. Een modelstudie heeft laten zien dat daardoor 90% van de huidige populaties in de komende 50 jaar kan gaan verdwijnen. In de komende 15 jaar kan de wierookproductie uit de bestaande bossen al met 50% afnemen.

De bedreigingen kunnen leiden tot een verlies van genetische diversiteit, een verlies van genetische structuur, en een toename van de genetische verschillen tussen populaties. Inzicht in het patroon van genetische diversiteit is van vitaal belang bij het bepalen van een strategie voor in situ conservering, en om bemonstering van de genetische hulpbronnen te plannen voor ex situ conservering en regeneratie.

Deze studie vormt de start van conserveringsgenetica gericht op het behoud van *B. papyrifera*. We ontwikkelden genetische hulpmiddelen en pasten ze toe om de ruimtelijke verdeling van de genetische diversiteit van de soort te bepalen. In de studie bemonsterden we 12 *B. papyrifera* populaties in Ethiopië die verschilden qua verstoringsregime, omvang, het optreden van regeneratie, en dichtheid. De verschillen tussen generaties bomen werd onderzocht door twee cohorten (volwassen bomen en zaailingen) te bestuderen in twee populaties in West-Ethiopië. Beide populaties waren uniek omdat ze een overvloed aan kleine zaailingen en jonge boompjes hadden.

De studie was specifiek gericht op 1) het sequencen van het genoom en de ontwikkeling van microsatelliet merkers; 2) het karakteriseren van de genetische diversiteit binnen en tussen *B. papyrifera* populaties in Ethiopië met behulp van deze merkers; 3) het karakteriseren van de ruimtelijke genetische structuur en verspreiding binnen twee *B. papyrifera* populaties; en 4) het beoordelen van de variatie binnen en tussen volwassen bomen en zaailing generaties.

Illumina paired-end next generation sequencing werd gebruikt om een deel van het *B. papyrifera* genoom te sequencen. Op basis van de korte reads ontwikkelden we 46 microsatelliet markers waarvan de meeste ook bruikbaar zijn in twee andere Boswellia soorten: *B. popoviana* en *B. pirrotae*. Dit suggereert dat de moleculaire

merkers ook bruikbaar kunnen zijn voor verdere genetische studies van nauw verwante soorten.

De genetische diversiteit werd gekarakteriseerd in 12 populaties uit het hele verspreidingsgebied in Ethiopië. De meeste bemonsterde populaties hadden een gebrek aan verjonging. In de Noord-oostelijke regio waren sommige populaties erg verarmd (AF-BIR en KTD), met slechts een paar bomen verspreid over heuvels en landbouwgrond. In tegenstelling tot de verwachtingen ontdekten we in de populaties in de westelijke regio (Benishangul-Gumuz) een hoge dichtheid aan zaailingen en jonge boompjes, wat aangeeft dat natuurlijke regeneratie er nog plaats vindt. We gaan ervan uit dat dit komt door minder agrarische activiteiten, een kleinere veestapel en meer geschikte ecologische omstandigheden voor de regeneratie qua bodem en regenval. We raden aan om gedetailleerde studies in dit gebied uit te voeren om de redenen voor de overvloedige regeneratie beter te begrijpen. Deze informatie kan vervolgens worden gebruikt voor toepassing in de andere delen van het verspreidingsgebied, waar de soort wordt bedreigd en daarmee het toekomstige gebruik van wierook.

Ondanks de bedreigingen die de populaties ondervinden, vonden we nog voldoende genetische variatie in de volwassen bomen in alle delen van Ethiopië, ook in de meest gedegradeerde populaties. Deze diversiteit is waarschijnlijk te danken aan de goede verspreiding van pollen en zaad van *B. papyrifera*, de lange levensduur van de soort, en het feit dat de bosfragmentatie en degradatieprocessen in Ethiopië relatief recente gebeurtenissen zijn. Dit betekent dat de *B. papyrifera* bomen die nu zijn bemonsterd, nog de diversiteit representeren van de voormalige grote en aaneengesloten bossen die bestonden in de 19e eeuw. De analyse van de cohorten suggereert dat er geen verlies van genetische diversiteit, noch selectie voor heterozygosity plaatsgevonden heeft tussen generaties, en dat de diversiteit van de volwassen bomen voldoende aanwezig is in hun nageslacht. Het hoge

niveau van de genetische diversiteit en het lage niveau van inteelt suggereert dat uitkruising de belangrijkste vorm van reproductie is in *B. papyrifera*.

Op basis van een analyse met STRUCTURE werden de 12 populaties onderverdeeld in vier genetische clusters: Noord-Oost (NE), West (W), Noord-west (NW) en Noord (N) in Ethiopië. De clusters komen overeen met verschillende milieuomstandigheden, met name temperatuur, neerslag en de bodemgesteldheid. We raden aan om geografische afstand tussen populaties als de belangrijkste factor te beschouwen bij de bemonstering voor conservering. Daarnaast moeten ook de ecologische verschillen tussen populaties een rol spelen bij de planning van conserveringsmaatregelen.

Op basis van de genetische clustering van de populaties adviseren wij een minimum van vier zones voor de conservering van *B. papyrifera*, die overeenkomen met de genetische clusters West, Noord-west, Noord, en Noord-oost. Elk van de instandhoudingsgebieden moet een populatie hebben die voldoende groot en beschermd is om geen genetische diversiteit te verliezen in de volgende generaties als gevolg van genetische drift en inteelt, de oorzaken voor het verlies van genetische diversiteit in kleine populaties. De beschermde gebieden kunnen wellicht ook dienen om andere soorten in Combretum-Terminalia bossen te beschermen, met name de Boswellia soorten die in hetzelfde geografische gebied voorkomen (bijv. *B. pirottae*).

Dit proefschrift onderzocht ook de genetische structuur op kleine schaal (FSGS) van twee populaties in het westen van Ethiopië, met name de mate van verspreiding van pollen en zaad binnen populaties en de verandering in de genetische diversiteit tussen generaties. We vonden een lage FSGS waarde in beide populaties; individuen waren significant meer gelijkend aan elkaar tot een afstand van 60-65 meter. Op basis van deze bevindingen raden wij aan om voor

zaadcollecties voor het conservering middels ex situ en plantage-programma's, zaden te verzamelen op ten minste 60 m, maar bij voorkeur 100 m van elkaar.

Uitkruisende soorten zoals *B. papyrifera* hebben over het algemeen last van inteelt als er beperkte gene flow is als gevolg van fragmentatie. We raden daarom aan dat beschermingsmaatregelen zorgen voor het genetische uitwisseling tussen de versnipperde populaties en verspreid overgebleven bomen, om ervoor te zorgen dat het bestaande niveau van de genetische diversiteit zal de toekomstige generaties.

De studie werpt een aantal kwesties op in verband met het behoud, die verder moeten worden onderzocht. Het probleem dat in de meeste populaties geen regeneratie door jonge zaailingen voorkomt, moet dringend worden opgelost om de continuïteit van de bestaande genetische diversiteit, het overleven van de soort, en de duurzame productie van wierook in de toekomst te garanderen. In het licht van de voorspelde klimaatverandering moet het potentiele habitat van de soort worden bepaald, en waar de soort zich kan uitbreiden.

De recente initiatieven om te komen tot 'community joint forest management initiatives' (PFM) vormen een veelbelovend concept voor bosbeheer samen met de mensen die leven in en rond de bossen en het bos feitelijk gebruiken en beheren. Het bevorderen van deze aanpak in verschillende bossen in het land, waaronder de droge bossen, zou kunnen leiden tot het bereiken van de doelen qua conservering en behoud van de bossen, terwijl tegelijk het bos kan bijdragen aan het levensonderhoud van de lokale gemeenschap middels producten uit het bos.

ACKNOWLEDGEMENTS

This thesis has been made possible with the contribution of many individuals. I am very grateful to you all. First I would like to thank my supervisors Frans Bongers, René Smulders and Kassahun Tesfaye. This thesis is the result of your unreserved inputs and commitments which also inspired me for a hard work. I have learnt a lot from your critical comments and suggestions. The lots of comments I received back in track change (sometimes around 0:00 to 1:00 hrs) were seeming frustrating at first glance, nevertheless shaping the thesis nicely. Beyond your scientific contribution and guidances, your continuous encouragement and support during my long time illness along this PhD journey inspired me a lot to stay strong and achieve this goal. Without the contribution of Frans as my promotor this thesis wouldn't have been started and have come to a successful end. Only one week was left before the deadline for NUFFIC funding application when I contacted you. Your very quick replies, the comments on the concept note, and instructions on the way forward in the application processes were the reasons behind this success. I am very grateful for your encouragement and confidence in me. I also thank you very much for all your efforts in arranging the financial supports from NUFFIC and other sources for the extended period of this study. With this respect, I am also grateful to FEM research group and the staff members who provided me their support. Rene's supervision and unreserved scientific instruction were the basis for the successful completion of this thesis. Your office was always open for me and I must mention that discussions were always fruitful and improved my understanding about molecular approaches and their application in different aspects of genetics. I benefitted a lot from your advises. I am also very grateful for the translation of the Summary to Samenvatting (in Dutch). I would like to thank Dr. Kassahun Tesfaye for his encouragement and contribution from the comments on the concept note (before the start of the PhD) to the completion of this thesis. I also thank you for your field visit at Kurmuk, Bensishangul-Gumuz as supervisor from Bioversity International. To have Jerome Duminil as co-supervisor from

Bioversity International was a valuable experience and his contribution helped me to broaden my views on Fine-scale genetic structure in tree populations.
Danny Esselink and Wendy van 't Westende taught me the molecular techniques and managing molecular data. I thank Danny also for managing the sequencing and analysing the data. I would like to thank Doret Wouters for doing the lab analysis of the samples for the fourth chapter and continuous support during my labwork. The laboratory staff of the Plant Breeding group were all so cooperative and always at disposal to help. I thank you all for the support you provided me. I thank my officemate Anne-marie Wolters for her help and encouragements. I also would like to thank you for the kindness I received from you as a friend specially in sharing my personal problems.

All district forest and natural resource management offices and their staff members deserve my gratitude for their support during my field work. Boswellia's favourite growing place is extremely hot and difficult to tolerate by humans and most tree populations were located in difficult-to-access areas. The field work was made successful with the support of the hardworking field assistants. I thank them all very much. My special thanks go to Jemal, Abebe, Tesfaye, Debela and Ayele who directly involved in the collection of the samples in different sites. Their commitments and the local tree climbing skills made the sample collection possible within the relatively short leafing season. The cooperation from the Benishangul-Gumuz Natural Gum processing and marketing Enterprise is highly appreciated. I would like to thank the staff members of the enterprise for all their supports. Bushra Mohamod, the head of the enterprise of the region, deserves special thanks for all his unreserved supports throughout my field work in the region. I am grateful to his commitments to help me in handling all the troubling field conditions I came across. Sisay Zewdie and Ashenafi Talew who without reservation were driving the field car from early in the morning to late in the evenings through almost inaccessible sampling sites contributed a lot to the

successful completion of the field work. I would like to extend my deepest gratitude and appreciation for their commitments. Sisay, the sudden temperature stroke you hardly survived in Afar-Birhale and the sliding land we stepped on in Limalimo were few examples of unfortunate encounterings you shared with me during the field work. Thankfully, they are now all adventurous stories. My heartfelt thanks for sharing with me all these burdens and for continuously working with me without hesitation. I am grateful to the Kurmuk Woreda, Benishangul-Gumuz region Ethiopian Army members and specially would like to thank Lieutenant Abiy, for all the supports during sample collection in such remote and insecure areas.

I am grateful to the staff members and PhD students the Forest Ecology and Management group of Wageningen University for the all the encouragements. Although I was not closely attached because of the collaboration of my research with the Plant Breeding group, I received a lots of encouragement from the occasional meetings I had with you. The meetings and refreshing group outings I shared with you were enjoyable moments. I would like to thank all the encouragements from: Madelon Lohbeck, Geovana Carreno, Catharina Jakovac, Estella Quintero, Mart Vlam, Peter Groenendijk, Paul Copini, Lucy Amissah, Monique Weemstra, Masha van der Sande, Carolina Levi. Madelon, thanks for your visit at the hospital. I extend my special thanks to Yvonne Geraedts for hosting me at her place, friendly character, supports, encouragement and especially for her concern for my health problem. Joke Jansen's administrative support in all aspects is highly appreciated. I would like to thank her for understanding the problem with communications from Ethiopia and always try to look for solutions even for delayed matters from my side. I also would like to thank Marion Rodenburg at PhD service for her unreserved help.

I am grateful to the whole FRAME team: Dr. Emiru Birhane, Dr. Abeje Eshete, Dr. Tefera Mengistu, Dr. Teshale Woldeamanuel, Atkilt Girma, Dr. Motuma Tolera and Dr. Mulugeta Lemenih. I would like to appreciate Dr. Mulugeta's comments on the conceptnote and continuous encouragement.

I want to thank families and friends from Wageningen for their support and encouragement during my illness. I want to extend my special thanks to Tilahun, Wossen, Bereket, Hibist, Etetu, Araya, Saba, Abate, Meron, Alemtsehay, Tewodros M. and Banchiayehu for standing close by my side and taking care of me throughout difficult times. Geni and Biruk, I enjoyed our friendship and our regular blessed prayers we had together in Wageningen. My thanks also go to my friends: Tewodros A., Dawit S. Adugna, Aregaw, Thomas, Bethelehem, Mahlet, Mindaye and Belay. Tewodros M. and Tewodros A., I thank you for your support in mapping my study area and helping me in managing computer technical matters.

I would like to thank my colleagues in Wondo Genet College and School of Forestry for understanding my situation and allowing me extended period of study. In particular, I would like to thank Bereket Roba for his support and highly appreciate him for his considerate and very efficient way of handling my request for extension as the head of the school. I would like to thank Dr. Melaku Bekele, Dr. Tsegaye Bekele, Dr. Abdella Gure, and Habtamu Degefa for their encouragement and supports. Seble Metaferia and Almaz Tessema deserve my special thanks for their supports and role as colleagues and friends. I would like to extend my thanks to Dr. Feyera Senbeta and Dr. Gete Tsegaye for their continuous encouragement and supports.

I thank my friends pastor Geremachew and his family, prophet Bogale and Marshet for their encouragement. I am also grateful to the encouragement from my

sisters and close relatives. I would like to thank my nephew Abiy, his wife Yewubdar and daughters Merry and Misgana for hosting me at their place during my arrival and departure from Addis.

I am very grateful to my late parents for their encouragement and love throughout my life. The encouragement from my father, who never educated himself except basic education was special. It inspired me a lot from my childhood for higher level educations and here is the fruit today. Throughout this PhD journey my mother was the only person who had been advising me to quit this PhD study and to take care of my health. This was a signal that she cares most for me and the feeling encouraged me more. I am very proud of you two and I dedicate this thesis to you.

Finally, I would like to acknowledge the financial support from the Netherlands Fellowship Program (NUFFIC), NWO-WOTRO (Netherlands Organization for Scientific Research-Science for Global Development) through the Integrated Program FRAME (Frankincense, myrrh and gum arabic: sustainable use of dry woodlands resources in Ethiopia), W01.65.220.00 and Bioversity International.

SHORT BIOGRAPHY

Addisalem Ayele Bekele was born on November 20, 1972 in Eastern Harar, Ethiopia. She attended elementary education in the same place at Boreda elementary school. She attended her high school at Harar Secondary school, Harar. She studied forestry from September 1993 to June 1997 at Haramaya University of Agriculture, Ethiopia. In July 1997, she was employed in Forestry Faculty as graduate assistant in the same University she graduated from (Haramaya University). In 1998 she moved to Wondo Genet College of Forestry, South of Addis Ababa and worked for 2 more years (until 2000) as graduate assistant at different levels. The three years exposure at this level provided her the opportunity to develop her teaching skills and good experiences in handling practical courses. In September 2000, she joined Dresden University of Technology, Germany for MSc study. She studied Tropical Forestry and Management and graduated in March 2003. Her MSc research dealt with assessing the impact of forest fragmentation on the genetic diversity of *Cordia africana* which was one of threatened tree species in Ethiopia. Back to Ethiopia, she re-joined Wondo Genet College of Forestry in September 2003 and she worked as a lecturer and researcher until August 2010. During this

time, she thought different courses related to forest conservation and management such as biodiversity conservation, plantation establishment tree genetics and improvement. She participated in forest resource management researches. Especially she was actively involved in community based researches (development-oriented interdisciplinary action research) which was jointly run by Wondo Genet College of Forestry and Natural Resources and Swedish University of Agricultural Sciences. In this project she actively involved as a team member of the organizing committee and a researcher. She initiated a research project targeting at assessing the potential of rainwater harvesting for homegardening use and contribution to local household's livelihood. She gained experience of closely working with the local community and team members of different academic backgrounds.

Addisalem won the Netherlands Fellowship Program (NUFFIC) PhD scholarship in 2010 and she joined the Forest Ecology and Forest Management chair group and the Plant Breeding Group in September 2010. She studied the conservation genetics of the frankincense tree *Boswellia papyrifera*.

LIST OF PUBLICATIONS

1. A.B. Addisalem, G.D. Esselink., F. Bongers, M.J.M. Smulders. 2015. Genomic sequencing and microsatellite marker development for *Boswellia papyrifera*, an economically important but threatened tree native to dry tropical forests. Aob Plants 7: plu086. doi 10.1093/aobpla/plu086.
2. A.B. Addisalem, F. Bongers, T. Kassahun, M.J.M. Smulders. 2016. Genetic diversity and differentiation of the frankincense tree (*Boswellia papyrifera* (Del.) Hochst) across Ethiopia and implications for its conservation. Forest Ecology and Management 360: 253–260
3. A.B. Addisalem, J. Duminil, D. Wouters, F. Bongers, M.J.M. Smulders 2016. Fine-scale spatial genetic structure in the frankincense tree *Boswellia papyrifera* andimplications for conservation. Accepted for publication in Tree Genetics & Genomes (with minor revisions).
4. A.B. Addisalem, F. Bongers, M.J.M. Smulders. Conservation genetics of frankincense tree, *Boswellia papyrifera* (Del.) Hochst). 2015. Abstract published in the Proceedings of the Annual Conference of the Society for Tropical Ecology, Resilience of Tropical ecosystems: future challenges and opportunities (Theme: Conservation Genetics and Genomics in the tropics), April 7-11, Zurich, Switzerland.
5. A.B. Addisalem, 2016. Conservation genetics of frankincense tree, *Boswellia papyrifera* (Del.) Hochst). Abstract published in the Proceedings of the 26th Annual Conference of the Biological Society of Ethiopia, Enhancing the Quality of Education in the Higher Education System of Ethiopia, March 2016, Mekele, Ethiopia.
6. A.B. Addisalem, 2003. Forest fragmentation and its impacts on genetic variation of *Cordia africana* Lam. population around Wondo Genet, Ethiopia. MSc thesis.

PE&RC Training and Education Statement

With the training and education activities listed below the PhD candidate has complied with the requirements set by the C.T. de Wit Graduate School for Production Ecology and Resource Conservation (PE&RC) which comprises of a minimum total of 32 ECTS (= 22 weeks of activities)

Review of literature (6 ECTS)

- Conservation genetics of frankincense tree

Writing of project proposal (4.5 ECTS)

- Conservation genetics of frankincense tree

Post-graduate courses (3.1 ECTS)

- Frame project - PhD workshop; Ethiopia (2011)
- Estimating mating system and gene flow in plants with emphasis on trees; Kazimierz Wielki university, Poland (2013)

Laboratory training and working visits (0.6 ECTS)

- Visit to Genetics and Biotechnology Laboratory; Holeta Agricultural Research Centre, Ethiopia (2011)
- Visit to the Molecular Laboratory of Tree Genetics and department; Kazimierz Wielki University, Poland (2013)

Deficiency, refresh, brush-up courses (7.5 ECTS)

- Basic statistics; PE&RC (2010)
- Population and quantitative genetics; Wageningen University (2015)

Competence strengthening / skills courses (2.7 ECTS)

- Techniques for writing and presenting scientific paper; WGS (2014)
- Information literacy including EndNote introduction; WGS (2015)
- Project and time management; WGS (2015)

PE&RC Annual meetings, seminars and the PE&RC weekend (1.5 ECTS)

- PE&RC Weekend (2010)
- PE&RC Day: optimization of science: pressure and pleasure (2014)
- WGS PhD Workshop carousel (2015)

Discussion groups / local seminars / other scientific meetings (6.4 ECTS)

- Netherlands annual ecology meeting (2011)
- Biodiversity discussion group; Wageningen University (2012-2015)
- Current themes in ecology symposium (2014)
- Recovery of tropical forests mini symposium (2014)
- Wood in the SPOTlight: an interactive evening on wood, its science and beauty (2015)
- Biotechnology Institute, PhD series seminars; Participation and presentation (2016)

International symposia, workshops and conferences (4.1 ECTS)

- Annual conference of the society for tropical ecology; oral presentation; Zürich, Switzerland (2015)
- Annual meeting of Ethiopian biological society; oral presentation (2016)

The FRAME Project

FRAME: Frankincense, Myrrh and gum arabic: sustainable use of dry woodland resources in Ethiopia

More than half of the total land area in Ethiopia is covered by arid to semiarid woodlands withmarginal agricultural potential. These woodlands are commonly overexploited for their natural resources, which reduces the local livelihood options for a rapidly expanding population. Climate change (e.g. drought) may intensify this negative trend. Consequently, there is an urgent need for improved land-use strategies that will make the vast arid and semiarid woodland resources optimally contribute to the livelihoods of local people and national development goals.

The dry woodlands in Ethiopia are not resource poor as they host several woody species that hold economically well recognized aromatic products such as gum arabic, frankincense and myrrh, which are widely used locally and in several of today's commercial industries such as cosmetic, pharmacological and food industries. Frankincense and myrrh are among the oldest internationally traded commercial tree products. Ethiopia is worldwide the main producer of frankincense and myrrh, and exports much gum arabic. Gum/resin production could significantly contribute towards sustainable development of these dry woodland areas. However, the overexploitation of natural resources by intensive grazing and intensive resin/gum harvesting and the lack of land management threatens the sustainability of the woody vegetation, and as a result of that also the long-term gum/resin production. Local communities may also enhance the productive capacity of the natural vegetation by establishing protected enclosures and by cultivation of trees. Such production systems may have a lower status regarding biodiversity and natural ecosystem functioning, but maintain ecological buffering capacity and improve production for human benefit.

The FRAME program addresses the following main research question: in what way dry land forests in Ethiopia can be made productive while maintaining ecosystem integrity in terms of sustainability of production and vegetation cover, with special attention to resin and gum resources?

FRAME uses a multidisciplinary approach involving scientific disciplines ranging from landscape-level geo-information studies to village-level socio-economic studies, plot level ecological and harvesting technology studies to tree-level ecophysiological studies with a strong contribution of local knowledge in answering the central research question. FRAME thus establishes a scientific basis for the sustainable management, including cultivation, of gum and resin yielding tree species and their habitat, the dry woodlands in the Horn of Africa. FRAME is actually involved in development of long-term scenarios for proper use and selection of suitable areas of dry woodland resources in Ethiopia.

The current PhD thesis is part of this FRAME program. A large part of this integrated FRAME research program was financially supported by NWO-WOTRO (Netherlands Organization for Scientific Research- Science for Global Development), grant W01.65.220.00.

The research described in this thesis was financially supported by the Netherlands Fellowship Program (NUFFIC), NWO-WOTRO (Netherlands Organization for Scientific Research-Science for Global Development) through the Integrated Program FRAME (Frankincense, myrrh and gum arabic: sustainable use of dry woodlands resources in Ethiopia), W01.65.220.00 and Bioversity International.

Printed by: Ridderprint BV