Plant Evolution in the African 'Sky Islands': Evidence from Fossil Calibrated Molecular Dating and Amplified Fragment Length Polymorphism

Abel Gizaw Seid

A Thesis Submitted to

The Department of Plant Biology and Biodiversity Management

Presented in Fulfillment of the Requirements for the Degree of Doctor of Philosophy (Biology: Botanical Science)

Addis Ababa University

Addis Ababa, Ethiopia

July 2012

Addis Ababa University

School of Graduate Studies

This is to certify that the thesis prepared by Abel Gizaw Seid, entitled: Plant Evolution in the African 'Sky Islands': Evidence from Fossil Calibrated Molecular Dating and Amplified Fragment Length Polymorphism and submitted in fulfillment of the requirements for the Degree of Doctor of Philosophy (Biology: Botanical Science) complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the Examining Committee:

Name	Signature	Date
1. Prof. Sebesbe Demissew (Examiner)	_____	_____
2. Prof. Richard Abott (Examiner)	_____	_____
3. Prof. Sileshi Nemomissa (Advisor)	_____	_____
4. Prof. Christian Brochmann (Advisor)	_____	_____

_____

Chair of Department/Graduate Program Coordinator

ABSTRACT

Plant Evolution in the African ‘Sky Islands’: Evidence from Fossil Calibrated Molecular Dating and Amplified Fragment Length Polymorphism

Abel Gizaw

Doctor of Philosophy in Plant Biology and Biodiversity Management

Addis Ababa University, 2012

The afro-alpine region in Eastern Africa encompasses the elevated plateaus and mountains in Ethiopia and the isolated high mountain peaks in Kenya, Tanzania, and Uganda, often referred to as biological 'sky islands'. The age of the flora on these mountains has been debated. Based on the high level of endemism, it has been suggested that the flora is old and might represent Tertiary relicts. Others have argued that some plant groups are young and have colonized these mountains during the Pleistocene. The impact of the glacial-interglacial oscillations of the Pleistocene on the origin and geographical structuring of the genetic diversity in many temperate plant species has been well documented, whereas only few studies for the afro-alpine plant speices have so far been conducted. Here we use a combination of molecular dating and phylogeographic analysis of selected plant species to provide insight into the evolution of plants in the afro-alpine region. A fossil calibrated multilocus species tree based on two nuclear and three plastid DNA regions was used to estimate the age of the afro-montane endemic *Lychnis* (Caryophyllaceae). Furthermore, Amplified Fragment Length Polymorphism (AFLP) fingerprinting was used to assess the

genetic diversity and population genetic structure of seven plant species: two heather species from the ericaceous zone, *Erica arborea*, *E. trimera* (Ericaceae), and five species from the afro-alpine zone proper: *Carex monostachya*, *C. runssoroensis* (Cyperaceae), *Lysimachia serpens* (Primulaceae), *Dicrocephala chrysanthemifolia* and *Haplocarpha rueppellii* (Asteraceae). The multilocus species tree analysis have placed the mean age of the afro-montane *Lychnis* to 2.28-5.8 Myr (million years) corresponding to late Miocene to the late Pliocene and thus indicating a Tertiary relicts. The results from the AFLP analysis showed that all investigated species have very low levels of overall gene diversity, suggesting bottlenecks in small refugial populations during unfavourable climatic periods. The populations of *E. trimera* and *H. rueppellii* showed relatively distinct geographical structuring of the genetic variation, suggesting long-term isolation of the former in at least three separate places and of the latter in two refugial populations. The lack of common phylogeographic patterns indicates that different species have responded individually to the Pleistocene climatic oscillations.

***Key words**:* AFLP, Afro-alpine, Afro-montane, *Lychnis*, *Erica arborea*, *Erica trimera*, *Carex monostachya*, *Carex runssoroensis*, *Lysimachia serpens*, *Dicrocephala chrysanthemifolia*, *Haplocarpha rueppellii*, Molecular dating, Phylogeography, Plio-Pleistocene, Tertiary-relicts.

ACKNOWLEDGMENTS

I owe my sincere gratitude for a number of people and institutions for making this PhD study a memorable and worthwhile academic experience. First of all, I am very grateful to my supervisors Prof. Sileshi Nemomissa and Prof. Christian Brochmann for accepting me twice, first as a Master student in the previous AFROALP I project and again as a PhD student in current AFROALP II project. Ever since I started working with you two and all the way to the completion of this PhD, you have played an immense role in developing my academic career as well as in strengthening my competence level, even at international levels. Without you two, this doctoral study would have never come to an end.

Dr. Magnus Popp, as a mentor and a friend, has always inspired me and opened my eyes to the world of molecular evolution from the time when I started collaborating with him as a Master student. Your continuous guidance and assistance since day one has been quite tremendous. You have already taken a huge part in tutoring me in both molecular techniques as well as in complex data analysis.

I am very much grateful to my fellow AFROALP II colleagues Tigist Wondimu, Felly Mugizi, Catherine Masao, and Ahmed Abdi Kaman with whom we conducted a huge and painstaking field work in various afro-alpine regions in East Africa. The courage and enthusiasm we had within us were nothing but the major key for the big success we accomplished in the field, what others haven't tried and achieved yet. Such enormous field work would have not

been possible without the kind help and assistance offered from the following organizations: the Simen Mountains National Park and the Bale Mountains National Park in Ethiopia, the National Museums of Kenya, the Tanzanian Commission for Science and Technology, and the Ruwenzori Mountains National Park and the Mgahinga Gorilla National Park in Uganda. The following people and others involved in the AFROALP II project are acknowledged for various kind of help I received in one way or another: Dr. Geoffrey Mwachala (Kenya), Dr. Gerald Eilu (Uganda), Prof. Pantaleo Munishi (Tanzania), Prof. Peter Linder (Switzerland), and Dr. Pierre Taberlet (France). I am also grateful to Dr. Beentje Henk for sharing his wide knowledge on the taxonomy of *Erica* and for providing valuable suggestions.

I wish to thank the following institutions for their kind support and assistance in facilitating my study in one way or another: the Plant Biology and Biodiversity Management Programme and the School of Graduate Studies of Addis Ababa University, Ethiopia, and the National Centre for Biosystematics, Natural History Museum, University of Oslo, Norway. I thank the Institute of Biodiversity and Conservation for permission to take samples to Norway for laboratory analyses. The study was funded by the NUFU Programme (The Norwegian Programme for Development, Research and Higher Education) under a project headed by S. Nemomissa and C. Brochmann (2007/1058: 'AFROALP-II - Afro-alpine ‘sky islands’: genetic versus taxonomic biodiversity, climate change, and conservation’).

I am very much indebted to members of the staff at the National Herbarium (ETH) of the Addis Ababa University, Ethiopia, and at the National Centre for Biosystematics, University of Oslo, Norway. My friends Getenet Simeneh, Hailu Atnafu, Fisseha Asmelash, Yonas Assefa, Zewge Abate, Ephrem Debebe, Solomon Gashaw, Amha Wudie, Endeshaw Wolde and friends in the Ethiopian Student Union in Norway, and many more in other places in the world are sincerely acknowledged for their friendship and assistance in the course of the study.

Finally, I owe my overwhelming gratitude to my beloved mother W/o Aster Fisseha and the rest of the family for their continuous encouragement, support, and belief in me to pursue my dreams and reach where I am right now.

TABLE OF CONTENETS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1 INTRODUCTION

1.1. The afro-alpine flora and Pleistocene range shift

The East African Rift System (EARS) is a prominent geological feature, extending from Syria in the north through Eastern Africa to Mozambique in the south. Its formation began around 45 million years ago and divided the region into the Western and Eastern Rift branches (Baker *et al.*, 1972; Logatche *et al.*, 1972; Trauth *et al.*, 2005). The volcanism associated with EARS resulted in the formation of elevated plateaus and mountains in Ethiopia and isolated high mountain peaks in East Africa in Kenya, Tanzania, and Uganda (such as Mt. Kenya, Mt. Meru, Mt. Kilimanjaro, and Ruwenzori). Most of these mountains, often referred to as biological 'sky islands', are of volcanic origin and have different ages, ranging from Miocene to late Pliocene (Griffiths, 1993).

The mountains are famous for harbouring the vast and unique afro-alpine ecosystem. The vegetation on these mountains forms three distinct zones: the lowermost afro-montane forest zone, the transitional ericaceous or heather zone, and the uppermost afro-alpine zone proper (Hedberg, 1951). The afro-alpine and afro-montane regions largely share the same history and a high proportion of endemic species, housing 80% and 75% of the vascular plants endemic to the Afro-alpine and Afro-montane regions, respectively (Hedberg, 1961, 1965, 1969). Based on the high level of endemism on these mountains, Hedberg (1969) suggested that the flora has long been isolated from other high mountains and temperate floras and hence may be a Tertiary relict.

Many plants growing in these mountains have their closest relatives not in their adjacent lowlands but in other temperate areas of the world (Hedberg, 1965, 1970). Many of these plants seem to have arrived in East Africa and colonized the 'sky islands' from Eurasia via the Arabian Peninsula (Koch *et al.*, 2006; Assefa *et al.*, 2007; Ehrich *et al.*, 2007; Popp *et al.*, 2008). Further diversification was mainly shaped by the climatic oscillation of the Plio-Pleistocene period (Livingstone, 1962; Hamilton, 1982; Mohammed and Bonnefille, 1998; Gottelli *et al.*, 2004; Kebede *et al.*, 2007). All of these findings suggest at least part of the contemporary afro-alpine flora is rather young. However, the relative age of this flora, the actual time of colonization, and the mechanism of speciation on these mountains are poorly understood and remain to be investigated.

It is now well established that the glacial-interglacial oscillations of the Pleistocene have played a major role in shaping the current genetic structure of species (Hewitt, 2000), including the contemporary African flora (Livingstone, 1962; Hamilton, 1982; Mohammed and Bonnefille, 1998; Gottelli *et al.*, 2004; Kebede *et al.*, 2007). The African tropics were colder and drier during the last glaciations (Bonnefille *et al.*, 1990) and the afro-alpine proper and ericaceous zones were pushed down by ~1000 m and covered larger areas than today (Flenley, 1979; Gottelli *et al.*, 2004). In contrast, the afro-montane forest zone was reduced and open vegetation expanded in response to aridification in its lower-lying part (Bonnefille, 1995; deMenocal, 1995). During the warm and more humid inter-glacial periods, on the other hand, the afro-alpine and ericaceous zones contracted to higher elevations whereas the afro-montane

forest expanded and may have connected originally isolated patches via temporary forest bridges between adjacent mountains (Mohammed and Bonnefille, 1998; Ryner *et al.*, 2006; Kebede *et al.*, 2007; Voje *et al.*, 2009).

Previous molecular studies on the African rainforest tree (Couvreur *et al.*, 2008), the African *Coccinia* (Holstein and Renner, 2011), African clawed frogs (Evans *et al.*, 2004), African birds (Roy *et al.*, 2001; Bowie *et al.*, 2006; Voelker *et al.*, 2010) and the African forest chameleons (Tolley *et al.*, 2011) incorporated climatic and/or geological history of the continent into a dated phylogeny and tested the various mechanisms of speciation that shaped the biogeographical patterns of the species. A climatically induced speciation model assumes that climatic oscillation of the Plio-Pleistocene period may have led to periodic expansion and contraction of the forest in East Africa following the corresponding interglacial and glacial periods in the northern hemisphere (Bonnefille *et al.*, 1990; deMenocal, 1995; Gottelli *et al.*, 2004; Ryner *et al.*, 2006; Kebede *et al.*, 2007; Voje *et al.*, 2009). Thus, such periodic isolation and fragmentation of the forest habitat may have ultimately led to allopatric speciation. A montane speciation model, on the other hand, states that despite the change in climatic conditions during Plio-Pleistocene some part of the African tropics remained climatically stable, with relatively highly-stable areas between 4° and 11°S (Lovett *et al.*, 1988; Fjeldså and Lovett, 1997). These mountains were likely providing a stable and moist habitat (Fjeldså 1994; Fjeldså and Lovett, 1997; Roy, 1997; Hewitt, 2000) which may have promoted speciation through ecological diversification. The third hypothesis, adaptive speciation, considers the varying ages of many of the isolated mountain peaks

and their formation through volcanic activity which may have left new and unoccupied niches that could potentially promote speciation when invaded by new arrival species (Jordan *et al.*, 2003; Coyne and Orr, 2004; Voje *et al.*, 2009). It follows that most of the speciation occurred almost over the same period that the volcanoes on which species grew were formed.

Despite the increasing number of molecular-based biogeographical studies of afro-alpine plants over the past few years (Koch *et al.*, 2006; Assefa *et al.*, 2007; Ehrich *et al.*, 2007; Popp *et al.*, 2008), only one study (Koch *et al.*, 2006) has dealt with time of colonization. In their phylogeographic study of *Arabis alpina* L. (Brassicaceae) based only on synonymous mutation rates and sequence distance of the plastid DNA, Koch *et al*. (2006) inferred a Pleistocene origin of the afro-alpine *A. alpina* and suggested twice colonization of the African mountains by two independent Eurasian lineages. However, no studies have so far used an explicit fossil-calibrated model of molecular evolution to infer a multilocus species-tree and estimated the absolute time of divergence for any afro-alpine/montane plants.

A recent biogeographical study (Popp *et al.*, 2008) revealed that the ancestor of the monophyletic group of the afro-montane *Lychnis* diverged from a joint ancestor with the Eurasian *L. flos-cuculi* and migrated to East Africa. However, the divergence time for this split, which later resulted in two strongly supported ploidy-specific lineages, as well as the absolute time of colonization and subsequent divergence in the African mountains, are still not known. Estimating the divergence date between the widespread and local endemic groups of the

monophyletic afro-montane *Lychnis* and the sister taxon, the temperate *L. flos-cuculii,* will shed some light on the relative age of the flora and the approximate time of colonization of the African 'sky islands'.

The effects of Pleistocene climatic oscillations on the range dynamics of the afro-alpine plants happened through repeated contractions and expansions, which ultimately resulted in the formation of highly fragmented populations, with species within the ericaceous and afro-alpine zones currently restricted to isolated high mountain interglacial refugia (Hedberg, 1970). Studies on the genetic consequences of such range shifts in response to climate change during the Pleistocene have provided new insights into the history of many species and identified important refugial areas. However, whereas many phylogeographic studies have addressed the histories of temperate, boreal and arctic species (Taberlet *et al.*, 1998; Hewitt, 2000; Brochmann *et al.*, 2003), only a few molecular studies have so far focused on plants from the afro-alpine region and their history in relation to past climate changes (Knox, 1993; Knox and Palmer, 1998; Koch *et al.*, 2006; Assefa *et al.*, 2007; Ehrich *et al.*, 2007; Kebede *et al.*, 2007; Popp *et al.*, 2008; Gehrke and Linder, 2009; Désamoré *et al.*, 2011). Only the studies of the afro-montane forest endemic *Lobelia giberroa* (Kebede *et al.*, 2007) and the widespread, afro-alpine and arctic-alpine *Arabis alpina* (Assefa *et al.*, 2007; Ehrich *et al.*, 2007) have used an explicit phylogeographic approach.

In the giant lobelia *L. giberroa*, three major, distinct lineages were identified based on AFLPs, suggesting old divergence and survival during several glacial periods in separate forest refugia. This study also inferred gradual expansion

within the main groups consistent with formation of montane forest bridges between adjacent mountain systems in previous interglacials, and possibly during the early period of the present interglacial (Kebede *et al.*, 2007). In *Arabis alpina*, plastid DNA sequence variation suggested that the afro-alpine region was colonized twice from the north by two divergent lineages. The latest-arriving, 'Asian lineage' is still confined to the high mountains closest to the Arabian Peninsula (the Ethiopian Simen and Gara Muleta), whereas the early-arriving, 'African lineage' is widespread and subdivided into two phylogeographic groups with distinct geographical distributions (Koch *et al.*, 2006; Assefa *et al.*, 2007). This result suggests long-term isolation of the African lineage in at least two inter-glacial refugia, although occurrences of some haplotypes are also consistent with episodes of long-distance dispersal (Assefa *et al.*, 2007; Ehrich *et al.*, 2007). A range-wide AFLP study of the same species revealed conspicuously different consequences of climate change among different geographic areas (Ehrich *et al.*, 2007). Whereas the arctic populations showed virtually no genetic diversity, consistent with large-scale latitudinal colonization after the last glaciation from a single southern refugial population, both the populations from the European Alps and those from the afro-alpine mountains were highly differentiated and diverse, suggesting long-term isolation and small-scale altitudinal range shifts in different mountain areas in response to climatic fluctuations.

CHAPTER 2 LITERATURE RIVEW

2.1 Molecular dating methods

The availability of molecular sequence data provides opportunities to answer several evolutionary questions including the timing of important evolutionary events that have occurred in the history of the organism. External information in the form of fossil or known paleo-geological events can potentially improve divergence time estimates in a phylogenetic tree. Inclusion of such external information into the phylogenetic tree requires different assumptions including the rate of molecular evolution. One of the assumptions in rate constancy is that molecular evolution occurs at constant rate over the whole part of the phylogenetic tree (Hedges and Kumar, 2003). Alternativelly, molecular rates can vary between and across lineages under the rate variation assumption (Nei and Kumar, 2000).

The rate at which DNA sequences are evolving and the time at which lineages are diverging can be addressed in different methods that employ various molecular-clock models and assumptions. Molecular dating using global clock model assumes that genetic distance between any two lineages is proportional to the amount of time spent since divergence from their common ancestor, and that the different parts of the tree evolve under a general global clock model (Zuckerkandl and Pauling, 1965). In another method, different parts of the phylogenetic tree are expected to display varying rates of molecular evolution, and thus under a different local clock model (Li and Tanimura, 1987). Alternatively, molecular-clock model can be relaxed by using either

aurtocorrelated or uncorrelated clock models that allow evolutionary rates to change over time. The autocorrelated relaxed molecular-clock approach assumes that closely related lineages/branches evolve under a more or less similar rate of molecular evolution than to distantly related lineages/branches in the phylogenetic tree (Gillespie, 1991). The uncorrelated clock model, on the other hand, assumes that evolutionary rates vary across the different parts of the phylogenetic tree. The program BEAST (Bayesian Evolutionary Analysis of Sampling Trees; Drummond and Rambaut, 2007), uses a Markov Chain Monte Carlo (MCMC) algorithms in Bayesian framework to co-estimate the phylogeny and divergence time using variable rate methods implemented in the program. More recently, an extended version of BEAST, *BEAST, uses a coalescent model approach to provide a joint inference of species tree and divergence time estimate from multiple genes sampled from multiple individuals across a set of closely related species (Heled and Drummond 2010).

2.2 Phylogeographic inference methods

Phylogeography can be defined as a field of study concerned with the principles and processes governing the geographical distribution of genealogical lineages, especially those within and among closely related species (Avise *et al*. 1987; Avise 2000). In a more comprehensive approach of comparative phylogeography efforts are excreted to identify whether species or group of species have responded to past climatic or environmental changes in a concerted manner or not.

Various approaches are available for inferring phylogeographic patterns observed in species or group of species. Hickerson *et al*. (2010) reviewed some of the main phylogeographic approaches currently used to infer phylogeographic data. In the descriptive phylogeographic approach, the history of species is inferred based on qualitative interpretation that combine the geographical distribution of the lineage and its gene genealogies. In this approach, for example, summary statistics such as the *F* statistics can be employed to test the differentiation of the genetic variation among and within populations (Hickersion *et al*. 2010; Nei and Li, 1979).

In the model-based statistical approach, on the other hand, phylogeographic hypotheses and parameters are rigorously tested under a statistical framework. In this approach models are generated under coalescent theory and a Bayesian and/or likelihood methods are then used to test a phylogeographic hypothesis (Hickerson, et al., 2010). For example, genetic mixture analysis of dominant data can be performed using a model based clustering approaches in STRUCTURE program (Pritchard *et al.* 2000). The program STRUCTURE uses a Bayesian MCMC algorithm to identify genetically homogenous groups. By comparing the likelihood of the data estimated in different runs for different numbers of groups (K) it is possible to identify the optimal K. Individuals are assigned probabilistically to one of the clusters defined by allele frequencies at each locus.

2.3 Objectives of the study

2.3.1 General objective

The general objective of the present study is to extend our limited knowledge of the evolution of afro-alpine plants and the processes that shaped the contemporary afro-alpine flora in the context of Pleistocene climatic oscillations using molecular tools.

2.3.2 Specific objectives

The specific objectives of the study are:

- To infer the age of the *Lychnis* stem group using a fossil calibrated molecular clock for the chloroplast *mat*K gene
- To co-estimate a multilocus species phylogenetic tree and divergence time of the afro-montane *Lychnis* using the estimated age of the *Lychnis* stem group from fossil calibrated *mat*K gene
- To identify possible mechanisms that contributed to speciation and diversification within the afro-montane *Lychnis*.
- To test whether the two morphologically and ecologically very similar species *E. arborea* and *E. trimera* of the ericaceous zone have similar phylogeographic histories in the afro-alpine region using AFLP and plastid DNA sequences.
- To describe the pattern of genetic variation and infer the phylogeographic histories of *Carex monostachya, C. runssoroensis, Lysimachia serpens, Dicrocephala chrysanthemifolia,* and *Haplocarpha rueppellii* of the afro-alpine zone proper using AFLP.

- To assess potential congruence in phylogeographic patterns and to test whether the investigated species have responded individually or in a concerted manner to the Pleistocene climatic oscillations.
- To identify potential refugial areas and provide baseline information for biodiversity conservation strategies based on the results from the present molecular studies.

CHAPTER 3 MATERIALS AND METHODS

3.1 The study species

The present study included a total of 13 species representing six genera and five different families distributed across the eastern Africa Mountains. These are the African representatives of *Lychnis* L. (afro-montane *Lychnis* L.; Caryophyllaceae) with currently six accepted species (Popp, 2008); *Erica arborea* L. and *E. trimera* (Engl.) Beentje (Ericaceae); *Carex monostachya* A. Rich., *C. runssoroensis* K. Schum. (Cyperaceae); *Lysimachia serpens* (Hochst. ex DC.) U.Manns & Anderb. (Primulaceae) following the nomenclature of Manns and Anderberg (2009); *Dicrocephala chrysanthemifolia* (Bl.) DC. and *Haplocarpha rueppellii* (Sch. Bip.) Beauv. (Asteraceae).

3.1.1 The afro-montane Lychnis L. (Caryophyllaceae)

The genus *Lychnis* L. (Caryophyllaceae) comprises about 30 species of which most occur in temperate areas of the northern hemisphere (Oxelman *et al.*, 2000). The African representatives of *Lychnis* are distributed in the Ethiopian Highlands, the high mountains of East Africa (Kenya, Tanzania, Uganda, Rwanda, and Burundi), the Democratic Republic of Congo, Cameroon and Nigeria. Fig. 1b depicts sampling sites of the different species of the genus from Africa. The afro-montane *Lychnis* (previously referred to the endemic afro-alpine genus *Uebelinia*) comprises two informal sections corresponding to two different ploidy levels, i.e., the diploid section '*Uebelinia*' [*L. rotundifolia, L. scottii, L. abyssinica* (Fig. 1a) and *L. kigesiensis*] and the tetraploid section '*Trigynuebelinia*' (*L. crassifolia* and *L. kiwuensis*) (Popp, 2008).

A recent biogeographical study (Popp *et al.*, 2008) revealed that the ancestor of the afro-montane *Lychnis*, which was found to be monophyletic, diverged from a Eurasian common ancestor with *L. flos-cuculi* which migrated to East Africa. However, the divergence time of this split, which later resulted in the formation of two strongly supported ploidy-specific lineages, as well as the absolute time of colonization and subsequent divergence in the African mountains are still not known.

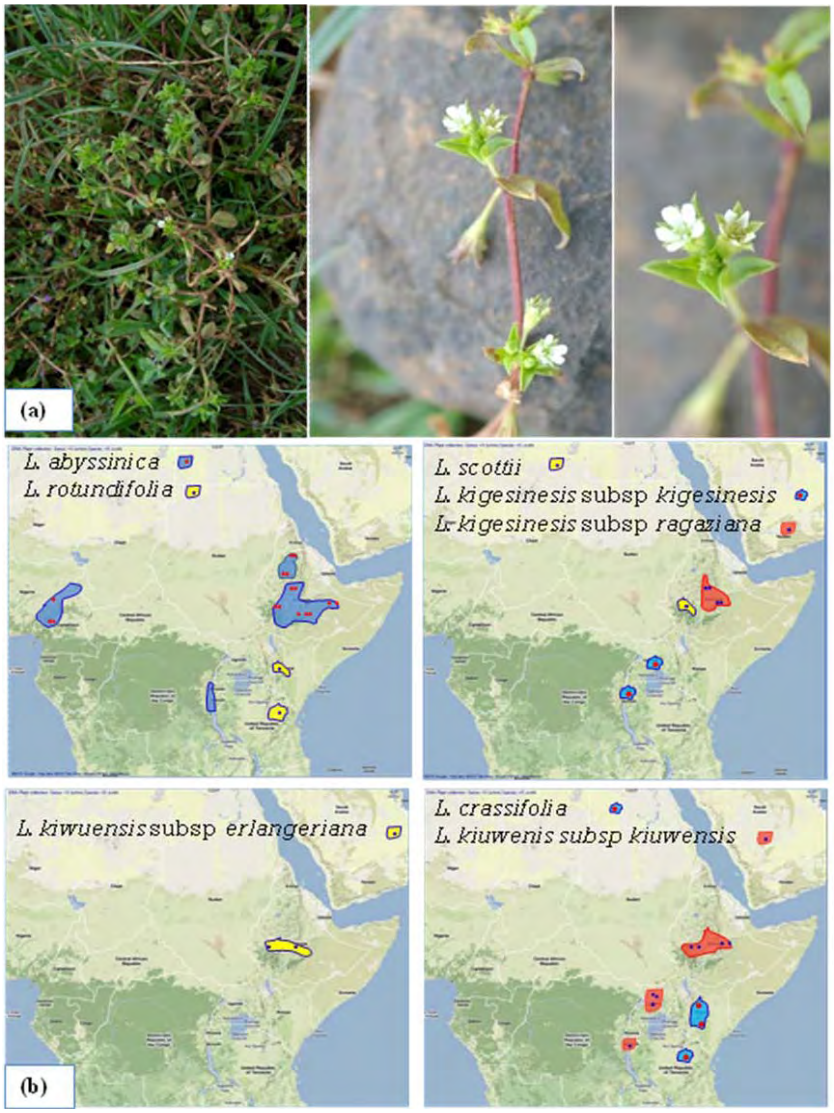


Fig. 1 (a) *Lychnis abyssinica*; (b) sampling localities of the afro-montane *Lychnis* in Cameroon and East Africa

3.1.2 Erica arborea L. and E. trimera (Engl.) Beentje (Ericaceae)

The genus *Erica* comprises more than 800 species and most of these are confined to the Cape Floristic Region in South Africa with a northward range extension reaching the high mountains of tropical Africa and into Europe (Beentje, 2006; Pirie *et al.*, 2011). Preliminary phylogenetic analysis based on sequences of the plastid DNA regions *atp*B-*rbc*L and *mat*K suggests that *E. arborea* (Fig. 2a) and *E. trimera* (Fig. 2b) are closely related and probably constitute sister species (Désamoré *et al.*, 2011). *Erica arborea* is widespread with a disjunct distribution. It occurs in Africa, the Mediterranean region including the Atlantic coast of Spain, the Pyrenees and the Middle East (Yemen). In Africa, it is distributed in North Africa including Morocco and Tunisia, Chad (Tibesti Mountains of Central Sahara), the Ethiopian Highlands, the mountains of East Africa and Malawi (Quezel, 1978; Messerli and Winiger, 1992; McGuire and Kron, 2005). *Erica arborea* is morphologically variable across its range, also within eastern Africa, but the morphological variation considered by Hedberg (1957) was too continuous to enable taxonomic distinction. In contrast, *E. trimera* is endemic to the afro-alpine region and occurs in several mountain systems of East Africa and the Ethiopian Highlands. This species is also variable morphologically and the complexity of this variation is reflected by the contrasting taxonomic treatments provided by different taxonomists. Five or six different subspecies, endemic to different mountains or mountain systems, were recognized by e.g. Hedberg (1957; as *Philippia trimera*), Beentje (2006), and Dorr (2006). The five subspecies accepted in Flora of Tropical East Africa (Beentje, 2006) are subsp. *trimera* (Ruwenzori Mts), subsp. *keniensis* (Mt. Kenya), subsp. *elgonensis* (Mt. Elgon),

subsp. *jaegeri* (Mt. Meru and Mt. Loolmalassin), subsp. *kilimanjarica* (Mt. Meru and Kilimanjaro), and a sixth subspecies, subsp. *abyssinica* (the Ethiopian mountains), is accepted in Flora of Ethiopia and Eritrea (Hedberg and Hedberg, 2003).

Both *E. arborea* and *E. trimera* are shrubs or small trees, 0.3-12 m high, and they are morphologically quite similar except in a few floral characters. *Erica arborea* has actinomorphic flowers and glabrous pedicels with two or three small bracteoles, whereas *E. trimera* has zygomorphic flowers and glandular pedicels without or with only one, recaulescent bracteole (Hedberg and Hedberg, 2003). Both species produce tiny wind-dispersed seeds less than 1 mm long (Huckerby *et al.*, 1972; Lind and Morrison, 1974), and can therefore be expected to have similar long-distance dispersal capacity. The similarities between these two closely related species in morphology as well as ecology often make it difficult to efficiently distinguish them in the field (Johansson *et al.*, 2010).

The two species often co-occur in the ericaceous zone, but *E. trimera* tends to dominate at higher elevations and *E. arborea* at lower elevations (Hedberg, 1971; Miehe and Miehe, 1994; Wesche *et al.*, 2000). In the Bale mountains, for example, *E. trimera* grows together with species of *Alchemilla* and *Helichrysum* in the upper part of the ericaceous zone, whereas *E. arborea* dominates the lower part of ericaceous zone and also occurs in the uppermost part of the afro-montane forest zone, growing together with *Hypericum revolutum* (Friis, 1986; Lissanwork Nigatu and Mesfin Tadesse, 1989; Bussman, 1997). According to

Hedberg (1957), *E. arborea* occurs in the afro-montane forests at altitudes down to 2000 m, occasionally also below this level, although it may in some mountains reach the afro-alpine zone proper up to 4700 m.a.s.l (meter above see level). On the other hand, *E. trimera* is virtually restricted to altitudes between 3000 and 4500 m.a.s.l. Sampling localities for the current study are given in Fig. 2c.

In a phylogenetic analysis of African and European Ericas based on nuclear (ITS) and plastid (*rbc*L-*atp*B and *mat*K) sequences, it was suggested that the African taxa were derived from European ones and that *E. arborea* was sister to the African clade (McGuire and Kron, 2005). They suggested that their common ancestor extended from Europe into Africa after the establishment of the connection of the two continents during the mid-Miocene. In contrast, a recent, range-wide phylogeographic study of *E. arborea* based on four plastid DNA loci (*atp*B-*rbc*L, *mat*K, *trn*H-*psb*A and *rpl*16) provided strong evidence for an origin of *E. arborea* in eastern Africa and expansion to other parts of the world during the Pleistocene (Désamoré *et al.* 2011). This study did not, however, address the species' phylogeographic history within eastern Africa.

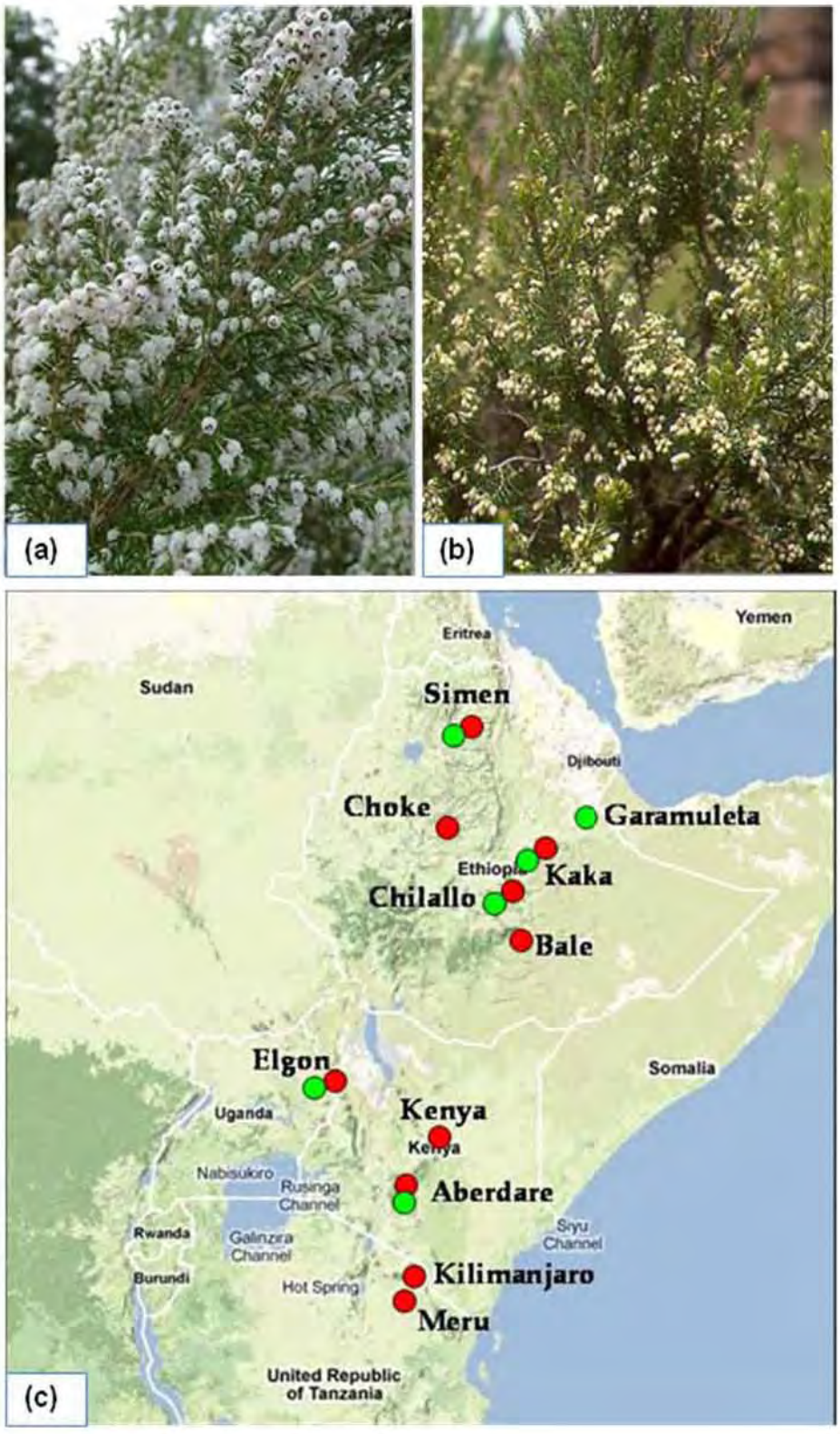


Fig. 2 (a) *E. arborea* (@ Flora. Cyberia); (b) *E. trimera* (Photographs courtesy of Christian Puff); (c) sampling localities in East Africa; circle colors indentify species (red - *E. trimera*; and green - *E. arborea*)

3.1.3 Carex monostachya A. Rich. and C. runssoroensis K. Schum. (Cyperaceae)

The genus *Carex* comprises 1600–2000 species (Reznicek, 1990; Govaerts *et al.*, 2010), making it one of the largest genera of vascular plants. *Carex* species are perennial, tufted, rhizomatous, stoloniferous or tussock-forming. The species of the genus commonly grow in moist to wet habitats and are also known to occur in drier habitats, such as montane or alpine grasslands, montane rocky habitats and forest understorey (Ball and Reznicek, 2002). Most of the *Carex* species are confined to the Northern Hemisphere temperate zone and some occur in warmer regions, such as in South-East Asia and eastern Africa (Gehrke, 2011).

Carex monostachya (Fig. 3a) is a creeping rhizomatous perennial plant, sometimes forming tussocks up to 0.5(-1) m tall. It has a distinct triangular stem (at least just below the inflorescence), and long and flat leaves. It mainly grows on moist ground in grasslands, in swamps, and along the edges of streams and lakes, and occurs mainly in the ericaceous and alpine belts (2750-4500 m). It occurs in Ethiopia, Kenya, Tanzania, and eastern Uganda (Mt. Elgon).

Carex runssoroensis is a tufted perennial (or creeping), forming hard tussocks up to 1 m high (Hedberg, 1957). It is characterized by the terete (round) culm below the spike (Gehrke, 2011). It mainly grows on flat, moist ground with poor drainage, usually as dominant species, occurring mainly in the ericaceous and alpine belts (2750-4300 m). It occurs in Kenya, Uganda, Rwanda and the

Democratic Republic of Congo (Virunga Volcanoes, Mt. Ruwenzori and Mt. Elgon).

A hybrid between *C. monostachya* and *C. runssoroensis* has been suggested by Hedberg (1957) and may be difficult to distinguish as characters can be intermediate between these two species (Gehrke, 2011). Furthermore, Hedberg (1957) has recognized varieties within *C. runssoroensis* and suggested that var. *aberdarensis*, endemic to Mt. Kenya and Aberdare, could be a hybrid between *C. monostachya* and *C. runssoroensis*. This variety can be distinguished from the typical variety by its thinner culms and the hyaline margin of its pistillate scale. On the other hand, var. *runssoroensis* has thicker culms and brown to blackish margins of the pistillate scale (Gehrke, 2011). Fig. 3b denotes sampling localities of the *Carex* taxa for the current study.

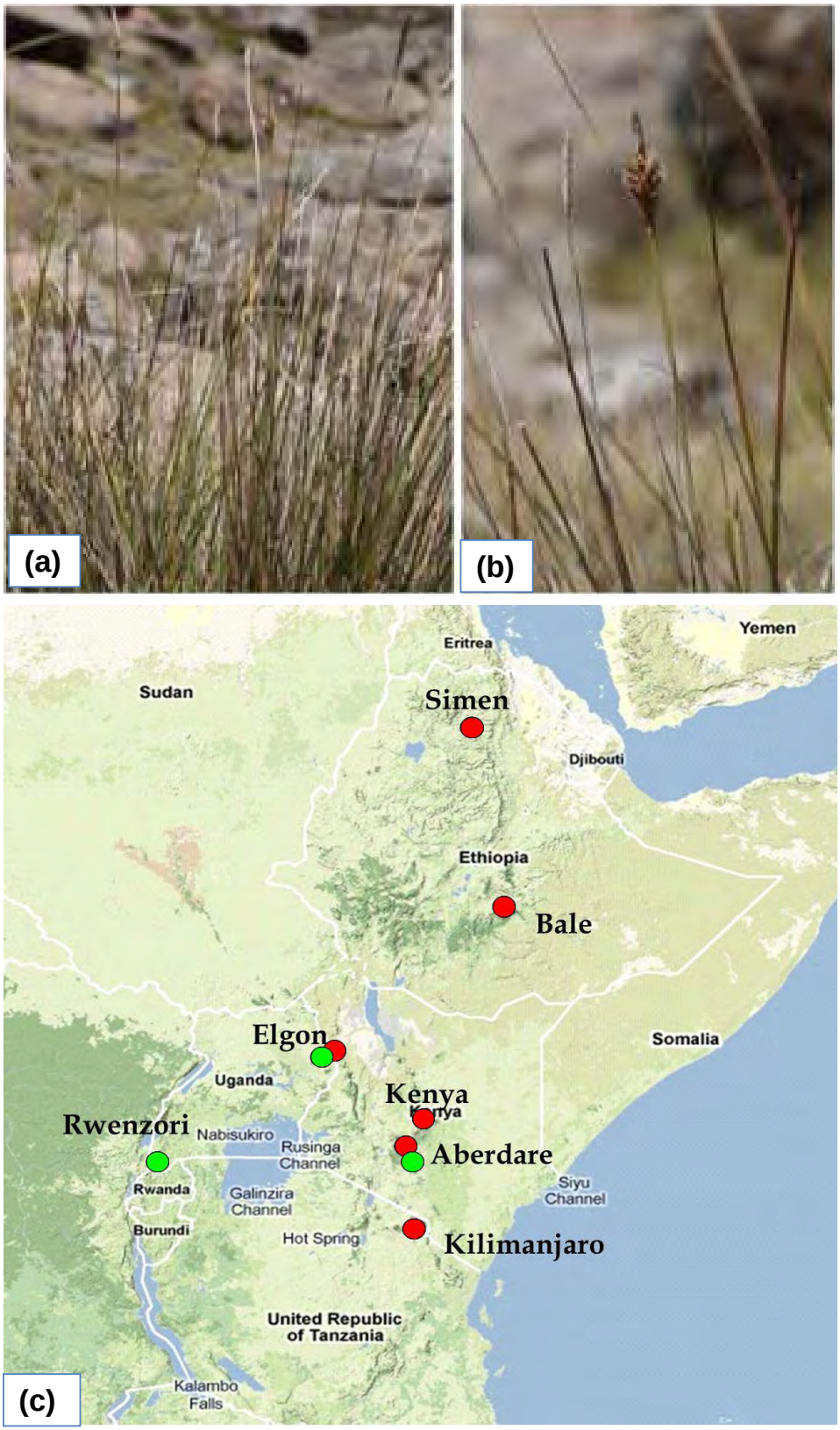


Fig. 3 (a and b) Growth form and inflorescence of *Carex monostachya*; (c) Sampling localities in East Africa.; circle colors indentify species (red - *C. monostachya*; and green - *C. runssoroensis*)

3.1.4 Lysimachia serpens (Hochst. Ex DC.) U.Manns & Anderb. (Primulaceae)

The genus *Anagallis* (Primulaceae) was reduced as a synonym of *Lysimachia* following the work of Manns & Anderberg (2009) and Anderberg *et al*. (2007). Following these findings, a new combination, *Lysimachia serpens*, was made for *Anagallis serpens. Lysimachia serpens* (Fig. 4a) is a perennial (or annual) herb, usually closely adpressed to the soil; stems rooting at the nodes throughout their length; leaves usually widest at or above the middle; pedicels fleshy; filaments glabrous. Taylor (1955) used opposite and alternate leaf arrangements as distinguishing morphological characters to recognize two subspecies; subsp. *serpens* and subsp. *meyeri-johannis*, respectively. Accordingly, in subsp. *serpens*, leaves are opposite, obovate to orbicular, 5-20 mm long, 3-12 mm broad; corolla lobes 8-10 mm long, whereas in subsp. *meyeri-johannis* leaves are alternate, oblanceolate to orbicular, usually obovate, 4-16 mm long, 3-10 mm broad; corolla lobes 3-13 mm long. The two subspecies were also reported to have allopatric distributions across the range of the species, extending from Sudan to Southern Zimbabwe in East Africa. Whereas subsp. *serpens* is distributed in Sudan, Ethiopia and Southern Zimbabwe, subsp. *meyeri-yohannis* occurs in the mountains of eastern Africa in Kenya, Tanzania, and Uganda (Taylor, 1955). In a recent taxonomic treatment, Manns & Anderberg (2009) reported however that leaves are always alternate in the species, and thus they did not recognize any subspecies.

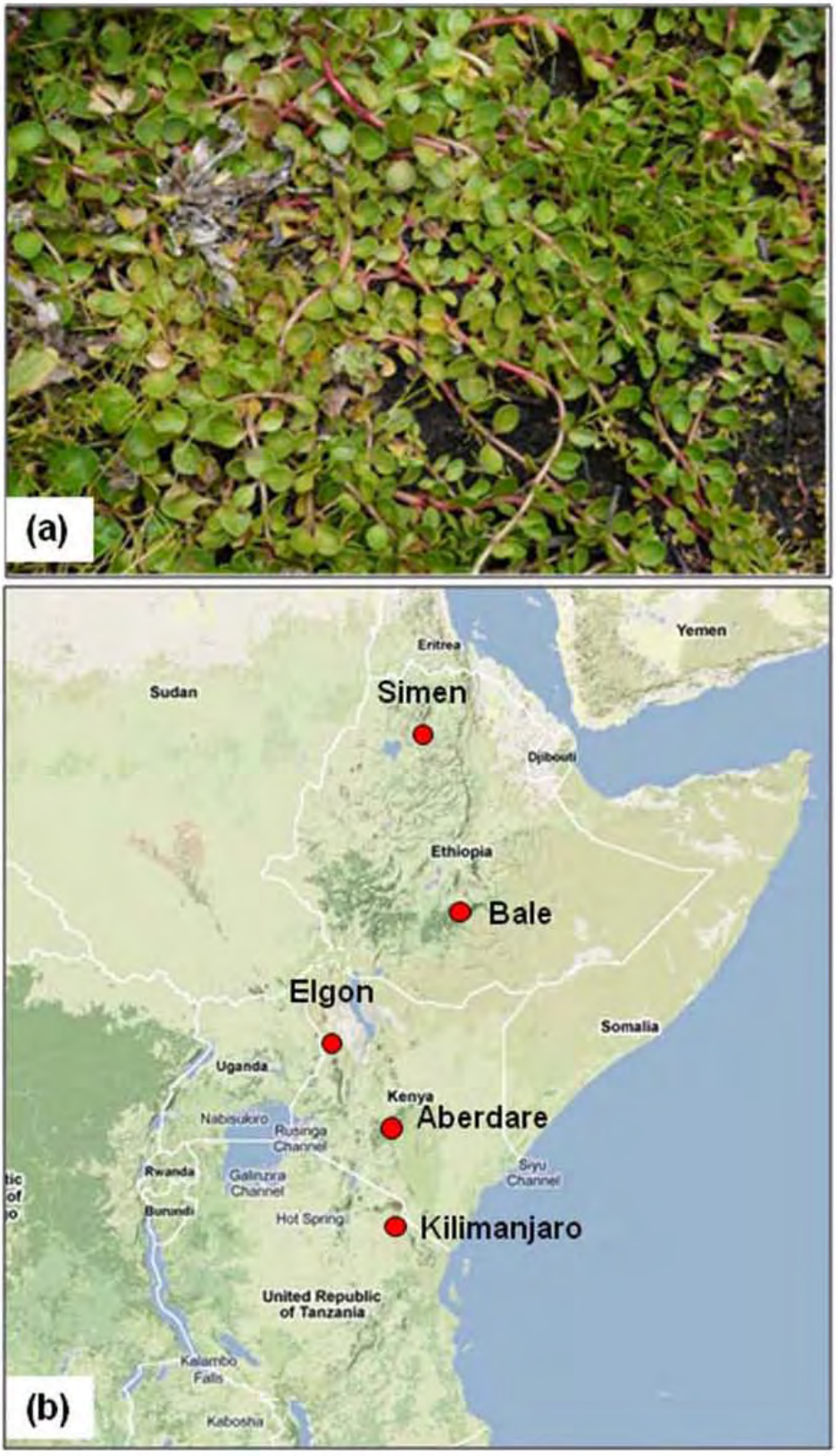


Fig. 4 (a) *Lysimachia serpens* and (b) sampling localities in Eastern Africa

3.1.5 Dicrocephala chrysanthemifolia (Bl.) DC. (Asteraceae)

Dicrocephala L'Hér. ex DC. is an annual or perennial herb native to Africa, Asia and the western Pacific region that has been introduced into Australia, Europe and Guatemala. Currently, three species are recognized in the genus (Beentje, 2002). In *Dicrocephala chrysanthemifolia* (Fig. 5a), capitula are usually 6–8 mm in diam.; phyllaries pilosulose to pilose; marginal pistillate florets with corollas funnelform to nearly salverform. In the other two species, *D. integrifolia* and *D. benthamii*, the capitulum is 2–4 mm in diam.; phyllaries glabrous or subglabrous; marginal pistillate florets with corollas tubular to urceolate. Different authors have suggested various taxonomical treatments as to whether to recognize two separate species or only two varieties/subspecies within the widely distributed *D. chrysanthemifolia* in the mountains of eastern Africa (Hedberg, 1957; Beentje, 2002; Tadesse, 2004). Hedberg (1957) and Tadesse (2004) used morphological characters and geographical distributions to recognize two closely related but separate species, *D. chrysanthemifolia* and *D. alpina*, in Eastern Africa. Accordingly, *D. chrysanthemifolia* is usually tall, 30-100 (-150) cm high, and occurs at mid-altitude, whereas *D. alpina* is short, 2-10(-20) cm, and dominates the afro-alpine habitat (Tadesse, 2004).

Dicrocephala chrysanthemifolia mainly grows in damp places in grasslands, forests, among grass along streams and river banks, and occasionally as a weed of cultivated and fallow fields or waste ground; 1500-3750 m.a.s.l. In Africa it is distributed in Ethiopia, Kenya, Tanzania, Uganda, Somalia, Rwanda, Sudan, Eritrea, in the west to Guinea and south to Malawi and Mozambique and in Madagascar. *Dicrocephala alpina* grows on rocky ground with thin soil in the

afro-alpine habitat under *Erica arborea*, *Helichrysum* spp., and *Lobelia rhynchopetalum*, 3500-4000 m.a.s.l. It occurs in Ethiopia, Kenya, Tanzania, Uganda, and DR Congo (Tadesse, 2004). Beentje (2002), however, reported that the differences in the morphological characters between the two species were few and gradual and suggested to reduce *D. alpina* to a variety of *D. chrysanthemifolia.* Some *D. alpina* specimens from Mt. Elgon were almost as tall as *D. chrysanthemifolia* and such continuity made Hedberg (1957) to acknowledge a varietal level of *D. alpina*, although he also thought it possible to treat them as separate species.

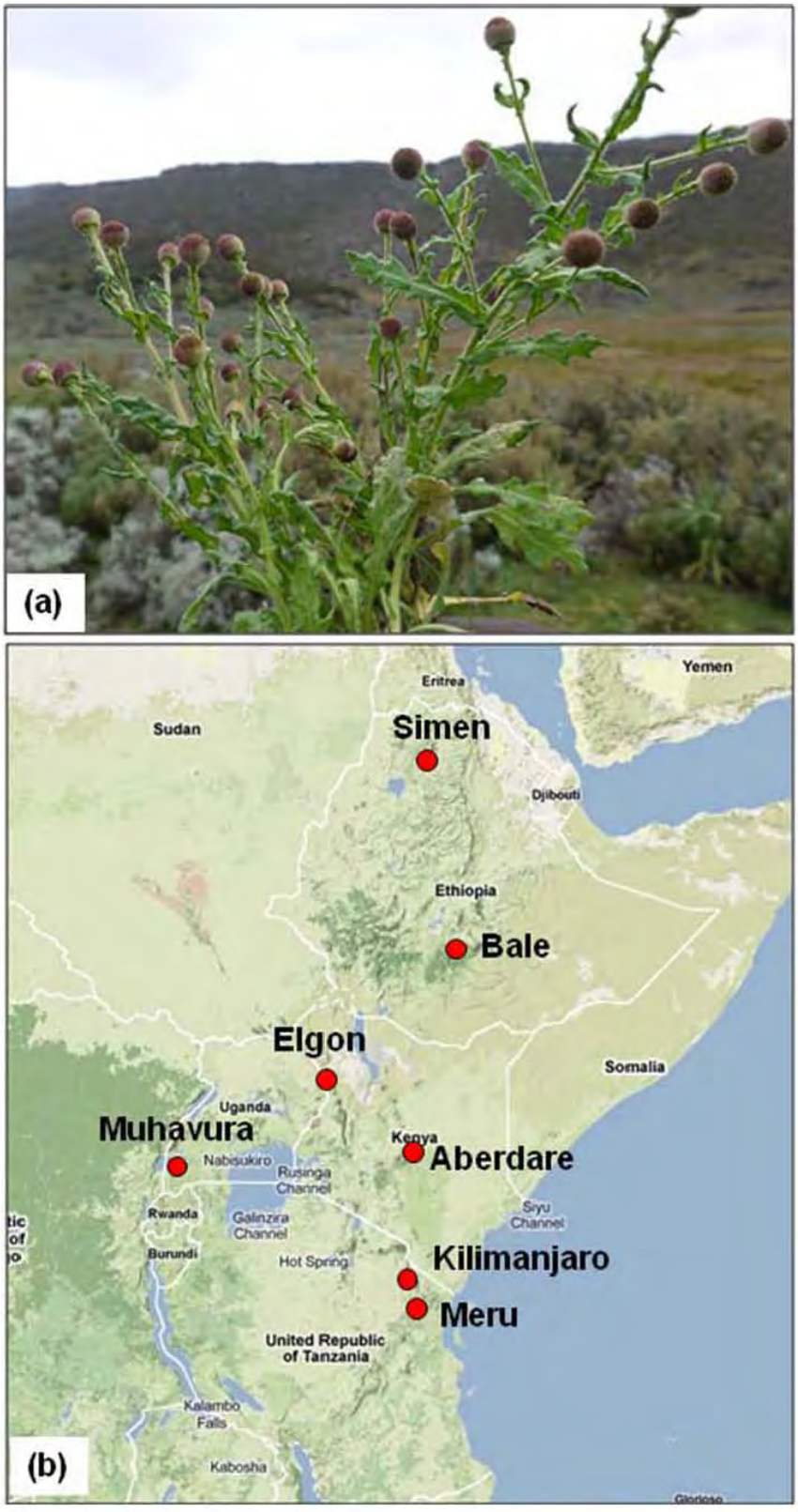


Fig. 5 (a) habit of *Dicrocephala chrysanthemifolia* and (b) sampling localities in Eastern Africa

3.1.6 Haplocarpha rueppellii (Sch. Bip.) Beauv. (Asteraceae)

The genus *Haplocarpha* comprises nine species, of which six are found in southern Africa and three in the afro-alpine region of East Africa as far north as Eritrea (Hedberg, 1957; Germishuizen and Meyer, 2003). *Haplocarpha rueppellii* (Fig. 6a) is a perennial herb with a short and stout caudex, crowned by a flat leaf rosette giving rise to a number of short-stalked or almost sessile capitula. Leaves are radical, few to numerous, spreading on ground to sub-erect. It may sometimes form large mats or cushions. It mainly occurs in open upland woodland or bushland with *Erica arborea*, *Helichrysum*, and *Alchemilla*, along streams and margins of forests on dark brown clay, eroded grassy slopes on red soils, and in dry *Hagenia – Juniperus* forest between altitudes of 2440 – 4400 (4700) m.a.s.l. It occurs in Ethiopia Kenya, Tanzania, and Uganda (Fig. 6b).

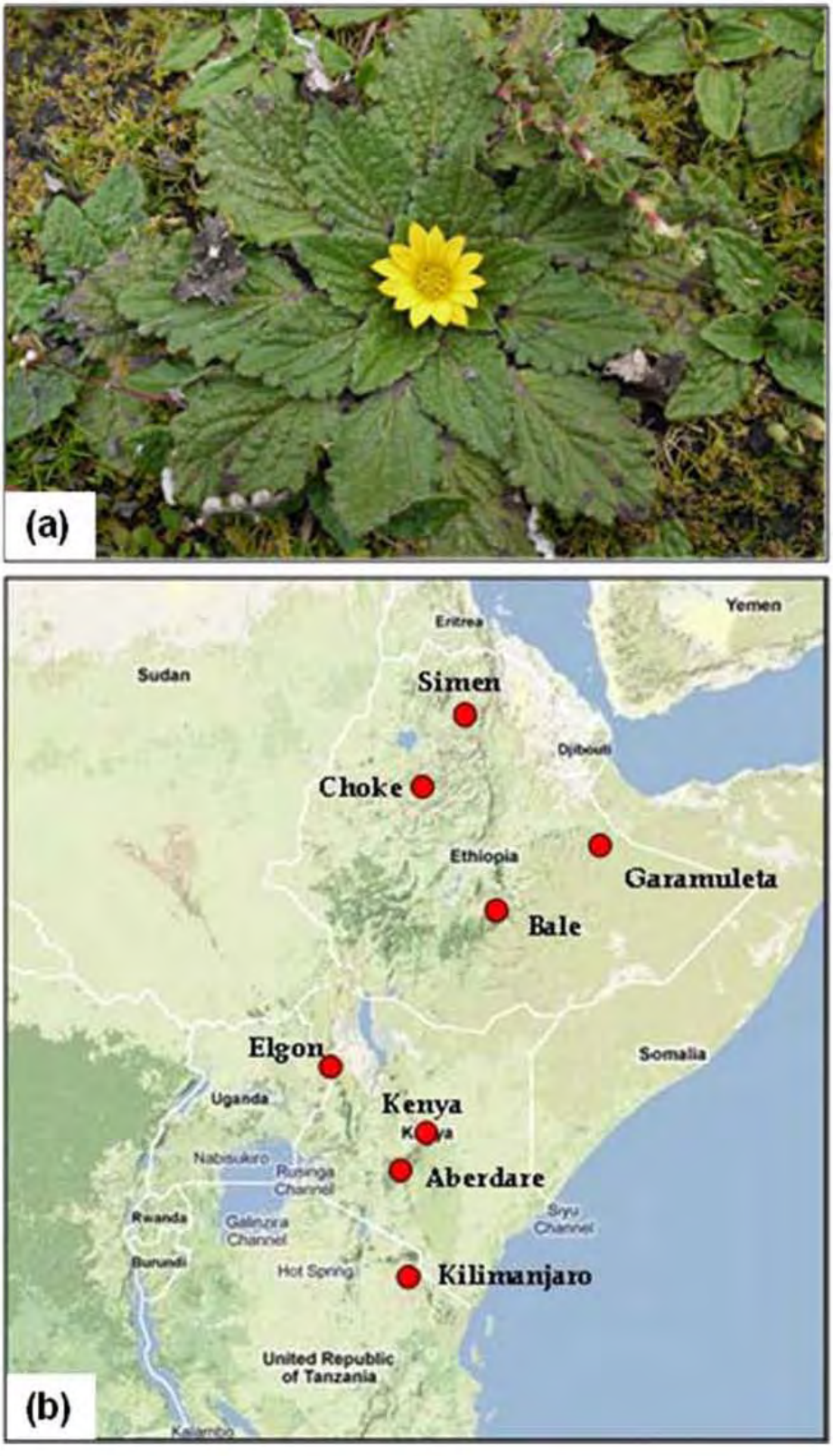


Fig. 6 (a) habit of *Haplocarpha rueppellii* and (b) sampling localities in Eastern Africa

3.2 Sampling strategies

Two different sampling strategies were employed in this study. The strategies were designed following their methodological suitability to address the two main evolutionary questions we set to investigate through molecular dating and phylogeographic inference.

3.2.1 Molecular dating

Leaf tissue of afro-montane *Lychnis* species was collected in the field from Ethiopia, Kenya, Uganda, Tanzania, and Cameroon in 2004-2009 (Table 1) and dried in silica gel. Voucher specimens are deposited at the National Herbarium, Addis Ababa University (Ethiopia); East African Herbarium, National Museum of Kenya (Kenya); Sokoine University of Agriculture (Tanzania); Makerere University Herbarium (Uganda) and Natural History Museum, University of Oslo (Norway). In addition, herbarium voucher samples of some closely related species, the Moroccan endemic *L. lagrangei*, the Eurasian *L. flos-cuculi*, and the Eurasian *L. subintegra*, were included in the study.

A total of 13 *mat*K sequences representing the afro-montane *Lychnis*, *L. lagrangei*, *L. flos-cuculi*, and *L. subintegra* were generated in this study. In addition, 33 *mat*K sequences representing the Alsionoideae and Caryophylloideae subfamilies as well as one outgroup sequence of *Polycarpon tetraphyllum* (Paronychioideae) were downloaded from GenBank (Table 2). The final data matrix of the *mat*K sequence, therefore, comprised forty-six terminal sequences. A middle to late Eocene (Jordan and Macphail, 2003) old inflorescence fossil of *Caryophylloflora paleogenica* (Caryophyllaceae) was

used to calibrate the rate of evolution of the *mat*K gene and infer the age of the *Lychnis* stem group.

The estimated age of the *Lychnis* stem group was then used for joint estimation of a multilocus species tree and divergence time of the afro-montane *Lychnis* using sequences of *mat*K, *rps*16, *psb*E-*pet*L, ITS, and *RPB2* genes. All the ITS and *RPB2* sequences as well as some sequences of *rps*16 and *psb*E-*pet*L were obtained from a previous study by Popp *et al.* (2008) and their GenBank accession numbers are indicated in Table 2.

.

Table 1 Materials used in the molecular dating analysis and sampling localities. All materials were collected in the field and dried in silica if not otherwise indicated

Database number	Population number	Taxon	Country	Locality	Altitude (m)	Latitude	Longitude
O-DP-13583	AFR522-4	*L. abyssinica*	Ethiopia	Arsi, Mt. Chilallo	3000	7.92686 N	39.1795 E
O-DP-13592	AFR528-5	*L. abyssinica*	Ethiopia	Arsi, Mt. Kaka	3400	7.36633 N	39.202 E
O-DP-13596	AFR533-1	*L. abyssinica*	Ethiopia	Arsi, 15 km from Kofele towards Asessa	2700	7.17354 N	38.8746 E
O-DP-13600	AFR536-1	*L. abyssinica*	Ethiopia	Bale Mountains	3200	6.93279 N	39.946 E
O-DP-13604	AFR540-1	*L. abyssinica*	Ethiopia	2 km away from Bale Mountains National Park Head Quarter towards Goba	3100	7.10558 N	39.8015 E
O-DP-13608	AFR545-1	*L. abyssinica*	Ethiopia	Shoa, Entoto	3000	9.08951 N	38.7648 E
O-DP-13616	AFR548-4	*L. abyssinica*	Ethiopia	Arsi, 12.5 km from Shashemene towards Kofele	2300	7.14924 N	38.699 E
O-DP-13619	AFR560-2	*L. abyssinica*	Ethiopia	Sidamo, 23 km from Ageremariam towards Dilla	2300	5.81814 N	38.265 E
O-DP-13621	AFR570-1	*L. abyssinica*	Ethiopia	Keffa, Masha	2200	7.76931 N	
O-DP-29620	ET098-2	*L. abyssinica*	Ethiopia	Gondar Simen Mountains, Gich	3574	13 15.996 N	38 06.469 E
O-DP-30300	ET271-2	*L. abyssinica*	Ethiopia	Gondar Simen Mountains, Gich	3652	13 16.183 N	38 06.353 E
O-DP-30734	ET391-2	*L. abyssinica*	Ethiopia	Gondar Simen Mountains, Silkii	3760	13 20 N	38 14 E
O-DP-31231	ET509-2	*L. abyssinica*	Ethiopia	Gondar Simen Mountains, Silkii	3596	13 19.218 N	38 14.108 E
O-DP-33637	ET1336-2	*L. abyssinica*	Ethiopia	Gojam, Choke	3960	10 38.52 N	37 50.14 E
O-DP-33712	ET1351-2	*L. abyssinica*	Ethiopia	Gojam, Choke	3943	10 39.36 N	37 49.54 E
O-DP-33907	ET1391-2	*L. abyssinica*	Ethiopia	Gojam, Choke	3919	10 39.45 N	37 49.32 E
O-DP-34660	ET1526-1	*L. abyssinica*	Ethiopia	Gara Muleta	3144	09 13.17 N	41 47.20 E
O-DP-34661	ET1526-2	*L. abyssinica*	Ethiopia	Gara Muleta	3144	09 13.17 N	41 47.20 E
O-DP-34662	ET1526-3	*L. abyssinica*	Ethiopia	Gara Muleta	3144	09 13.17 N	41 47.20 E

	AFR582-1	*L. abyssinica*	Cameroon				
	AFR582-2	*L. abyssinica*	Cameroon				
	AFR611-1	*L. abyssinica*	Cameroon				
	AFR611-2	*L. abyssinica*	Cameroon				
O-DP-27285	KN0443-2	*L. crassifolia*	Kenya	Cherangani Hills	3184	01 10.618 N	35 31.103 E
O-DP-27499	KN0493-2	*L. crassifolia*	Kenya	Aberdare	3086	00 32.55 S	36 43.05 E
O-DP-27812	KN0563-2	*L. crassifolia*	Kenya	Aberdare, Mutumbio Gate			
	180610-1	*L. floscuculi*					
O-DP-41977	UG2599-2	*L. kigesiensis* subsp. *kigesiensis*	Uganda	Kanaba Swamp	2300	01 15.32S	29 48.56E
O-DP-41982	UG2600-2	*L. kigesiensis* subsp. *kigesiensis*	Uganda	Kanaba Swamp	2300	01 15.32S	29 48.56E
O-DP-13628	AFR542-4	*L. kigesiensis* subsp. *ragazziana*	Ethiopia	2 km away from Bale Mountains National Park Head Quarter towards Goba	3100	7.10552 N	39.8015 E
O-DP-13629	AFR546-1	*L. kigesiensis* subsp. *ragazziana*	Ethiopia	Shoa, Entoto	3000	9.09496 N	38.7662 E
O-DP-13637	AFR534-3	*L. kiwuensis* subsp. *eralngeriana*	Ethiopia	Arsi, 15 km from Kofele towards Asessa	2700	7.17354 N	38.8746 E
O-DP-13648	AFR550-4	*L. kiwuensis* subsp. *eralngeriana*	Ethiopia	Arsi, 26 km from Shashemene towards Goba	2600	7.09851 N	38.7579 E
O-DP-13654	AFR555-1	*L. kiwuensis* subsp. *eralngeriana*	Ethiopia	Sidamo, 17 km from Agereselam towards Aleta Wondo	2500	6.57492 N	38.4891 E

O-DP-13657	AFR527-4	*L. kiwuensis* subsp. *kiwuensis*	Ethiopia	Arsi, Mt. Chilallo	3100	7.92063 N	39.184 E
O-DP-13662	AFR538-4	*L. kiwuensis* subsp. *kiwuensis*	Ethiopia	Bale, Rira Village	2800	6.75306 N	39.7204 E
O-DP-13664	AFR551-3	*L. kiwuensis* subsp. *kiwuensis*	Ethiopia	Gamo Gofa, Chencha	2600	6.2296 N	37.5678 E
O-DP-13671	AFR561-2	*L. kiwuensis* subsp. *kiwuensis*	Ethiopia	Illubabur, 15 km away from Tepi towards Gorie	1800	7.31531 N	35.3692 E
O-DP-13674	AFR566-1	*L. kiwuensis* subsp. *kiwuensis*	Ethiopia	Keffa, Masha	2200	7.76327 N	35.4745 E
O-DP-39472	UG2001-2	*L. kiwuensis* subsp. *kiwuensis*	Uganda	Mgahinga	2694	01 21.599 S	29 38.089 E
O-DP-39477	UG2002-2	*L. kiwuensis* subsp. *kiwuensis*	Uganda	Mgahinga	2352	01 21.165 S	29 37.206 E
O-DP-40181	UG2167-2	*L. kiwuensis* subsp. *kiwuensis*	Uganda	Kabaragnuma swamp	3058	01 22.049 S	29 40.279 E
O-DP-40186	UG2168-2	*L. kiwuensis* subsp. *kiwuensis*	Uganda	Kabaragnuma swamp	3058	01 22.049 S	29 40.279 E
	UP16925	*L. lagrangei**	Morocco	Friedhof in Oulad-el-Arbi, südlich Asilah, auf Küstenland		35.462597	-6.031698
	UP16926	*L. lagrangei**	Morocco	Habib, near the road from Arba-Ayacha (Soko Arbaa) to Dar Chaoui		35.47	-5.80
O-DP-27519	KN0497-2	*L. rotundifolia*	Kenya	Aberdare, on the way to Satima peak	3686	00 19.31 S	36 38.44 E
O-DP-36976	TZ0016-2	*L. rotundifolia*	Tanzania	Kilimanjaro	3636	03 02.055 S	37 14.58 E
O-DP-37105	TZ0043-2	*L. rotundifolia*	Tanzania	Kilimanjaro	3536	03 00.336 S	37 14.493 E
O-DP-13679	AFR554-4	*L. scottii*	Ethiopia	Gamo Gofa, Mt. Guge	3000	6.19919 N	37.3333 E

O-DP-13680	AFR554-5	*L. scottii*	Ethiopia	Gamo Gofa, Mt. Guge	3000	6.19919 N	37.3333 E
O-DP-13682	AFR554-1	*L. scottii*	Ethiopia	Gamo Gofa, Mt. Guge	3000	6.19919 N	37.3333 E
	UP16917	*L. subintegra**	Greece	Ioanninon, Metsovou, above Metsovo, N of the road to the Katara pass		39.776	21.175

* Herbarium material

Table 2 DNA sequences and their accession number retrieved from GenBank

Taxon	Population number	GenBank accession number				
		*mat*K	*rps*16	*psb*E-*pet*L	ITS	*RPB2*
L. abyssinica	21568		EF602350	EF602323	EF602379	-
	21571		EF602351	EF602324	EF602380	EF602407
	21591		EF602352	EF602325	EF602381	EF602408
	AFR522_4		EF602353	EF602326	-	EF602409
	AFR528_5		EF602354	EF602327	-	EF602410
	AFR533_1		EF602355	EF602328	EF602382	EF602411
	AFR536_4		EF602356	EF602329	EF602383	EF602412
	AFR540_1		EF602357	EF602330	EF602384	EF602413
	AFR545_1		EF602358	EF602331	EF602385	EF602414
	AFR548_4		-	-	EF602386	EF602415
	AFR560_2		EF602360	EF602333	EF602387	EF602416
	AFR570_1		EF602361	EF602334	EF602388	EF602417
L. chalcedonica			_	-	X86894	AJ634068
L. coronaria			EF674193	FJ376841	X86891	AJ634069
			-	-	AY857966	-
			-	-	SCU30953	-

		-	-	SCU30979	-
L. crassifolia	21567	EF602362	EF602335	EF602389	EF602418
L. flos-cuculi		Z83163	EF602320	X86893	AJ634070
		-	-	EF407939	AJ634071
		-	-	-	FJ376910
		-	-	SFU30957	-
L. flos-jovis		Z83166	EF602321	X86892	-
		-	-	AY936261	-
		-	-	EF407940	-
L. kigesiensis subsp. *kigesiensis*	21595	EF602366	EF602339	EF602393	EF602422
L. kigesiensis subsp. *ragazziana*	AFR542_4	EF602367	EF602340	EF602394	EF602423
	AFR546_1	EF602368	EF602341	EF602395	EF602424
L. kiwuensis subsp. *eralngeriana*	AFR534_3	EF602363	EF602336	EF602390	EF602419
	AFR550_4	EF602364	EF602337	EF602391	EF602420
	AFR555_1	EF602365	EF602338	EF602392	EF602421
L. kiwuensis subsp. *kiwuensis*	21598	-	-	EF602396	-
	AFR527_4	EF602370	EF602342	EF602397	EF602425
	AFR538_4	EF602371	EF602343	EF602398	-
	AFR551_3	EF602372	EF602344	EF602399	EF602426
	AFR561_2	EF602373	EF602345	EF602400	EF602427
	AFR566_1	EF602374	EF602346	EF602401	EF602428
L. lagrangei	LAG01	-	-	EF602377	EF602404
	LAG02	EF602349	EF602322	EF602378	EF602405

L. rotundifolia	21573		EF602375	EF602347	EF602402	EF602429
L. scottii	AFR554_1		EF602376	EF602348	EF602403	EF602430
L. coronaria		FJ589507	-	-	-	-
L. flos-jovis		AY936313	-	-	-	-
Polycarpon tetraphyllum		AY936287	-	-	-	-
Scleranthus perennis		AY514847	-	-	-	-
Arenaria serpylloides		FJ404826	-	-	-	-
Moehringia macrophylla		FJ404852	-	-	-	-
Stellaria media		FJ404877	-	-	-	-
Lepyrodiclis holosteoides		FJ404840	-	-	-	-
Minuartia nuttallii		FJ404847	-	-	-	-
Bufonia paniculata		FJ404827	-	-	-	-
Schiedea globosa		DQ907818	-	-	-	-
Geocarpon minimum		FJ404836	-	-	-	-
Gypsophila paniculata		FJ404838	-	-	-	-
Petrorhagia saxifraga		FJ404857	-	-	-	-
Dianthus armeria		FJ404832	-	-	-	-
Arenaria bryophylla		FJ404815	-	-	-	-
Arenaria gypsophiloides		FJ404818	-	-	-	-
Agrostemma githago		FJ589503	-	-	-	-
Petrocoptis pyrenaica		FJ589508	-	-	-	-
Silene hawaiiensis			-	-	-	-
Silene uniflora		FJ589565	-	-	-	-
Silene douglasii		EF547238	-	-	-	-
Silene noctiflora		EF547240	-	-	-	-
Silene repens		FJ589552	-	-	-	-

Silene schafta	FJ404873	-	-	-	-
Silene ciliata	FJ589519	-	-	-	-
Silene muscipula	FJ589543	-	-	-	-
Silene antirrhina	FJ589512	-	-	-	-
Silene armena	FJ589514	-	-	-	-
Silene acaulis	EF547235	-	-	-	-
Silene campanula	AY936311	-	-	-	-
Silene yemensis	FJ589567	-	-	-	-
Silene sordida	FJ589559	-	-	-	-
Silene odontopetala	FJ589546	-	-	-	-

3.2.2 Phylogeography

Samplings for phylogeography study was conducted in 11 high mountain systems in Ethiopia and Tropical East Africa (Kenya and Tanzania) in 2003/2004 and 2007-2006. Whenever possible four sample plots with 100 m x 100 m dimension were randomly selected in each mountain and fresh leaf tissues from five individuals were collected and dried in silica gel. The silica material and voucher specimens were deposited at the National Herbarium of Addis Ababa University, Addis Ababa, Ethiopia, and duplicates of the silica material were deposited at the Natural History Museum, University of Oslo (O).

Samplings of *E. arborea* and *E. trimera* speices (Table 8, Fig. 2) were however difficult because of the very similar morphology which are difficult to distinguish except by some floral characters. As many of our populations for both species were not in flower or fruit at the time of collection, their taxonomic assignment had to be re-evaluated after the genetic analysis was carried out. This procedure resulted in consistent separation of the species, since we identified two distinct genetic groups (cf. Results section) to which those specimens collected with flowers or fruits could be assigned. Although this approach resulted in a somewhat uneven representation of the two species in our sampling from the different mountains, it was sufficient to address the broad histories of both species. For the AFLP analysis we used a total of 169 individual plant samples collected in the field of which 59 were belonged to *E. arborea* and 110 to *E. trimera*. Of the total of 169 samples used in the AFLP analysis, nine of the same samples of *E. arborea* and 14 of *E. trimera* were re-used for the plastid DNA sequencing. Additional samples for plastid DNA

sequencing of *E. arborea* and *E. trimera* were also obtained from silica or voucher specimens (Table 8).

3.3 DNA extraction, PCR amplification and DNA sequencing

Total genomic DNA was extracted from silica gel-dried leaf samples and herbarium material (for DNA sequencing) using an automated GeneMole® robot (Qiagen Nordic, Oslo, Norway) or DNeasy™ Plant Mini Kit (Qiagen, Valencia, CA). Leaf tissue was mechanically ground in 2.0 mL tubes with two tungsten carbide beads during c. 2 min at 15 Hz in a mixer mill (MM301, Retsch GmbH & Co., Haan, Germany), 300 µL of lysis buffer was added to the crushed material, vortexed, centrifuged at 3400 Hz for 20 sec, incubated on a heat block for 15 min at 65°C, and centrifuged at 14000 Hz for 3 min. 200 µL of the lysate was transferred into new tubes and loaded to the GeneMole® robot.

For the afro-montane *Lychnis*, the amplification and DNA sequencing protocols described in Popp *et al*. (2008) were used to amplify and sequence the two non-coding chloroplast DNA regions of *rps*16 intron and *psb*E-*pet*L interspacer. An approximately 1566 base pairs (bp) length of the *mat*K region was amplified and sequenced in three overlapping fragments using primers (matK-F2b-matK-547.rc), (matK-526-matK-R2), and (matK-1066- matK-1463.rc) (Mower *et al.*, 2007; Sloan *et al.*, 2009). Each PCR reaction included Ready-To-Go™ PCR Beads (Amhersham Pharmacia Biotech, UK), 0.5 µL of each 10 µM forward and reverse primers, and 1 µL of genomic DNA of 100 ng/µL concentration.

For the *Erica* species, four geographically distant individuals of each species were selected for a preliminary test of sequence variation in five non-coding cpDNA regions, *psb*E-*pet*L, *trn*D-*trn*T, *rps*16, *rpl*32-*trn*L$^{(UAG)}$ and *trn*G$^{\boldsymbol{UUC}}$-*trn*G2G. Based on amplification success and levels of intraspecific variation, we selected *rpl*32-*trn*L$^{(UAG)}$ and *trn*G$^{\boldsymbol{UUC}}$-*trn*G2G as final sequencing regions. Amplifications of the two regions were performed using primers (rpL32-F and trnL^{UAG}) for *rpl*32-*trn*L$^{(UAG)}$ (Shaw *et al.*, 2007) and (3'trnG^{UUC} and 5'trnG2G) for *trn*G$^{\boldsymbol{UUC}}$-*trn*G2G (Shaw *et al.*, 2005). The total PCR reaction volume of 10 µL contained 0.4 U *Taq* DNA polymerase (Applied Biosystems, Foster City, CA, USA), 1 X PCR buffer supplied with the enzyme, 2.5 mM Mg^{2+}, 0.4 µL of each primer, 1mM of each dNTP (Applied Biosystems), 0.04% BSA and 3 µL template DNA of unknown concentration. PCR cycling was performed with a GeneAmp 3700 (Applied Biosystems, Foster City, CA, USA) or a PTC100 or PTC200 (MJ Research, Watertown, MA, USA) thermocycler with initial denaturation for 3 min at 95 °C followed by 35 cycles of 30 sec at 95 °C, 30 sec at 55 °C, 2 min at 72 °C, and ending with 10 min at 72 °C with subsequent cooling at 6 °C.

PCR products were purified using 2 µL of Exosap-IT (USB Corp., Cleveland, OH, USA), diluted 1:10 for 5 µL PCR products, and incubated for 45 min at 37°C and 15 min at 80°C. The same primers utilized during the PCR were also used in the DNA sequencing. DNA sequencing was performed using BigDye V 1.1 (Applied Biosystems) according to the manufacturer's manual, except for using 10 µL reaction volumes, and visualized with an ABI 3100 capillary sequencer (Applied Biosystems). The sequencing program started with initial

denaturing at 96°C for 1 min followed by 25-35 cycles of 96°C for 10 sec, 50°C for 5 sec, 60°C for 4 min and a final extension at 60°C for 10 min.

3.4 AFLP fingerprinting

AFLP was performed to assess the genome-wide genetic variation following the protocol in Gaudal *et al*. (2000) with a few modifications. Genomic DNA was digested with two restriction enzymes of 6-base cutters (*Eco*RI; 5'-G↓AATTC-3') and 4-base cutters (*Mse*I; 5'-T↓TAA-3'). The double-stranded *Eco*RI adaptors and *Mse*I adaptors were ligated to the sticky ends created by the restriction enzymes to generate a template DNA for PCR. A restriction-ligation reaction was incubated at 37° C for 3 hr and consisted of 1.1 µL of 10X T4 ligase buffer, 1.1 µL of 0.5 M NaCl, 0.55 µL of 1 mg/mL bovine serum albumin (BSA), 0.02 µL of 50 U/µL MseI, 0.125 µL of 40 U/µL EcoRI, 0.2 µL of 5 U/µL T4- DNA ligase, 1 µL of 10 µM *Mse*I-adapter,1 µL of 10 µM *Eco*RI-adapter and 0.41 µL of double distilled sterile water (ddH2O), 5 µL template DNA of unknown concentration in a final reaction volume of 10 µL. Prior to their addition into the reaction mixture, the two adapter pairs (*Mse*I and *Eco*RI) were denatured at 94°C for 5 min and annealed by slow cooling to room temperature. The final restriction-ligation product was diluted 10 times with ddH2O before the pre-selective reaction.

Restriction-ligation products were diluted 10 times with purified water prior to the pre-selective reaction. Pre-selective PCR was performed using 1.5 µL of the diluted restriction-ligation product and 11µL of pre-selective reactions consisting of 1.25 µL of 10X PCR buffer, 0.75 µL of 25 mM MgCl2, 1 µL of10

mM dNTPs, 0.25 µL of10 µM *Eco*RI-A primer, 0.25 µL of 10 µM *Mse*I-C primer, 0.075 µL of 5 U/µL AmpliTaq DNA Polymerase (Perkin-Elmer), and 7.45 µL of ddH2O having a total reaction volume of 12.5 µL. PCR reactions were carried out in a thermocycler GeneAmp PCR System 9700 (PE Biosystems) with the following program: 2 min. at 72°C and 30 cycles of 30 sec. denaturing at 94°C, 30 sec. annealing at 56°C, and 2 min. extension at 72°C, and completing of the extension for 10 min. at 72°C. Successive amplification was checked with 1% agarose gel electrophoresis by loading 5 µL of pre-selective product, 2.5 µL of loading buffer and 1 µL of ethidium bromide. Pre-selective PCR products were diluted 20 times in ddH2O and stored at -20°C or used immediately.

An initial screening of 12 primer combinations was performed on two individuals of each species from two different mountains. Six primer pairs resulting in AFLP profiles with many polymorphic markers and well separated fragments were used in a second primer test. This extended primer test was performed with six new primer pairs (i.e. with a total of 12 primer combinations) on eight individuals from seven different mountains and one duplicate. Three selective primer combinations were then selected for the final analysis of all samples of each species (Table 3). Whenever four selective nucleotides were selected in the *Mse*I primer, pre-selective primers with two selective bases were used to avoid mismatching amplification (Vos *et al.*, 1995).

Selective PCR was performed using the three best primer combinations selected for each plant species (Table 3). The selective PRC reaction was composed of

2.5 µL diluted pre-selective PCR products, 1.25µL 10X PCR buffer, 1.25µL 2.5mM MgCl2, 0.10µL 1mg/ml BSA, 1.0µL 10mM dNTPs, 0.1 µL of 10 µM *Eco*RI + A*NN* primer and 0.25 µL of 10 µM *Mse*I + C*NN* primer marked with a fluorescent dye specific for each plant species (*N* in the primer indicates selective nucleotides), and 0.10 µL 5U AmpliTaqGold Polymerase (Perkin-Elmer) and 5.95 µL ddH_2O was added to make the total reaction volume to 12.5 µL. PCR reactions were performed with the following selective PCR parameters: 10 min at 95°C, 13 cycles of 30 sec denaturing at 94°C, 1 min. annealing temperature by decreasing from 65-56°C and 1 min elongation at 72°C, followed by 23 cycles for 30 sec at 94°C, 1 min at 56°C, 1 min at 72°C; and 10 min final extension at 72°C.

Table 3 Selective primer combinations used in the AFLP analysis

Species	Primer combinations
Erica arborea	6FAM-*Eco*RI-AGA/*Mse*I-CTC
	VIC-*Eco*RI-ACA/*Mse*I-CAT
	NED-*Eco*RI-AAC/*Mse*I-CAG
Erica trimera	6FAM-*Eco*RI-AGA/*Mse*I-CTC
	VIC-*Eco*RI-ACA/*Mse*I-CAT
	NED-*Eco*RI-AAC/*Mse*I-CAG
Carex monostachya	6FAM-*Eco*RI-ATG/*Mse*I-CG
	VIC-*Eco*RI-AAG/*Mse*I-CC
	NED-*Eco*RI-AAC/*Mse*I-CG
Carex runssoroensis	6FAM-*Eco*RI-ATG/*Mse*I-CG
	VIC-*Eco*RI-AAG/*Mse*I-CC
	NED-*Eco*RI-AAC/*Mse*I-CG
Lysimachia serpens	6FAM-*Eco*RI-ACA/*Mse*I-CAC
	VIC-*Eco*RI-AGG/*Mse*I-CAA
	NED-*Eco*RI-ACC/*Mse*I-CTT
Dicrocephala chrysanthemifolia	6FAM-*Eco*RI-AGA/*Mse*I-CAC
	6FAM-*Eco*RI-AGA/*Mse*I-CTG
	VIC-*Eco*RI-ACA/*Mse*I-CAC
Haplocarpha rueppellii	6FAM-*Eco*RI-AGA/*Mse*I-CTCC
	VIC-*Eco*RI-ACA/*Mse*I-CTAG
	NED-*Eco*RI-ATG/*Mse*I-CTAG

For each sample, 2 μL 6-FAM, 2 μL VIC and 3 μL NED labeled selective PCR products were pooled together and diluted in 14 μL distilled water. A total volume of 3.5 μL of the diluted reaction was transferred into SEPHADEX® G-50 Superfine Resin (GE Healthcare Bio-Science, Uppsala, Sweden) and mixed with 11.7 μL HiDi formamide and 0.3 μL GeneScan^{TM} 500 ROX^{TM} internal-lane size standards, denatured at 95°C for 5 min, cooled on ice and run on an ABI 3100 sequencer (Applied Biosystems). For each species eight to ten percent of the total numbers of individuals subjected in the final run were duplicated by re- extracting DNA and repeating the whole AFLP procedure. The duplicated individual samples were then used to test the reproducibility of the markers and also to calculate an error rate according to Bonin *et al*. (2004).

3.5 Sequence alignment and analysis

The sequences were edited using CodonCode aligner v.3.5.7 (CodonCode Corporation, Dedham, MA). Multiple sequence alignment was done using ClustalW in SeaView v.4 program (Gouy *et al.*, 2010) and adjusted manually.

3.5.1 Plastid DNA sequence analysis for phylogeography

For phylogeographic analysis of the *Erica* dataset, sequences of the two regions were concatenated and then divided into two partitions for analysis. Best-fit DNA substitution models were determined for each data partitions using PAUP* 4.10b (Swofford, 2002) and the Akaike information criterion (AIC) as implemented in MrModeltestv.3.1 (Nylander, 2004). Bayesian phylogenetic analyses were performed for the concatenated dataset using MrBayes v.3.1.2

(Huelsenbeck and Ronquist, 2001; Ronquist and Huelsenbeck, 2003). Two independent analyses were run simultaneously, each starting from different random trees. Four Markov Chain Monte Carlo (MCMC) chains, i.e., three 'heated' and one 'cold' chain, were run for 2 x 10^6 generations. Trees were sampled every 1000^{th} generation and resulted in a total of 2 x 10^3 trees. The first 500 were discarded as the burn-in phase. Convergence between the two independent analyses was evaluated and confirmed using the standard deviation of split frequencies. The majority rule consensus phylogenetic tree with posterior probabilities was visualized with FigTree v.3.1 (Rambout, 2009). The tree was midpoint rooted. Haplotypes were identified using DnaSP v.5 (Librado and Rozas, 2009) and a maximum parsimony median-joining haplotype network was constructed with the NETWORK v.4.6.0 program (Bandelt *et al.*, 1999). Gaps were treated as missing data and each indel was scored as a single binary character irrespective of its length. Haplotype and nucleotide diversity for the total samples of each species was estimated using DnaSP v.5 (Librado and Rozas, 2009).

3.5.2 Molecular dating analysis

3.5.2.1 Phylogenetic analyses and fossil calibration of the mat*K phylogeny*

The input files for Bayesian Evolutionary Analysis Sampling Trees (BEAST) and 'Species Tree Ancestral Reconstruction' (*BEAST) were prepared in XML format using BEAUti v. 1.6.1 (Drummond and Rambaut, 2007; Heled and Drummond, 2010). A monophyletic group was defined to include all taxa except *Polycarpon tetraphyllum*. A single inflorescence fossil record of *Caryophylloflora paleogenica* (Caryophyllaceae) (Jordan and Macphail, 2003)

was used to calibrate the tree. The fossil is inferred to be from middle to late Eocene, which corresponds to 48.6-33.9 Myr (http://www.paleodb.org). A lognormal distribution prior with a logmean = 0.0, standard deviation = 1.37 and offset of 33.9 Myr was assigned to the Alsinoideae/Caryophylloideae node. The standard deviation was estimated by assigning a 95% probability distribution lying between the minimum age of the oldest fossil (33.9 Myr) and a soft-maximum bound corresponding to the lower boundary of the middle Eocene period (48.6 Myr) (Ho and Phillips, 2009). A lognormal distribution for priors fixes the minimum age of a calibrated node and allows the maximum age to be sampled following a lognormal distribution with no hard limit (Ho and Phillips, 2009).

A preliminary analysis was set up to determine the right clock model used in the BEAST analyses of *mat*K dataset. A substitution model of GTR+G was tested against the codon position based model of SRD06 both in the strict and uncorrelated lognormal relaxed clock model. The results from the preliminary analysis rejected the strict molecular clock, thus all subsequent analyses of the *mat*K dataset were modelled using a relaxed molecular clock. Both a Yule and a birth-death speciation process were used as a tree prior in separate analyses. All other defaults in BEAUti were left unchanged for all parameters. Four independent Markov Chains Monte Carlo (MCMC) were run for 10 million generations with tree and parameter values saved every 10,000 generations. To test the influence of the prior on the posterior estimate, one additional chain was run for 10 million generations without data, sampling only the prior. Convergence of the chains to stationary distributions was confirmed by

inspection of the MCMC samples in each analysis using the program Tracer 1.5 (Rambaut and Drummond, 2007) and joint estimates were produced using LogCombiner v 1.6.1 (Drummond and Rambaut, 2007).

3.5.2.2 Phylogenetic analyses and inference of a species tree

A species tree and divergence time for the afro-montane *Lychnis* was co-estimated from *mat*K, *rps*16, *psb*E-*pet*L, ITS, and *RPB2* genes using *BEAST (Heled and Drummond, 2010). All substitution models, clock models and tree models were unlinked between the datasets, except the *rps*16 and *psb*E-*pet*L tree model that were linked. The sampling for *mat*K DNA sequences lacked a large proportion of sequences compared to the *rps*16 and *psb*E-*pet*L genes and thus the tree topology is left unlinked. Best fitting model of sequence evolution for each gene was estimated following the same method used for the fossil calibrated *mat*K dataset.

The posterior age estimate of *Lychnis* from fossil calibrated *mat*K BEAST analysis was used as the prior age in the *BEAST analysis of the species tree. The XML file was manually edited to include priors for the species tree root age with a normal distribution having a mean value = 7.87 and a standard deviation = 1.25 which approximate the 95% HPD from the four BEAST runs of the calibrated *mat*K dataset (5.34-10.55). A preliminary *BEAST analysis was run using the uncorrelated lognormal clock model to determine the clock-like model of the dataset and test the performance of the default strict clock model for each of the datasets. If the 95% HPD of the posterior distribution of the standard deviation of the substitution rate (ucld.stdev) includes 0 then the data are quite

clock-like and a strict molecular clock model cannot be rejected. Based on the results from the preliminary analysis, a strict clock could not be rejected for ITS and *rps*16. The remaining datasets were analysed using an uncorrelated lognormal clock model.

The final *BEAST analysis was performed using a Yule tree prior, constant population size model with an autosomal nuclear ploidy type set for ITS and *RPB2* and mitochondrial for the linked *rps*16 and *psb*E-*pet*L as well as *mat*K. Four independent analyses were run for 100 million generations and trees and parameter values sampled every 10,000 generations. To test the influence of the prior on the posterior estimate one additional chain was run for 100 million generations without data, sampling only the prior. Convergence of the chain to stationary distributions was confirmed by inspection of the MCMC samples in each analysis using the program Tracer 1.5 (Rambaut and Drummond, 2007) and joint estimates were produced using LogCombiner v 1.6.1 (Drummond and Rambaut, 2007).

3.6 AFLP data scoring and phylogeographic analysis

AFLP bands were aligned and visualized with the internal size standard using ABI prism GENESCAN v.3.7 analysis software (Applied Biosystems) and imported for scoring into GeneMapper version 4.0 (Applied Biosystems). AFLP fragments in the size range of 50 - 500 base pairs (bp) for each primer combinations were scored automatically either as present (1) or absent (0), and corrected manually before exporting the AFLP marker as presence/absence matrix. The error rate was calculated for each primer combinations for each

species as recommended by Bonin *et al.* (2004) by taking the ratio of mismatches (scoring of 0 versus 1) over matches (1 versus 1) in AFLP profiles of the duplicated individuals. The error rate was further used to evaluate the quality of the fingerprints.

3.6.1 Genetic distance

Similarity between AFLP multilocus phenotypes was inferred using Dice's coefficient of similarity in NTSYSpc v.2.11 (Rohlf, 1998), and principal coordinate analyses (PCoAs) were used to visualize pair-wise similarities among AFLP phenotypes. A neighbour-joining analysis of Nei-Li distance measurement (Nei and Li, 1979) based on a genetic distance matrix was computed using the software TREECON 1.3b (Van de Peer and De Wachter, 1994). The trees were midpoint rooted and support for the branches was estimated with 1,000 bootstrap replicates.

3.6.2 Genetic structuring and differentiation

Genetic mixture analyses were performed to identify the two species in the total dataset of *Erica* (*E. arborea* and *E. trimera*) and *Carex* (*C. monostachya* and *C. runssoroensis*), and to identify genetically homogeneous groups within each species. The appropriate number of groups (*K*) and the group to which each individual most likely belonged to were estimated using STRUCTURE (Pritchard *et al.*, 2000). The program implements a model based clustering method using Bayesian MCMC estimation. By comparing the likelihood of the data estimated in different runs for different numbers of groups (*K*), it is possible to identify the optimal *K*. The best value of *K* was chosen based on two

criteria: the estimated posterior log probability of the data, L (K) and the stability of assignment patterns across runs. Since L (K) continued to grow slightly with increasing values of *K*, the most likely number of clusters was determined also by taking into account the rate of change in the probability between successive *K*, ΔK (Evanno *et al.*, 2005). Individuals are assigned probabilistically to one of the clusters defined by allele frequencies at each locus. Our data were analyzed with the STRUCTURE program v.2.3.3 at the Bioportal, University of Oslo (http://www.bioportal.uio.no) with $K = 1–10$ and 10 replicates per *K*, a burn-in period of 2×10^5 and 10^6 iterations. We used the recessive allele model to take into account the dominant nature of AFLP data (Pritchard *et al.*, 2000). For analysis including both species of *Erica* (*E. arborea* and *E. trimera*) and *Carex* (*C. monostachya* and *C. runssoroensis*), we used a no admixture model with uncorrelated allele frequencies, as well as a model allowing for admixture and correlated allele frequencies. Assuming that populations are more closely related within species, we used only the latter for the within-species analyses. The output files from the STRUCTURE run were summarized and visualized using STRUCTURE HARVESTER, a web-based program (available at http://taylor0.biology.ucla.edu/structureHarvester/; Earl, 2011). Population genetic structure and differentiation were also assessed using Analyses of Molecular Variance (AMOVAs) with Arlequin v.3.01 (Excoffier and Schneider, 2005). AMOVAs were run with different hierarchical levels based on the genetic groupings inferred from the STRUCTURE analyses. Non-hierarchical AMOVAs to estimate among-populations and among-individuals-within-populations variation components were also performed for each species.

3.6.3 Genetic diversity and distinctiveness

Genetic diversity was estimated for each species in total, for each mountain massif, for each sampling locality (population) within mountain massifs and for the identified genetic groups. We used AFLPdat to estimate Nei's gene diversity (*D*; corresponding to the average proportion of pairwise differences between AFLP profiles, i.e. phenotypes (Kosman, 2003). In addition we estimated genetic distinctiveness for each mountain massif and population as frequency-down-weighted marker values (DW or rarity) following Schönswetter and Tribsch (2005), using AFLPdat R-script (Ehrich, 2006).

3.6.4 Hybrid detection

The preliminary PCoA of the total dataset for *C. monostachya* and *C. runssoroensis* identified seven individuals of the two species from Elgon (one population of each species with 3 or 4 individuals) placed in an intermediate position (along the first axis) between the two species; indicating a possible hybrid population. In order to further test these possible hybrids between *C. monostachya* and *C. runssoroensis*, a multilocus assignment test was performed using AFLPOP 1.2 (Duchesne and Bernatchez, 2002). In AFLPOP, the differences in frequencies among polymorphic loci are used to assign individuals to their potential source populations or species based on the likelihood of observing their AFLP genotype in the source populations, given the population frequencies of AFLP bands. Individuals are also assigned to potential (synthetic) hybrid populations of different genealogy (F1, F2, and first backcrosses to each parental population), based on estimation of allele

frequencies in the parental populations by assuming Hardy–Weinberg (H-W) equilibrium.

Prior to the analysis, the total dataset of the two species was pruned for redundant loci (bands) with the '*Loci filter'* option in the AFLPOP program. Three different sets of separate analysis were performed to identify the most likely parents of the possible hybrid individuals from Elgon. First all individuals in the total dataset were assigned to each of the two species datasets (source species). All individuals were then again assigned to each source species and to the synthetic first generation hybrid between the two source species (F1: *C. monostachya* x *C. runssoroensis*). In the third analyses, the possible-hybrid individuals identified in the second analysis were assigned to each species' source population, F1, second generation hybrid (F2: F1 x F1), and first backcrosses to each species' source population (F1B1: F1 x *C. monostachya,* and F1B2: F1 x *C. runssoroensis*). The analyses were performed with 1000 randomly generated genotypes for each of the six categories (two source species, F1, F1B1, F1B2, and F2). Those 6000 simulated individuals were then blindly reassigned to their most probable category using '*Allocation (unknown origin)*' option in the AFLPOP. Detection of hybrids was confirmed by using two criteria. In the minimal log likelihood difference, a hybrid was assigned to any of the six categories if the likelihood was ten times higher for belonging to that group than to any other group, otherwise the hybrid was not assigned and in the second criteria the assignment of the hybrid should not be statistically rejected ($P < 0.01$).

CHAPTER 4 RESULTS

4.1 Molecular dating of the afro-montane *Lychnis* L. (Caryophyllaceae)

Details about gene specific information including the total number of characters and total taxa included in the study are presented in Table 4. The best fit models estimated for each gene according to MrAIC include the generalized time reversible model with gamma rate distribution (GTR+G) for ITS and *mat*K, the generalized time reversible model with invariant sites (GTR+I) for *rps*16, the Hasegawa, Kishino and Yano model with invariant sites (HKY+I) for *psb*E-*pet*L, and the Hasegawa, Kishino and Yano model with gamma rate distribution (HKY+G) for *RPB2*.

Table 4 Gene specific information used in molecular dating

	*mat*K	*rps*16	*psb*E-*pet*L	ITS	*RPB2*
Number of total taxa	46	57	57	35	29
Total number of characters	1566	666	1253	602	1842
Conserved sites	778	616	1160	500	1567
Variable sites	741	41	90	100	231
Parsimony informative sites	429	28	69	63	134

*4.1.1 Fossil calibration of mat*K

The result from the preliminary analyses using an uncorrelated lognormal relaxed molecular clock showed that the marginal distribution of the standard deviation of the substitution rate did not include 0.0 (0.24 to 0.54), thus a strict molecular clock was rejected and a relaxed molecular clock was used for the *mat*K dataset. The posterior estimates of all parameters in the joint estimate for each of the four BEAST analyses showed high effective sample sizes greater

than 1800. The result from the MCMC inspection indicated that all analyses had converged effectively. The maximum clade credibility tree with all estimated node ages based on the combined result from the four separate analyses of the *mat*K dataset is presented in Fig. 7. The tMRCA of the genus *Lychnis* was estimated to be 7.87 Myr ago [5.38-10.55, 95% HPD] (Fig. 7, Node B, Table 5), and the age for the afro-montane *Lychnis* crown group was inferred to be 3.86 Myr [2.43-5.35, 95% HPD] (Fig. 7, Node C, Table 5).

Table 5 Node mean to the most recent ancestor (million years), 95% HPD interval of lower and upper limits from fossil-calibrated *mat*K gene tree based on Fig.7 as estimated from BEAST

Node	Mean	95% HPD lower	95% HPD upper
A	35.7	34	40.6
B	7.87	5.38	10.55
C	3.86	2.43	5.35

Table 6 Node mean to the most recent ancestor (million years), 95% HPD interval of lower and upper limits from species tree and individual gene tree of the genus *Lychnis* based on Fig.8, 9a-d as estimated from *BEAST

Node	Mean *Species tree/mat*K/*rps*16&*psb*E-*pet*L/ITS/*RPB2*	95% HPD lower *Species tree/mat*K/*rps*16&*psb*E-*pet*L/ITS/*RPB2*	95% HPD upper *Species tree/mat*K/*rps*16&*psb*E-*pet*L/ITS/*RPB2*
A	4.95/5.11/5.08/6.19/5.5	2.86/3.04/3.01/3.73/3.15	7.15/7.43/7.39/8.8/8.06
B	4.21/4.36/4.38/4.64/4.63	2.39/2.45/2.42/2.63/2.62	6.32/6.48/6.46/6.89/6.87
C	3.95/4.1/4.09/4.58/4.22	2.28/2.37/2.36/2.62/2.39	5.8/6/6/6.92/6.2
D	2.44/2.54/2.56/2.74/2.9	1.38/1.43/1.42/1.55/1.62	3.64/3.77/3.78/4.15/4.43
E	1.28/1.41/1.52/1.52/2.43	0.58/0.67/0.75/0.75/0.98	2.13/2.32/2.5/2.51/4.41

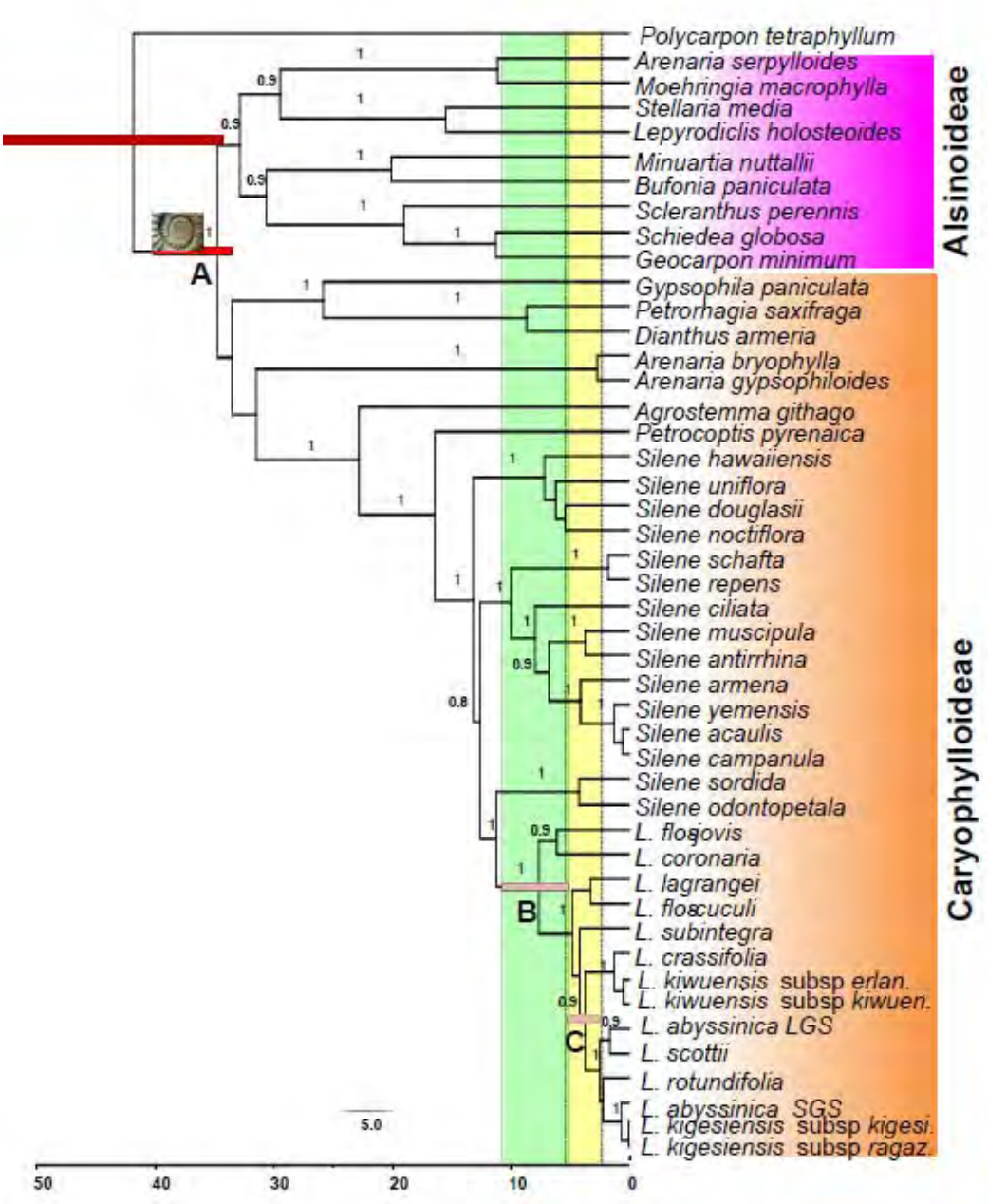


Fig. 7 Maximum clade credibility tree of *mat*K sequences (obtained with BEAST) and the 95% HPD age interval estimates of nodes inferred from forty-six sequences representing species of Alsionoideae and Caryophylloideae subfamilies and the outgroup *Polycarpon tetraphyllum* (Paronychioideae). Node bars and their corresponding letter indicate the 95% HPD interval estimates for the fossil calibrated node (Node A), tMRCA of the genus *Lychnis* and the afro-montane *Lychnis* (Node B and C, respectively). Numbers associated with nodes indicate posterior probabilities. Scale bar indicates million years (Myr).

4.1.2 Phylogenetic dating of species tree

The results from the preliminary *BEAST analysis suggested a strict clock model for ITS, *rps*16 and *mat*K and an uncorrelated relaxed lognormal clock for the *RPB2* and *psb*E-*pet*L regions. The effective sample size values of all parameters were greater than 270 in the joint estimate of the five separate *BEAST analyses of the afro-montane *Lychnis*. The split between the ancestral lineage of the afro-montane *Lychnis* and its sister lineage comprising the Eurasian *L. flos-cuculi* and the Moroccan endemic *L. lagrangei* was estimated to 4.95 Myr ago [2.86-7.15, 95% HPD] (Node A, Fig. 8, Table 6) in the species tree and ranged from 5.08 to 6.19 Myr ago (Node A, Table 6) in the gene trees. The tMRCA for the afro-montane *Lychnis* was estimated to 3.95 Myr ago [2.28-5.8, 95%HPD] (Node C, Fig. 8, Table 6) in the species tree and ranged from 4.09 to 4.58 Myr ago (Node C, Table 6) in the gene trees. The tMRCA for the diploid 'section *Uebelinia'* was estimated to 2.44 Myr ago [1.38-3.64, 95%HPD] (Node D, Fig. 8, Table 6) in the species tree and ranged between 2.54 to 2.9 Myr ago (Node D, Table 6) in the gene trees. The tMRCA for the tetraploid 'section *Trigynuebelinia'* was estimated to 1.28 Myr ago [0.58-2.13, 95%HPD] (Node E, Fig. 8, Table 6) in the species tree and ranged between 1.41 to 2.43 Myr ago (Node E, Table 6) in the gene trees.

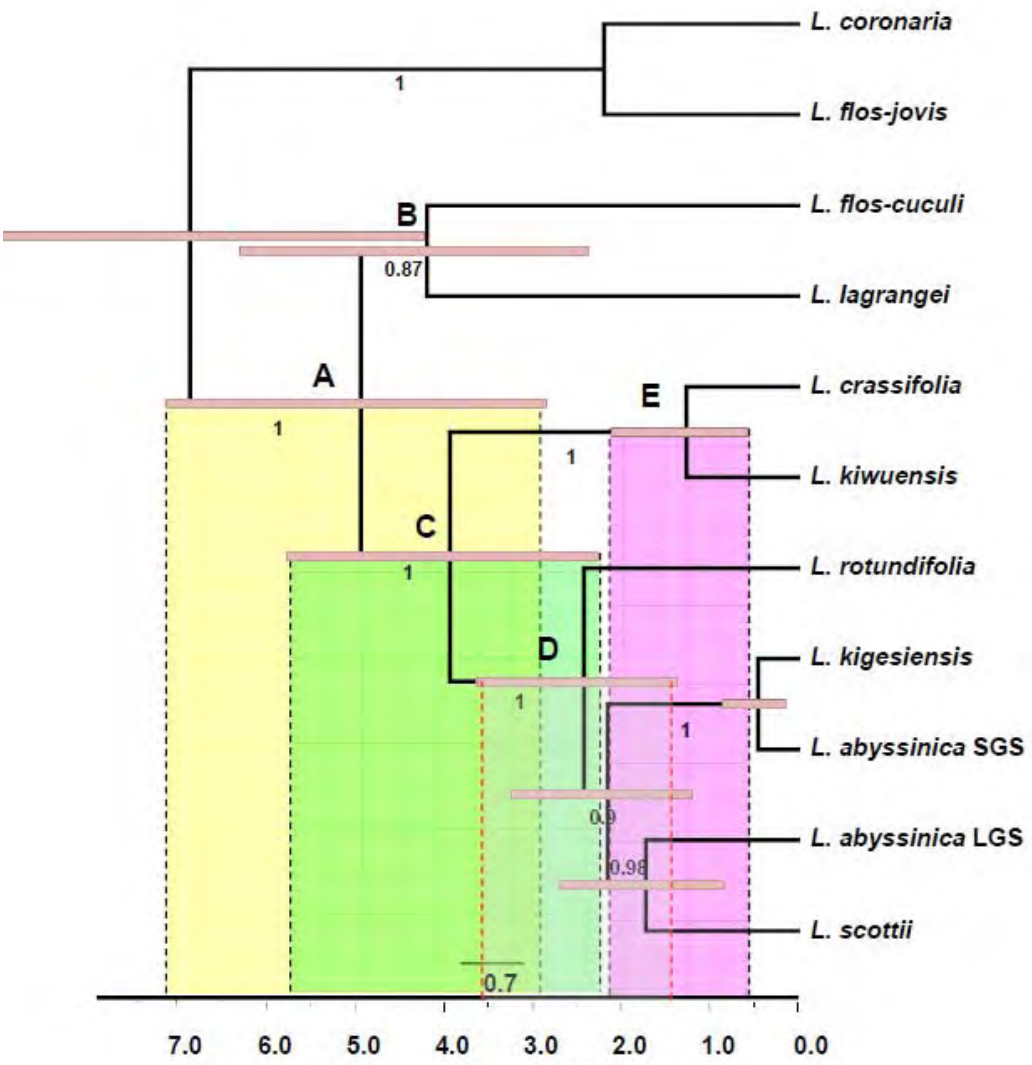
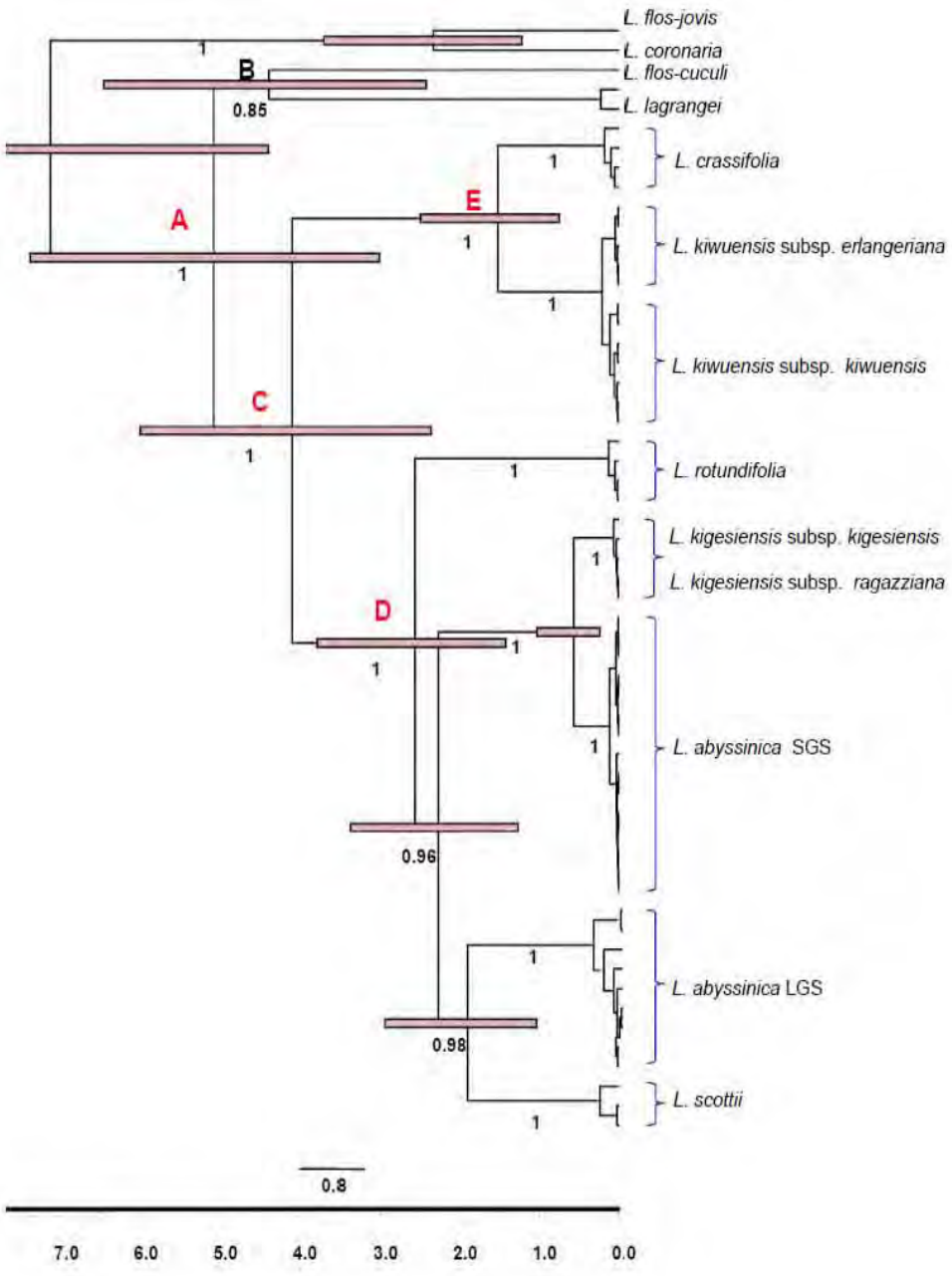
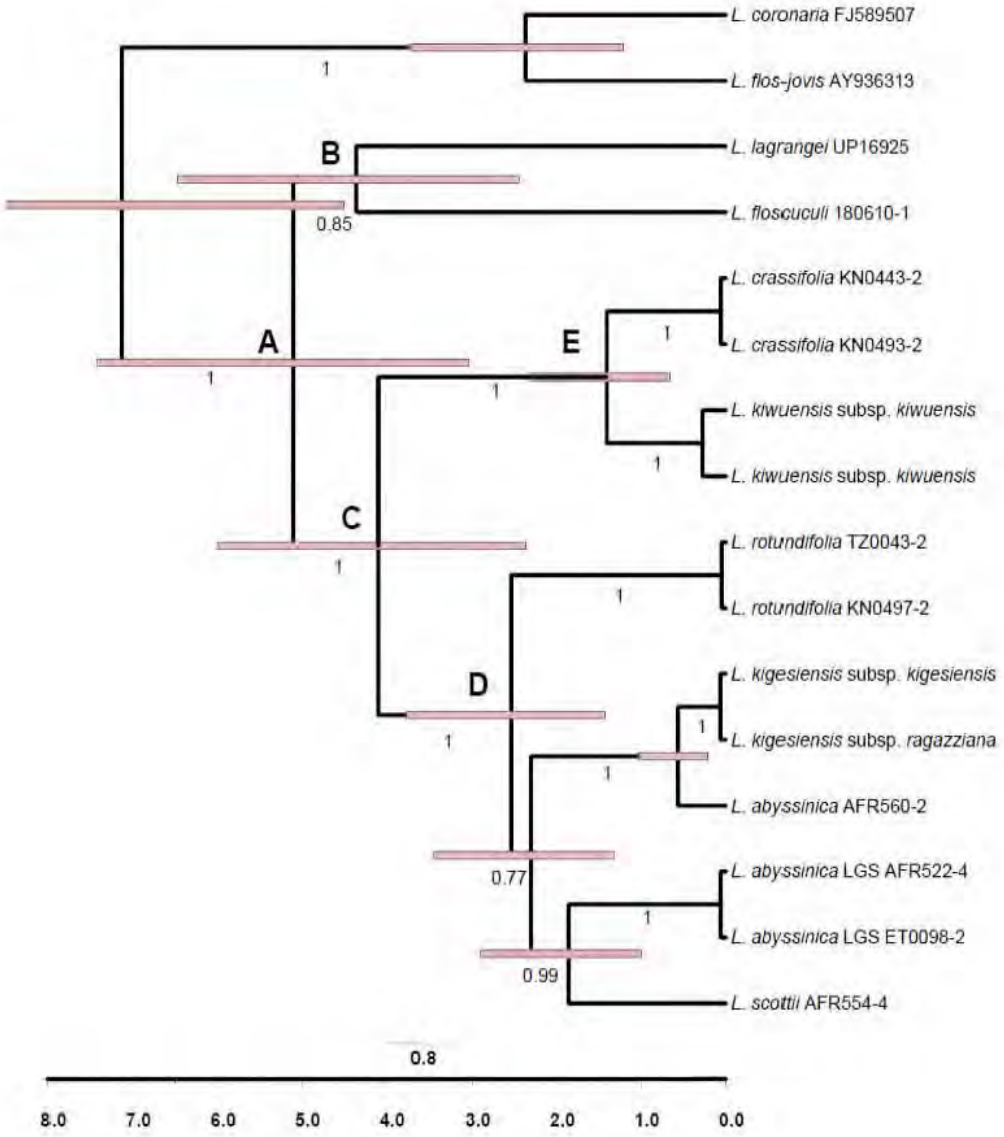


Fig. 8 The multilocus species tree of the genus *Lychnis* (obtained with *BEAST) inferred from five DNA regions (*mat*K, *rps*16, *psb*E-*pet*L, ITS and RPB2). Node bars and the shaded color indicate 95% HPD age interval estimates. Nodes letters A-E and their estimated ages are discussed in the text and Table 6. Numbers associated with nodes indicate posterior probabilities. Scale bar indicates million years (Myr). Terminal names indicate species of the genus *Lychnis* and *L. abyssinica* LGS and *L. abyssinica* SGS are used to indicate the large-genome-size and small-genome-size variants of *L. abyssinica*, respectively

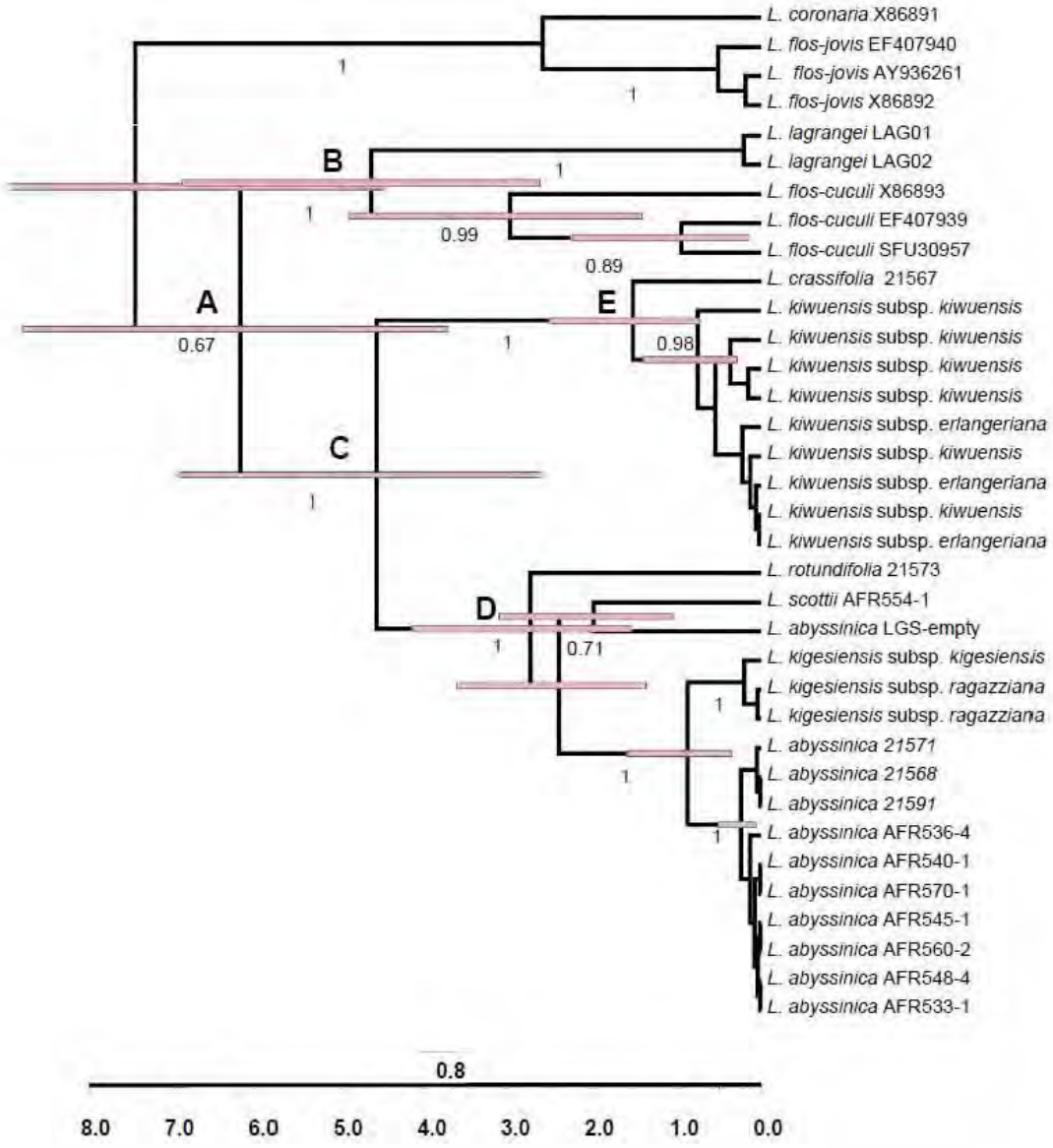
(a)


L. flos-jovis
L. coronaria
L. flos-cuculi
L. lagrangei
L. crassifolia
L. kiwuensis subsp. erlangeriana
L. kiwuensis subsp. kiwuensis
L. rotundifolia
L. kigesiensis subsp. kigesiensis
L. kigesiensis subsp. ragazziana
L. abyssinica SGS
L. abyssinica LGS
L. scottii
A
B
C
D
E
1
0.85
0.96
0.98
0.8
7.0
6.0
5.0
4.0
3.0
2.0
1.0
0.0

(b)


L. coronaria FJ589507
L. flos-jovis AY936313
L. lagrangei UP16925
L. floscuculi 180610-1
L. crassifolia KN0443-2
L. crassifolia KN0493-2
L. kiwuensis subsp. kiwuensis
L. kiwuensis subsp. kiwuensis
L. rotundifolia TZ0043-2
L. rotundifolia KN0497-2
L. kigesiensis subsp. kigesiensis
L. kigesiensis subsp. ragazziana
L. abyssinica AFR560-2
L. abyssinica LGS AFR522-4
L. abyssinica LGS ET0098-2
L. scottii AFR554-4
A
B
C
D
E
1
0.85
0.77
0.99
0.8
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0
0.0

(c)


L. coronaria X86891
L. flos-jovis EF407940
L. flos-jovis AY936261
L. flos-jovis X86892
L. lagrangei LAG01
L. lagrangei LAG02
L. flos-cuculi X86893
L. flos-cuculi EF407939
L. flos-cuculi SFU30957
L. crassifolia 21567
L. kiwuensis subsp. kiwuensis
L. kiwuensis subsp. kiwuensis
L. kiwuensis subsp. kiwuensis
L. kiwuensis subsp. kiwuensis
L. kiwuensis subsp. erlangeriana
L. kiwuensis subsp. kiwuensis
L. kiwuensis subsp. erlangeriana
L. kiwuensis subsp. kiwuensis
L. kiwuensis subsp. erlangeriana
L. rotundifolia 21573
L. scottii AFR554-1
L. abyssinica LGS-empty
L. kigesiensis subsp. kigesiensis
L. kigesiensis subsp. ragazziana
L. kigesiensis subsp. ragazziana
L. abyssinica 21571
L. abyssinica 21568
L. abyssinica 21591
L. abyssinica AFR536-4
L. abyssinica AFR540-1
L. abyssinica AFR570-1
L. abyssinica AFR545-1
L. abyssinica AFR560-2
L. abyssinica AFR548-4
L. abyssinica AFR533-1
A
B
C
D
E
1
1
1
0.99
0.89
0.67
1
0.98
1
0.71
1
1
1
0.8
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0
0.0

(d)

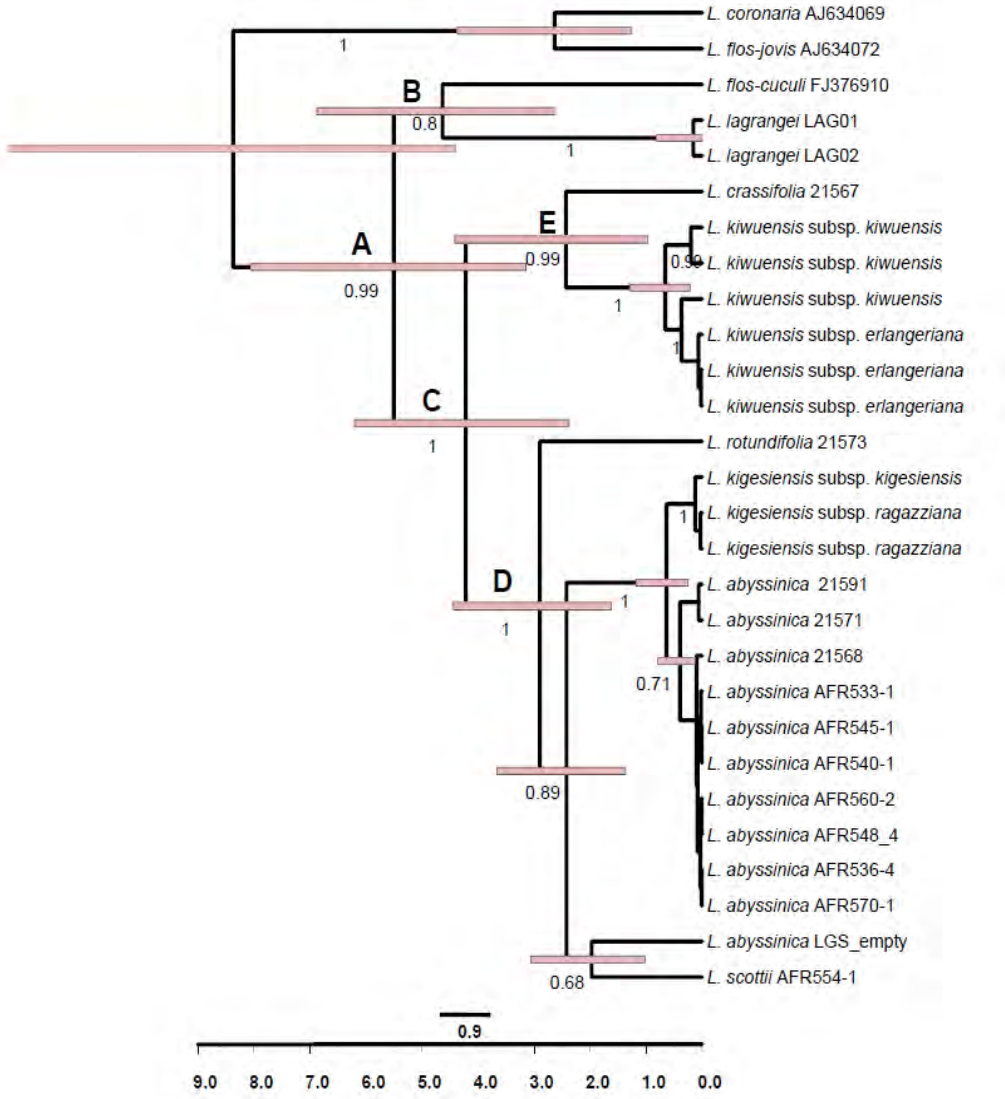


Fig. 9 Maximum clade credibility gene trees of (a) the concatenated chloroplast DNA sequences of *rps*16 and *psb*E-*pet*L, (b) *mat*K, (c) ITS, and (d) RPB2. Node bars indicate the 95% HPD age interval estimates of Nodes (A-E). Nodes letters A-E and their estimated ages are discussed in the text and Table 6. Numbers associated with nodes indicate posterior probabilities. Scale bar indicates million years (Myr). Terminal names indicate species of the genus *Lychnis* and *L. abyssinica* LGS and *L. abyssinica* SGS are used to indicate the large-genome-size and small-genome-size variants of *L. abyssinica*, respectively

4.2 Phylogeographic analysis

4.2.1 Erica arborea and *E. trimera* (Ericaceae)

4.2.1.1 Plastid DNA variation

The final concatenated data matrix consisted of a total of 33 sequences representing 16 plants of *E. arborea* and 17 plants of *E. trimera* and 1632 nucleotide sites. The *rpl32-trn*L$^{(UAG)}$ region consisted of 924 sites, and the *trn*G^{UUC}-*trn*G2G region consisted of 708 sites. A 52 bp indel was identified in the *rpl32-trn*L$^{(UAG)}$ region. Only a single plastid DNA haplotype was observed in *E. arborea.* Excluding the indel observed in the *rpl32-trn*L$^{(UAG)}$ region, the *E. trimera* matrix contained 1580 sites, of which four were parsimony informative and one singleton. Eight different plastid DNA haplotypes (H-1 to H-8) with total haplotype diversity of 0.868 and nucleotide diversity of 0.0013 were observed in *E. trimera* (Table 7). The Ethiopian haplotype H-3 was most frequently observed in our samples (frequency 0.294), whereas rare haplotypes occurred in Kilimanjaro (H-4), Meru (H-5) and Elgon (H7 and H8).

The best-fit models estimated for each partitioned dataset were the Hasegawa, Kishino and Yano model (HKY) for the *rpl32-trn*L$^{(UAG)}$ region and the generalized time reversible with invariant sites model (GTR+I) for the *trn*G^{UUC}-*trn*G2G region. Both species were strongly supported as monophyletic groups with posterior probability (PP) value of 1. The relationships among the eight plastid DNA haplotypes observed in *E. trimera* were poorly resolved with some exceptions. The Ethiopian ones (H-1 - H-3) were grouped together with strong support (PP = 1; Fig. 11). The haplotypes from Mt Elgon (H-7 and H-8) were placed in a basal polytomy with the Ethiopian clade in a strongly supported

clade (PP = 0.99; Fig. 11). In the plastid DNA haplotype network (Fig. 10), the Ethiopian haplotypes were placed at one end and the East African haplotypes at the other, with the two Mt Elgon haplotypes in the middle. All haplotypes of *E. trimera* were connected by single mutational steps, whereas the single haplotype observed in *E. arborea* was separated by eight mutational steps from the nearest haplotype of *E. trimera*.

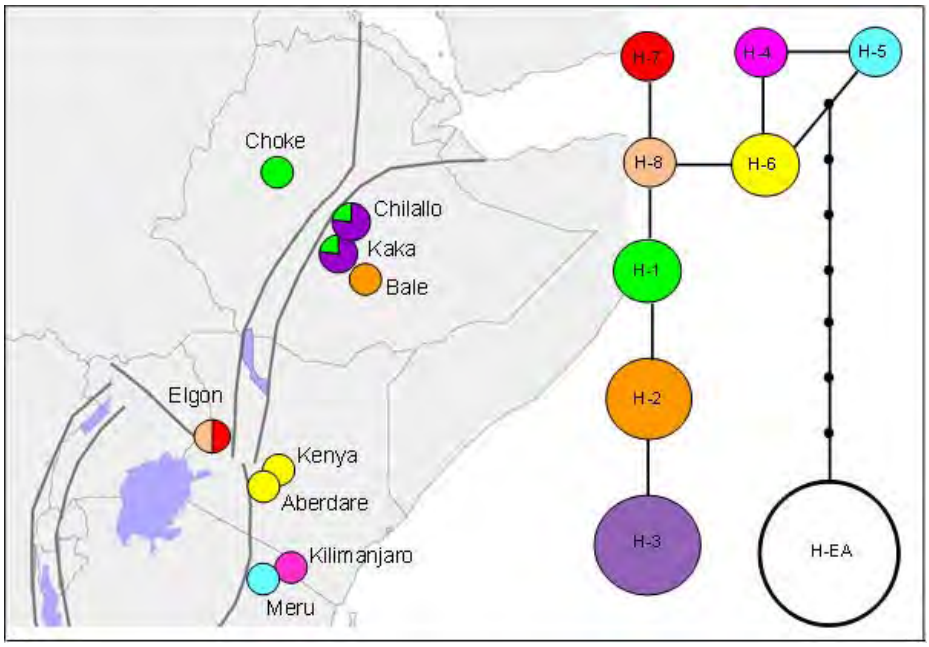


Fig. 10 Plastid DNA sequence variation and the seven haplotypes (H1 – H7) in *E. trimera*. The network also includes the single haplotype observed in *E. arborea* (H-EA), with black dots indicating hypothetical intermediate haplotypes (extinct or unsampled) and circle size reflecting sample size. The pie diagrams on the map show haplotype frequencies; colours are as in the haplotype network.

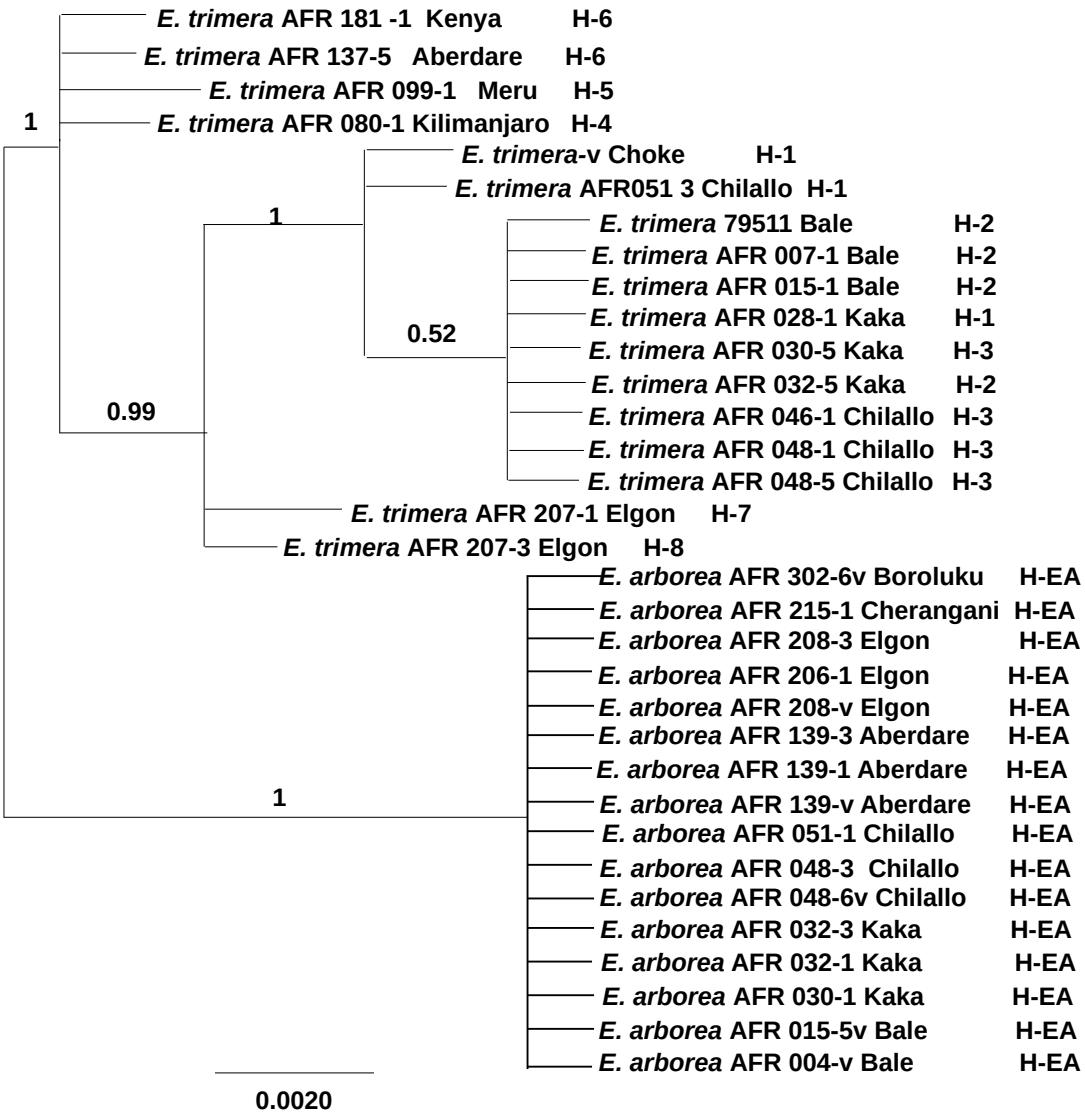


Fig. 11Midpoint rooted Bayesian tree inferred from the concatenated sequences of the two non-coding plastid DNA regions (*rpl*32-*trn*L$^{(UAG)}$ and *trn*G^{UUC}-*trn*G2G) for the widespread *E. arborea* and for the afro-alpine endemic *E. trimera.* Posterior probability values are indicated above the branches. Plastid DNA haplotypes are indicated for each accession.

4.2.1.2 AFLP variation

The average reproducibility of the AFLP markers was 96%. The total dataset contained 169 successfully genotyped individuals, of which 59 belonged to *E. arborea* and 110 to *E. trimera*. Among the 106 markers retained in the final dataset, 58% and 73% were polymorphic in *E. arborea* and *E. trimera*, respectively. In the PCoA of the total dataset, the two species were clearly separated along the first axis, which extracted 49.1% of the variation (Fig. 12). The second axis, which extracted only 6.3% of the variation, mainly depicted intraspecific variation in *E. trimera*. Similarly, the STRUCTURE analysis of the total dataset identified $K = 2$ as the most likely number of clusters, corresponding to the two species (Fig. 13).

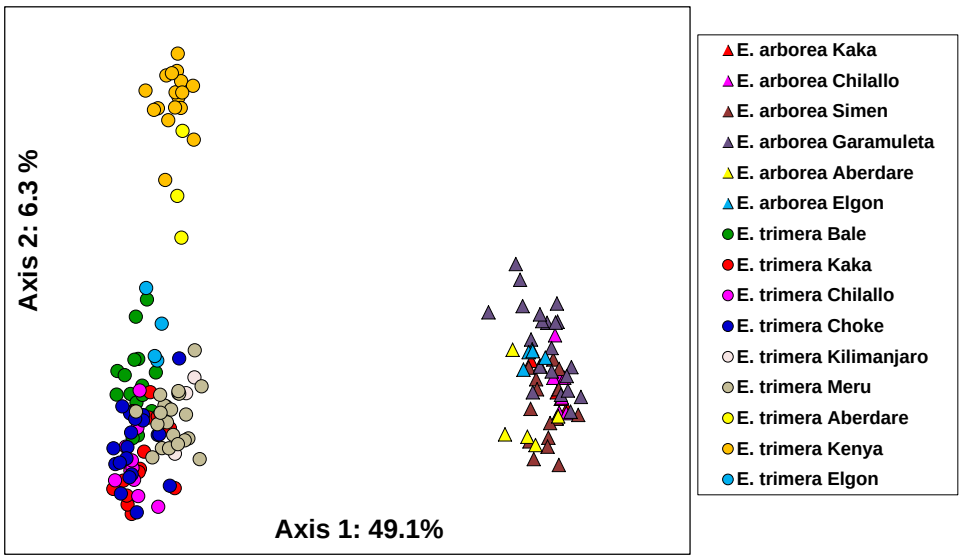


Fig. 12 PCoA based on Dice similarity among AFLP phenotypes of *E. arborea* and *E. trimera*. The percentages of the variance explained are indicated on each axis. Symbols indicate species (*E. arborea*, triangles; *E. trimera*, circles), colours represent mountains

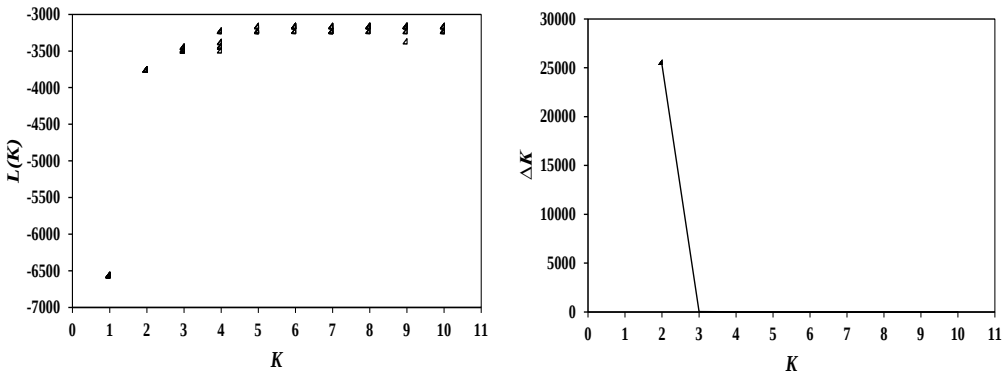


Fig. 13 Genetic structure of the total dataset for *E. arborea* and *E. trimera* inferred from Bayesian clustering of AFLP data. Log probability of the data L (*K*) as function of K ranging from one to 10 and rate of change in the probability between successive runs, Δ *K* as a function of *K*, calculated according to Evanno *et al*. (2005)

In *E. arborea*, both gene diversity and genetic distinctivity showed a clear geographical trend with high values in Ethiopia and low values in East Africa (Fig. 14; Table 9). This was most pronounced for gene diversity. The total gene diversity in this species was 0.1181 (Table 8), whereas gene diversity estimated as average within-population diversity was 0.0804 (*SD* = 0.234). Within-population gene diversity varied from 0.1000 in one Ethiopian population (Simen AFR 065) to 0.0528 in one East African population (Elgon AFR 206). Genetic distinctiveness was highest in one population from Elgon and lowest in one Ethiopian population from Chilallo (Table 8). Within-mountain estimates pooled over all individuals from each mountain ranged from 0.0940 in the Ethiopian mountain Simen to 0.0528 in the East African mountain Elgon for gene diversity, and from 1.36 in Elgon to 0.66 in another Ethiopian mountain, Kaka, for genetic distinctivity (mean 0.98, *SD* = 0.256; Table 9).

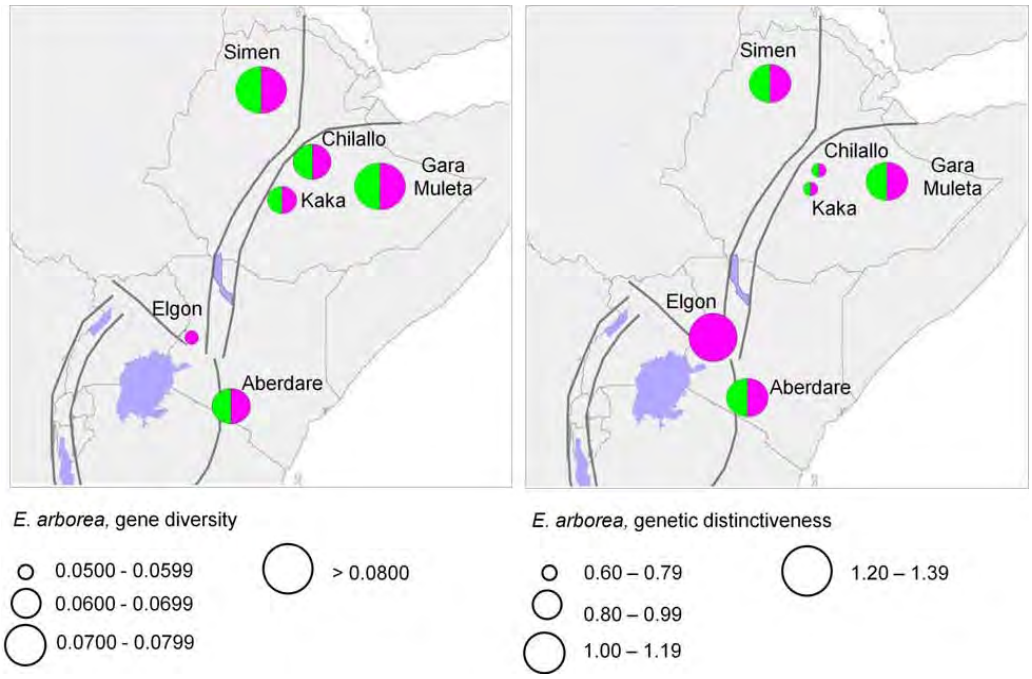


Fig. 14 Genetic structure, diversity and distinctiveness in *E. arborea* based on AFLP variation, with circle colours reflecting genetic (STRUCTURE) groups and circle size reflecting within-mountain gene diversity or genetic distinctiveness (DW). The thick dark grey lines depict the Great Rift Valley

Table 7 Polymorphic sites, number of individuals observed within each haplotype (*n*), haplotype and nucleotide diversity for 17 sequenced individuals of *E. trimera* based on the concatenated matrix of the two non-coding plastid DNA regions with1580 sites (after excluding gaps and missing data). The polymorphic sites at 221, 318, 974, and 1213 were parsimony informative.

		*rpl*32-*trn*L$^{(UAG)}$			*trn*G^{UUC}-*trn*G2G			Haplotype	Haplotype	Nucleotide
Species	Haplotype	178	221	318	974	1213	*n*	frequency	diversity	diversity
E. trimera	H-1	G	T	A	C	A	2	0.118		
E. trimera	H-2	-	-	-	A	-	4	0.235		
E. trimera	H-3	-	-	-	A	C	5	0.294		
E. trimera	H-4	-	A	C	-	C	1	0.059	0.868	0.0013
E. trimera	H-5	-	A	C	A	C	1	0.059		
E. trimera	H-6	-	A	C	-	-	2	0.118		
E. trimera	H-7	T	-	C	-	-	1	0.059		
E. trimera	H-8	-	-	C	-	-	1	0.059		

Table 8 Sampling information and genetic data for the analyzed populations of *Erica arborea* and *E. trimera*. Database ID are given for individuals with plastid DNA sequence which in most cases were taken from the same populations analyzed for AFLPs, but indicated separately below for specification of the individual sequenced. Country: ET = Ethiopia, KN = Kenya, TZ = Tanzania). Only a single plastid DNA haplotype (H-EA) was observed in *E. arborea*; eight plastid DNA haplotypes (H-1 to H-8) were observed in *E. trimera*. AFLP analysis: *n* - number of individuals successfully analyzed, *P* - percentage of polymorphic loci, *D* - gene diversity, DW – the relative frequency-down-weighted marker value as a measure of genetic distinctiveness. Silica-dried leaf material was used except in a few cases where voucher material was used for plastid DNA sequencing (indicated with ‡).

Species	Database ID	Population/-indiv. ID	Country	Mountain	Latitude/ Longitude	Altitude (m)	Plastid DNA haplotype	AFLP analysis			
								n	*P* (%)	*D*	DW
E. arborea		AFR 048	ET	AR, Chilallo	N07.93724/ E039.21493	3500		4	12.26	0.0676	0.50
E. arborea	O-DP-10889	AFR 048-3	ET	AR, Chilallo	N07.93724/ E039.21493	3500	H-EA				
E. arborea	O-DP-10885	AFR 048-6	ET	AR, Chilallo	N07.93724/ E039.21493	3500	H-EA				
E. arborea		AFR 051	ET	AR, Chilallo	N07.91853/ E039.18356	3090		3	13.21	0.0881	0.95
E. arborea	O-DP-10897	AFR 051-1	ET	AR, Chilallo	N07.91853/ E039.18356	3090	H-EA				
E. arborea		AFR 108	ET	HA, Gara Muleta	N09.22841/ E041.79417	2960		11	27.36	0.0885	0.90
E. arborea		AFR 114	ET	HA, Gara Muleta	N09.20987/ E041.79477	2770		5	17.92	0.0792	0.78
E. arborea		AFR 116	ET	HA, Gara Muleta	N09.20955/	2860		5	20.75	0.0906	1.01

					E041.79327						
E. arborea		AFR 032	ET	AR, Kaka	N07.36624/	3690		5	13.21	0.0698	0.66
					E039.18676						
E. arborea	O-DP-10864	AFR 032-1	ET	AR, Kaka	N07.36624/	3690	H-EA				
					E039.18676						
E. arborea	O-DP-10869	AFR 032-2	ET	AR, Kaka	N07.36624/	3690	H-EA				
					E039.18676						
E. arborea	O-DP-10846	AFR 030-1	ET	AR, Kaka	N07.36773/	3760	H-EA				
					E039.18508						
E. arborea		AFR 061	ET	GD, Simen	N13.26066/	3710		4	18.87	0.0975	1.07
					E038.13575						
E. arborea		AFR 065	ET	GD, Simen	N13.23181/	3250		5	20.75	0.1000	1.29
					E038.03970						
E. arborea		AFR 066	ET	GD, Simen	N13.21222/	3190		7	19.81	0.0791	1.12
					E038.00089						
E. arborea		AFR 004-v ‡	ET	BA, Bale	N06.77974/	3380	H-EA				
					E039.75373						
E. arborea	O-DP-10832	AFR 015-5	ET	BA, Bale	N07.04575/	3330	H-EA				
					E039.75152						
E. arborea	O-DP-11223	AFR 302-6	ET	AR, Boroluku	N08.38194/	3400	H-EA				
					E039.5954						
E. arborea		AFR 139	KN	Aberdare	S00.33389/	3670		5	16.04	0.0717	1.13
					E036.65130						
E. arborea	O-DP-11100	AFR 139-1	KN	Aberdare	S00.33389/	3670	H-EA				
					E036.65130						
E. arborea	O-DP-11107	AFR 139-3	KN	Aberdare	S00.33389/	3670	H-EA				

					E036.65130						
E. arborea		AFR 139-v ‡	KN	Aberdare	S00.33389/	3670	H-EA				
					E036.65130						
E. arborea	O-DP-11216	AFR 215-1	KN	Cherengani	N01.21259	2530	H-EA				
					E035.29301						
E. arborea		AFR 206	KN	Elgon	N01.08708/	3380		5	12.26	0.0528	1.36
					E034.63153						
E. arborea	O-DP-11175	AFR 206-1	KN	Elgon	N01.08708/	3380	H-EA				
					E034.63153						
E. arborea	O-DP-11195	AFR 208-1	KN	Elgon	N01.09678/	3620	H-EA				
					E034.62244						
E. arborea		AFR 208-v ‡	KN	Elgon	N01.09678/	3620	H-EA				
					E034.62244						
E. arborea total								59	54.72	0.1181	
E. trimera		AFR 004	ET	BA, Bale	N06.77974/	3380		4	11.32	0.0582	0.42
					E039.75373						
E. trimera		AFR 007	ET	BA, Bale	N06.91311/	3630		10	19.81	0.0774	0.61
					E039.92304						
E. trimera	O-DP-10821	AFR 007-1	ET	BA, Bale	N06.91311/	3630	H-2				
					E039.92304						
E. trimera		AFR 015	ET	BA, Bale	N07.04575/	3330		2	8.49	0.0849	0.44
					E039.75152						
E. trimera	O-DP-10831	AFR 015-1	ET	BA, Bale	N07.04575/	3330	H-2				

					E039.75152						
E. trimera		Bale79511-v ‡	ET	BA, Bale	N06.77217/	3420	H-2				
					E039.75533						
E. trimera		AFR 046	ET	AR, Chilallo	N07.93842/	3710		8	20.75	0.0714	0.57
					E039.22580						
E. trimera	O-DP-10879	AFR 046-1	ET	AR, Chilallo	N07.93842/	3710	H-3				
					E039.22580						
E. trimera		AFR 048	ET	AR, Chilallo	N07.93724/	3500		2	3.77	0.0377	0.48
					E039.21493						
E. trimera	O-DP-10882	AFR 048-1	ET	AR, Chilallo	N07.93724/	3500	H-3				
					E039.21493						
E. trimera	O-DP-10883	AFR 048-5	ET	AR, Chilallo	N07.93724/	3500	H-3				
					E039.21493						
E. trimera	O-DP-10903	AFR 051-3	ET	AR, Chilallo	N07.91853/	3090	H-1				
					E039.18356						
E. trimera		AFR 068	ET	Choke	N10.63408/	3770		10	24.53	0.0853	0.50
					E037.84431						
E. trimera		AFR 068-v ‡	ET	Choke	N10.63408/	3770	H-1				
					E037.84431						
E. trimera		AFR 069	ET	Choke	N10.63755/	3710		5	13.21	0.0604	0.49
					E037.82961						
E. trimera		AFR 074	ET	Choke	N10.68319/	3790		5	16.98	0.0849	0.95
					E037.83739						
E. trimera		AFR 028	ET	AR, Kaka	N07.36887/	3820		5	18.87	0.0925	0.73
					E039.18170						
E. trimera	O-DP-10837	AFR 028-1	ET	AR, Kaka	N07.36887/	3820	H-3				

					E039.18170						
E. trimera		AFR 030	ET	AR, Kaka	N07.36773/	3760		2	7.55	0.0755	0.47
					E039.18508						
E. trimera	O-DP-10855	AFR 030-5	ET	AR, Kaka	N07.36773/	3760	H-3				
					E039.18508						
E. trimera		AFR 032	ET	AR, Kaka	N07.36624/	3700		7	20.75	0.0800	0.46
					E039.18676						
E. trimera	O-DP-10859	AFR 032-5	ET	AR, Kaka	N07.36624/	3700	H-2				
					E039.18676						
E. trimera		AFR 137	KN	Aberdare	S00.33997/	3560		3	16.04	0.1069	1.31
					E036.65958						
E. trimera	O-DP-11090	AFR 137-5	KN	Aberdare	S00.33997/	3560	H-6				
					E036.65958						
E. trimera		AFR 207	KN	Elgon	N01.10912/	3790		4	12.26	0.0676	0.83
					E034.60957						
E. trimera	O-DP-11182	AFR 207-1	KN	Elgon	N01.10912/	3790	H-7				
					E034.60957						
E. trimera	O-DP-11188	AFR 207-3	KN	Elgon	N01.10912/	3790	H-8				
					E034.60957						
E. trimera		AFR 166	KN	Kenya	S00.16791/	3930		5	16.98	0.0792	0.84
					E037.25395						
E. trimera		AFR 170	KN	Kenya	S00.16765/	3850		10	24.53	0.0830	0.86
					E037.24820						
E. trimera		AFR 181	KN	Kenya	S00.05820/	3530		2	8.49	0.0849	0.61
					E037.29133						
E. trimera	O-DP-11160	AFR 181-1	KN	Kenya	S00.05820/	3530	H-6				

					E037.29133						
E. trimera		AFR 080	TZ	Kilimanjaro	S03.13998/	3710		5	15.09	0.0717	0.78
					E037.44009						
E. trimera	O-DP-11008	AFR 080-1	TZ	Kilimanjaro	S03.13998/	3710	H-4				
					E037.44009						
E. trimera		AFR 094	TZ	Meru	S03.21762/	3610		5	16.04	0.0717	0.64
					E036.77106						
E. trimera		AFR 098	TZ	Meru	S03.21697/	3610		11	19.81	0.0731	0.58
					E036.76657						
E. trimera		AFR 099	TZ	Meru	S03.21817/	3630		5	15.09	0.0811	0.81
					E036.76739						
E. trimera	O-DP-11038	AFR 099-1	TZ	Meru	S03.21817/	3630	H-5				
					E036.76739						
E. trimera total								110	68.87	0.1282	

Table 9 Average gene diversity and genetic distinctiveness in different mountains and genetic (STRUCTURE) groups based on the AFLP data for *E. arborea* and *E. trimera*. *n* - number of individuals successfully analyzed, *D*- gene diversity, DW - frequency-down-weighted marker value as a measure of genetic distinctiveness. Numbers in parentheses are standard deviations

Species	Mountain/ Genetic group	*n*	*D*	DW
E. arborea	Chilallo	7	0.0710	0.69
E. arborea	Gara Muleta	21	0.0928	0.90
E. arborea	Kaka	5	0.0698	0.66
E. arborea	Simen	16	0.0940	1.16
E. arborea	Aberdare	5	0.0717	1.13
E. arborea	Elgon	5	0.0528	1.36
				0.98 (0.256)
E. trimera	Bale	16	0.0831	0.54
E. trimera	Chilallo	10	0.0713	0.56
E. trimera	Choke	20	0.0856	0.61
E. trimera	Kaka	14	0.0832	0.56
E. trimera	Aberdare	3	0.1069	1.31
E. trimera	Elgon	4	0.0676	0.83
E. trimera	Kenya	17	0.0832	0.82
E. trimera	Kilimanjaro	5	0.0717	0.78
E. trimera	Meru	21	0.0739	0.65
				0.74 (0.228)
E. trimera	'Ethiopia' group	60	0.0944	0.57
E. trimera	'Aberdare-Kenya' group	20	0.0900	0.89
E. trimera	'Elgon-Kilimanjaro-Meru' group	30	0.0834	0.70

In the STRUCTURE analyses of the *E. arborea* dataset, the largest increase in the log posterior probability occurred at $K = 2$ (Fig. 16). The rate of change in the probability between successive *K*s, ΔK, identified $K = 2$ as the most likely number of clusters (Fig. 16). The PCoA of *E. arborea* (Fig. 15) showed limited geographic structuring, with plants from individual mountains (e.g. Mt Gara Muleta, Mt Kaka and Mt Chilallo) spanning much of the variation along axis 1 (17.7%). The two STRUCTURE groups were not distinctly separated on the PCoA plot, and plants from the same mountain were often divided between the

two groups. The non-hierarchical AMOVAs assigned 33.46% of the overall genetic variation to variation among the 11 populations (Table 10). The hierarchical AMOVAs assigned 31.32% of the variation to among-mountain variation, 5.22% to among-populations-within-mountains variation and 63.46% to within-population variation (Table 10). In another analysis, 23.67% of the variation was found among the two genetic groups (Table 10).

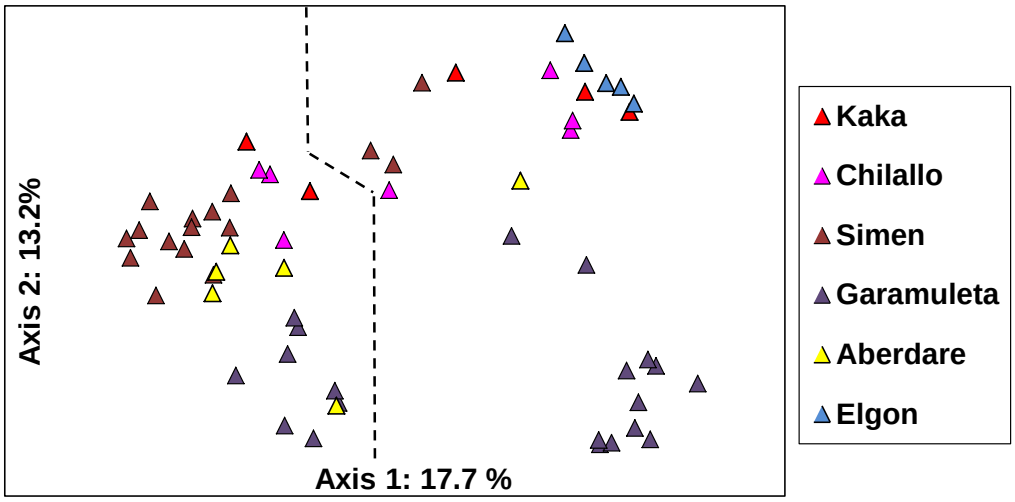


Fig. 15 PCoA based on Dice similarity among AFLP phenotypes of *E. arborea*. The percentages of the variance explained are indicated on each axis. Colours of the symbols indicate mountains. The broken line separates the two genetic groups inferred by STRUCTURE.

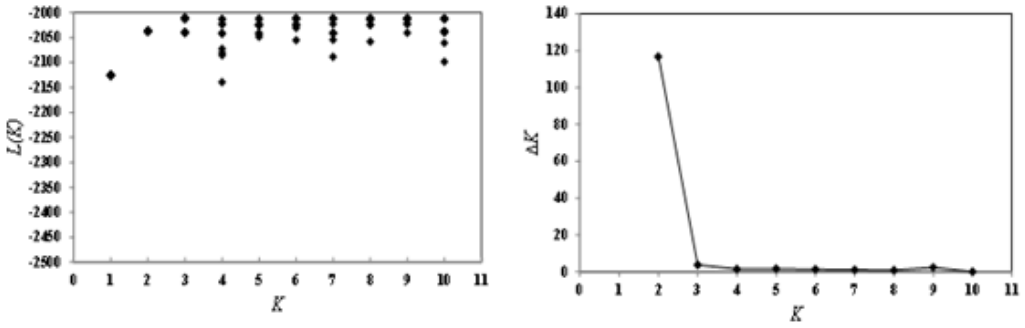


Fig. 16 Genetic structure of *E. arborea* inferred from Bayesian clustering of AFLP data. Log probability of the data L (*K*) as function of *K* ranging from one to 10 and rate of change in the probability between successive runs, Δ *K* as a function of *K*, calculated according to Evanno *et al*. (2005)

Table 10 Analysis of molecular variance (AMOVA) for *E. arborea* and *E. trimera* based on the AFLP data. *P*-values were estimated in a permutation test (10 000 permutations).

Source of Variation	d.f.	Sum of squares	Variance components	Percentage of variation	P-value
E. arborea					
Among populations	10	157.08	2.16	33.46	*P < 0.0001*
Within populations	48	206.04	4.30	66.54	
Among mountains	5	126.22	2.12	31.32	*P < 0.0001*
Among populations	5	30.87	0.35	5.22	*P = 0.2248*
Within populations	48	206.04	4.30	63.46	
Among two genetic groups	1	54.77	1.68	23.67	*P < 0.0001*
Among mountains	57	308.36	5.41	76.33	
E. trimera					
Among populations	19	371.83	2.85	41.03	*P < 0.0001*
Within populations	90	368.99	4.10	58.97	
Among mountains	8	311.09	2.81	39.03	*P < 0.0001*
Among populations	11	60.74	0.28	3.96	*P < 0.0001*
Within populations	90	368.99	4.10	57.01	
Among three genetic groups	2	226.99	3.16	39.02	*P < 0.0001*
Among populations	17	144.84	0.85	10.4	*P < 0.0001*
Within populations	90	368.99	4.10	50.58	
'Ethiopia'					
Among mountains	3	51.48	0.67	12.87	*P = 0.0078*
Among populations	7	43.89	0.47	9.03	*P < 0.0001*
Within populations	49	199.68	4.08	78.11	
'Elgon-Kilimanjaro-Meru'					
Among mountains	2	23.90	1.20	23.71	*P = 0.1036*
Among populations	2	7.16	-0.05	-0.94	*P = 0.6491*
Within populations	25	97.08	3.88	77.23	
'Aberdare-Kenya'					
Among mountains	1	8.73	0.78	14.47	*P = 0.2443*
Among populations	2	9.69	0.07	1.31	*P = 0.2727*
Within populations	16	72.23	4.52	84.22	

In *E. trimera*, on the other hand, we observed no distinct geographical trend in gene diversity or distinctivity (Fig. 17 & Table 9). The Ethiopian and East Africa populations of this species tended to be quite similar, particularly for

gene diversity, except that a single population from Mt Aberdare with three individuals showed relatively high level of diversity and distinctivity (Fig. 17 & Table 9). The total gene diversity estimate in *E. trimera* was 0.1282 (Table 8), and gene diversity estimated as average within-population diversity was 0.0764 (*SD* = 0.142). Within-population gene diversity varied from 0.1069 in the Aberdare population (Aberdare AFR 137) to 0.0377 in one Ethiopian population (Chilallo AFR 048). Genetic distinctiveness was again highest in the Aberdare population and lowest in one Ethiopian population from Bale (Table 8). Within-mountain estimates pooled over all individuals from each mountain ranged from 0.1069 again in the Aberdare populaiton to 0.0676 in another East African mountain population, Mt Elgon, for gene diversity and from 1.31 in the Aberdare to 0.54 in the Ethiopian Bale mountain for genetic distinctivity (mean 0.74, *SD* = 0.228; Table 9).

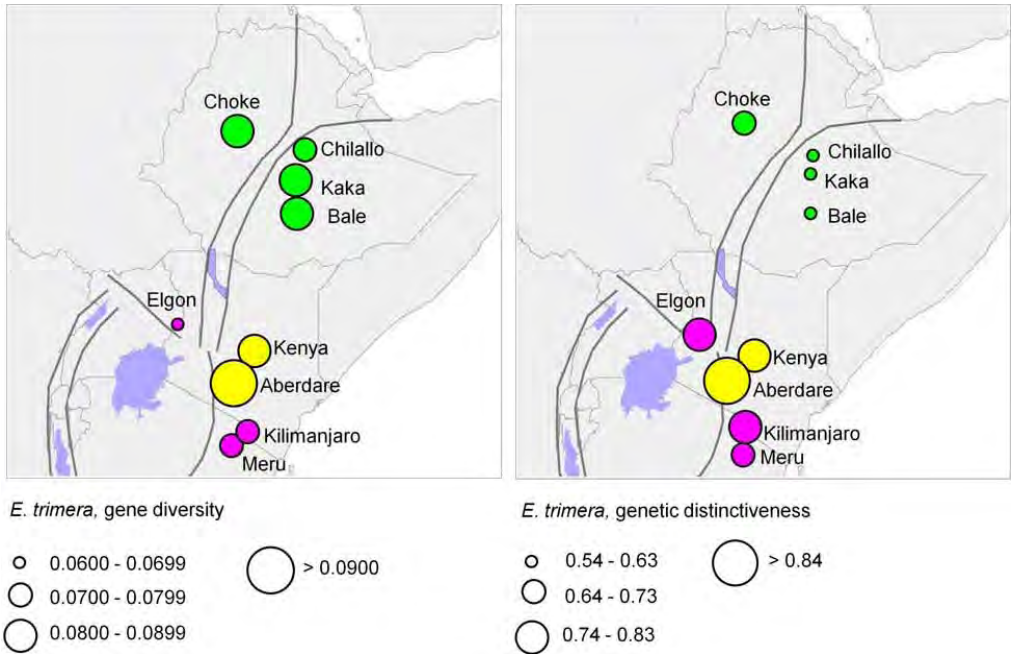


Fig. 17 Genetic structure, diversity and distinctiveness in *E. trimera* based on AFLP variation. Circle colours reflecting genetic (STRUCTURE) groups and circle size reflecting within-mountain gene diversity or genetic distinctiveness (DW). The thick dark grey lines depict the Great Rift Valley

In the STRUCTURE analysis of the *E. trimera* dataset, the largest increase in the log posterior probability of the data occurred at $K = 3$ (Fig. 18). The rate of change in the probability between successive *K*s, ΔK, identified $K = 3$ as the most likely number of clusters (Fig. 18). In this species, the genetic groups inferred from the STRUCTURE analysis turned out to be quite distinct in the PCoA, in which the first and second axis explained 26.2% and 10.4% of the variance, respectively, and the groups were geographically distinct with no mixture among mountains (Fig. 19). The 'Ethiopia' group included all plants from the Ethiopian mountains, the 'Kenya-Aberdare' group included all plants from the East African mountains Kenya and Aberdare, and the 'Elgon-Kilimanjaro-Meru' group included all plants from the remaining East African mountains, Elgon, Kilimanjaro and Meru. The non-hierarchical AMOVA assigned 41.03% of the overall genetic variation to variation among the 20 populations (Table 10). The hierarchical AMOVAs assigned 39.03% of the variation to among-mountain variation and 3.96% to among-populations-within-mountains variation. 39.02% of the total variation was assigned to variation among the three genetic groups (Table 10). The AMOVAs for each genetic group attributed the highest proportions (78-84%) to variation within populations and low levels of differentiation among mountains (13-24%; Table 10).

Within-genetic group gene diversity estimates in *E. trimera* ranged from 0.0944 in the 'Ethiopia' group to 0.0834 in the 'Elgon-Kilimanjaro-Meru' group, with 0.0900 in the 'Aberdare-Kenya' group (Table 9). Genetic distinctiveness was highest in the 'Aberdare-Kenya' group and lowest in the 'Ethiopia' group

(Table 3). The three genetic (AFLP) groups in *E. trimera* were characterized by different plastid DNA haplotypes, and they were consistent with the plastid DNA haplotype network in separating the Ethiopian populations from the East African ones.

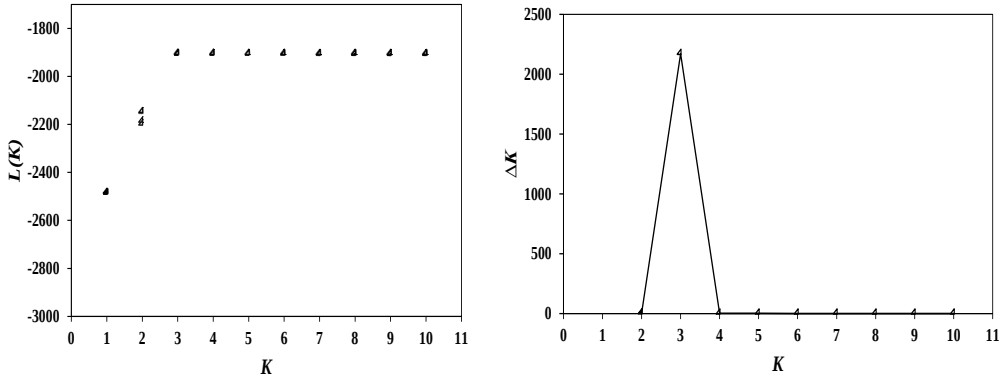


Fig. 18 Genetic structure of *E. trimera* inferred from Bayesian clustering of AFLP data. Log probability of the data L*(K)* as function of *K* ranging from one to 10 and rate of change in the probability between successive runs, Δ *K* as a function of *K*, calculated according to Evanno *et al*. (2005)

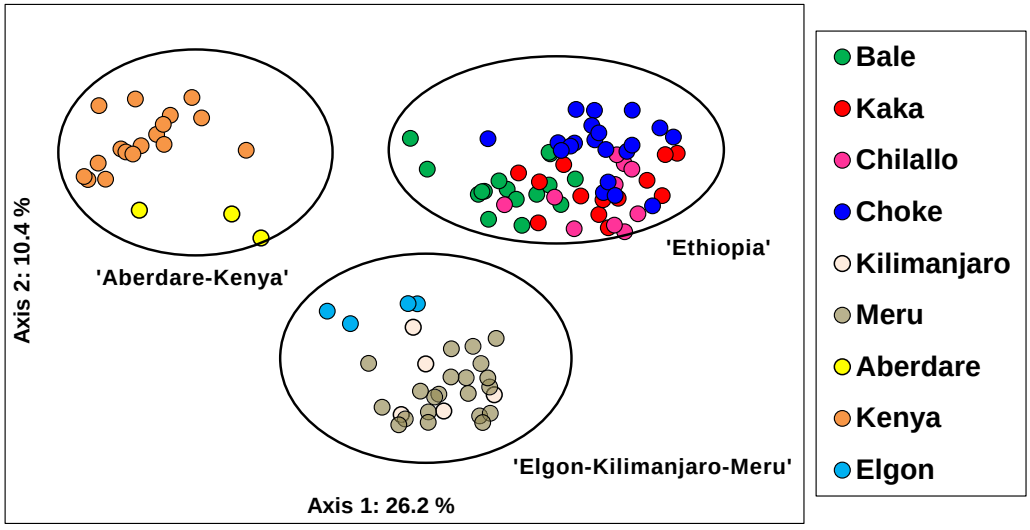


Fig. 19 PCoA based on Dice similarity among AFLP phenotypes of *E. trimera*. The percentages of the variance explained are indicated on each axis. Colours of the symbols indicate mountains. The three genetic groups inferred by STRUCTURE are encircled.

4.2.2 Carex monostachya and *C. runssoroensis* (Cyperaceae)

A total of 120 plants, of which 88 and 32 belonged to *C. monostachya* and to *C. runssoroensis*, respectively, were successfully genotyped. AFLP analysis of the total dataset provided a total of 262 markers, of which 95.2% were polymorphic. The mean scoring error of the total dataset calculated from nine duplicated individuals was 4.5%. In the PCoA of the total dataset, the first three axes altogether explained 41% of the total variation. Although 17% of the total variation along the first axis mainly separated the two species, plants of the two species from Elgon were closer to each other and situated in an intermediate position between the two species (Fig. 20).

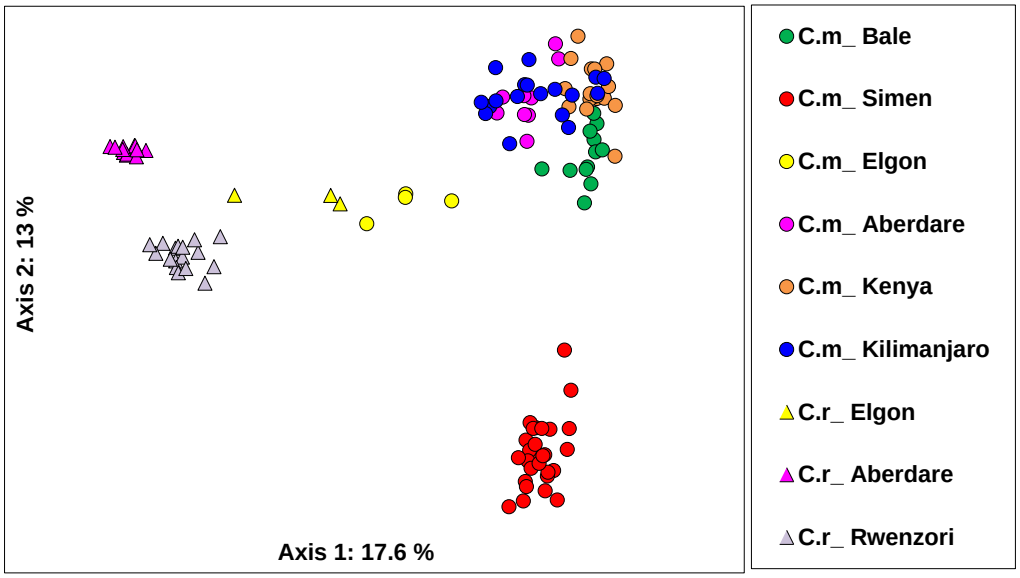


Fig. 20 PCoA based on Dice similarity among AFLP phenotypes of *C. monostachya* and *C. runssoroensis*, abbreviated in the legend as *C.m* and *C.r*, respectively. The percentages of the variance explained are indicated on each axis. Symbols indicate species (*C. monostachya*, circles; *C. runssoroensis*, triangles), colours represent mountains.

The STRUCTURE analysis of the total dataset revealed the highest likelihood of the data at $K = 3$ suggesting the presence of three main groups (Fig. 21). The *C. monostachya* collections were divided into two groups which were clearly

separated along the second axis of the PCoA (Fig. 20); while all plants from Simen formed a distinct genetic group, the remaining *C. monostachya* plants except those from Mt. Elgon formed the second group. The third group comprised all *C. runssoroensis* collections and four plants of *C. monostachya* collected from Mt. Elgon. On the NJ tree, only two of the three main STRUCTURE groups had high bootstrap support. A close relationship between the Elgon populations of *C. monostachya* and the remaining *C. runssoroensis* plants was supported with 92% bootstrap value, and all plants from Simen formed a supported group (95%; Fig. 22). No strong genetic relationships were detected among the remaining *C. monostachya* populations of the STRUCTURE group. Thus, all analyses (STRUCTURE, PCoA, and NJ) indicated that plants from Mt. Elgon displayed a different pattern of relationship to that of either species, indicating possible hybridisation events.

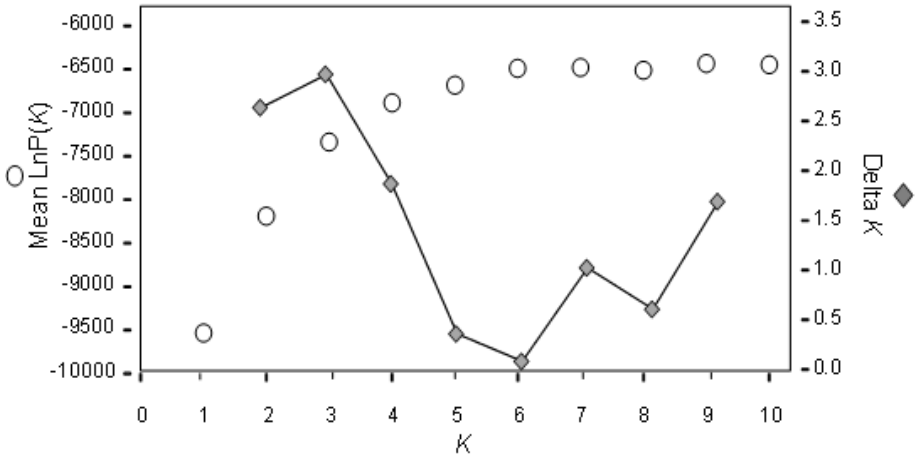


Fig. 21 Genetic structure of the total dataset for *C. monostachya* and *C. runssoroensis* inferred from Bayesian clustering of AFLP data. Mean value for log probability of the data LnP(*K*) and rate of change in the log probability between successive runs, Delta *K* (calculated according to Evanno *et al*. 2005) are shown as a function of *K*, ranging from 1 to 10.

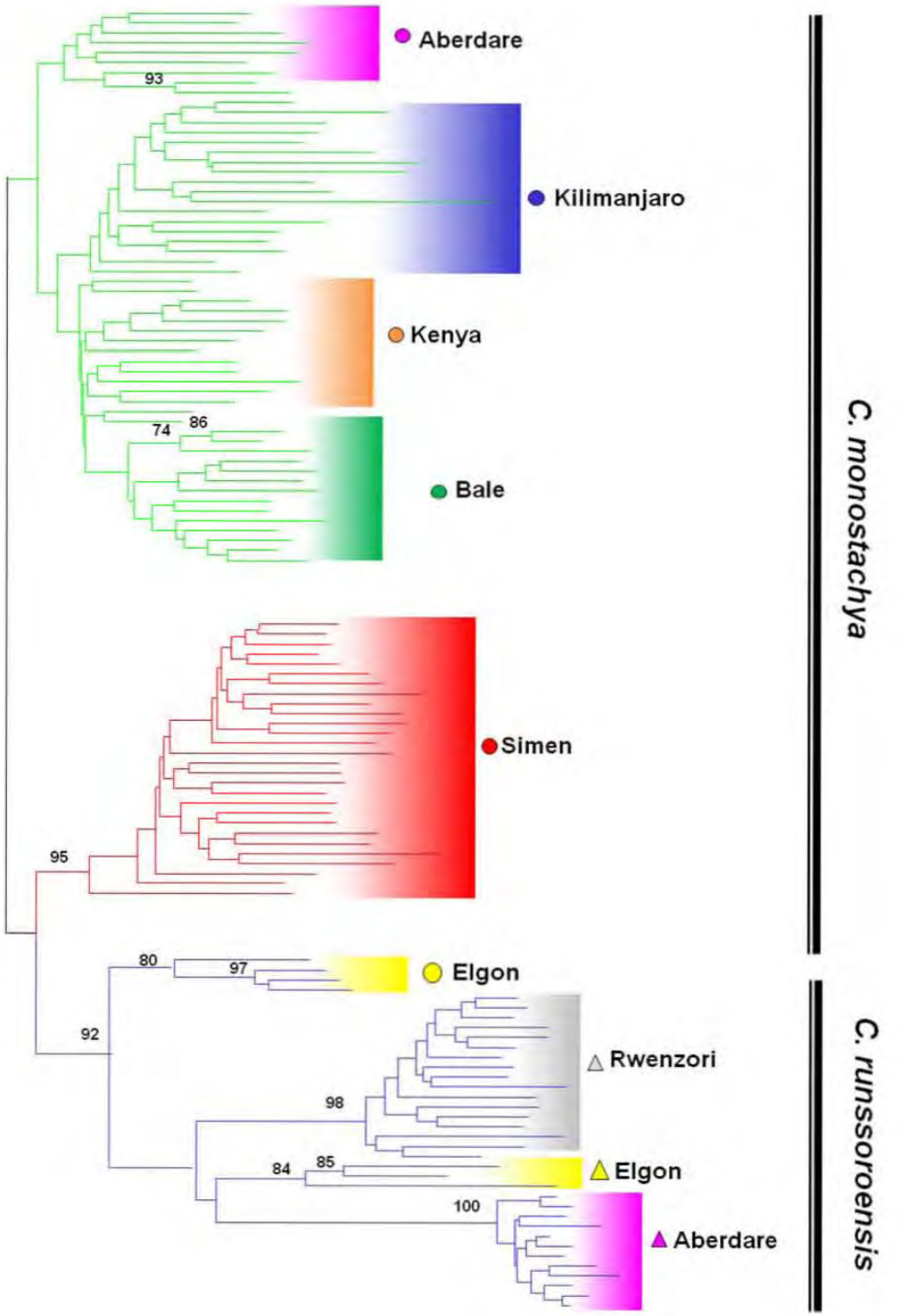
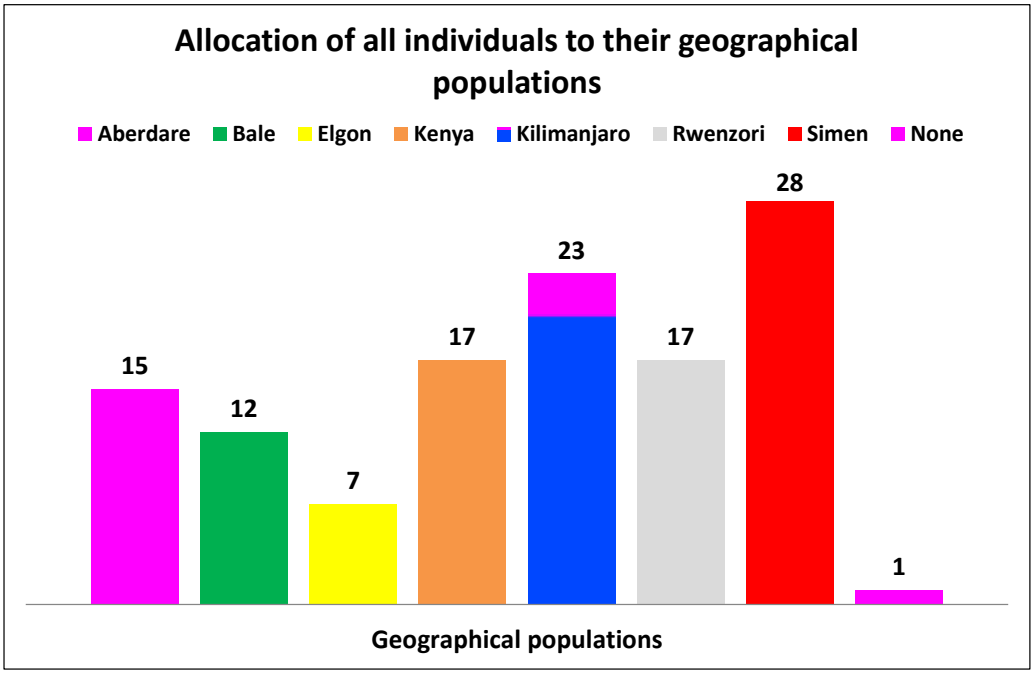


Fig. 22 Neighbour-joining tree of the AFLP phenotypes for 120 individual plants of *C. monostachya* and *C. runssoroensis* based on Nei and Li (1979) genetic distance. Symbols and colours identify species and mountains, respectively, as shown in the PCoA (Fig. 20). Branch colours indicate the three major genetic groups identified in the STRUCTURE analysis. Bootstrap values above 50% (estimated from 1000 replicates) are shown above the branches.

The assignment of all the individuals to their geographical populations resulted in 95% of the total individuals correctly reallocated to their original geographical populations. A few plants of *C. monostachya* from Mt. Aberdare were either allocated to Kilimanjaro (five individuals) or not assigned at all (one individual; Fig. 23a). When all individuals were assigned to the two source species, all of the 88 collections of '*C. monostachya'* and a single individual of '*C. runssoroensis'* from Elgon (KN0391_3) were allocated to the '*C. monostachya*' source species (Fig. 23b). In the remaining 31 collections of *C. runssoroensis* all individuals were reassigned to their respective source species, except a single individual from Elgon (KN0391_4) which was not assigned to either source species. All the assignments were statistically significant except for six of the seven collections from Elgon and two individuals (TZ0046_4, and TZ0303_2) of '*C. monostachya*' from Kilimanjaro.

(a)



(b)

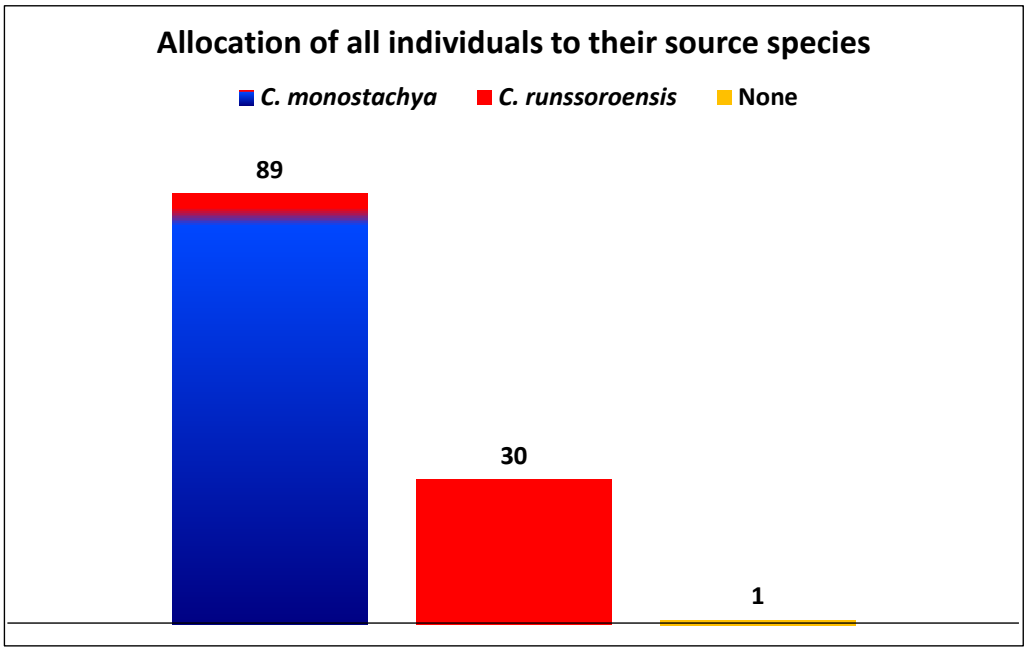


Fig. 23 Assignment of all individuals of *C. monostachya* and *C. runssoroensis* (a) to their geographical populations and (b) to source species populations

When allocated to the source species and F1 hybrid using AFLPOP, 96.6% (85 of 88) of the '*C. monostachya*' and 90.63% (29 of 32) of the '*C. runssoroensis*'

collections were reallocated to their respective source species (Fig. 24). All five detected F1 hybrids were allocated to collections from Mt. Elgon. A single individual (KN0391_2) of '*C. runssoroensis* from Elgon was not assigned to any of the three categories (Fig. 24). All the assignments were significant except the allocation of two individuals from Kilimanjaro (TZ0046_4, and TZ0303_2) and a single collection from Elgon (KN0391_4) to the '*C. monostachya*' and '*C. runssoroensis'* source species, respectively.

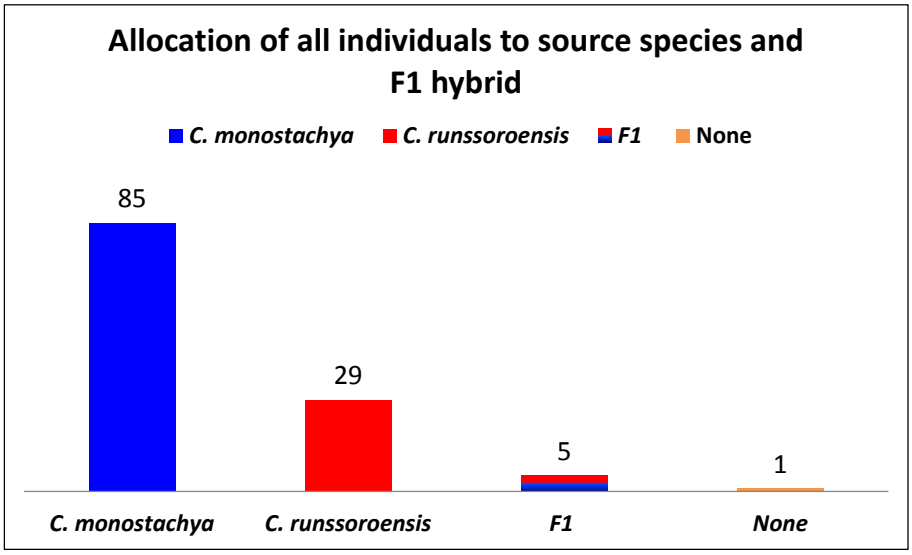
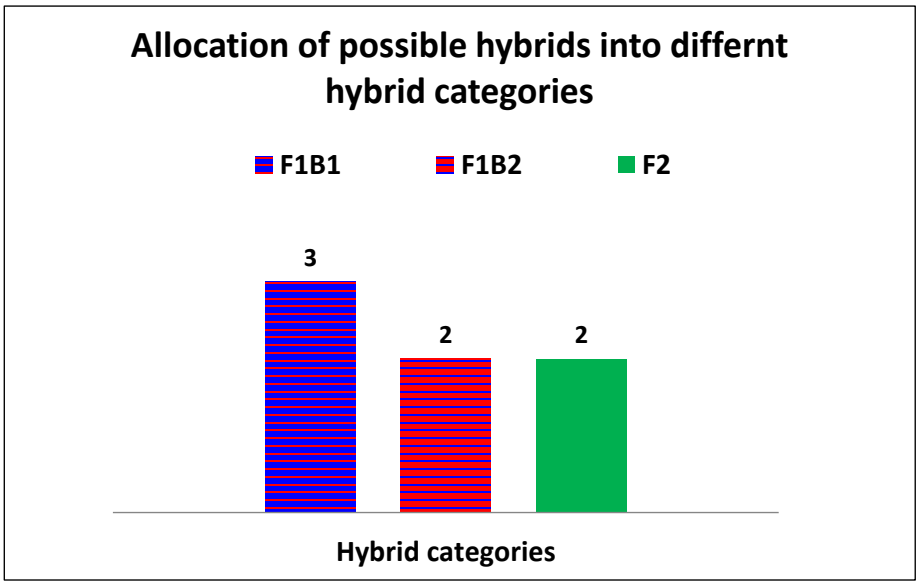


Fig. 24 Assignments of all individuals of *C. monostachya* and *C. runssoroensis* to source species population and F1 hybrid

An extended assignment of the detected hybrids (all seven plants from Mt. Elgon) into the six categories (two source species, F1, F1-B1, F1-B2, and F2), resulted in significant allocation into three hybrid categories (Fig. 25a & b). Three individuals of the '*C. monostachya*' collections were assigned to the F1B1 hybrid category while two of the '*C. runssoroensis*' collections were allocated to the F1B2 backcross hybrid class. The remaining two individuals of each species were assigned to the F2 hybrid category (Fig. 25a & b).

(a)



(b)

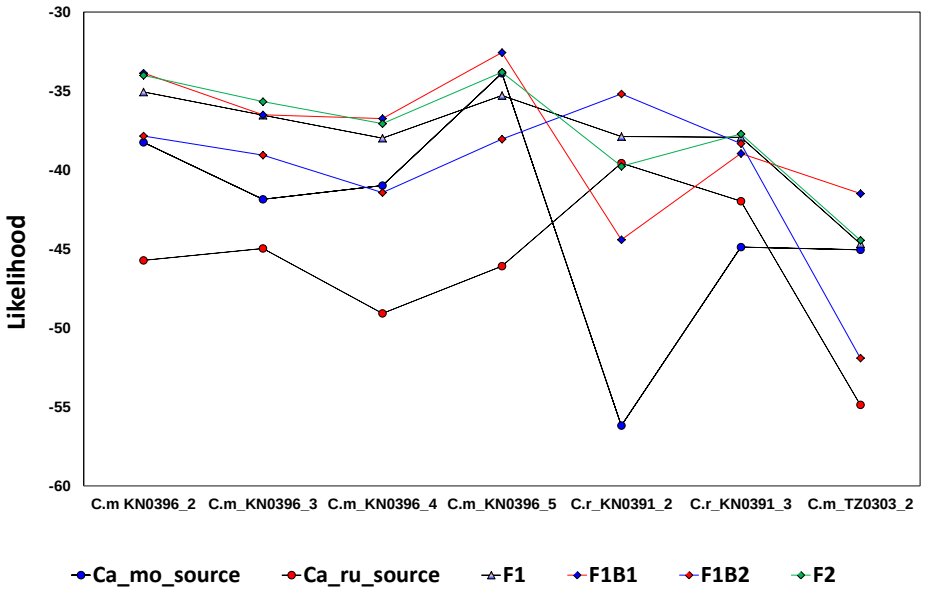


Fig. 25 (a) Assignments of possible hybrids into different hybrid categories (b) Likelihood of the assignment plotted against the seven hybrids.(KN & TZ in the hybrid names identify samples from Elgon and Kilimanjaro, respectively; C.m & C.r indicate *C. monostachya* and C. *runssoroensis*, respectively; as shown in Fig. 12). Ca_mo_source and Ca_ru_source indicate *C. monostachya* and *C. runssoroensis* source species, respectively. See text for definition of the different hybrid categories.

C. monostachya

In *C. monostachya*, a total of 88 individuals from 20 populations were successfully genotyped. The AFLP analysis provided a total of 233 markers of which 211 (90.6%) were polymorphic. The mean scoring error calculated from nine duplicated individuals was 4.5%. After removing the four hybrid individuals detected from Elgon, the final dataset of *C. monostachya* comprised 84 individuals and 19 populations. The first three main axes of the PCoA explained 35.7% of the total variation. Along the first axis, which explained 22 % of the total variation, the plants were divided into two distinct groups, distributed on opposite side of the Rift Valley (Fig. 26). Populations from Simen, on the Western side of the Valley (Western-rift lineage), were genetically more distinct and clearly separated from plants from Mt. Kilimanjaro, Mt. Aberdare, Mt. Kenya, and Bale on the Eastern side of the Valley (Eastern-rift lineage). Within the Eastern-rift lineage, plants from geographically distant mountains appeared to be genetically more closely related but formed subgroups involving plants from Kilimanjaro and Aberdare in one subgroup and those from Kenya and Bale in other additional subgroup (Figs. 26 & 27). The NJ tree identified a close relationship among all plants from the Western-rift lineage (95%, Fig. 22), whereas no strong support was obtained for the Eastern-rift lineage (Fig. 22).

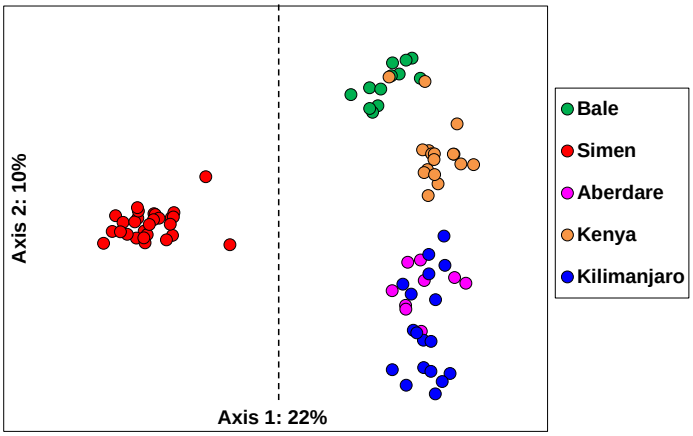


Fig. 26 PCoA based on Dice similarity among AFLP phenotypes of *C. monostachya*. The percentages of the variance explained are indicated on each axis. Colours represent mountains. The broken line separates the two genetic groups inferred by STRUCTURE.

The substructuring pattern was further detected in a separate PCoA for the 'Eastern-rift' lineage in which the first three axes explained 29.2% of the total variation. The first axis explained 18% of the total variation and mainly separated Mt. Kilimanjaro and Mt. Aberdare from Bale and Mt. Kenya (Fig. 27). Interestingly, plants from adjacent mountains, such as Mt. Kenya and Aberdare, were genetically more similar to plants from distant mountains, Bale and Mt. Kilimanjaro, than they were to each other.

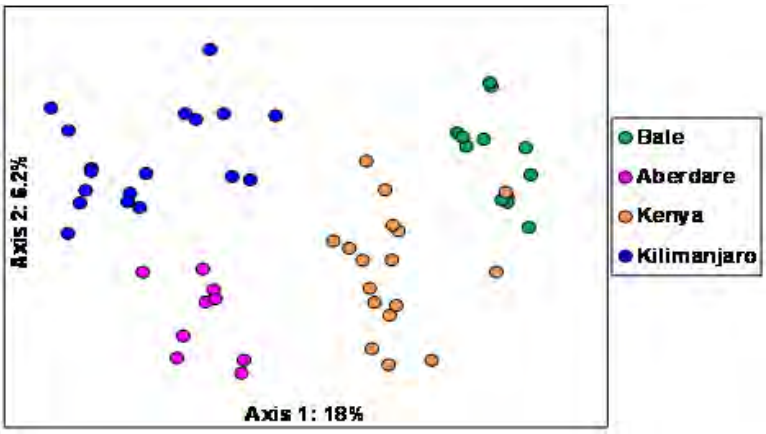


Fig. 27 PCoA based on Dice similarity among AFLP phenotypes for the 'Eastern-rift' lineage of *C. monostachya*. The percentages of the variance explained are indicated on each axis. Colours represent mountains

The STRUCTURE analysis, excluding the hybrid individuals from Mt. Elgon, identified the largest increase in the log probability of the data at $K = 2$ (Fig. 28). The rate of change in the probability of the data between the successive *K*s, ΔK, also identified $K = 2$ (Fig. 28) as the most likely number of groups. Furthermore, at $K = 2$, all 10 replicates provided a stable assignment of individuals and suggested the split of the data into two major lineages corresponding to the same Eastern-rift and Western-rift lineages as detected along the first axis of the PCoA in Fig. 26. Further analysis within the Eastern-rift lineage identified $K = 2$ as the most likely number of clusters corresponding to the subgrouping structure detected in the PCoA of Fig. 27. In the first subgroup, plants from Mt. Kilimanjaro were genetically more similar to Mt. Aberdare and clearly separated from the other subgroup which comprised plants from Mt. Kenya and Bale.

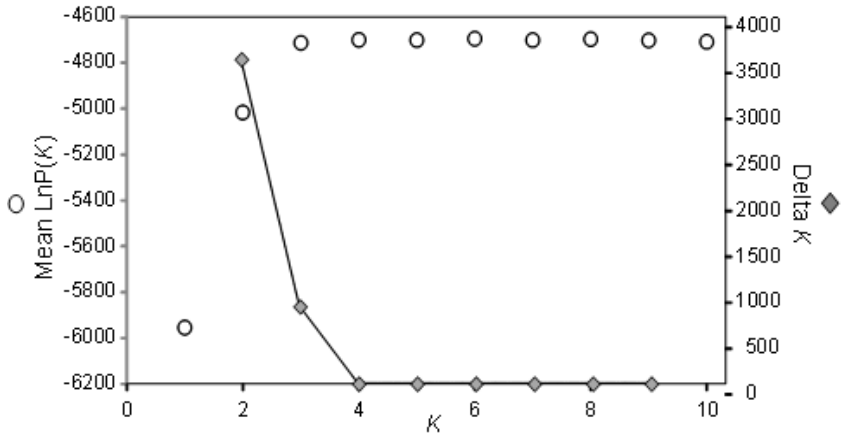


Fig. 28 Genetic structure of *C. monostachya* inferred from Bayesian clustering of AFLP data. Mean value for log probability of the data LnP (*K*) and rate of change in the log probability between successive runs, Delta *K* (calculated according to Evanno *et al*. 2005) are shown as a function of *K*, ranging from 1 to 10.

The total gene diversity estimate averaged over all investigated individuals of *C. monostachya* was $D = 0.1001$. The average within-population gene diversity estimate varied from 0.0572 in the ET0009 Bale population from Ethiopia to

0.1402 in the TZ0046 Kilimanjaro population from Tanzania. The average genetic distinctiveness value calculated over all populations was 2.39 (*SD* = 0.625) and it varied from 1.59 in the TZ0243 population from Mt. Kilimanjaro to 4.34 in KN0396, population from Mt. Kenya. The highest values of regional gene diversity and distinctiveness were observed in populations from the two East African mountains of Kilimanjaro (*D* = 0.1385, DW = 2.67) and Aberdare (*D* = 0.1330, DW = 2.79; Table 12). The lowest value of the two estimates were observed in Bale (*D* = 0.0920, and DW = 2.03). Comparison between the two major lineages revealed a more or less similar level of average diversity and distinctiveness in which the Western-rift lineage (Simen) displayed a relatively lower diversity and higher distinctiveness (*D* = 0.1122, DW = 2.48) compared to the relatively more diverse but less distinct Eastern-rift lineage (*D* = 0.1165, DW = 2.40; Table 12).

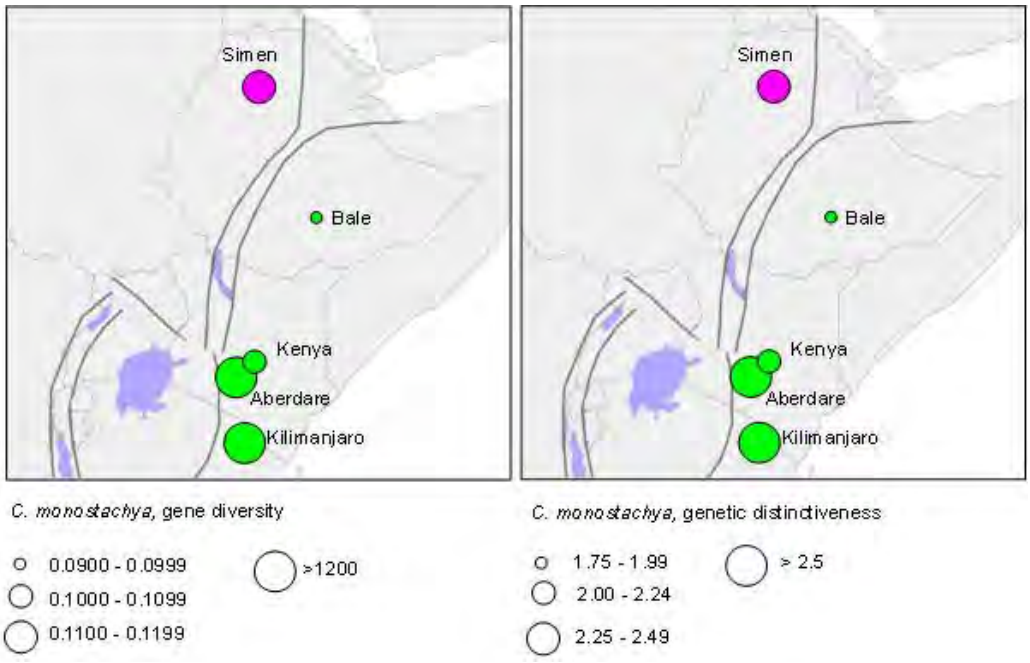


Fig. 29 Genetic structure, diversity and distinctiveness in *C. monostachya* based on AFLP variation. Circle colours reflecting genetic (STRUCTURE) groups and circle size reflecting within-mountain gene diversity or genetic distinctiveness (DW). The thick dark grey lines depict the Great Rift Valley

Table 11 Geographic origin, collection data, and genetic diversity for 20 populations of *C. monostachya*; *n* - number of individuals successfully analyzed *D*- gene diversity, DW - frequency-down-weighted marker value as a measure of genetic distinctiveness.

Pop ID	Country	Mountain	Lat/Long	Altitude	*n*	*D*	DW
ET0009	Ethiopia	Bale	N6.8547/E39.8933	4059	3	0.0572	1.72
ET0782	Ethiopia	Bale	N6.8687/E39.8822	4017	4	0.0858	1.69
ET0849	Ethiopia	Bale	N6.8554/E39.8965	4019	5	0.0687	2.16
ET0083	Ethiopia	Simen	N13.2827/E38.1108	3711	5	0.1013	2.97
ET0344	Ethiopia	Simen	N13.2007/E38.2667	3748	5	0.1047	1.90
ET0474	Ethiopia	Simen	N13.3491/E38.2625	3912	5	0.1124	2.55
ET0510	Ethiopia	Simen	N13.3284/E38.2430	3681	5	0.0936	1.89
ET0563	Ethiopia	Simen	N13.3285/E38.2409	3643	4	0.0944	3.15
ET0628	Ethiopia	Simen	N13.3271/E38.2424	3682	4	0.0594	1.79
KN0568	Kenya	Aberdare	S0.3338/E36.6418	3591	5	0.1236	2.49
KN0734	Kenya	Aberdare	S0.3341/E36.4019	3214	4	0.1237	2.96
KN0396	Kenya	Mt. Elgon	N1.1012/E34.6073	3839	4	0.0787	4.34
KN0846	Kenya	Mt. Kenya	S0.1392/E37.3143	4230	5	0.0987	2.46
KN0902	Kenya	Mt. Kenya	S0.1214/E37.2956	4044	5	0.0833	1.81
KN0997	Kenya	Mt. Kenya	S0.1503/E37.3310	4379	5	0.0987	1.80
KN1078	Kenya	Mt. Kenya	S0.1461/E37.3471	4019	2	0.1159	1.99
TZ0046	Tanzania	Kilimanjaro	S3.0056/E37.2415	3536	3	0.1402	2.37
TZ0243	Tanzania	Kilimanjaro	S3.1010/E37.4211	4109	5	0.0927	1.59
TZ0271	Tanzania	Kilimanjaro	S3.1467/E37.4420	3818	5	0.1082	2.33
TZ0303	Tanzania	Kilimanjaro	S3.135/E37.4337	3817	5	0.1399	3.86

Table 12 Genetic diversity for five regional populations and two structure groups of *C. monostachya* (after excluding Elgon); *n* - number of individuals successfully analyzed *D* - gene diversity, DW - frequency-down-weighted marker value as a measure of genetic distinctiveness.

Structure group	Mountain	*n*	*D*	DW
Western-rift lineage	Simen	28	0.1122	2.48
Eastern-rift lineage	Aberdare	9	0.1330	2.79
	Bale	12	0.0920	2.03
	Kenya	17	0.1023	2.07
	Kilimanjaro	18	0.1385	2.67
Mean		56	0.1165	2.4

The non-hierarchical AMOVA assigned 42% of the total genetic variance to variation among the 19 populations and the remaining to within-population variation (Table 13). The three-level hierarchical AMOVAs assigned 35% and 10% of the total variation to variation among mountains and among populations within mountains, respectively. In a separate hierarchical analysis, a similar trend of diversity partitioning was detected in which 33% and 18.5% of the total variation was assigned between the two major lineages and among populations within genetic groups, respectively. Separate AMOVAs for each main lineage assigned 16.5% and 31.8% of the total variation to among-population variation within the Western-rift and Eastern-rift lineages, respectively. In all AMOVAs, a high percentage of the genetic variability in *C. monostachya* was allocated to within-population variation.

Table 13 Analysis of molecular variance (AMOVA) of the AFLP data for *C. monostachya*. *p*-values were estimated in a permutation test (10000 permutations).

Source of Variation	d.f.	Sum of squares	Variance components	Percentage of components	*P*-value
Among populations	18	876.29	8.39	41.86	*P < 0.0001*
Within populations	65	757.70	11.66	58.14	
Among mountains	4	574.10	7.52	35.06	*P < 0.0001*
Among population	14	302.18	2.27	10.58	*P < 0.0001*
Within populations	65	757.70	11.66	54.36	
Between two STRUCTURE groups	1	339.94	8.23	33.71	*P < 0.0001*
Among populations	17	536.35	4.52	18.53	*P < 0.0001*
Within populations	65	757.70	11.66	47.75	
Western-rift lineage					
Among populations	5	107.18	2.20	16.48	*P < 0.0001*
Within populations	22	245.75	11.17	83.52	
Eastern-rift lineage					
Among populations	12	429.18	5.56	31.85	*P < 0.0001*
Within populations	43	511.95	11.91	68.15	

C. runssoroensis

In *C. runssoroensis*, a total of 32 individuals from seven populations were successfully genotyped. The AFLP analysis provided a total of 156 markers of which 133 (85.3%) were polymorphic. The mean scoring error calculated from nine duplicated individuals was 4.5%. After removing the three hybrid individuals detected from Elgon, the final dataset of *C. runssoroensis* comprised 29 individuals and seven populations from Aberdare and Ruwenzori mountains. A very strong geographic structuring of genetic variability was detected between the two mountains. The first three axes extracted 67.7% of the overall variation. The first axis alone explained 58.7% of the variation and clearly separated the plants into two distinct groups corresponding to the two mountains (Fig. 30). The NJ tree also showed strong support for mountain-specific groups (bs = 98%, 100 %; Fig. 22).

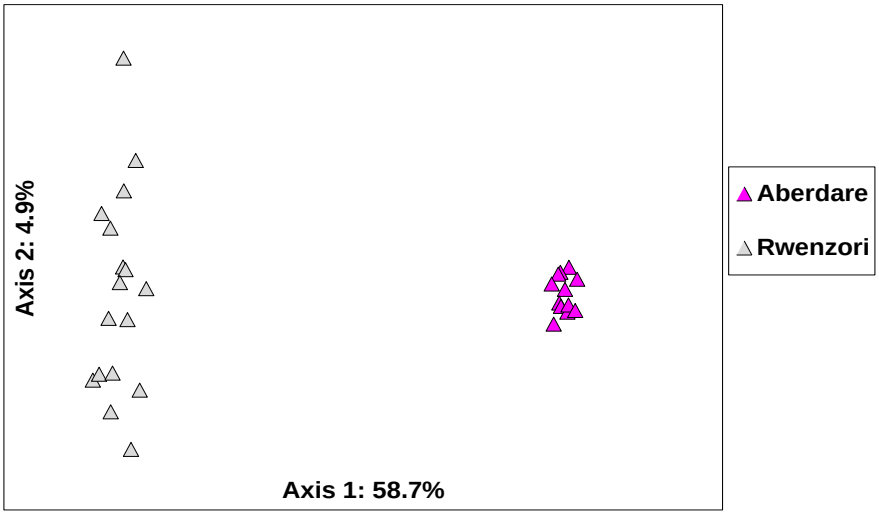


Fig. 30 PCoA based on Dice similarity among AFLP phenotypes of *C. runssoroensis*. The percentages of the variance explained are indicated on each axis. Colours represent mountains.

In the STRUCTURE analysis, the largest increase in the log probability of the data occurred at $K = 2$ (Fig. 31). The rate of change in the probability of the data

between the successive *K*s, Δ *K*, also identified *K* = 2 (Fig. 31) as the most likely number of clusters. Furthermore, at *K* = 2, all 10 replicate thus provided a stable assignment of individuals and suggested a split of the data into two clusters corresponding to the same groups as revealed in the PCoA (Fig. 30).

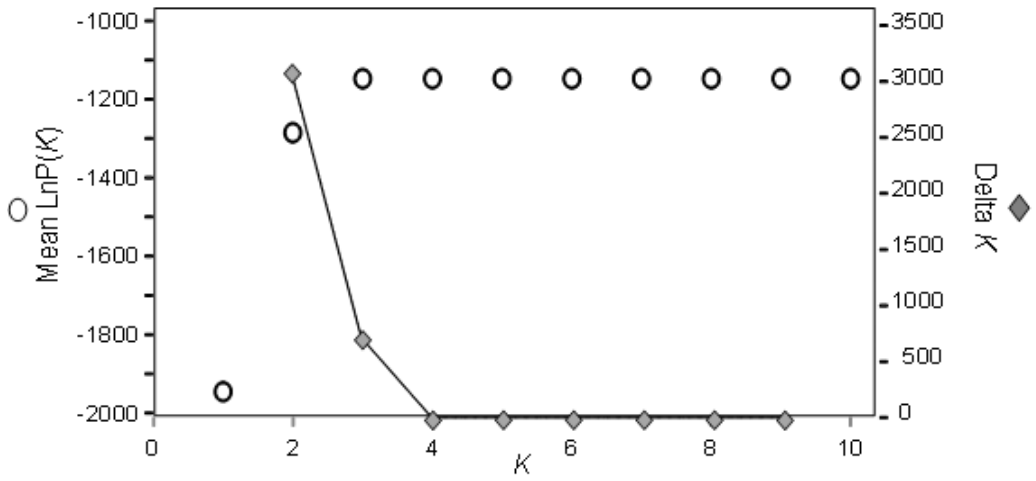


Fig. 31 Genetic structure of *C. runssoroensis* inferred from Bayesian clustering of AFLP data. Mean value for log probability of the data LnP(*K*) and rate of change in the log probability between successive runs, Delta *K* (calculated according to Evanno *et al*. 2005) are shown as a function of *K*, ranging from 1 to 10.

The total gene diversity estimate averaged over all investigated individuals of *C. runssoroensis* was *D* = 0.0820. The average within-population gene diversity varied from 0.0385 in the KN0672 population from Mt. Aberdare to 0.1186 in UG2314 from Ruwenzori (Table 15). The average genetic distinctiveness overall populations was 3.86 (*SD* = 0.543). The highest (DW = 4.68) and the lowest (DW = 3.22) values were observed in the KN0539 Aberdare and UG2282 Ruwenzori populations, respectively. Regional gene diversity was almost two-fold as high in Ruwenzori (*D* = 0.1183) as compared to Aberdare (*D*

= 0.0662). However, the plants in Aberdare possessed higher level of rare markers (DW = 4.11) than those in Ruwenzori (DW = 3.42).

Table 14 Geographic origin, collection data, and genetic diversity for seven populations of *C. runssoroensis*; *n* - number of individuals successfully analyzed, *D* - gene diversity, DW - frequency-down-weighted marker value as a measure of genetic distinctiveness

Pop ID	Country	Mountain	Lat/Long	Altitude	*n*	*D*	DW
KN0539	Kenya	Aberdare	N00.3037/E36.6177	3997	4	0.0716	4.68
KN0671	Kenya	Aberdare	N00.271/E36.6235	3873	3	0.0556	3.88
KN0672	Kenya	Aberdare	N00.3182/E36.6353	3813	5	0.0385	3.79
UG2239	Uganda	Ruwenzori	N00.3850/E29.9273	3425	5	0.0872	3.42
UG2282	Uganda	Ruwenzori	N00.3852/E29.9137	3574	5	0.1000	3.22
UG2314	Uganda	Ruwenzori	N00.4008/E29.9365	3795	4	0.1186	3.59
UG2534	Uganda	Ruwenzori	N00.3823/E29.8884	3932	3	0.1026	3.55

Table 15 Genetic diversity for two regional populations of *C. runssoroensis. n* - Number of individuals successfully analyzed *D* - gene diversity, DW - frequency-down-weighted marker value as a measure of genetic distinctiveness.

Mountain	*n*	*D*	DW
Aberdare	12	0.0662	4.11
Rwenzori	17	0.1183	3.42

The non-hierarchical AMOVA assigned 67.5% of the overall genetic variation to variation among the seven populations (Table 16). The hierarchical AMOVAs assigned 71.8% of the variation to between the two STRUCTURE groups/mountains (P = 0.0303; Table 16), 5.7% to among-populations within genetic group, and 22.4% to within-population variation (Table 16). In separate analyses for each of the two STRUCTURE groups, 18 and 25% of the total variation were assigned to among-population variation within Ruwenzori and Aberdare, respectively.

Table 16 Analysis of molecular variance (AMOVA) of the AFLP data for *C. runssoroensis*. *p*-values were estimated in a permutation test (10000 permutations).

Source of Variation	d.f.	Sum of squares	Variance components	Percentage of components	P-value
Among population	6	364.57	13.22	*67.58*	*P < 0.0001*
Within populations	22	139.57	6.34	*32.42*	
Between the two mountains	1	299.66	20.36	*71.86*	*P = 0.0303*
Among population	5	64.91	1.63	*5.74*	*P < 0.0001*
Within populations	22	139.57	6.34	*22.4*	
Ruwenzori					
Among population	3	45.50	1.74	*18.14*	*P < 0.0001*
Within populations	13	102.15	7.86	*81.8*	
Aberdare					
Among population	2	19.42	1.42	*25.42*	*P < 0.0001*
Within populations	9	37.42	4.16	*74.5*	

4.2.3 Lysimachia serpens (Primulaceae)

In *Lysimachia serpens*, a total of 110 individuals representing 22 populations were successfully genotyped. The AFLP analysis provided a total of 483 markers, of which all were polymorphic. The mean scoring error calculated from ten duplicate samples was 2.5%. In the PCoA of *L. serpens*, 35.7% of the total variation was explained by the first three axes. Two divergent groups (Structure-groups 1 & 2) were detected along the first axis which explained 23.6% of the total variation (Fig. 32). Some of the plants from Simen, Bale and Mt. Kilimanjaro split between the two lineages and formed a separate group (Structure-group 2). Hereafter these subgroups are named as Simen-2, Bale-2 and Kilimanjaro-2. The remaining plants from these mountains and all the

plants from Mt. Aberdare and Mt. Elgon formed the second group (Structure-group 1). The second axis which extracted 7.4 % of the variation mainly separated plants from Mt. Kilimanjaro from the remaining plants within Structure-group 1. (Fig. 32).

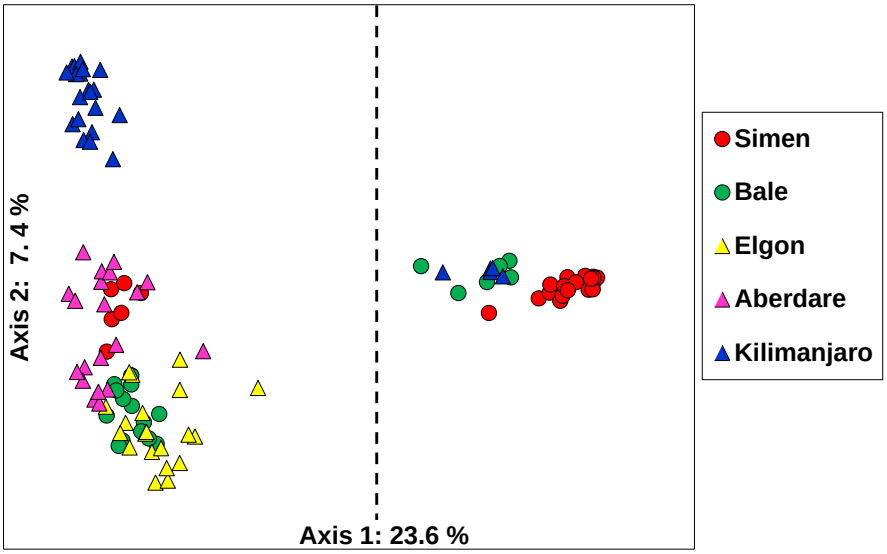


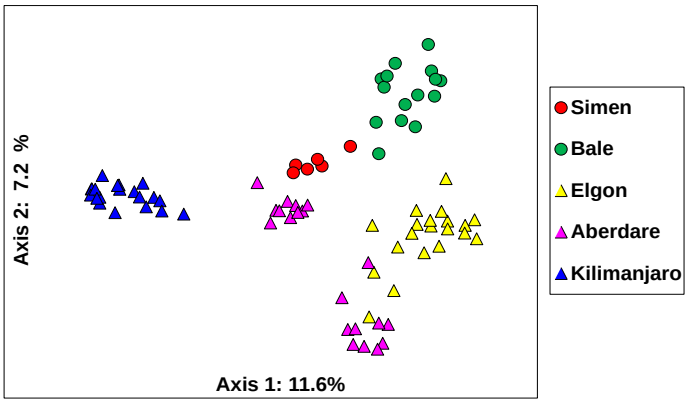
Fig. 32 PCoA based on Dice similarity among AFLP phenotypes of *L. serpens*. The percentages of the variance explained are indicated on each axis. Symbols identify regions (circles Ethiopian and, triangles East African), and colours represent mountains. The broken line separates the two genetic groups inferred by STRUCTURE.

In a separate PCoA of Structure-group 1, the first three axes extracted 25% of the total variance. Plants from Mt. Kilimanjaro were separated from the rest along axis 1 and 2, which extracted 11.6 % and 7.2 % of the total variation, respectively (Fig. 33a). Except for the plants from Mt. Kilimanjaro, the plants were not geographically structured. Plants from geographically close mountains (such as Bale/Simen and Mt. Aberdare/Elgon) were genetically more similar to each other than to plants from distant mountains. The first three main axes of

the PCoA for Structure-group 2 explained 58.6% of the overall variation. Within Structure-group 2, three geographically structured genetic subgroups were identified along the first and second axes, which extracted 32% and 15% of the total variation, respectively (Fig. 33b). Each of the three genetic subgroups, from Bale, Simen, and Mt. Kilimanjaro, were distinctly separated from each other and comprised plants only from a single mountain with no mixture. In the NJ analysis of the total dataset, Structure-group 2 formed a strongly supported cluster (100%, Fig. 35). All three subgroups identified within this lineage also formed three geographically isolated separate clusters with high bootstrap values (87%, 99% and 100%; Fig. 35). The genetic clustering within Structure-group 1 was rather weak and only three geographically separate populations were supported with 75% for Bale and 96% for each of Simen and Kilimanjaro (Fig. 35).

In the STRUCTURE analysis of the total dataset, the largest increase in the log probability of the data occurred at $K = 2$ (Fig. 34). The rate of change in the probability of the data between the successive Ks, ΔK, also identified $K = 2$ (Fig. 34) as the most likely number of clusters. Furthermore, at $K = 2$, all 10 replicates provided a stable assignment of individuals and suggested a split of the data into two clusters corresponding to the same two groups revealed in the PCoA; Structure-group 1 and Structure-group 2.

(a)



(b)

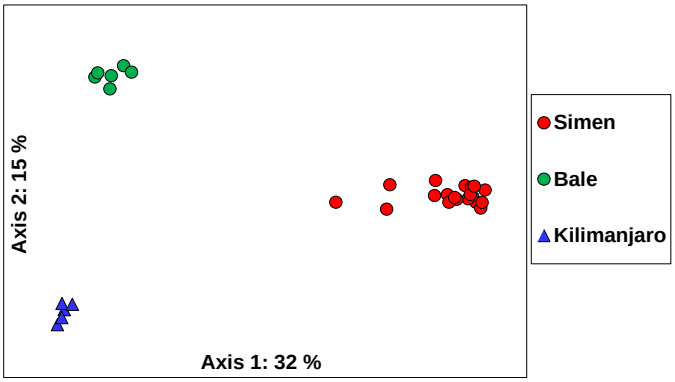


Fig. 33 PCoAs based on Dice similarity among AFLP phenotypes of (a) 'Structure-group 1', (b) 'Structure-group 2', lineages of *L. serpens*. The percentages of the variance explained are indicated on each axis. Symbols identify regions (circles Ethiopia, triangles East Africa), and colours represent mountains.

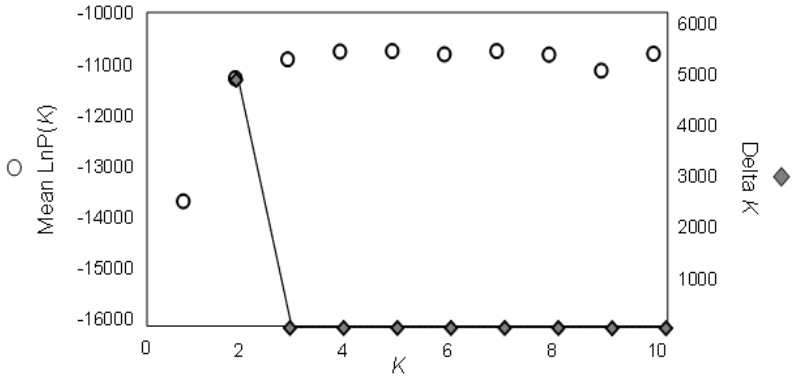


Fig. 34 Genetic structure of *L. serpens* inferred from Bayesian clustering of AFLP data. Mean value for log probability of the data LnP(*K*) and rate of change in the log probability between successive runs, Delta *K* (calculated according to Evanno *et al*. 2005) are shown as a function of *K*, ranging from 1 to 10.

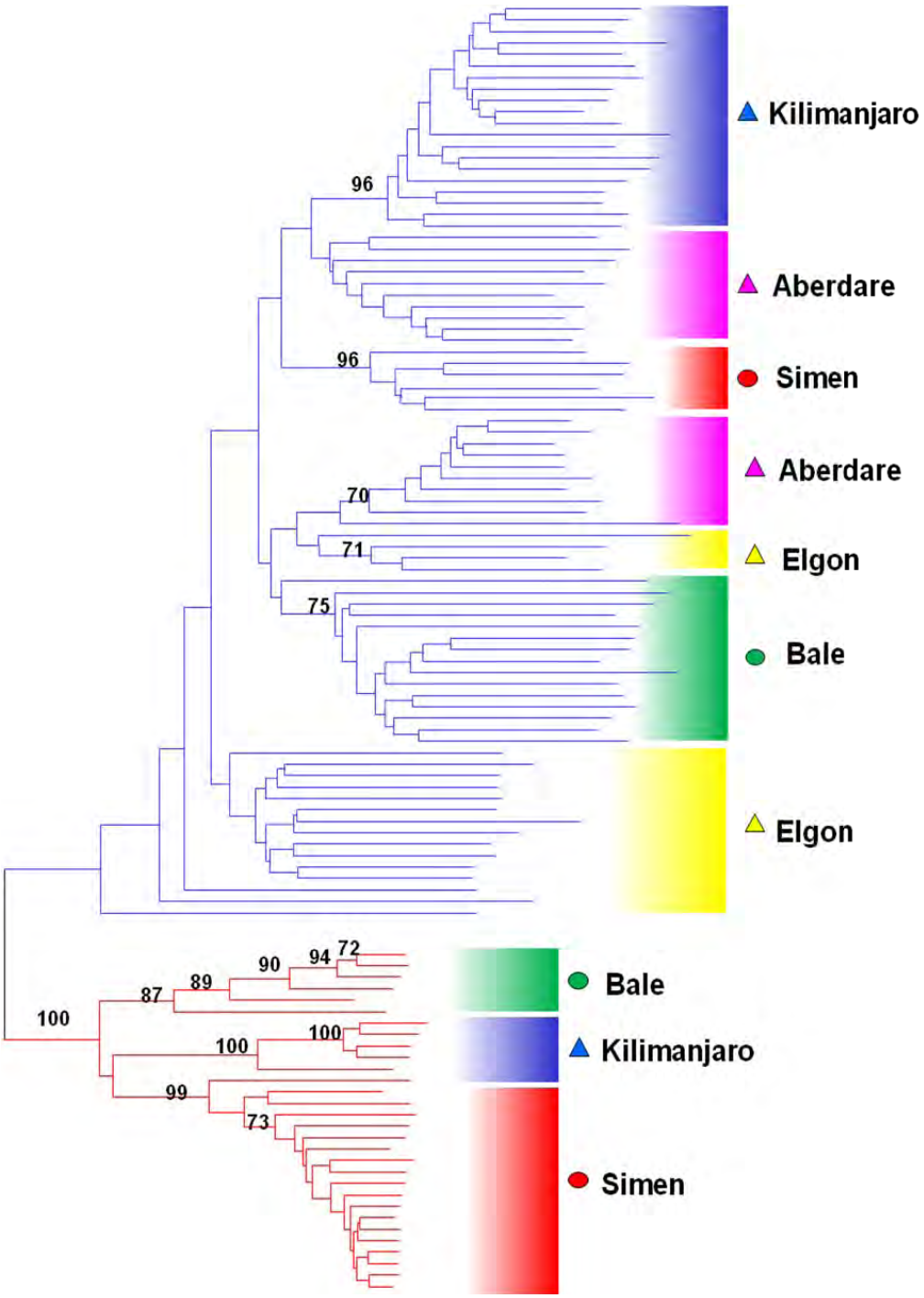


Fig. 35 Neighbour-joining tree of the AFLP phenotypes for 110 individual plants of *Lysimachia serpens* based on Nei and Li (1979) genetic distance. Symbols and their colours identify species and mountains respectively, as shown in the PCoA (Fig. 32). Branch colours indicate the two major genetic groups identified in the STRUCTURE analysis. Bootstrap values above 50% (estimated from 1000 replicates) are shown above the branches.

The total gene diversity estimate averaged over all investigated individuals of *L. serpens* was 0.0800. Both the highest and the lowest values of the within-population gene diversity estimates, D = 0.1238 and D = 0.0290, were observed from the two genetically divergent populations in Simen (ET0629 and ET0460, Table 17). The average within-population genetic distinctiveness value was 4.39 (SD = 1.10) and varied from 2.59 in the ET0460 (Simen) population to 7.09 in the TZ0013 (Kilimanjaro-2) population (Table 17). The populations from Mt. Elgon were most diverse (D = 0.1185, DW = 5.25; Table 18) followed by Mt. Aberdare (D = 0.1040, DW = 4.11). The divergent Kilimanjaro-2 population possessed the lowest level of genetic diversity and many rare markers (D = 0.0547, and DW = 7.09; Fig. 36 & Table 18).

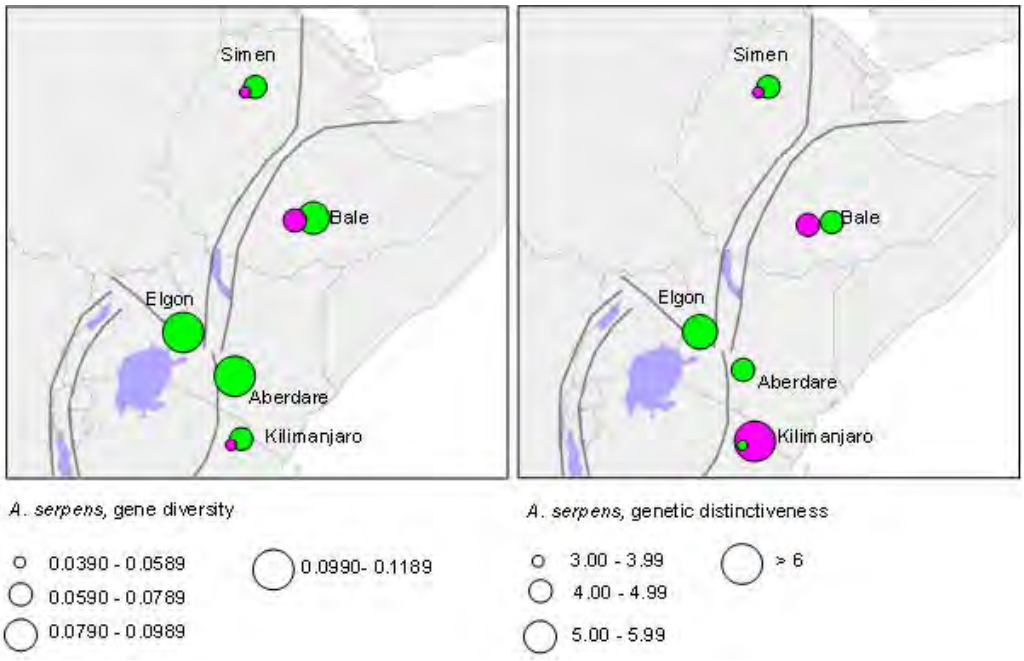


Fig. 36 Genetic structure, diversity and distinctiveness in *L. serpens* based on AFLP variation. Circle colours reflecting genetic (STRUCTURE) groups and circle size reflecting within-mountain gene diversity or genetic distinctiveness (DW). The thick dark grey lines depict the Great Rift Valley

Table 17 Geographic origin, collection data, and genetic diversity for 22 populations of *L. serpens*; *n* - number of individuals successfully analyzed, *D* - gene diversity, DW - frequency-down-weighted marker value as a measure of genetic distinctiveness

Pop ID	Country	Mountain	Lat/Long	Altitude	*n*	*D*	DW
ET0692	Ethiopia	Bale	N06.8793/E39.8689	3948	5	0.0563	4.37
ET0781	Ethiopia	Bale	N06.8687/E39.8822	4017	5	0.1230	5.69
ET0896	Ethiopia	Bale	N06.9049/E39.9046	3875	5	0.0741	3.89
ET1011	Ethiopia	Bale	N06.8931/E39.8974	3875	5	0.0795	5.18
ET0090	Ethiopia	Simen	N13.2827/E38.1108	3711	5	0.0621	4.23
ET0460	Ethiopia	Simen	N13.3491/E38.2625	3912	5	0.0290	2.59
ET0490	Ethiopia	Simen	N13.3285/E38.2430	3681	5	0.0319	2.74
ET0550	Ethiopia	Simen	N13.3285/E38.2409	3643	5	0.0948	3.33
ET0629	Ethiopia	Simen	N13.3271/E38.2425	3682	5	0.1238	4.23
KN0489	Kenya	Aberdare	S00.5573/E36.7175	3130	5	0.0932	5.34
KN0490	Kenya	Aberdare	S00.5403/E36.7092	3007	5	0.0733	3.58
KN0532	Kenya	Aberdare	S00.3068/E36.626	3858	5	0.0907	3.91
KN0623	Kenya	Aberdare	S00.3338/E36.6417	3590	5	0.0600	3.60
KN0298	Kenya	Mt. Elgon	N01.1007/E34.6215	3629	5	0.1201	5.15
KN0377	Kenya	Mt. Elgon	N01.1016/E34.6144	3772	5	0.1176	5.59
KN0380	Kenya	Mt. Elgon	N01.1017/E34.6047	3618	5	0.0928	4.56
KN0389	Kenya	Mt. Elgon	N01.1038/E34.6217	3617	5	0.1139	5.71
TZ0013	Tanzania	Kilimanjaro	S03.0343/E37.243	3636	5	0.0547	7.09
TZ0060	Tanzania	Kilimanjaro	S03.0056/E37.2416	3536	5	0.0729	4.70
TZ0149	Tanzania	Kilimanjaro	S03.0862/E37.3234	4155	5	0.0696	4.10
TZ0256	Tanzania	Kilimanjaro	S03.1099/E37.4211	4109	5	0.0538	3.04
TZ0302	Tanzania	Kilimanjaro	S03.135/E37.4337	3817	5	0.0737	3.98

Table 18 Genetic diversity for eight regional populations of *L. serpens*; *n* - number of individuals successfully analyzed, *D* - gene diversity, DW - frequency-down-weighted marker value as a measure of genetic distinctiveness. Divergent populations from Bale Kilimanjaro and Simen within Structure-group 2 lineage are indicated as Bale-2, Kilimanjaro-2 and Simen-2

Mountain	*n*	*D*	DW
Aberdare	20	0.1040	4.11
Bale	14	0.0831	4.91
Bale-2	6	0.0595	4.47
Elgon	20	0.1185	5.25
Kilimanjaro	20	0.0759	3.96
Kilimanjaro-2	5	0.0547	7.09
Simen	6	0.0679	4.69
Simen-2	19	0.0398	3.02

The non-hierarchical AMOVA for *L. serpens* allocated 45% of the variation among the 22 populations (Table 19). In the hierarchical AMOVA, 36% of the variation was assigned between the two structure groups (*P = 0.0244;* Table 19). The remaining 21% and 43% of the variation were assigned to among the seven mountain populations and within-mountain variation, respectively. In separate analyses, a high level of differentiation (58%; Table 19) was detected among the three mountain populations within Structure-group 2, whereas, differentiation among mountains within Structure-group 1 was low (26%).

Table 19 Analysis of molecular variance (AMOVA) of the AFLP data for *L. serpens. p*-values were estimated in a permutation test (10000 permutations).

Source of Variation	d.f.	Sum of squares	Variance components	Percentage of components	P-value
Among populations	21	2067.16	15.82	45.01	*P < 0.0001*
Within populations	88	1700.80	19.33	54.99	
Between two STRUCTURE groups	1	893.13	16.70	36.33	*P = 0.0244*
Among mountains	6	865.13	9.57	20.82	*P < 0.0001*
Within mountains	102	2009.69	19.70	42.8	
Structure-group 1					
Among mountains	4	597.28	8.15	26.3	*P < 0.0001*
Within mountains	75	1712.01	22.83	73.7	
Structure-group 2					
Among mountains	2	267.85	15.43	58.32	*P < 0.0001*
Within mountains	27	297.69	11.03	41.68	

4.2.4 Dicrocephala chrysanthemifolia (Asteraceae)

In *Dicrocephala chrysanthemifolia*, a total of 151 individuals (31 populations) were successfully genotyped. The AFLP analysis provided a total of 509 markers, of which all were polymorphic. Average error rate calculated over the 24 duplicated samples was 3.4%. The first three main axes in the PCoA of *D. chrystanthemifolia* explained 18.5% of the variation. No clear geographic structuring was detected in the PCoA, however, two divergent lineages were identified along the first and second axes, which explained 8.7 % and 6.4% of the variation, respectively (Fig. 37). Four populations from Bale were separated from the rest of the plants and formed a distinct group (Bale-2), whereas the remaining plants from Bale were genetically closer to plants from Simen and from other mountains in Kenya, Tanzania and Uganda (Structure-group 1) (Fig. 37).

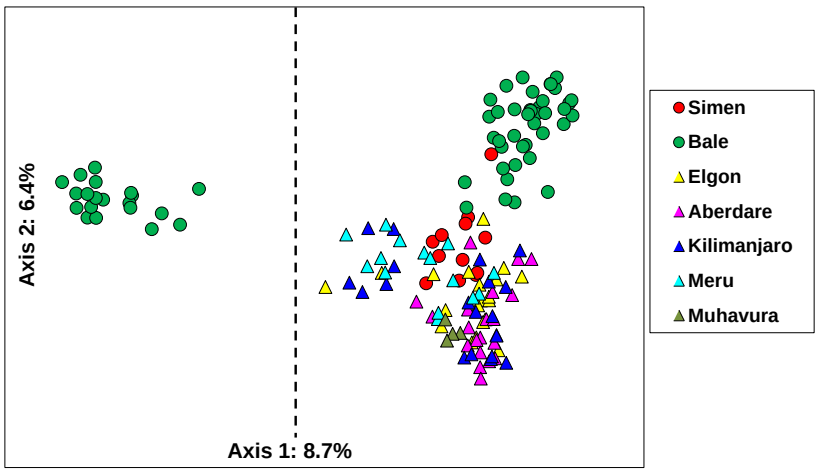


Fig. 37 PCoA based on Dice similarity among AFLP phenotypes of *D. chrysanthemifolia*. The percentages of the variance explained are indicated on each axis. Symbols identify regions (circles-Ethiopian and, triangles-East African), and colours represent mountains. The broken line separates the two genetic groups inferred by STRUCTURE.

In a separate PCoA for Structure-group 1 without the four divergent Bale populations, a geographical pattern of the genetic variability was not very clear along the first and second axes, which extracted 7.5% and 2.4 % of the total variation, respectively (Fig. 38). However, there was some genetic structuring. Along the first axis, the plants from the Bale mountain were slightly separated from the remaining plants. Individual plants from Mt. Kilimanjaro, Meru, and Elgon were scattered along this axis. On the contrary, individual plants from Mt. Aberdare, Mt. Muhavura, and Simen were closer to each other. In the NJ tree, the four divergent populations from Bale formed a cluster with bootstrap support of 98% (Fig. 40). The relationships within the other lineage were weak and individual plants from different mountains grouped together mostly without bootstrap support. Only a few individuals from Simen and Kilimanjaro grouped together with bootstrap support value > 53% (Fig. 40).

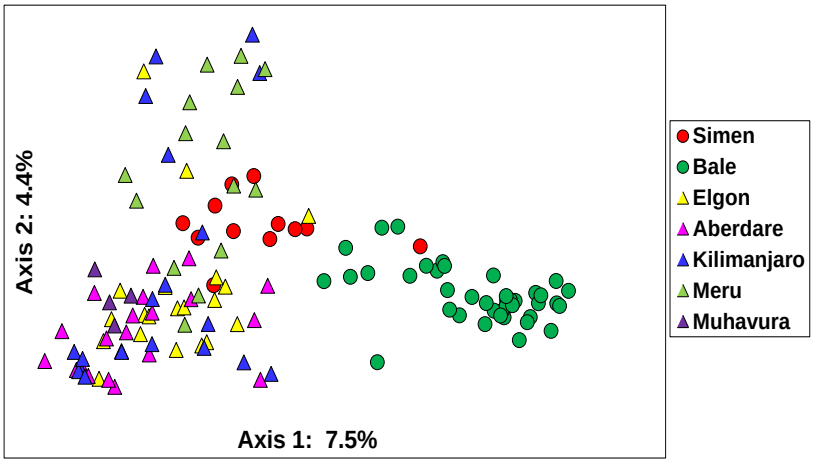


Fig. 38 PCoA based on Dice similarity among AFLP phenotypes of *D. chrysanthemifolia*, without the four divergent populations from Bale. The percentages of the variance explained are indicated on each axis. Symbols identify regions (circles Ethiopian and, triangles East African), and colours represent mountains. The broken line separates the two genetic groups inferred by STRUCTURE.

In the STRUCTURE analysis of the total dataset, the largest increase in the log posterior probability of the data occurred at $K = 2$ (Fig. 39). The rate of change in the probability of the data between the successive Ks, ΔK, also identified $K = 2$ (Fig. 39) as the most likely number of groups, corresponding to the same groups detected along the first two axes in the PCoA of Fig. 37.

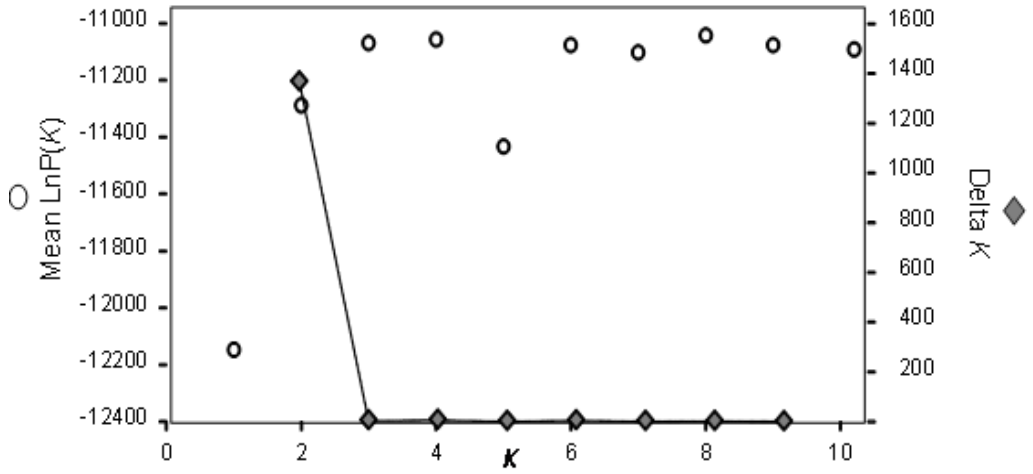


Fig. 39 Genetic structure of *D. chrysanthemifolia* inferred from Bayesian clustering of AFLP data. Mean value for log probability of the data LnP (*K*) and rate of change in the log probability between successive runs, Delta *K* (calculated according to Evanno *et al*. 2005) are shown as a function of *K*, ranging from 1 to 10.

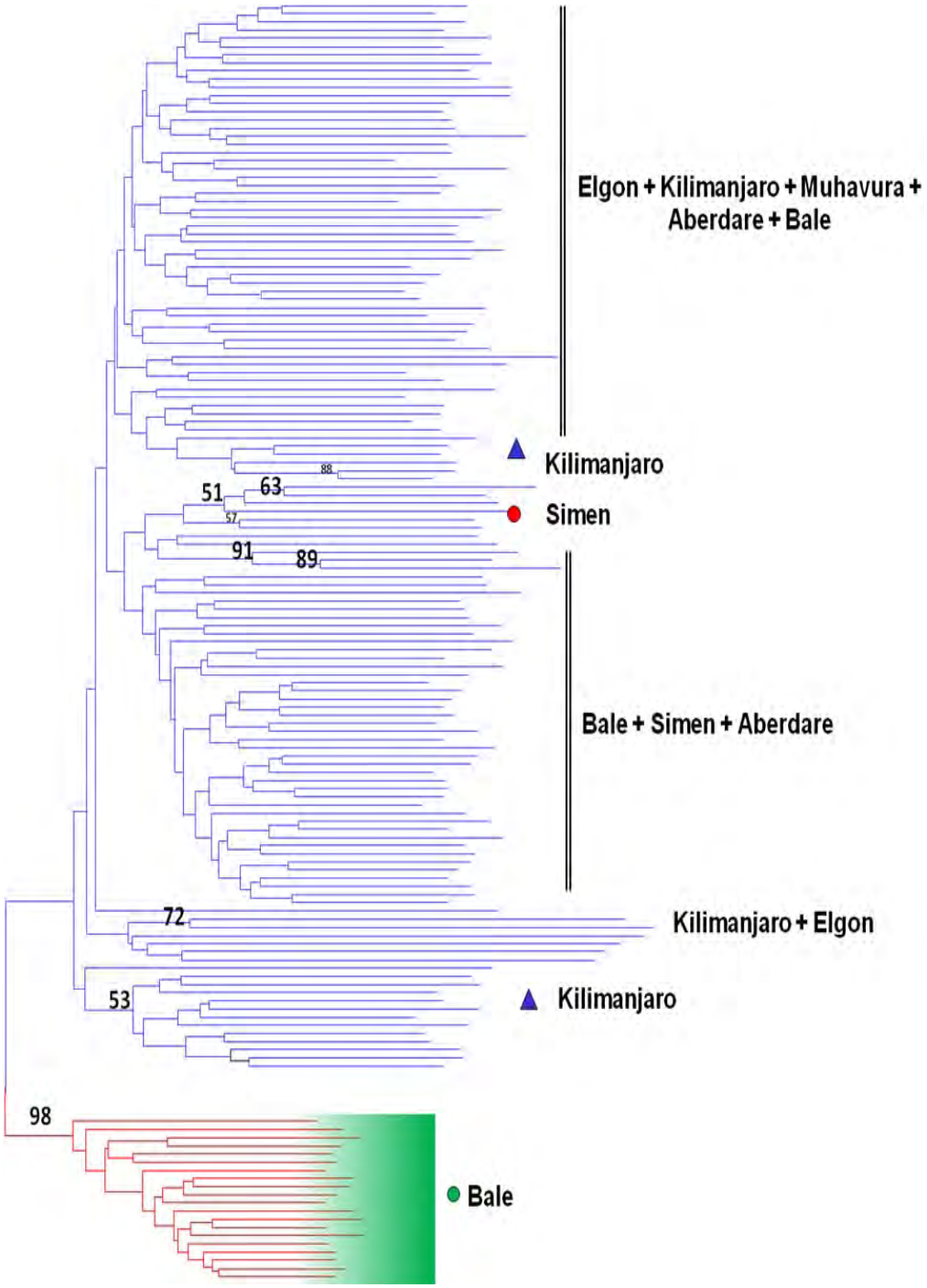


Fig. 40 Neighbour-joining tree of the AFLP phenotypes for 151 individual plants of *D. chrysanthemifolia* based on Nei and Li (1979) genetic distance. Symbols and their colours identify species and mountains respectively, as shown in the PCoA (Fig. 37 & 38). Branch colours indicate the two major genetic groups identified in the STRUCTURE analysis. Bootstrap values above 50% (estimated from 1000 replicates) are shown above the branches.

The total gene diversity estimate averaged over all investigated *D. chrysanthemifolia* individuals was *D* = 0.0580. The average within-population gene diversity varied from 0.0310 in the ET0871 population from Bale to 0.1014 in the KN0310 population from Elgon (Table 20). Within-population genetic distinctiveness value was highest in one population from Simen and lowest in one population from Bale (Table 20). The plants from the Simen mountain possessed the highest level of gene diversity as well as genetic distinctiveness (*D* = 0.0760, DW =5.60; Table 21), while low levels of genetic diversity and distinctivity were observed among the plants from Bale (*D* = 0.0459, DW =2.23; Fig. 41 & Table 21).

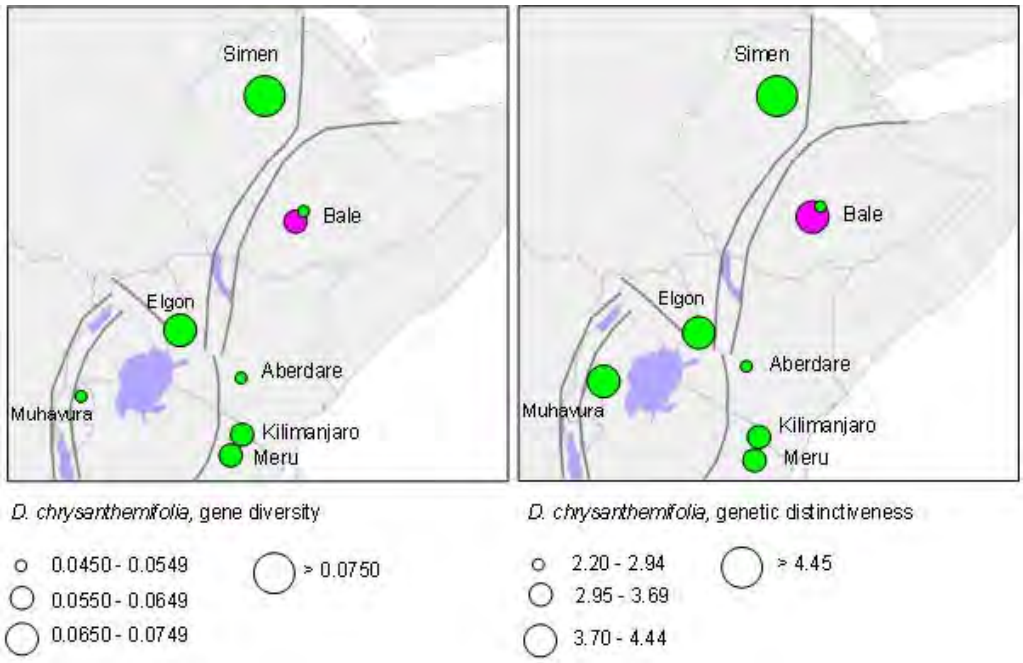


Fig. 41 Genetic structure, diversity and distinctiveness in *D. chrysanthemifolia* based on AFLP variation. Circle colours reflecting genetic (STRUCTURE) groups and circle size reflecting within-mountain gene diversity or genetic distinctiveness (DW). The thick dark grey lines depict the Great Rift Valley

Table 20 Geographic origin, collection data, and genetic diversity for 31 populations of *D. chrysanthemifolia*; *n* - number of individuals successfully analyzed, *D* - gene diversity, DW - frequency-down-weighted marker value as a measure of genetic distinctiveness

Pop ID	Country	Mountain	Lat/Long	Altitude	n	*D*	DW
ET0151	Ethiopia	Simen	N13.2823/E38.1108	3711	2	0.0766	4.57
ET0164	Ethiopia	Simen	N13.28800/E38.1188	3716	5	0.0656	4.91
ET0269	Ethiopia	Simen	N13.2697/E38.1059	3652	5	0.0857	7.32
ET0654	Ethiopia	Bale	N6.8550/E39.8780	4143	5	0.0420	2.02
ET0758	Ethiopia	Bale	N6.8703/E39.8678	4101	5	0.0413	2.18
ET0804	Ethiopia	Bale	N6.8448/E39.8804	4129	5	0.0365	1.44
ET0871	Ethiopia	Bale	N6.8694/E39.8947	4068	5	0.0310	1.26
ET0937	Ethiopia	Bale	N6.8822/E39.8883	3986	5	0.0566	3.25
ET0956	Ethiopia	Bale	N6.8500/E39.8531	4116	5	0.0633	3.91
ET0978	Ethiopia	Bale	N6.8931/E39.8973	3875	5	0.0558	3.38
ET0989	Ethiopia	Bale	N6.90487/E39.9046	3875	5	0.0813	5.66
ET1424	Ethiopia	Bale	N7.0187/E39.7207	3484	5	0.0538	3.01
ET1450	Ethiopia	Bale	N7.0073/E39.7098	3482	5	0.0511	2.40
ET1458	Ethiopia	Bale	N6.9907/E39.69	3499	5	0.0566	2.57
ET1488	Ethiopia	Bale	N6.9897/E39.703	3520	5	0.0550	2.92
KN0310	Kenya	Mt. Elgon	N1.1007/E34.6215	3629	5	0.1014	6.35
KN0364	Kenya	Mt. Elgon	N1.0905/E34.6207	3584	5	0.0479	2.55
KN0397	Kenya	Mt. Elgon	N1.1002/E34.6098	3724	5	0.0550	2.73
KN0399	Kenya	Mt. Elgon	N1.0923/E34.6183	3664	5	0.0656	4.08
KN0608	Kenya	Aberdare	S0.339/E36.668	3466	5	0.0593	2.86
KN0609	Kenya	Aberdare	S0.339/E36.668	3466	5	0.0460	2.06
KN0612	Kenya	Aberdare	S0.339/E36.668	3466	5	0.0566	2.65
KN0658	Kenya	Aberdare	S0.3388/E36.6515	3702	5	0.0420	1.82
TZ0067	Tanzania	Kilimanjaro	S3.0056/E37.2416	3536	4	0.0586	4.03
TZ0336	Tanzania	Kilimanjaro	S3.1581/E37.4906	3063	5	0.0487	2.72
TZ0339	Tanzania	Kilimanjaro	S3.1518/E37.4758	3237	7	0.0709	4.60
TZ0361	Tanzania	Kilimanjaro	S3.1375/E37.4368	3710	4	0.0566	2.74
TZ0494	Tanzania	Mt. Meru	S3.2201/E36.7952	3383	5	0.0519	3.03
TZ0497	Tanzania	Mt. Meru	S3.2178/E36.7707	3589	5	0.0562	2.35
TZ0504	Tanzania	Mt. Meru	S3.2178/E36.7707	3589	5	0.0750	5.55
UG2214	Uganda	Muhavura	S1.3813/E29.6760	4050	4	0.0527	3.94

Table 21 Genetic diversity for eight regional populations of *D. chrysanthemifolia*; *n* - number of individuals successfully analyzed, *D* - gene diversity, DW - frequency-down-weighted marker value as a measure of genetic distinctiveness. The divergent populations from Bale are indicated as Bale-2

Mountain	*n*	*D*	DW
Aberdare	20	0.0510	2.35
Bale	40	0.0459	2.23
Bale-2	20	0.0642	4.05
Elgon	20	0.0675	3.93
Kilimanjaro	20	0.0587	3.52
Meru	15	0.0610	3.64
Muhavura	4	0.0527	3.94
Simen	12	0.0760	5.60

The non-hierarchical AMOVA for the total dataset of *D. chrysanthemifolia* attributed 26.3% of the overall variation to among-population variation (Table 22). In the three-level hierarchical analysis, 35.3% of the total variation was attributed to variation between the two divergent STRUCTURE groups, and 10.8 and 54% to variation among populations and within populations, respectively. In a different hierarchical analysis, AMOVA assigned the highest percentage of variation, 72.5%, to within-population variation and 19% among populations within mountains. The plants from different mountains were genetically similar, as indicated by low level of genetic differentiation among mountains (8.4%, *P = 0.00196;* Table 22).

Table 22 Analysis of molecular variance (AMOVA) of the AFLP data for *D. chrysanthemifolia.* *p*-values were estimated in a permutation test (10000 permutations).

Source of Variation	d.f	Sum of squares	Variance components	Percentage of variation	*P*-value
Among populations	30	1205.79	5.24	26.27	*P < 0.0001*
Within populations	120	1764.04	14.70	73.73	
Among STRUCTURE groups	1	363.44	9.62	35.29	*P < 0.0001*
Among population	29	842.35	2.95	10.82	*P < 0.0001*
Within populations	120	1764.04	14.70	53.90	
Among mountains	6	398.88	1.71	8.41	*P* = 0.00196
Among populations	24	806.91	3.87	19.11	*P < 0.0001*
Within populations	120	1764.04	14.70	72.48	

4.2.5 Haplocarpha rueppellii (Asteraceae)

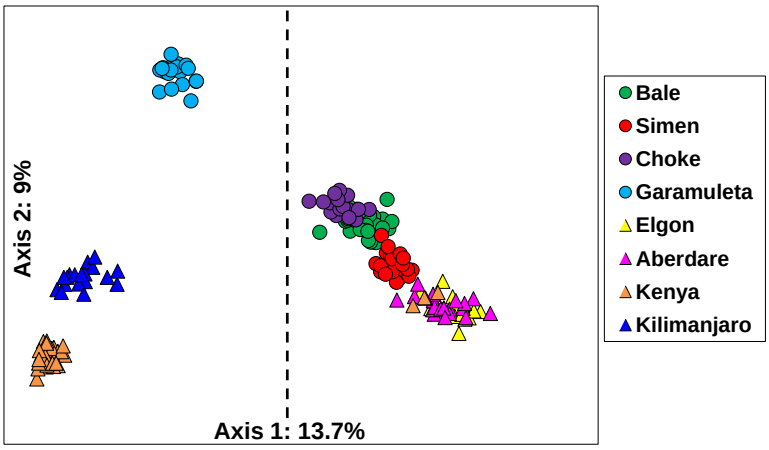
In *Haplocarpha rueppellii*, a total of 191 individuals from 38 populations were successfully genotyped. The AFLP analysis provided a total of 637 markers, of which 99.84% were polymorphic. The mean scoring error calculated from 28 duplicated samples was 3.8%. Clear geographical structuring of the genetic variation was observed in the total dataset of *H. rueppellii*. The first three main axes of the PCoA explained 30% of the total genetic variation. Along the first axis, which explained 13.7% of the variation (Fig. 42a), two main lineages were observed. In the first lineage (Structure-group 1), the plants from Gara Muleta, Mt. Kilimanjaro and Mt. Kenya formed three distinct clusters and were clearly separated from the other lineage (Structure-group 2), which comprised plants from Ethiopia (Bale, Simen, and Choke) and East Africa (Aberdare, Elgon and five individuals from the Kenya788 population; Fig. 42a & b). The second axis, which explained 9% of the variation, mainly separated the Gara Muleta plants from the remaining plants. Further geographical structuring was observed along the first and third axes (which explained 7.3% of the variation) and mainly

separated Mt. Aberdare, Mt. Elgon and one Mt. Kenya population (Kenya788) from the Ethiopian plants from Bale, Simen and Choke (Fig. 42b). In the NJ analysis, the two main lineages inferred from the PCoAs were also resolved and formed two strongly supported clusters with bootstrap value of 94% (Fig. 46). Furthermore, strong genetic structuring was detected within each lineage. Within the first lineage all plants from Gara Muleta, Mt. Kilimanjaro and Mt. Kenya formed separate clusters with bootstrap support of 99-100% (Fig. 46). In the other lineage, plants from the three East African mountains of Elgon, Aberdare and Kenya (Kenya788) formed a cluster with 98% bootstrap support. Similarly, all plants from Simen formed a strongly supported (bs = 100%) cluster which again was supported with 75% bootstrap value to form a cluster with plants from the East African mountains of Mt. Elgon, Mt. Aberdare and Mt. Kenya (Kenya788).

In the STRUCTURE analysis of the total dataset, the largest increase in the log posterior probability of the data occurred at $K = 2$ (Fig. 44). The rate of change in the probability of the data between the successive Ks, ΔK, also identified $K = 2$ (Fig. 44) as the most likely number of groups. Furthermore, at $K = 2$, all 10 replicates provided a stable assignment of individuals and suggested a split of the total dataset into two major Structure-groups corresponding to the split observed along the first axis of the PCoA. Structure-group 1 comprised plants from Mt. Kilimanjaro, Gara Muleta, and Mt. Kenya, and Structure-group 2 included plants from Ethiopia (Bale, Simen, and Choke) and other East African mountains (Mt. Aberdare, Mt. Elgon and Kenya0788). Separate analyses for

each of the two main STRUCTURE groups resulted in six subgroups, three in each group (Fig. 43a & b).

(a)



(b)

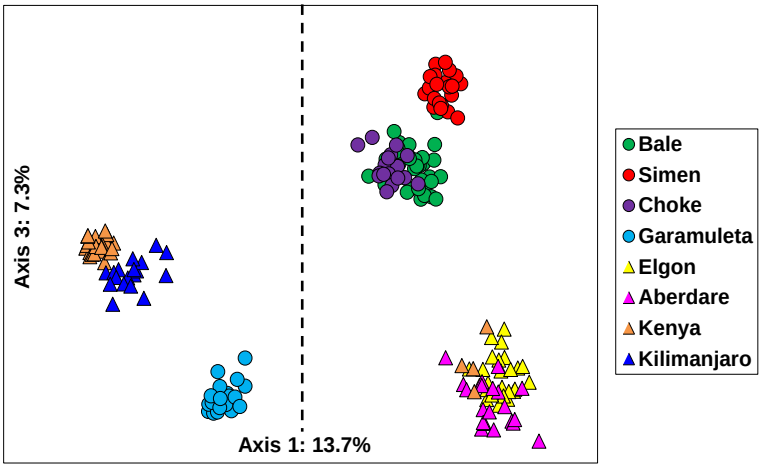
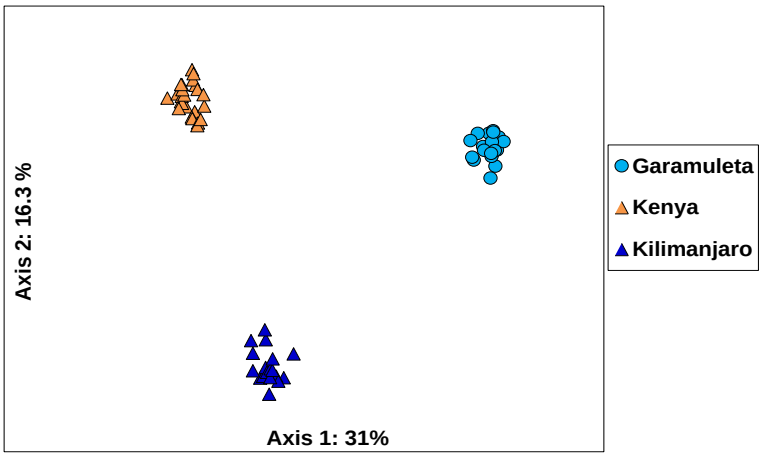


Fig. 42 PCoAs based on Dice similarity among AFLP phenotypes of *H. rueppellii*. (a) Axis 1 & 2 (b) Axis 1 & 3. The percentages of the variance explained are indicated on each axis. Symbols identify regions (circles-Ethiopian and, triangles-East African), and colours represent mountains. The broken line separates the two genetic groups inferred by STRUCTURE.

Two separate PCoAs were run, one for each of the main STRUCTURE groups (Fig. 43a & b). In the PCoA of Structure-group 1 (Fig. 43a), the first three axes explained half of the total genetic variation (50%). Strong geographical structuring of the genetic variation was observed, among the plants from the

three geographically isolated mountains of Mt. Kilimanjaro, Gara Muleta, and Mt. Kenya. Axis 1, which explained 31% of the total variation, separated the Ethiopian plants (Gara Muleta) from the East African plants (Mt Kilimanjaro and Kenya). Plants from Mt. Kilimanjaro were mainly separated from the rest of the plants along the second axis, which depicted 16.3% of the total variation.

(a)



(b)

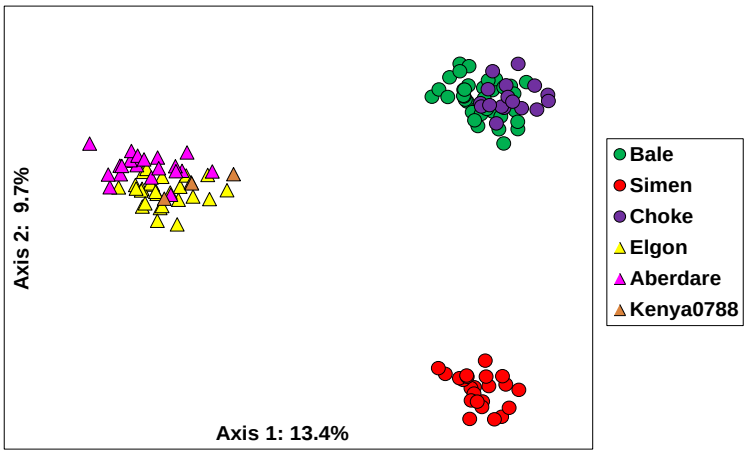


Fig. 43 PCoA based on Dice similarity among AFLP phenotypes of (a) Structure-group 1 (b) Structure-group 2 for *H. rueppellii* populations. The percentages of the variance explained are indicated on each axis. Symbols identify regions (circles-Ethiopia, triangles-East Africa), and colours represent mountains.

In the PCoA for Structure-group 2, the first three axes explained 30% of the total variation. A strong geographical structuring was observed with three

distinct clusters along the first two axes. While plants from the East African mountains (Mt. Aberdare, Mt. Elgon and Kenya788) formed a distinct cluster, the plants from Ethiopia formed two distinct groups. In Ethiopia, the plants from geographically close mountains such as Simen and Choke in the northern part of the country were not genetically similar. Rather, the plants from Choke were genetically most similar to plants in the south, Bale. Along the first axis of the PCoA, which explained 13.4% of the variation, the Ethiopian plants (Bale, Simen, and Choke) were clearly separated from East African plants (Fig. 43b). The second axis explained 9.7% of the total variation and mainly separated Simen from the rest of the mountains.

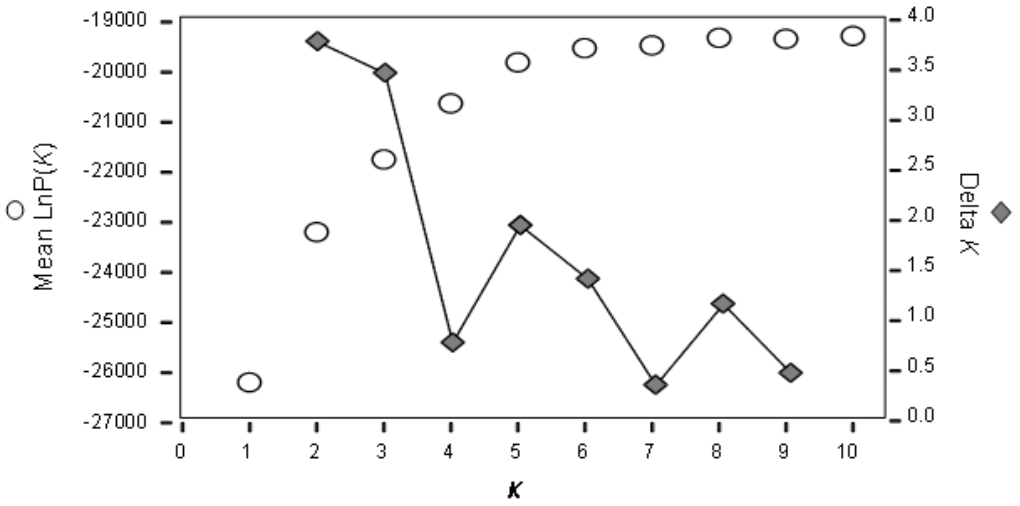


Fig. 44 Genetic structure of *H. rueppellii* inferred from Bayesian clustering of AFLP data. Mean value for log probability of the data LnP(*K*) and rate of change in the log probability between successive runs, Delta *K* (calculated according to Evanno *et al*. 2005) are shown as a function of *K*, ranging from one to 10.

The total gene diversity estimate averaged over all investigated individuals of *H. rueppellii* was 0.0619 (*SD* = 0.016). The average within-population gene diversity varied from 0.0390 in the ET0299 Ethiopian population from Simen to 0.1074 in the ET0882 Ethiopian population from Bale (Table 23). The within-

population genetic distinctiveness value varied from 1.83 in the KN0141 population from Elgon to 6.8 in the ET1038 population from Bale. The average distinctiveness value was 3.26 (*SD* = 1.082). Gene diversity and distinctiveness was highest in Bale (*D* = 0.0972, DW = 4.44) and Choke (*D* = 0.0761, DW = 4.27; Table 24, Fig. 45), whereas lowest levels of gene diversity and distinctivity were observed in the population Kenya788 (*D* = 0.0542; Fig. 45 & Table 24).

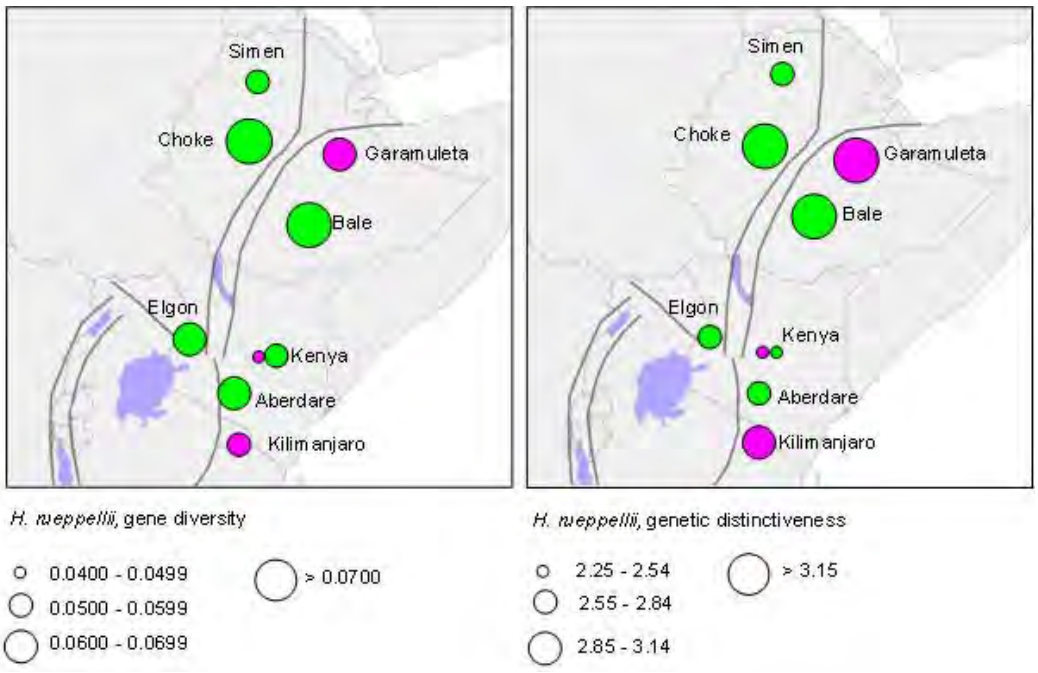


Fig. 45 Genetic structure, diversity and distinctiveness in *H. rueppellii* based on AFLP variation. Circle colours reflecting genetic (STRUCTURE) groups and circle size reflecting within-mountain gene diversity or genetic distinctiveness (DW). The thick dark grey lines depict the Great Rift Valley

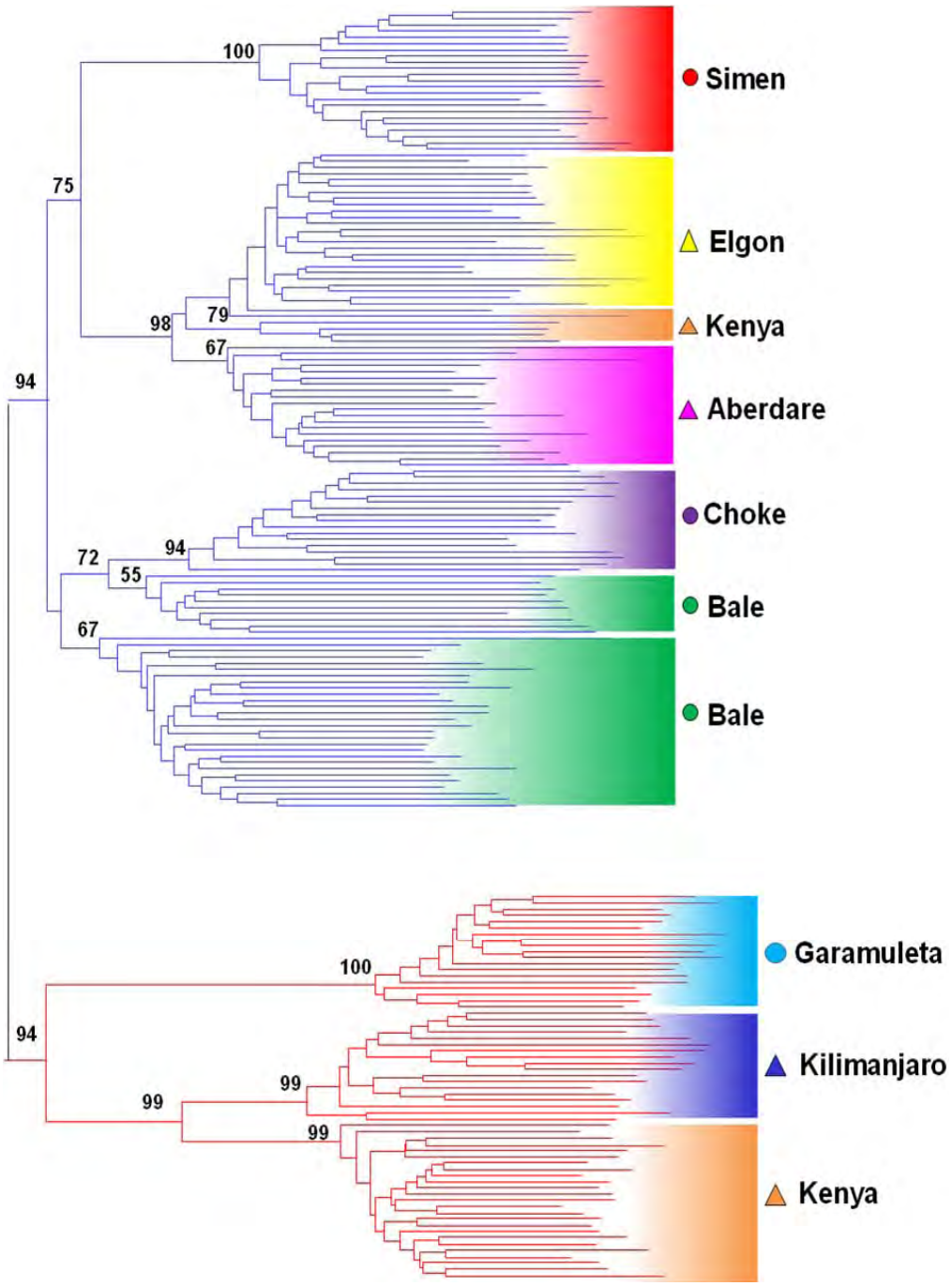


Fig. 46 Neighbour-joining tree of the AFLP phenotypes for 191 individual plants of *H. rueppellii* based on Nei and Li (1979) genetic distance. Symbols and their colours identify species and mountains respectively as shown in the PCoA (Fig. 42 & 43). Branch colours indicate the two major genetic groups identified in the STRUCTURE analysis. Bootstrap values above 50% (estimated from 1000 replicates) are shown above the branches.

Table 23 Geographic origin, collection data, and genetic diversity for 38 populations of *H. rueppellii*; *n* - number of individuals successfully analyzed, *D* - gene diversity, DW - frequency-down-weighted marker value as a measure of genetic distinctiveness

Pop ID	Country	Mountain	Lat/Long	Altitude	n	*D*	DW
ET0001	Ethiopia	Bale	N06.8547/E39.8933	4059	4	0.0633	3.10
ET0636	Ethiopia	Bale	N06.8551/E39.8781	4143	5	0.0706	3.23
ET0686	Ethiopia	Bale	N06.8793/E39.8681	3948	5	0.0754	3.31
ET0777	Ethiopia	Bale	N06.8687/E39.8822	4017	5	0.0791	4.02
ET0855	Ethiopia	Bale	N06.8694/E39.895	4068	4	0.0795	3.84
ET0882	Ethiopia	Bale	N06.8822/E39.8883	3986	5	0.1074	5.31
ET0968	Ethiopia	Bale	N06.9049/E39.9046	3875	5	0.0932	5.51
ET1038	Ethiopia	Bale	N06.8931/E39.8974	3875	5	0.1055	6.80
ET1339	Ethiopia	Choke	N10.642/E37.8357	3960	5	0.0744	4.04
ET1352	Ethiopia	Choke	N10.656/E37.8257	3943	3	0.0733	3.53
ET1366	Ethiopia	Choke	N10.6382/E37.8392	3908	5	0.0766	4.73
ET1389	Ethiopia	Choke	N10.6575/E37.822	3919	4	0.0798	4.55
ET1521	Ethiopia	Gara Muleta	N09.2195/E41.7867	3144	19	0.0620	4.05
ET0101	Ethiopia	Simen	N13.2827/E38.1108	3711	4	0.0403	2.01
ET0299	Ethiopia	Simen	N13.2513/E38.2022	4035	4	0.0390	2.61
ET0442	Ethiopia	Simen	N13.3284/E38.2421	3681	5	0.0509	3.13
ET0545	Ethiopia	Simen	N13.3285/E38.2401	3643	5	0.0537	2.85
ET0590	Ethiopia	Simen	N13.3271/E38.2425	3682	5	0.0490	2.96
KN0508	Kenya	Aberdare	S00.3069/E36.6257	3843	5	0.0512	2.29
KN0596	Kenya	Aberdare	S00.3338/E36.6417	3594	5	0.0584	2.78
KN0600	Kenya	Aberdare	S00.3353/E36.643	3605	5	0.0480	1.92
KN0601	Kenya	Aberdare	S00.3345/E36.641	3561	5	0.0681	4.09
KN0047	Kenya	Mt. Elgon	N01.1025/E34.6058	3864	3	0.0513	2.80
KN0122	Kenya	Mt. Elgon	N01.1232/E34.607	3995	5	0.0527	2.41
KN0141	Kenya	Mt. Elgon	N01.1185/E34.5857	4029	5	0.0484	1.83
KN0187	Kenya	Mt. Elgon	N01.118/E34.5867	4043	5	0.0650	3.76
KN0261	Kenya	Mt. Elgon	N01.1083/E34.6062	3864	4	0.0720	2.56
KN0275	Kenya	Mt. Elgon	N01.1029/E34.6131	3979	5	0.0703	3.04
KN0788	Kenya	Mt. Kenya	S00.0432/E37.2855	3395	4	0.0542	2.25
KN0852	Kenya	Mt. Kenya	S00.1392/E37.3143	4230	5	0.0414	2.26
KN0895	Kenya	Mt. Kenya	S00.1214/E37.2956	4044	5	0.0414	2.04
KN0933	Kenya	Mt. Kenya	S00.1336/E37.2765	4193	5	0.0427	2.59
KN1008	Kenya	Mt. Kenya	S00.1503/E37.3301	4379	5	0.0512	3.18
KN1046	Kenya	Mt. Kenya	S00.1461/E37.3471	4019	5	0.0477	2.44
TZ0028	Tanzania	Kilimanjaro	S02.9866/E37.2224	3406	4	0.0576	3.37
TZ0128	Tanzania	Kilimanjaro	S03.0621/E37.2815	3600	5	0.0556	3.49
TZ0233	Tanzania	Kilimanjaro	S03.1091/E37.4211	4109	5	0.0534	2.47
TZ0306	Tanzania	Kilimanjaro	S03.135/E37.4337	3817	4	0.0495	2.60

Table 24 Genetic diversity for nine regional populations of *H. rueppellii*; n - number of individuals successfully analyzed, *D*- gene diversity, DW - frequency-down-weighted marker value as a measure of genetic distinctiveness. The divergent populations from Kenya are indicated as Kenya788

Mountain	*n*	*D*	DW
Aberdare	20	0.0600	2.77
Bale	38	0.0972	4.44
Choke	17	0.0761	4.27
Elgon	27	0.0630	2.73
Gara Muleta	19	0.0620	4.05
Kenya	25	0.0463	2.50
Kenya788	4	0.0542	2.25
Kilimanjaro	18	0.0590	2.98
Simen	23	0.0537	2.75

The non-hierarchical AMOVA for the whole dataset assigned 48.2% of the total genetic variance to variation among the 38 populations and a relatively high percentage of the variation (51.8%) to within-population variation (Table 25). A three-level hierarchical AMOVA assigned a higher percentage of variance to among mountains (42.8%) than to among populations within mountains (7.7%). In a different hierarchical analysis, 21.4% of the total variation was assigned between the two STRUCTURE groups, and 32% of the variance to among populations within the two groups. Separate analyses of each of the STRUCTURE groups assigned higher percentages of total variance to among mountains (54.6%) and within populations (34 %) than to among populations within mountains (2% and 6.5%; Table 25).

Table 25 Analysis of molecular variance (AMOVA) of the AFLP data for *H. rueppellii*. *p*-values were estimated in a permutation test (10000 permutations).

Source of variation	d.f.	Sum of squares	Variance components	Percentage of variation	P-value
Among populations	37	4132.70	18.40	48.21	*P < 0.0001*
Within populations	153	3024.35	19.77	51.79	
Among mountains	7	3112.17	17.14	42.88	*P < 0.0001*
Among populations	30	1020.53	3.07	7.69	*P < 0.0001*
Within populations	153	3024.35	19.77	49.44	
Between two STRUCTURE groups	1	902.34	9.23	21.36	*P < 0.0001*
Among population	36	3230.35	14.22	32.90	*P < 0.0001*
Within populations	153	3024.35	19.77	45.74	
Structure-group 1					
Among mountains	2	924.99	21.42	54.63	*P = 0.00196*
Among populations	7	146.01	0.81	2.08	*P < 0.0001*
Within populations	52	882.42	16.97	43.29	
Structure-group 2					
Among mountains	5	1454.70	12.63	34.90	*P < 0.0001*
Among populations	22	704.66	2.35	6.48	*P < 0.0001*
Within populations	101	2141.93	21.21	58.62	

CHAPTER 5 DISCUSSION

5.1 Pliocene origin and intensive diversification during the Pleistocene: the evolution of the afro-montane *Lychnis* L. (Caryophyllaceae)

The incorporation of a fossil calibrated relaxed molecular clock and a multilocus species tree approach used in this study provided strong support for a late-Miocene to late-Pliocene origin of the afro-montane *Lychnis* [2.28-5.8 Myr, 95% HPD, Table 6] and initial diversification into two ploidy specific lineages. In East Africa, there was increased aridity at around 6-7 Myr ago (Cerling *et al.*, 1993) and also in the late-Pliocene around 3 Myr ago (Bobe, 2006). Furthermore, much of the climate during the last 5 Myr was known to be warm (Wara, 2005), and in combination these conditions could have promoted diversification.

Several studies have shown that trees and shrubs dominated in Eastern Africa during the Pliocene. Evidence from the Turkana Basin of Kenya indicates that there was a closed and humid environment present at about 4 Myr ago (Bonnefille, 1995; Bobe, 2006) during which rainforest expanded into regions of Eastern Africa (Williamson, 1985; Pickford *et al.*, 2004), where it persisted until the late-Pliocene at around 3.4 Myr ago in the Ethiopian highlands (Wolde Gabriel *et al.*, 1994; Bonnefille, 1995). The warm and moist climatic conditions may have created an environment that promoted the expansion of forest in the region. The 95% HPD mean estimates for the origin and initial diversification of the afro-montane *Lychnis* clearly correspond to this warm and moist period. It is, therefore, likely that the origin of the afro-montane *Lychnis* and its initial

diversification into two lineages in Eastern Africa may have occurred simultaneously with the expansion of forest during more favourable climatic conditions. Similar trends of pre-Pleistocene origin have been suggested for other plant taxa of the East African rainforest, *Isolona* and *Monodora* (Couvreur *et al.*, 2008) and for the African genus *Coccinia* (Holstein and Renner, 2011).

Interspecific divergence within each of the two ploidy specific lineages commenced around 2.44 Myr [1.38-3.64, 95% HPD, 'section *Uebelinia*'] and 1.28 Myr [0.58-2.13, 95% HPD, 'section *Trigynuebelinia*']. These estimates clearly fit well in time with the well-known periods of aridification in East Africa, which intensified near 2.8, 1.7, and 1.0 Myr ago (Cane and Molnar, 2001; deMenocal, 2004; Bobe, 2006; Sepulchre *et al.*, 2006). Following aridification in response to glaciations in the northern hemisphere, the once widely distributed ancestral lineages of 'section *Uebelinia*' and 'section *Trigynuebelinia'* may have become fragmented and isolated in different places and ultimately diversified into two lineages. Accordingly, climatically driven ecological fragmentation during the cold and dry period of the Pleistocene (deMenocal, 1995) may appear to be a driving force promoting speciation in the afro-montane *Lychnis*. Such a hypothesis would have been most parsimonious if the 95% HPD divergence time estimates within each 'section' had overlapped with the narrow time window of the known aridification periods in Eastern Africa. However, the 95% HPD divergence time estimates [1.38-3.64 Myr 'section *Uebelinia*', and 0.58-2.13 Myr 'section *Trigynuebelinia'*] included the time period when more humid interglacial periods occurred in Africa from

about 2.7-2.5, 1.9-1.7 and 1.1-0.9 Myr ago (Trauth *et al.*, 2005). This may suggest that the afro-montane *Lychnis* started diversifying during the warm and moist period, probably at the end of early-Pliocene around 3.64 Myr ago and intensified during the dry and cold glacial period as well as during the warm and humid interglacial periods until mid-Pleistocene. The inferred distributions of divergence times in our analysis showed that early-Pliocene diversification is equally likely as mid-Pleistocene diversification, and thus suggests that speciation within the afro-montane *Lychnis* may not have been solely triggered by climatic oscillations of the Plio-Pleistocene period.

Alternatively, it has been suggested that some parts of the Eastern African region were under direct climatic influence from the Indian Ocean and that most of the mountain areas had a stable climate despite the climatic oscillations of the Pleistocene that shaped the distribution of many species (Fjeldså and Lovett, 1997, Hewitt, 2000 #79). Ecological diversification, unlike climatically induced speciation, may thus play a major role in promoting speciation in areas that remained climatically stable during the Pleistocene (Fjeldså *et al.*, 1997; Fjeldså and Lovett, 1997; Tolley *et al.*, 2011). Given the overlapping geographical distribution of the afro-montane *Lychnis* in Eastern Africa to those areas which are assumed to be under the direct influence from Indian Ocean, it seems likely that diversification within the afro-montane *Lychnis* may also have been driven by ecological fragmentation in climatically stable areas. Nevertheless, based on the results from our study, we cannot rule out climatically induced speciation of the afro-montane *Lychnis* during the Pleistocene, but rather we acknowledge both mechanisms of speciation may have played a role in different time periods.

The co-occurrence of recently derived and old relict species in climatically stable montane areas may suggest that these areas served both as centres of rapid speciation as well as areas of species persistence (Fjeldså and Lovett, 1997; Fjeldså and Bowie, 2008). Whereas our time estimate for the origin of the afro-montane *Lychnis* evidently predates the Pleistocene period, a previous study by Koch et al. (2006) suggested a rather young Pleistocene origin of the afro-alpine *A. alpina*. Consequently, the ages of the present-day species and their actual time of colonization in these mountains may vary significantly over a wide temporal scale for different plant groups. The spatial coincidence of both young and old relic plant species, such as *A. alpina* and the afro-montane *Lychnis* in the contemporary afro-alpine flora may support that montane areas in general, and the afro-alpine area in particular, represent tropical 'biodiversity hotspots' as suggested previously (Fjeldså 1994; Fjeldså and Lovett, 1997).

The phylogenetic inference of the afro-montane *Lychnis* and its sister group taxa in the present study is largely in agreement with Popp *et al*. (2008), except that the sister relationship between the Eurasian *L. flos-cuculi* and the Moroccan endemic *L. lagrangei* is moderately supported in our *mat*K dataset (Fig. 9b) as well as in our multilocus species tree (Fig. 8). Furthermore, the phylogenetic analyses of the two datasets provided strong support for a sister relationship between the afro-montane *Lychnis* and the *L. flos-cuculi-L. lagrangei* group. The improved resolution of a sister relationship among the aforementioned species group from this study may be partly attributed to the inclusion of the *mat*K dataset in addition to what Popp *et al*. (2008) used in their previous study.

The species tree largely corroborates the phylogeny inferred by Popp *et al*. (2008) where the afro-montane *Lychnis* form a monophyletic group consisting of one diploid and one tetraploid lineage (Fig. 8). The tetraploid lineage consists of only two species, *L. crassifolia* restricted to Mt. Kenya and Mt. Aberdare in Kenya and Mt. Morogoro in Tanzania, and *L. kiwuensis*, which is more widely distributed in Ethiopia and the Western Rift Mountains (Fig. 1b & Fig. 8). The diploid lineage consists of two local endemics, the Ethiopian *L. scottii* and *L. rotundifolia*, which is known from Mt. Kenya, Mt. Aberdare, and Mt. Kilimanjaro, and the widespread *L. kigesiensis* (Ethiopia and the Western Rift Mountains in Uganda; Fig 1b & Fig. 8). The two distinct lineages identified as *L. abyssinica* differing in genome size (Popp *et al*., 2008) also belong to the diploid lineage. Popp *et al*. (2008) suggested hybridization as one possibility to explain the incongruence between plastid and nuclear DNA gene trees in the position of the large genome size *L. abyssinica* also shown in our multilocus species tree (*L. abyssinica* LGS; Fig. 8). To avoid violating the assumption of non-reticulate evolution in *BEAST, we only included plastid data for the large genome size *L. abyssinica*, which together with *L. scotti* formed a sister group to *L. kigesiensis* and the smaller genome size *L. abyssinica* (Fig. 8). Whereas the smaller genome size *L. abyssinica* (*L. abyssinica* SGS; Fig. 8) is known from Ethiopia and Cameroon, the large genome size *L. abyssinica* is only known from Ethiopia. Unfortunately, no material could be obtained from Mt. Elgon or the Western Rift Mountains and it is therefore not possible to test the biogeographical hypotheses of Popp *et al*. (2008).

A sister group relationship between the afro-montane *Lychnis* and the Moroccan endemic *L. lagrangei* is rejected in our analyses, thus corroborating the hypothesis of at least two African colonizations by *Lychnis* in Popp *et al.* (2008). However, the position of *L. lagrangei* differs between the analyses of Popp *et al.* (2008) and the present study. Whereas the former inferred *L. lagrangei* as a sister to the afro-montane *Lychnis* and *L. flos-cuculi*, the present study resolves *L. lagrangei* and *L. flos-cuculi* as the sister group to the afro-montane *Lychnis*. One explanation may be that the present study includes more data, both with respect to individuals and loci. This is particularly true for *L. lagrangei* which was represented by short and fragmented sequences, which may have caused phylogenetic artifacts. It has also been demonstrated that phylogenetic analysis of concatenated sequence data can be misleading, in particular if the internal branches are short (Kubatko and Degnan, 2007). Popp *et al.* (2008) used maximum parsimony to analyse a concatenated dataset consisting of sequences from both plastid and nuclear DNA regions, whereas the present study used a coalescent-based approach that may be more robust to incomplete lineage sorting. Both hypotheses, however, reject a single colonization of Africa. The time estimates from our multilocus species tree show that colonization by the Moroccan endemic *L. lagrangei* occurred at around 4.21 Myr ago [2.39-6.32, 95% HPD] (Node B; Fig.8, and Table 6) while that by the other the afro-montane *Lychnis* occurred at around 4 Myr ago [2.28-5.8, 95% HPD] (Node C, Fig. 8 and Table 6).

5.2 Confusingly similar morphology, ecology and distribution, but different history in the afro-alpine 'Sky Islands': the heathers *Erica arborea* and *E. trimera*

Based on new extensive field sampling in most of the high mountains of Ethiopia and East Africa, we have shown that although the two heathers *Erica arborea* and *E. trimera* are closely related and easily confused in the field, they are distinctly divergent at AFLP loci as well as at plastid DNA loci. We have also demonstrated that their genetic structuring and inferred histories are conspicuously different although both taxa have very tiny seeds and potentially are capable of long-distant dispersal by wind. They have a similar distribution pattern and ecology in the afro-alpine mountain systems and do frequently co-occur at the same sites. Both of them do, however, harbor low levels of overall AFLP diversity suggesting similar histories in terms of loss of diversity by bottlenecking caused by isolation in small refugial populations during unfavorable climatic periods.

In this study, we initially addressed whether the genetic variation of these two species is geographically structured, consistent with long-term isolation in different mountain areas, and/or whether their potentially large capacity for long-distant wind dispersal has resulted in weak genetic structuring among mountains. We found evidence for both of these main alternatives. *Erica trimera* consisted of three distinct AFLP groups which also had different plastid DNA haplotypes and different geographic distributions, suggesting long-term isolation in different areas, with at least one situated in Ethiopia and two in East

Africa (Figs. 10, 18, & 19). In marked contrast to the main pattern observed in *E. trimera*, *E. arborea* showed virtually no geographic structuring at AFLP loci and only a single, widespread plastid DNA haplotype, suggesting that this species recently colonized the entire afro-alpine region from a single source population. This source population of *E. arborea* seems to be situated in Ethiopia as inferred from the distinctly higher genetic diversity and distinctiveness observed in this area (Fig. 14).

It is difficult to pinpoint ecologically or historically based differences between the two species that can fully account for their conspicuously different patterns of genetic structuring in the afro-alpine region. However, part of the explanation may be found in their different altitudinal preferences, which led us to predict more distinct structuring in *E. trimera*. Since this species is restricted to the upper part of the ericaceous belt on the transition to the afro-alpine proper, occasionally occurring also above this zone but not extending downwards into the afro-montane forest, it has probably relied on intermountain colonization solely via long-distance dispersal. It is likely that dispersal of *E. arborea*, on the other hand, which has a larger altitudinal span in the mountains and extends further down into the afro-montane forest zone, also may have been facilitated by formation of afro-montane forests bridging neighbouring mountains during the warm and humid interglacial, as invoked for the forest species *Lobelia giberroa* (Kebede *et al*., 2007). Gradual expansion of *E. arborea* via recent montane forest bridges, possibly early in the present interglacial, may in particular have contributed to the genetic connectivity observed among some of the mountains in Ethiopia. Nevertheless, forest bridges cannot explain the

connectivity observed across the Great Rift Valley in Ethiopia, nor between the Ethiopian and the East African mountains, which must be explained by long-distant dispersal**.** Thus, although both *E. arborea* and *E. trimera* have small wind-dispersed seeds, it is possible that the former nevertheless has higher long-distance dispersal ability than the latter.

Our estimates of the total gene diversity pooled over all genotyped individuals of each species indicated very low levels of diversity in both of them (0.1181 for *E. arborea* and 0.1282 for *E. trimera*). In her survey of reported intra-specific genetic diversity estimates for plants Nybom (2004) indicated that 0.23 as the average diversity value based on AFLP markers and thus identified that the upper and lower limits of AFLP diversity in plants often ranged between 0.31 and 0.15, respectively. The highest level of gene diversity was found for long-lived perennial (0.25), outcrossing (0.27), and late successional taxa (0.30), whereas lowest level of diversity was found in the annual (0.13), selfing (0.12), and early successional taxa (0.17). The two *Erica* species, which both are long-lived perennials and most likely outcrossing species, thus displayed much lower level of gene diversity than was expected according to their life-history traits. Their low levels of diversity hence is best explained by survival through climatically unfavorable periods in small refugial populations where genetic drift reduced their diversity and further reduction by founder-effect during long-distant population expansion in more favorable climatic periods.

Notably, the Mt Elgon populations of both species displayed particularly low level of gene diversity (D = 0.0528 for *E. arborea* and D = 0.0676 for *E.*

trimera). Both species could have recently colonized this mountain via long-distant dispersal of a few founder individuals. Such an explanation would be supported if both species harboured relatively few numbers of private markers as it is expected in recently diverged populations. However, genetic distinctiveness of Mt Elgon populations for both species suggested different histories wherein the DW value in *E. arborea* population was markedly highest (1.36) and in *E. trimera* was above the total average value (0.83; Table 9). It is therefore likely that Mt Elgon populations of both species, once founded by few individuals, may have remained isolated and survived at least several unfavorable climatic periods in small populations, and that gene flow has been minimal between populations in different mountains. Such a hypothesis fits well to the observed low diversity and relatively high distinctiveness detected in Mt. Elgon populations of both species. This is also further supported by plastid DNA sequence data in which the Mt Elgon population of *E. trimera* harboured two private plastid DNA haplotypes, which is less likely in recently diverged populations (Fig. 10 & 11).

In *E. trimera*, both the PCoA and STRUCTURE analyses of the AFLP data provided evidence for presence of three or four distinct genetic entities with different geographic distributions, and these were also characterized by different plastid DNA haplotypes. These entities correspond to some extent to previously recognized subspecies (Hedberg and Hedberg, 2003; Beentje, 2006; Dorr, 2006). Hedberg (1957) analyzed both quantitative and qualitative morphological characters to group the populations of *E. trimera* (as *Philippia trimera*) from different mountains or mountain groups into six subspecies, of which several

had previously been recognized as separate species. He regarded the nominal subspecies, subsp. *trimera*, to be endemic to Mt Ruwenzori, from where we had no material available for analysis. Our AFLP-based 'Ethiopia' and 'Aberdare-Kenya' groups correspond to his subsp. *abyssinica* and subsp. *keniensis*, respectively. Subspecies *abyssinica* was distinguished from all other subspecies, which have glandless corollas and usually glandless calyces, by having numerous glands on the corolla and calyx. Hedberg's conclusion that all Ethiopian populations belong to a single subspecies is further supported by our AFLP data. The plants from different Ethiopian mountains were to a large degree intermingled in our PCoA analysis (Fig. 3c), and the two STRUCTURE groups identified in this region were much less divergent than the other groups.

The remaining subspecies (i.e. all except subsp. *abyssinica*) were mainly distinguished by the amount and type of pubescence on the leaves and stems and partly by the size of floral parts, but Hedberg (1957) emphasized that he found only very small differences in measurable characters to differentiate them. We nevertheless found subsp. *keniensis* to correspond to a distinct genetic group, although the morphological characteristics given for this subspecies by Hedberg (1957) also seem very minor. We also found that the populations from Mt Aberdare, which is situated very close to Mt Kenya and from which Hedberg (1957) had no material available, most likely also belong to subsp. *keniensis*, although they are slightly separated from the Mt Kenya populations in our PCoA analysis (Fig 3c).

The third genetic group recognized in *E. trimera* based on our AFLP analysis, the 'Elgon-Kilimanjaro-Meru' group, thus corresponds to all three remaining subspecies recognized by Hedberg (1957) in his morphological analysis. These subspecies were described as single-mountain endemics (subsp. *elgonensis* from Mt Elgon and subspecies *kilimanjarica* from Mt Kilimanjaro) or as endemic to two neighbouring mountains (subsp. *jaegeri* from Mt Meru and Loolmalasin). It was not surprising that we found two of these to belong to the same genetic group, as they occur on very closely situated mountains (ssp. *kilimanjarica* and ssp. *jaegeri*), and the plants from these mountains were intermingled in the PCoA analysis (Fig. 3c). This finding is in agreement with the merging of these two subspecies by Dorr (2006). More surprisingly, however, the Mt Elgon populations also belonged to this genetic group; although this mountain is situated much further away (Fig. 1d-e). The Elgon populations were however somewhat divergent in the PCoA analysis (Fig. 3c).

To summarize, we found some, but not perfect, correspondence between the degree of genetic differentiation, the degree of reported morphological (subspecies) differentiation, and geographic separation of the mountains, suggesting that morphological and genetic differentiation within *E. trimera* may proceed at different speeds and that the overall patterns are obscured by occasional long-distant colonizations. Further studies combining morphological and genetic data and inclusion of the type locality of *E. trimera* (Mt Ruwenzori) are needed to obtain a refined treatment of its intraspecific taxonomy.

5.3 Introgressive hybridization and phylogeographic pattern in the afro-alpine *Carex monostachya* and *C. runssoroensis* (Cyperaceae) inferred from AFLP

Our multilocus assignment tests provided evidence for the presence of hybridization between the afro-alpine *C. monostachya* and *C. runssoroensis*. Possible hybridization between the two species has been suggested previously by Hedberg (1957). However, the distinguishing morphological characters between the two parental species were too continuous to clearly identify hybrids. Although our sampling did not cover all distributional areas of *C. runssoroensis*, particularly in DR Congo and Rwanda, the present molecular study clearly indicates that all the analysed individuals from Mt. Elgon are of hybrid origin. Further assignment of all individuals into six different hybrid categories showed the existence of hybrids beyond the first generation, indicating introgressive hybridization. Based on the fact that all investigated individuals from Mt. Elgon are of hybrid origin beyond the first generation, it is likely that hybrid fitness in the F1 generation was high enough to promote repeated backcrossing with the parental species and thus leading to introgressive hybridization. This is in agreement with previous reports of genetic introgression in *Carex* (Reznicek, 1990; Choler *et al.*, 2003; Korpelainen *et al.*, 2010).

Although hybridization appears to be geographically localized to Mt. Elgon and that Mt. Elgon may represent a hybrid zone, genetic introgression beyond Mt. Elgon cannot be ruled out based on our analysis. This is shown by the rejection

of the assignment for two of the *C. monostachya* collections from Kilimanjaro (TZ0046_4, and TZ0303_2) into their own source species, and the significant allocation of a single individual plant (TZ0303_2) into the F1B1 hybrid class. The highest proportion of introgressive hybridization (75%; of which F1B1= 50%, F1B2 = 25%; Fig. 25a & b) among the different hybrid classes may indicate that introgressive hybridization is common and that hybrids between *C. monostachya* and *C. runssoroensis* are viable, fertile, and even capable of dispersing beyond the hybrid zone and backcrossing with the parental species, such as the *C. monostachya* population on Kilimanjaro.

The comparisons of the overall gene diversity between the two species, after excluding the hybrid populations from Mt. Elgon, show that populations of *C. monostachya* in the Eastern African region are genetically more diverse than *C. runssoroensis*. Inference of the phylogeographic history of *C. runssoroensis* populations appears not to be possible due to our limited samples from two mountains, Aberdare and Ruwenzori. However, our results of genetic mixture analyses and PCoA for *C. runssoroensis* clearly demonstrate that the two mountains harbour genetically very distinct populations. The first axis of the PCoA which extracted 57% of the variance and the non-hierarchical AMOVAs which allocated almost twofold of the variation to differentiation among-mountain/population (67.5%; Table 16) suggest an old divergence between the two mountain populations. Estimates for the average gene diversity and distinctiveness in the two mountain populations provided contrasting patterns. Populations in Ruwenzori appeared to be genetically less distinct and twice more variable (D = 0.1183, DW = 3.42; Table 15) than populations on Mt.

Aberdare (*D* = 0.0662, DW = 4.11), which on the other hand were more distinct. The geographical distribution of the two populations on each side of the Rift Valley and the contrasting patterns of genetic diversity and distinctiveness detected between the two mountain populations may suggest that they have different histories and that populations on Mt. Aberdare were isolated and survived the unfavourable climatic periods as small populations. This needs to be further tested with both molecular and morphological studies based on more intensive sampling throughout the species entire distributional ranges in Central and East Africa.

Hedberg (1957) and Verdcourt (2010) recognized a variety, var. *aberdarensis*, within *C. runssoroensis* and it is possible that our populations from Mt. Aberdare represent this variety. Var. *runssoroensis* is characterized by its thicker culms and brown to blackish margins of the pistillate scale, whereas var. *aberdarensis* has thinner culms and a hyaline margin of its pistillate scale (Gehrke, 2011). Detection of such morphological variation was not possible in the present study because of the few and poorly developed specimens available. Hedberg (1957) suggested that hybridization between *C. monostachya* and *C. runssoroensis* may have resulted in a hybrid, described as var. *aberdarensis*. Var. *runssoroensis* is distributed on the Virunga volcanoes, Ruwenzori, and Mt. Elgon, while var. *aberdarensis* is endemic to Mt. Aberdare and Kenya. Our multilocus assignment test, however, provided no evidence for the presence of hybrid individuals in either of these two mountain populations. We thus found no evidence for hybrid origin of var. *aberdarensis*. We rather show that

populations from Mt. Elgon, which were supposedly representing var. *runssoroensis,* are of hybrid origin and the result of introgressive hybridization.

Our analyses for *C. monostachya* clearly show the existence of two main lineages distributed on different sides of the Rift Valley. The lineages on the western side of the Rift Valley (Western-rift lineage) appears to be geographically localized and only occurs in Simen. The second lineage from the eastern side of Valley (Eastern-rift lineage) is widely distributed throughout the other mountains, i.e., Mt. Kilimanjaro, Mt. Aberdare, and Mt. Kenya in East Africa, and Bale in Ethiopia. These two lineages may represent refugial populations that survived harsh climatic conditions during the Pleistocene in two isolated areas on opposite sides of the Rift Valley. During the early period of the warm and humid interglacial, contraction of the afro-alpine habitat may have led a once widely distributed population of *C. monostachya* in Eastern Africa to have become fragmented and isolated into two separate refugia populations, one located in Ethiopia (Western-rift lineage) and the other in East Africa (Eastern-rift lineage). Accordingly, the long term isolation on the opposite sides of the Valley followed by genetic differentiation may have resulted in the formation of these two distinct lineages.

Additional PCoA and genetic mixture analyses within the Eastern-rift lineage provided evidence for the presence of two subgroups; one including Mt. Kilimanjaro and Mt. Aberdare, and the other comprising Bale and Mt. Kenya. Within the Eastern-rift lineage a distinct geographical trend was observed for both gene diversity and distinctivness, with high values on Mt. Aberdare and

Mt. Kilimanjaro (one of the subgroups) and the lowest value for Bale (Table 12). The highest level of gene diversity and distinctiveness observed in Mt. Kilimanjaro and Aberdare populations may suggest that the source population for the Eastern-lineage was located in this area. Alternatively, the high value of gene diversity observed in Mt. Kilimanjaro may be explained by introgressive hybridization, detected with the allocation of one *C. monostachya* individual (TZ0303_2) to an F1B1 hybrid class (Fig. 25b). We also found evidence for recent colonization of the Bale mountains, not from populations in the closer Mt. Simen, but from East African populations, probably from Mt. Kenya, via long-distance dispersal.

Interestingly, the substructuring of genetic variation within the Eastern-rift lineage show that geographically distant populations of *C. monostachya* (such as those on Mt. Kilimanjaro and Mt. Aberdare) are genetically more similar to each other than to any of the populations occurring on adjacent mountains (such as Mt. Kenya). It seems likely that the cumulative effects of gradual dispersal among populations on adjacent mountains as well as recent long-distance seed dispersal from East Africa, particularly from Kenya to Bale in Ethiopia, played a major role in shaping the present geographical distribution of genetic variation among populations of *C. monostachya* in Eastern Africa. It is also possible that dispersal in two ways (gradual and long-distance dispersal) may have occurred independently and at different times in the past. During the early period of the last glaciation, range expansion in the afro-alpine habitats was probably accompanied by gradual seed dispersal between adjacent mountains which may have also contributed to the colonization of Mt. Kenya by a few founder

individuals from the source population, perhaps located in the adjacent mountain of Mt. Aberdare. Likewise, the genetic similarity detected between populations from Mt. Kenya and Bale may be explained by recent long-distance seed dispersal from Mt. Kenya to Bale.

5.4 Colonization of the Ethiopian Highlands by two lineages of the afro-alpine *Lysimachia serpens* (Primulaceae) from Elgon and Kilimanjaro

The geographical structuring of the genetic variation in *Lysimachia serpens* across the different mountains comprising the afro-alpine region was found to be relatively weak. Nevertheless two main lineages were identified both in the PCoA as well as in the genetic mixture analysis. One of the lineages, structure-group1, comprises all populations from Mt. Aberdare, Mt. Elgon, and some populations from Bale, Simen, and Mt. Kilimanjaro. Structure-group 2 is composed of a single population from Mt. Kilimanjaro (Kilimanjaro-2) and populations or individuals of populations from Bale (Bale-2) and Simen (Simen-2). Despite the overlapping geographic distribution of the two lineages in three of the mountain systems, 36% of the total variation is due to divergence between the two lineages (P = 0.0244; Table 19), indicating limited gene flow. Both the patterns of genetic differentiation and geographical distribution of the two lineages may indicate long-term isolation in separate places followed by limited gene flow. The highest value of gene diversity observed within the Structure-group 1 lineage (D = 0.1185) was found on Mt. Elgon and was almost twice that of the highest value estimated within the Structure-group 2 lineage (D = 0.0547, Kilimanjaro-2 or TZ0013; Table 18) on Mt. Kilimanjaro. It is possible that isolation of these two populations during periods of different population size may explain their observed difference in genetic diversity. It is also likely that the currently overlapping geographical distribution of the two lineages in the three mountains Bale, Simen and Mt. Kilimanjaro reflects recent long-

distance dispersal events, occurring after initial isolation of the two lineages in different geographic areas.

Structure-group 1: High levels of genetic diversity (*D* = 0.1185; Table 18) as well as distinctiveness (DW = 5.25; Table 18) were observed in populations from Mt. Elgon, suggesting isolation with relatively large population size, and that the source population for this lineage was probably located on Mt. Elgon. Of the total genetic variation recovered within this lineage, genetic differentiation among regional-populations (26%) was three times less than the variation within regional-populations (74%; Table 19), indicating frequent gene flow among and between populations on different mountains in Eastern Africa. Interestingly, the pattern of genetic diversity and distinctiveness within Structure-group 1 lineage shows a of gradual decrease for both estimates with increasing distance from the source population on Mt. Elgon, both towards the Ethiopian highlands (Bale and Simen) in the North, and Mt. Kilimanjaro, through the Mt. Aberdare, in the Southeast.

At the onset of the northern hemisphere glaciations during the early period of the Pleistocene afro-alpine habitats were larger than they are today. Hence, ancient long-distance seed dispersal from Mt. Elgon towards the North may have founded the Bale population in the Ethiopian Highlands. Then gradual dispersal of a few immigrants from Bale across the rift Valley, probably by long-distance dispersal in Ethiopia, may have founded the Simen populations. The comparable levels of genetic distinctiveness values observed in the Bale (DW = 4.91; Fig 36 & Table 18) and Simen (DW = 4.69) populations as well as

the gradual decrease of the gene diversity (D = 0.0831 in Bale, and D = 0.0679, in Simen) may suggest dispersal from Mt. Elgon towards the Ethiopian Highlands occurred in the form of leading age colonization (Hewitt, 1996; Petit *et al.*, 2002). The colonization of Mt. Kilimanjaro through Mt. Aberdare by a few immigrant individuals from the source population in Mt. Elgon, would be expected to result in a similar pattern of gradual decrease of both gene diversity and distinctivity, suggesting a similar pattern of leading edge colonization. The observed distinctiveness values in Mt. Aberdare (DW = 4.11; Fig. 36 & Table 18) and Mt. Kilimanjaro (DW = 3.96) were below the average value of the total populations and also differed greatly from both Bale and Simen (which have values above the average), indicating that after colonization the Ethiopian populations remained relatively isolated from the East African populations on Mt. Kilimanjaro and Aberdare.

Structure-group 2: Despite the relatively low level of gene diversity (D = 0.0.547), considerably high level of genetic distinctiveness was detected in the Kilimanjaro-2 population of Structure-group 2 lineage (DW = 7.09; Table 18). This could indicate the long-term isolation of the Kilimanjaro-2 population on Mt. Kilimanjaro. The observed discrepancy in level of genetic divergence, within Structure-group 2 lineage, (such as, DW = 7.09 in Kilimanjaro-2, DW = 4.47, 3.02, in Bale-2 and Simen-2 populations respectively; Table 18), possibly indicates that the source population of this lineage was located on Mt. Kilimanjaro. Furthermore, the three regional populations within this lineage displayed a high level of genetic differentiation (58%; Table 19), indicating long-term isolation and limited gene flow among the regional mountains.

Moreover, the lowest level of total regional diversity was observed in all three of these regional populations (Table 18). The general pattern of a reduced regional genetic diversity in all three regional populations as well as the discrepancy in the level of genetic distinctiveness between the Kilimanjaro-2 population, and Bale-2 and Simen-2 populations may suggest genetic bottlenecking due to isolation in a small populations and/or founder effects associated with recolonization. Both scenarios may explain the reduced level of gene diversity among regional populations within Structure-group 2. Consequently, low diversity and high distinctiveness in Kilimanjaro-2 populations likely resulted from genetic bottlenecking in a small isolated refugial population. Likewise, the low levels of genetic diversity and distinctiveness detected in Bale-2 and Simen-2 populations may have resulted from genetic bottlenecking due to founder effects associated with recent recolonization.

Taxonomic implication: Taylor (1955) used opposite and alternate leaf arrangements as a distinguishing morphological character for recognizing two subspecies of *Lysimachia serpens;* subsp. *serpens* and *meyeri-yohannis*, respectively. The two subspecies were also reported to have allopatric distributions across the range of the species, extending from Sudan to Southern Zimbabwe in East Africa. Whereas subsp. *serpens* is distributed in Sudan, Ethiopia and Southern Zimbabwe, subsp. *meyeri-yohannis* is distributed in the mountains of Eastern Africa in Kenya, Tanzania, and Uganda (Taylor, 1955). Despite the huge sampling efforts made in this study to include samples from most of the distributional areas in East Africa, the result from our molecular

study provides no evidence either for the presence of morphologically distinct genetic groups or for differentiation between samples from the two suggested allopatric distributional ranges of the two subspecies. Recently, morphological studies of the two subspecies yielded similar results (Manns and Anderberg, 2009). The two main lineages identified in our study appear to have overlapping geographical distributions in at least three of the mountains. For example, different populations from Kilimanjaro represent the two lineages and some individuals from local populations in Bale (ET00781) and Simen (ET0550, and ET0629) were split between the two lineages.

Our attempt to assign our specimens into two subspecies based on Taylor's (1955) leaf arrangement character was unsuccessful, in agreement with the report of Manns & Anderberg (2011). Although seemingly opposite, the arrangement of the leaves always turns out to be alternate when closely examined (Manns & Anderberg, 2011). We nevertheless identified to two different genetic lineages within the species, probably reflecting different histories. The sympatric distribution of the two lineages in parts of their range may be caused by recent dispersal after initial isolation in different geographical areas.

5.5 Confusing geographical structure in the Afro-alpine *Dicrocephala chrysanthemifolia* (Asteraceae): initial divergence confounded by recent gene flow

Our results clearly show the presence of two lineages in *D. chrysanthemifolia* in the afro-alpine region of Eastern Africa. Whereas one of the lineages solely comprised four of the populations from Bale (Bale-2), another population from Bale formed together with the populations from Simen and the mountains in Kenya, Tanzania and Uganda another lineage (eastern Africa-lineage). Distinct genetic differentiation (35 %; Table 22) was detected between the two lineages. This pattern may reflect initial differentiation of populations in the Bale mountains vs other mountains, followed by recent immigration of the other lineage into the Bale mountains.

We found high and similar levels of genetic distinctivity and diversity in the Simen mountains (DW = 5.6, D = 0.0760; Fig. 41 & Table 21) and in the Bale-2 lineage (DW = 4.05, D = 0.0642), indicating long-term isolation on different sides of the Rift Valley and corroborating the hypotheses that the Rift can represent a major barrier to gene flow. On the other hand, the low levels of diversity and distinctiveness observed in the Bale population of Structure-group 1 (D = 0.0459; DW = 2.23) support that this widespread lineage recently colonized the Bale mountains. Our results may also indicate long-distance dispersal in a stepping stone manner from Simen to Mt. Elgon and then to Mt. Muhavura, explaining the observed pattern of gradual decrease in gene diversity among the three mountain-populations on the western side of the Valley. This may have been followed by colonization of Mt. Meru and Kilimanjaro by a few

immigrants from Mt. Elgon via long-distance dispersal, and subsequently by descendants from Mt. Meru/Kilimanjaro to Mt. Aberdare. More recently, migration from East Africa, probably from Mt. Aberdare, resulted in re-colonization of the Bale mountains in Ethiopia via long-distance dispersal, thus explaining the observed pattern of gradual decrease in level of diversity and distinctiveness during colonization of the East Africa mountains.

Taxonomic implications: A relatively recent split of the ancestral population of *D. chrysanthemifolia* in the Ethiopian highland resulted in the formation of two genetically differentiated but morphologically similar lineages with different distribution ranges in East Africa. Several authors in the past have suggested various taxonomic treatments, i.e., whether to recognize two separate species or variety/subspecies within the widely distributed *D. chrysanthemifolia* in the mountains of Eastern Africa (Hedberg, 1957; Beentje, 2002; Mesfin Tadesse, 2004). Hedberg (1957) and Tadesse (2004) recognized two closely related but distinct species, *D. chrysanthemifolia* s.str. and *D. alpina* in East Africa based on morphological characters and geographical distributions. *D. chrysanthemifolia* was described as 30-100(-150) cm high from mid-altitude areas whereas *D. alpina* was described as 2-10(-20) cm high and dominating the afro-alpine proper habitat (Mesfin Tadesse, 2004). Whereas *Dicrocephala chrysanthemifolia* s.str. was reported as widely distributed in Africa, i.e., in Ethiopia, Kenya, Tanzania, Uganda, Somalia, Rwanda, Sudan, Eritrea, in the west to Guinea and south to Malawi and Mozambique and in Madagascar, *D. alpina* was reported to have a more restricted distribution in Ethiopia, Kenya, Tanzania, Uganda, and DR Congo (Mesfin Tadesse, 2004).

Beentje (2002), however, reported that the differences in the morphological characters between the two species were few and gradual and suggested to reduce *D. alpina* to a variety of *D. chrysanthemifolia.* Some of the specimens referred to as *D. alpina* from Mt. Elgon were almost as tall as *D. chrysanthemifolia*, and such continuity in the morphological characters also made Hedberg (1957) acknowledge a varietal level of *D. alpina*. This is in agreement with our observations of continuous morphological variation in Eastern Africa. The two lineages observed in our genetic dataset do not reflect the previously reported morphological or geographical split. We therefore suggest that the populations scattered across the mountains of Eastern Africa should be referred to a single species, *D. chrysanthemifolia.*

5.6 Interglacial refugia in the Ethiopian Highlands: the phylogeography of the afro-alpine *Haplocarpha rueppellii* (Asteraceae)

Our analyses clearly distinguished two main genetic lineages in *H. rueppellii* distributed in the Ethiopian Highlands as well as in the mountains of tropical East Africa. The first lineage, Structure-group 1, contains plants from Mt Kilimanjaro, Gara Muleta, and Mt. Kenya, and Structure-group 2 contains plants from Ethiopia (Bale, Choke, and Simen) and other East African mountains (Mt. Aberdare, Mt. Elgon and the KN0788-a population from Mt. Kenya). The high genetic similarity between the populations in Bale and Choke in the Ethiopian highlands suggests frequent gene flow among these two mountains, in spite of the fact that they are located on opposite sides of the Rift valley. Thus, it appears that the Rift Valley can represent a major obstacle to gene flow in some taxa, but not in others.

The populations from Mt. Kenya were genetically different and contained low levels of diversity, which may be due to genetic bottlenecks caused in small refugial populations or resulting from founder effects during leading edge colonisation (Hewitt, 1996; Petit *et al.*, 2002). The former would appear to be a plausible explanation if the populations in Mt. Kenya harboured rare markers as expected in long-term isolated populations. However, their low DW values rather support recent colonization of Mt. Kenya by two independent lineages originating from source populations located in the Ethiopian Highlands.

Within Structure-group 1, three geographically structured subgroups were identified in separate PCoA and STRUCTURE analyses. The three subgroups included the population from Gara Muleta, and those from the two mountains (Mt. Kilimanjaro and Mt. Kenya) in East Africa (Fig. 43a). Within this lineage, the highest level of diversity was observed in Gara Muleta (0.0620, Fig. 45 & Table 24). The populations in Gara Muleta were also genetically distinct (DW = 4.05), suggesting that Gara Muleta harboured the source population of this lineage. The populations on Mt. Kilimanjaro contained more rare markers than the other populations. It is therefore possible that the contemporary populations in Mt. Kilimanjaro may have been founded relatively early by a few immigrants from Gara Muleta and that they have remained isolated in small populations, allowing genetic drift to cause a reduction in level of genetic diversity.

In a separate PCoA (Fig. 43b) and STRUCTURE analysis, three distinct subgroups were identified: the Bale-Choke subgroup, the Aberdare-Elgon-Kenya788 subgroup and the Simen subgroup. Interestingly, within Structure-group 2, the observed genetic divergence estimate among the three regional populations in Simen, Mt. Aberdare and Mt. Elgon were very similar to each other (DW = 2.73 – 2.77; Fig. 45 & Table 24). Furthermore, genetic diversity and distinctiveness varied greatly among the three subgroups with highest levels observed in the Bale-Choke subgroup. It is therefore likely that recent colonization of the other mountains in the two subgroups occurred from a source refugial population located in the Bale-Choke mountains probably by long-distance dispersal. Our analysis also showed the possibility of the

colonization of Mt. Kenya by two independent lineages. This was indicated by the occurrence of two divergent populations on Mt. Kenya.

The observed genetic differentiation among geographically close populations (e.g., between Mt. Aberdare and Mt. Kenya), and between two divergent populations on Mt. Kenya indicates absence of gene flow in spite of geographical proximity. On the contrary, populations on different sides of the Rift valley (e.g., between Elgon and Aberdare) were not well differentiated genetically, suggesting occasional long-distance dispersal across the Rift. This was also suggested in the AFLP analysis of *Arabis alpina* in which populations from different sides of the Rift valley such as Mt. Aberdare, Mt. Kenya and Mt. Kilimanjaro were genetically similar to plants from Mt. Elgon on the opposite side (Ehrich *et al.*, 2007).

5.7 Conclusions

The afro-alpine flora, which is distributed across the high and isolated mountains of eastern Africa, have recently become a topic of interest for addressing various evolutionary questions. One of the hot topics in the recent past concerned the age of the afro-alpine flora. While some argued that the flora is typically composed of Tertiary relicts, because of the high level of endemism it encompasses, others suggested a rather young flora mainly originating through colonization by afro-temperate elements during the Pleistocene. To our knowledge, the present study is one of the first to use a fossil calibrated relaxed molecular clock to estimate the approximate age of the flora by considering afro-montane *Lychnis* as a model group. A multilocus species tree analysis based on six different genes from nuclear and plastid DNA sequences placed the age of the Afro-montane *Lychnis* between 2.28 and 5.8 Myr, i.e. during the late Tertiary. Our analysis also showed intensive diversification and dispersal among mountains of tropical East Africa during the Plio-Pleistocene period. The present molecular dating analysis thus sheds some light on the age of the afro-alpine flora and also suggests that the contemporary afro-alpine flora in the mountains of eastern African region is complex, harbouring both old and young taxa.

The results from the phylogeographic analysis show a lack of common phylogeographic patterns among the different investigated plant species of the afro-alpine as well as of the ericaceous zone, suggesting that different species have responded individually to the Pleistocene climatic oscillations. The

responses even varied among closely related species from the same vegetation zone, as in the case of the two species of *Erica* and *Carex* from the ericaceous and afro-alpine zones, respectively. Despite the lack of common phylogeographic patterns, all of the investigated species have low levels of overall gene diversity, suggesting genetic bottlenecking in small refugial populations during unfavourable climatic periods and/or founder effects during recent colonizations. The lack of geographical structuring of the genetic variation and the relatively low levels of gene diversity observed in some of the species, particularly in *E. arborea* and *D. chrysanthemifolia,* may indicate recent colonization of the entire afro-alpine region, probably from a single source population.

The non-hierarchical AMOVAs in some cases revealed low levels of genetic differentiation among populations, suggesting higher levels of gene flow than traditionally believed among different mountains of Eastern Africa. In other cases, such as in *C. runssoroensis,* a high level of differentiation (67%) was observed among populations on Mt. Aberdare and Ruwenzori, which represent different branches of the Rift Valley in East Africa. In several cases, our results suggest that long-distance dispersals have occurred recently among mountains.

The present study provides strong evidence for introgressive hybridization between two afro-alpine *Carex* species, *C. monostachya* and *C. runssoroensis*. We showed that all *Carex* populations from Mt. Elgon were of hybrid origin and resulting from backcrossing to the two parental species.

5.8 References

Anderberg, A.A., Manns, U., Källersjö, M. (2007). Phylogeny and floral evolution of the *Lysimachieae* (*Ericales*, *Myrsinaceae*). Evidence from *ndh*F sequence data. *Wildenowia* **37:** 407-421.

Assefa, A., Ehrich, D., Taberlet, P., Nemomissa, S., Brochmann, C. (2007). Pleistocene colonization of afro-alpine 'sky islands' by the arctic-alpine *Arabis alpina*. *Heredity* **99:** 133-142.

Avise, J.C. (2000). *Phylogeography:* The History and Formation of Species. Harvard University Press, Cambridge.

Avise, J.C., Arnold, J., Ball, R.M., Bermingham, E., Lamb, T., Neigel, J.E., Reeb, C.A., Saunders, N.C. (1987). Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Annual Review of Ecology and Systematics* **18:** 489-522.

Baker, B.H., Mohr, P.A., Williams, L.A.J. (1972). Geology of the Eastern Rift System of Africa. *The Geological Society of America. Special Paper* **136:** 67.

Ball, P.W., Reznicek, A.A. (2002). *Carex*. In**:** Editorial Committee (Ed.), *Flora of North America*. Oxford University Press, New York, pp. 254–572.

Bandelt, H.-J., Forster, P., Röhl, A. (1999). Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution* **16:** 37-48.

Beentje, H.J. (2002). Compositae (part 2). In**:** Beentje, H.J. (Ed.), *Flora of tropical East Africa*. Royal Botanic Gardens, Kew (UK), pp. 315–546.

Beentje, H.J. (2006). *Flora of Tropical East Africa-Ericaceae*. Royal Botanical Gardens, Kew (UK).

Bobe, R. (2006). The evolution of arid ecosystems in eastern Africa. *Journal of Arid Environments* **66:** 564-584.

Bonin, A., Bellemain, E., Eidesen, P.B., Pompanon, F., Brochmann, C., Taberlet, P. (2004). How to track and assess genotyping errors in population genetics studies. *Molecular Ecology* **13:** 3261–3273.

Bonnefille, R. (1995). A reassessment of the Plio-Pleistocene pollen record of East Africa. In**:** Vrba, E.S., Denton, G.H., Partridge, T.C., Burckle, L.H. (Eds.), *Paleoclimate and Evolution with Emphasis on Human Origins*. Yale University Press, New Haven, pp. 299-310.

Bonnefille, R., Roeland, J.C., Guiot, J. (1990). Temperature and rainfall estimates for the past 40 000 years in equatorial Africa. *Nature* **346:** 347-349.

Bowie, R.C.K., Fjeldså, J., Hackett, S.J., Bates, J.M., Crowe, T.M. (2006). Coalescent models reveal the relative roles of ancestral polymorphism, vicariance and dispersal in shaping phylogeographical structure of an African montane forest robin. *Molecular Phylogenetics and Evolution* **38:** 171-188.

Brochmann, C., Gabrielsen, T.M., Nordal, I., Landvik, J.Y., Elven, R. (2003). Glacial survival or tabula rasa? The history of North Atlantic biota revisited. *Taxon* **52:** 417-450.

Cane, M.A., Molnar, P. (2001). Closing of the Indonesian seaway as a precursor to east African aridification around 3-4 million years ago. *Nature* **411:** 157-162.

Cerling, T.E., Wang, Y., Quade, J. (1993). Expansion of C4 ecosystems as an indicator of global ecological change in the Late Miocene. *Nature* **361:** 344-345.

Choler, P., Erschbamer, B., Tribsch, A., Gielly, L., Taberlet, P. (2003). Genetic introgression as a potential to widen a species' niche: Insights from the alpine *Carex curvula*. *Proceedings of the National Academy of Science of the United States of America* **101:** 171-176.

Couvreur, T.L.P., Chatrou, L.W., Sosef, M.S.M., Richardson, J.E. (2008). Molecular phylogenetics reveal multiple tertiary vicariance origins of the African rain forest trees. *BMC Evolutionary Biology* **6:** 54.

Coyne, J.A., Orr, H.A. (2004). *Speciation*. Sinauer & Associates, Sunderland, Massachusetts.

deMenocal, P.B. (1995). Plio-Pleistocene African climate. *Science* **270:** 53-59.

deMenocal, P.B. (2004). African Climate Change and Faunal Evolution during the Pliocene-Pleistocene. *Earth and Planetary Science Letters* **220:** 3-24.

Désamoré, A., Laenen, B., Devos, N., Popp, M., Gonzalez-Mancebo, J.M., Carine, M.A. (2011). Out of Africa: north-westwards Pleistocene expansions of the heather *Erica arborea*. *Journal of Biogeography* **38:** 164-176.

Dorr, L.J. (2006). New combinations in *Erica* (Ericaceae: Ericoideae) from the high mountains of East Africa. *Novon* **16:** 56-58.

Drummond, A.J., Ho, S.Y., Phillips, M.J., Rambaut, A. (2006). Relaxed phylogenetics and dating with confidence. *PLoS Biology* **4:** e88.

Drummond, A.J., Rambaut, A. (2007). BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology* **7:** 214.

Duchesne, P., Bernatchez, L. (2002). aflpop: a computer program for simulated and real population allocation, based on AFLP data. *Molecular Ecology Notes* **2:** 380-383.

Ehrich, D. (2006). AFLPdat: a collection of R functions for convenient handling of AFLP data. *Molecular Ecology Notes* **6:** 603-604.

Ehrich, D., Gaudeul, M., Assefa, A., Koch, M.A., Mummenhoff, K., Nemomissa, S., Brochmann, C. (2007). Genetic consequences of Pleistocene range shifts: contrast between the Arctic, the Alps and the East African mountains. *Molecular Ecology* **16:** 2542-2559.

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.

Evans, B.J., Kelley, D.B., Tinsley, R.C., Melnick, D.J., Cannatella, D.C. (2004). A mitochondrial DNA phylogeny of African clawed frogs: phylogeography and implications for polyploid evolution. *Molecular Phylogenetics and Evolution* **33:** 197-213.

Excoffier, L.L.G., Schneider, S. (2005). ARLEQUIN version 3.0: an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online* **1:** 47-50.

Fjeldså , J. (1994). Geographical patterns of relict and young species of birds in Africa and South America and implications for conservation priorities. *Biodiversity and Conservation* **3:** 107-126.

Fjeldså, J., Bowie, R.C.K. (2008). New perspectives on the origin and the diversification of Africa's Forest Avifauna. *African Journal of Ecology* **46:** 235-247.

Fjeldså, J., Ehrlich, D., Lambin, E., Prins, E. (1997). Are biodiversity 'hotspots' correlated with current ecoclimatic stability? A pilot study using the NOAA-AVHR remote sensing data. *Biodiversity and Conservation* **6:** 401-422.

Fjeldså, J., Lovett, J.C. (1997). Geographical patterns of old and young species in African forest biota: The significance of specific montane areas as evolutionary centres. *Biodiversity and Conservation* **6:** 325-346.

Flenley, J. (1979). *The Equatorial Rain Forest: A Geological History*. Butterworth, London.

Friis, I. (1986). Zonation of the forest vegetation on the south slopes of Bale Mountains, south Ethiopia. *Sinet* **9:** 29-44.

Gaudeul, M., Taberlet, P., Till-Bottraud, I. (2000). Genetic diversity in an endangered alpine plant, *Eryngium alpinum* L. (Apiaceae), inferred from amplified length polymorphism markers. *Molecular Ecology* **9:** 1625-1637.

Gehrke, B. (2011). Synopsis of *Carex* (Cyperaceae) from sub-Saharan Africa and Madagascar. *Botanical Journal of the Linnean Society* **166:** 51-59.

Gehrke, B., Linder, H.P. (2009). The scramble for Africa: pan-temperate elements on the African high mountains. *Proceedings of the Royal Society of London, Series B: Biological Sciences* **276:** 2657-2665.

Germishuizen, G., Meyer, N.L. (2003). *Plants of Southern Africa: An Annotated Checklist*. Strelitzia 14 National Botanical Institute, Pretoria.

Gillespie, J.H. (1991). *The causes of molecular evolution*. Oxford University Press, New York, USA.

Gottelli, D., Marino, J., Sillero-Zubiri, C., Funk, S.M. (2004). The effect of the last glacial age on speciation and population genetic structure of the endangered Ethiopian wolf (*Canis simensis*). *Molecular Ecology* **13:** 2275-2286.

Gouy, M., Guindon, S., Gascuel, O. (2010). SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Molecular Biology and Evolution* **27:** 221-224.

Govaerts, R., Koopman, J., Simpson, D., Goetghebeur, P., Wilson, K., Egorova, T., Bruhl, J. (2010). *World checklist of Cyperaceae* The Board of Trustees of the Royal Botanic Gardens, Kew (UK) Available at http://apps.kew.org/wcsp/monocots/.

Griffiths, C.J. (1993). The geological evolution of East Africa. In**:** Lovett, J.C., Wasser, S.K. (Eds.), *Biogeography and Ecology of the Rain Forests of Eastern Africa*. Cambridge University Press, Cambridge, UK, pp. 19-23.

Hamilton, A.C. (1982). *Environmental History of East Africa: A Study of the Quaternary*. Academic Press, London.

Hedberg, I., Hedberg, O. (2003). Ericaceae. In**:** Hedberg, I., Edwards, S., Sileshi, N. (Eds.), *Flora of Ethiopia and Eritrea Volume 4, Part 1. Apiaceae to Dipsaceae*. The National Herbarium, Addis Ababa University, Addis Ababa, pp. 46-48.

Hedberg, O. (1951). Vegetation belts of the East African mountains. *Svensk Botanisk Tidskrift* **45:** 140-202.

Hedberg, O. (1957). Afroalpine vascular plants. a taxonomic revision. *Symbolae Botanicae Upsalienses* **15:** 1-411.

Hedberg, O. (1961). The phytogeographical position of the afroalpine flora. *Recent Advances in Botany* **1:** 914-919.

Hedberg, O. (1965). Afroalpine flora elements. *Webbia* **19:** 519-529.

Hedberg, O. (1969). Evolution and speciation in a tropical high mountain flora. *Botanical Journal of the Linnean Society* **1:** 135-148.

Hedberg, O. (1970). Evolution of the Afroalpine Flora. *Biotropica* **2:** 16-23.

Hedberg, O. (1971). The high mountain flora of the Galama Mountain in Arussi Province, Ethiopia. *Webbia* **26:** 101-128.

Hedges, S.B., Kumar, S. (2003). Genomic clocks and evolutionary timescales. *Trends in Genetics* **19:** 200-206.

Heled, J., Drummond, A.J. (2010). Bayesian inference of species trees from multilocus data. *Molecular Biology and Evolution* **27:** 570-580.

Hewitt, G. (2000). The genetic legacy of the Quaternary ice ages. *Nature* **405:** 907-913.

Hewitt, G.M. (1996). Some genetic consequences of ice ages, and their role in divergence and speciation. *Biological Journal of the Linnean Society* **58:** 247-276.

Hickerson, M.J., Carstens, B.C., Cavender-Bares, J., Crandall, K.A., Graham, C.H., Johnson, J.B., Rissler, L., Victoriano, P.F., Yoder, A.D. (2010). Phylogeography's past, present, and future: 10 years after Avise, 2000. *Molecular Phylogenetics and Evolution* **54:** 291-301.

Ho, S.Y., Phillips, M.J. (2009). Accounting for calibration uncertainty in phylogenetic estimation of evolutionary divergence times. *Systematic Biology* **58:** 367-380.

Holstein, N., Renner, S.S. (2011). A dated phylogeny and collection records reveal repeated biome shifts in the African genus *Coccinia* (Cucurbitaceae). *BMC Evolutionary Biology* **11:** 28.

Huckerby, E., Marchant, R., Oldfield, F. (1972). Identification of fossil seeds of *Erica* and *Calluna* by scanning electron microscopy. *New Phytologist* **71:** 387-392.

Huelsenbeck, J.P., Larget, B., Swofford, D. (2000). A compound Poisson process for relaxing the molecular clock. *Genetics* **154:** 1879-1892.

Huelsenbeck, J.P., Ronquist, F. (2001). MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* **17:** 754-755.

Johansson, M., Rooke, T., Fetene, M., Granstrostrom, A. (2010). Browser selectivity alters post-fire competition between *Erica arborea* and *E. trimera* in the sub-alpine heathlands of Ethiopia. *Plant Ecology* **207:** 149-160.

Jordan, G.J., Macphail, M.K. (2003). A Middle-Late Eocene inflorescence of Caryophyllaceae from Tasmania, Australia. *American Journal of Botany* **90:** 761-768.

Jordan, S., Simon, C., Polhemus, D. (2003). Molecular systematics and adaptive radiation of Hawaii's endemic damselfly genus megalagrion (Odonata: Coenagrionidae). *Systematic Biology* **52:** 89-109.

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.

Knox, E.B. (1993). The conservation status of the giant senecios and giant lobelias in Eastern Africa. *Opera Botanica* **121:** 195-216.

Knox, E.B., Palmer, J.D. (1998). Chloroplast DNA evidence on the origin and radiation of the giant Lobelias in Eastern Africa. *Systematic Botany* **23:** 109-149.

Koch, M.A., Kiefer, C., Ehrich, D., Vogel, J., Brochmann, C., Mummenhoff, K. (2006). Three times out of Asia Minor: the phylogeography of *Arabis alpina* L. (Brassicaceae). *Molecular Ecology* **15:** 825-839.

Korpelainen, H., Virtanen, V., K., K., Väre, H. (2010). Hybridization and introgression in *Carex aquatilis* and *C. paleacea*. *Plant Systematics and Evolution* **287:** 141-151.

Kosman, E. (2003). Nei's gene diversity and the index of average differences are identical measures of diversity within populations. *Plant Pathology* **52:** 533-535.

Kubatko, L.A., Degnan, J.H. (2007). Inconsistency of Phylogenetic Estimates from Concatenated Data under Coalescence. *Systematic Biology* **56:** 17-24.

Li, W.-H., Tanimura, M. (1987). The molecular clock runs more slowly in man than in apes and monkeys. *Nature* **326:** 93-96.

Librado, P., Rozas, J. (2009). DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **25:** 1451-1452.

Lind, E.M., Morrison, M.E.S. (1974). *East African vegetation*. Longman, London.

Lisanework Nigatu and Mesfin Tadesse (1989). An ecological survey of the vegetation of the Harenna forest, Bale, Ethiopia. *Sinet* **12:** 63-93.

Livingstone, D.A. (1962). Age of deglaciation in the Ruwenzori range,Uganda. *Nature* **194:** 859-860.

Logatche, N.A., Beloussov, V.V., Milanovsky, E.E. (1972). East African Rift Development. *Tectonophysics* **15:** 71-81.

Lovett, J.C., Bridson, D.M., Thomas, D.W. (1988). A preliminary list of the moist forest angiosperm flora of the Mwanihana Forest Reserve, Tanzania. *Annals of the Missouri Botanical Garden* **75:** 874-888.

Manns, U., Anderberg, A.A. (2009). New combinations and names in *Lysimachia* (Myrsinaceae) for species of *Anagallis*, *Pelletiera* and *Trientalis*. *Wildenowia* **39:** 49-54.

Manns, U., Anderberg, A.A. (2011). Biogeography of 'tropical *Anagallis'* (Myrsinaceae) inferred from nuclear and plastid DNA sequence data. *Journal of Biogeography* **38:** 950-961.

McGuire, A.F., Kron, K.A. (2005). Phylogenetic relationships of European and African ericas. *International Journal of Plant Sciences* **166:** 311-318.

Mesfin Tadesse (2004). *Asteraceae* (Compositae). In**:** Hedberg, I., Friss, I., Edwards, S. (Eds.), *Flora of Ethiopia and Eritrea Volume 4, Part 2*. The National Herbarium, Addis Ababa University, Addis Ababa.

Messerli, B., Winiger, M. (1992). Climate, environmental change, and resources of the African mountains from the Mediterranean to the equator. *Mountain Research and Development* **12:** 315-336.

Miehe, G., Miehe, S. (1994). *Ericaceous forests and heathlands in the Bale Mountains of South Ethiopia: ecology and man's impact*. Stiftung Walderhaltung in Africa, Hamburg.

Mohammed, M.U., Bonnefille, R. (1998). A late Holocene pollen record from a highland peak at Tamsaa, Bale Mountains, South Ethiopia. *Global and Planetary Change* **16-17:** 121–129.

Mower, J.P., Touzet, P., Gummow, J.S., Delph, L.F., Palmer, J.D. (2007). Extensive variation in synonymous substitution rates in mitochondrial genes of seed plants. *BMC Evolutionary Biology* **7:** 135.

Nei, M., Kumar, S. (2000). *Molecular Evolution and Phylogenetics*. Oxford University Press, New York.

Nei, M., Li, W.H. (1979). Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proceedings of the National Academy of Science of the United States of America* **76:** 5269–5273.

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

Nylander, J.A.A. (2004). MrModeltest v2. Program distributed by the author. *Evolutionary Biology Centre, Uppsala University*.

Oxelman, B., Liden, M., Rabeler, R.K., Popp, M. (2000). A revised generic classification of the tribe Sileneae (Caryophyllaceae). *Nordic Journal of Botany* **20:** 743-748.

Petit, R.J., Brewer, S., Bordacs, S., Burg, K., Cheddadi, R., Coart, E., Cottrell, J., Csaikl, U.M., van Dam, B., Deans, J.D., Espinel, S., Fineschi, S. (2002). Chloroplast DNA variation in European white oaks phylogeography and patterns of diversity based on data from over 2600 populations. *Forest Ecology and Management* **156:** 5-26.

Pickford, M., Senut, B., Mourer-Chauvire, C. (2004). Early Pliocene Tragulidae and peafowls in the Rift Valley, Kenya: evidence for rainforest in East Africa. *Comptes Rendus Palevol* **3:** 179–189.

Pirie, M.D., Oliver, E.G.H., Bellstedt, D.U. (2011). A densely sampled ITS phylogeny of the Cape flagship genus *Erica* L. suggests numerous shifts in floral macro-morphology. *Molecular Phylogenetics and Evolution* **61:** 593-601.

Popp, M. (2008). New combinations in *Lychnis* (Caryophyllaceae) from Africa. *Novon* **18:** 99-100.

Popp, M., Gizaw, A., Nemomissa, S., Suda, J., Brochmann, C. (2008). Colonization and diversification in the African 'sky islands' by Eurasian *Lychnis* L. (Caryophyllaceae). *Journal of Biogeography* **35:** 1016-1029.

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

Quezel, P. (1978). Analysis of the flora of Mediterranean and Sahara Africa. *Annals of the Missouri Botanical Garden* **65:** 479-534.

Rambaut, A., Drummond, A.J. (2007). *Tracer, version 1.5.* http://beast.bio.ed.ac.uk/Tracer.

Rambout, A. (2009). FigTree v.3.1. http://tree.bio.ed.ac.uk/software/figtree/.

Reznicek, A.A. (1990). Evolution in sedges (*Carex*, Cyperaceae). *Canadian journal of botany* **68:** 1409-1432.

Richards, C.L., Carstens, B.C., Knowles, L.L. (2007). Distribution modeling and statistical phylogeography: an integrative framework for generating and testing alternative biogeographic hypotheses. *Journal of Biogeography* **34:** 1833-1845.

Rohlf, F.J. (1998). *NTSYS-PC Numerical Taxonomy and Multivariate Analysis System*. Exeter Publications, Setauket, New York.

Ronquist, F., Huelsenbeck, J.P. (2003). MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19:** 1572-1574.

Roy, M.S. (1997). Recent diversification in African Greenbuls (Pycnonotidae: Andropadus) supports a montane speciation model. *Proceedings of the Royal Society B: Biological Sciences* **264:** 1337-1344.

Roy, M.S., Sponer, R., Fjeldså, J. (2001). Molecular systematics and evolutionary history of akalats (genus Sheppardia): a pre-Pleistocene radiation in a group of African forest birds. *Molecular Phylogenetics and Evolution* **18:** 74–83.

Ryner, M.A., Bonnefille, R., Holmgren, K., Muzuka, A. (2006). Vegetation changes in Empakaai Crater, northern Tanzania, at 14 000–9300 cal yr BP. *Review of Palaeobotany and Palynology* **140:** 163-174.

Sanderson, M.J. (1997). A nonparametric approach to estimating divergence times in the absence of rate constancy. *Molecular Biology and Evolution* **14:** 1218-1231.

Schönswetter, P., Tribsch, A. (2005). Vicariance and dispersal in the alpine perennial *Bupleurum stellatum* L. (Apiaceae). *Taxon* **54:** 725-732.

Sepulchre, P., Ramstein, G., Fluteau, F., Schuster, M., Tiercelin, J.J., Brunet, M. (2006). Tectonic uplift and eastern Africa aridification. *Science* **313:** 1419-1423.

Shaw, J., Lickey, E., Beck, J.T., Farmer, S.B., Liu, W., Miller, J., Siripun, K.C., Winder, C.T., Schilling, E.E., Small, R.L. (2005). The tortoise and the hare II: Relative utility of 21 noncoding chloroplast sequences for phylogenetic analysis. *American Journal of Botany* **92:** 142–166.

Shaw, J., Lickey, E.B., Schilling, E.E., Small, R.L. (2007). Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: The tortoise and the hare III. *American Journal of Botany* **94:** 275–288.

Sloan, D.B., Oxelman, B., Rautenberg, A., Taylor, D.R. (2009). Phylogenetic analysis of mitochondrial substitution rate variation in the angiosperm tribe Sileneae. *BMC Evolutionary Biology* **9:** 1-16.

Swofford, D.L. (2002). *PAUP* Phylogenetic Analysis Using Parsimony (*and Other Methods), version 4.* . Sinauer Associates, Sunderland, Massachusetts.

Taberlet, P., Fumagalli, L., Wust-Saucy, A.-G., Cosson, J.-F. (1998). Comparative phylogeography and postglacial colonization routes in Europe. *Molecular Ecology* **7:** 453-464.

Taylor, P. (1955). The genus *Anagallis* in tropical and South Africa. *Kew Bulletin* **10:** 321-350.

Tolley, K.A., Tilbury, C.R., Measey, G.J., Menegon, M., Branch, W.R., Matthee, C.A. (2011). Ancient forest fragmentation or recent radiation? Testing refugial speciation models in chameleons within an African biodiversity hotspot. *Journal of Biogeography* **38:** 1-13.

Trauth, M.H., Maslin , M.A., Deino, A., Strecker, M.R. (2005). Late cenozoic moisture history of East Africa. *Science* **309:** 2051-2053.

Van de Peer, Y., De Wachter, R. (1994). TREECON for Windows: a software package for the construction and drawing of evolutionary trees for the Microsoft Windows environment. *Computer Applications in the Biosciences* **10:** 569-570.

Verdcourt, B. (2010). *Carex.* In**:** Beentje, H.J., Ghazanfar, S.A. (Eds.), *Flora of tropical East Africa*. Richmond, Kew Publishing, Kew (UK).

Voelker, G., Outlaw, R.K., Bowie, R.C.K. (2010). Pliocene forest dynamics as a primary driver of African bird speciation. *Global Ecology and Biogeography* **19:** 111-121.

Voje, K.L., Hemp, C., Flagstad, O., Saetre, G.P., Stenseth, N.C. (2009). Climatic change as an engine for speciation in flightless Orthoptera species inhabiting African mountains. *Molecular Ecology* **18:** 93-108.

Vos, P., Hogers, R., Bleekwe, M., Reijians, M., Lee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J., Kuiper, M., Zabeau, M. (1995). AFLP a new techinique for DNA fingerprinting. *Nucleic Acids Research* **23:** 4407-4414.

Wara, M.W. (2005). Permanent El Nino-like conditions during the Pliocene warm period. *Science* **313:** 1739-1739.

Wesche, K., Miehe, G., Kaeppeli, M. (2000). The significance of fire for afroalpine ericaceous vegetation. *Mountain Research and Development* **20:** 340-347.

Williamson, P. (1985). Evidence for an early Plio-Pleistocene rainforest expansion in East Africa. *Nature* **315:** 487-489.

Wolde Gabriel, G., White, T.D., Suwa, G., Renne, P., Heinzelin, J.D., Hart, W.K., Heiken , G. (1994). Ecological and temporal placement of early Pliocene hominids at Aramis, Ethiopia. *Nature* **371:** 330-333.

Yang, Z. (2004). A heuristic rate smoothing procedure for maximum likelihood estimation of species divergence times. *Acta Zoologica Sinica* **50:** 645-656.

Zuckerkandl, E., Pauling, L. (1965). Evolutionary divergence and convergence in proteins. In**:** Bryson, V., Vogel, H. (Eds.), *Evolving Genes and Proteins*. Academic Press, New York, pp. 97–166.