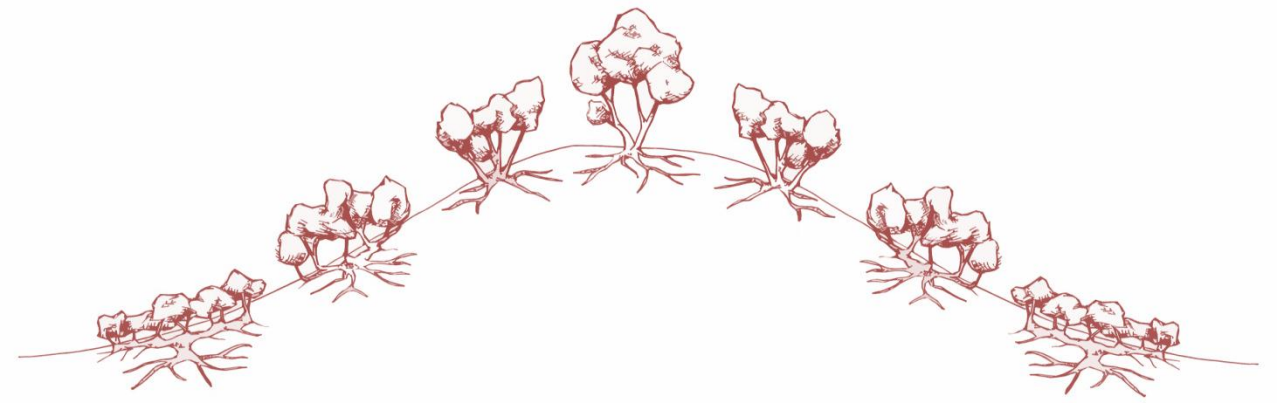


*Master Thesis*

# Molecular biogeography of the *Syzygium guineense* complex: How environment and genetics push suffrutication in Africas Miombo region


Paulina Zigelski neé Meller

Born 10.04.1989

Universität Hamburg 2016


Supervisors:

Prof. Dr. Norbert Jürgens

Dr. Barbara Rudolph

# Erklärung

Hiermit bestätige ich an Eides statt , dass die vorliegende Arbeit von mir selbständig verfasst wurde und ich keine anderen als die angegebenen Hilfsmittel – insbesondere keine im Quellenverzeichnis nicht benannten Internet–Quellen – benutzt habe und die Arbeit von mir vorher nicht in einem anderen Prüfungsverfahren eingereicht wurde. Die eingereichte schriftliche Fassung entspricht der auf dem elektronischen Speichermedium.

Ich bin damit einverstanden, dass die Masterarbeit veröffentlicht wird.

Datum Unterschrift

# Table of contents

# Abstract

The forest-grassland-mosaic of the Miombo region in Southern Africa is home to a peculiar phenomenon: Many woody species from different plant families that occurr in the forests as regular trees have close relatives in the open grasslands, growing as dwarf shrubs with huge underground woody biomasses. These “underground trees”, the geoxylic suffrutices, resemble their aboveground tree partners very much with regard to leave, flower and fruit traits, but it is unknown how close they are related to their respective tree partners and whether they have to be considered as separate species or rather separate ecotypes of one species. Furthermore environmental drivers of their evolution, e.g. fire, frost or edaphic conditions, are still controversely discussed in literature and in this study. The Myrtaceae *Syzygium guineense* (Willd.) DC s.l. was chosen as a model species to investigate the genetic relatedness between tree and suffrutex and to assess the influence of genetics and environment on the growth form and the evolutionary path. Seven distinct phenotypes of *Syzygium* have been identified, ecologically characterized and genotyped via sequence and SSR analyses. Genetically all phenotypes were very close though three distinct genotypes, concurrently also distinct broader ecotypes, were determinable: one water associated tree ecotype growing along rivers, one disturbance tolerant tree/shrub that occurs in Miombo woodlands and ecotones on ferralitic soils and finally one suffrutex that grows in sandy grasslands as an underground tree. The genetic patterns and their similarity, contrasting with the ecological distinctiveness, indicate a common progenitor of humid, tropical origin which evolved into different ecological directions induced by a changing environment. A major climate change around 2.5 mya ago led to an increased precipitation seasonality, drier conditions and therefore habitat loss for the biota. New adaptation through evolution was facilitated by *Syzygium’*s polyploidy, a genetic preadaptation, which led to a heightened responsiveness to environmental drivers. Of these particularly frost is considered as a shaper of the suffrutex as it is a consequence of increased seasonality. Therefore both genetic preadaptation as well as environmental changes led to the evolution of suffrutices, but not enough genetic divergence has happened yet for them to evolve into separate species, at least in the case of *Syzygium guineense* s.l.

# Zusammenfassung

Das Wald-Grasland-Mosaik der Miombo Region im Süden Afrikas beherbergt ein merkwürdiges Phänomen: Viele hölzerne Arten aus unterschiedlichen Pflanzenfamilien wachsen als Bäume in den dortigen Wäldern, haben aber nahe verwandte Taxa in den offenen Grasflächen, die dort als Zwergsträucher mit großer unterirdischer, hölzerner Biomasse wachsen. Diese „unterirdischen Bäume“, die geoxylischen Suffrutices, ähneln ihren „konventionallen“ Baumpartnern sehr in Hinsicht auf Blätter, Blüten oder Früchte. Es ist allerdings in den meisten Fällen unklar, wie nah sie wirklich mit ihren Baumpartnern verwandt sind, ob es sich um verschiedene Arten handelt oder eher um unterschiedlich ausgeprägte Ökotypen derselben Art. Außerdem wird in der wissenschaftlichen Gemeinschaft sowie der vorliegenden Studie noch immer über die Umweltfaktoren diskutiert, die diese Suffrutices geformt haben könnten (z.B. Feuer, Frost, Bodenbeschaffenheit).

Die Myrtaceaea *Syzygium guineense* (Willd.) DC s.l. wurde deshalb als Modellart ausgewählt, um die Frage nach dem genetischen Verwandtschaftsgrad zwischen Baum und Suffrutex zu untersuchen sowie um den Einfluss von Genetik und Umwelt auf die Wuchsform und seine Evolution zu beurteilen. Sieben unterschiedliche Phänotypen wurden identifiziert, besammelt und ökologisch und genetisch mittels Sequenz- und SSR-Analyse charakterisiert. Alle Phänotypen waren genetisch sehr ähnlich, wobei dennoch drei distinkte Genotypen, die breiteren Ökotypen entsprechen, ausgemacht werden konnten: ein wasser-gebundener Baum-Ökotyp, welcher entlang von Flüssen wächst; ein störungstoleranter Baum/Strauch-Ökotyp, der auf ferralitischen Böden der Miombowälder wächst sowie der Suffrutex-Ökotyp - der unterirdische Baum - der offenen, sandigen Grasflächen. Die genetischen Muster und ihre Ähnlichkeit, im Kontrast zu den deutlich unterschiedlichen ökologischen Eigenschaften, deuten auf einen gemeinsamen Vorfahren aus feuchten, tropischen Regionen, welcher sich durch sich verändernde Umweltbedingungen in verschiedene ökologische Richtungen entwickelte. Der Klimawandel vor ungefähr 2.5 mio Jahren führte zu einer gesteigerten Niederschlags-Saisonalität, jahreszeitlich trockeneren Bedingungen und damit einhergehend Habitatverlusten für die feuchttropischen Biota. Neue Anpassung durch Evolution wurde für *Syzygium* durch seine Polyploidie ermöglicht, eine genetische Vorangepasstheit, die eine erhöhte Reaktionsfähigkeit auf Umweltfaktoren bedingt. Von diesen ist insbesondere Frost ein bedeutender Former des Suffrutex, da häufigere Frostereignisse eine Konsequenz der zunehmend ausgeprägten Trockenzeiten sind. Insofern sind es die Kombination von genetischer Vorangepasstheit und als Filter wirkenden Umweltveränderungen, die die Evolution des Suffrutex vorangetrieben haben. Es ist aber noch nicht genug Zeit vergangen, um ihn als scharf getrennte eigene Art zu bezeichnen, zumindest nicht im Falle von *Syzygium guineense* s.l.

# Abbreviation Index

| | |
|---|---|
| ABI | Applied Biosystems (but refers to the Genetic Analyzer 3500 from Applied Biosystems) |
| AIC | Akaikes Information Criterion |
| AIS | Alleles in Space |
| AMOVA | Analysis of Molecular Variance |
| ANOVA | Analysis of Variance |
| DMSO | Dimethyl Sulfoxide |
| DNA | Desoxyribonucleic Acid |
| dNTP | Desoxynucleoside Triphosphate |
| DRC | Democratic Republic of Congo |
| DTT | Dithiotreitol |
| EDTA | Ethylenediaminetetraacetic Acid |
| EtOH | Ethanol |
| FAMD | Factor Analysis on Mixed Data |
| HCPC | Hierarchical Clustering on Principle Components |
| HT | Haplotype |
| ISCED | Instituto Superior de Ciências de Educação do Lubango |
| ITS | Internal Transcribed Spacer |
| MCMC | Monte Carlo Markov Chain |
| NCBI | National Center of Biotechnology Information |
| NJ | Neighbor Joining |
| NP | National Park |
| PCoA | Principle Coordinate Analysis |
| PCR | Polymerase Chain Reaction |
| PERMANOVA | Permutational Multivariate Analysis of Variance |
| PFT | Plant Functional Trait |
| SAA | Spatial Autocorrelation |
| SB | *Syzygium benguellense* (phenotype) |
| SC | *Syzygium cordatum* (phenotype) |
| SGB | *Syzygium guineense barotsense* (phenotype) |
| SGBx | *Syzygium guineense* (intermediate phenotype SGH-WZG/SGB) |
| SGG | *Syzygium guineense guineense* (phenotype) |
| SGH-Mio | *Syzygium guineense huillense* from the Angolan Miombo ecoregion (phenotype) |
| SGH-WZG | *Syzygium guineense huillense* from the Western Zambezian Grasslands ecoregion (phenotype) |
| SGL | *Syzygium guineense ssp littorale* |
| SGM | *Syzygium guineense macrocarpum* (phenotype) |
| SGMx | *Syzygium guineense* (intermediate phenotype SGH-Mio/SGM) |
| SGS | *Syzygium guineense ssp staudtii* |
| SNP | Single Nucleotide Polymorphism |
| SS | Sum of Squares |
| SSR | Simple Sequence Repeat |
| ST | *Syzygium "tundavalense"* (phenotype) |
| TBE | Tris – Boric Acid - EDTA |
| TE | Tris - EDTA |
| Tris | Tris(hydroxymethyl)aminomethane |
| UV | Ultra violet light |
| WZG | Western Zambezian Grasslands (biome) |

# I. Introduction

The Miombo biome covers large parts of Southern Africa, stretching between -5° to -20° latitude from Angola in the West, over Zambia to Mozambique and Tanzania in the East. It is a tropical, dry and semi-deciduous forest ecosystem dominated by Caesalpinioid tree species (particularly *Brachystegia*, *Isoberlinia* and *Julbernardia*), occurring in regions with pronounced precipitation seasonality. Today it represents with approx. 2.7 mio km² (Frost, 1996) one of the largest near-intact ecosystems, but growing human population and conversion to agriculture, particularly the emergence of an African soybean frontier, threaten its nativeness (Leadley, 2010; Gasparri et al., 2016). The tree cover that defines forests already declined in the past years within the Miombo region (Hansen et al., 2013; Global Forest Change), a course that might have devastating effects on the regional and global scale as the Miombo represents one of the tipping points regarding climate, biodiversity or ecosystem services (Leadley, 2010). A thorough understanding of this forest system is therefore a crucial step for its conservation.

But what do we imagine when we think of a forest? We see a dense congregation of woody species, of trees forming a (sometimes multilayered) canopy with their branches, spread over their stems and the understorey. This basic structural concept also applies to the Miombo woodlands of Africa. One of its main characteristics, however, is the mosaic appearance. Miombo forests cover in many parts undulating landscapes where forests, growing on hills and upper slopes, are interlocked with open lands at sites where trees do not grow, e. g. bogs, depressions or valleys (Huntley, 1974; White, 1983).

These seemingly grasslands are home to a high diversity of woody dwarf shrub species with huge subterraneous woody biomass, termed ‘suffrutices’. These suffrutices come from many different tropical plant families (Maurin et al., 2014). More than 100 suffrutex species from 55 genera in 30 families occur throughout the Zambezian Miombo Region, including the WWF ecoregions Angolan Miombo woodlands, the Zambezian Miombo woodlands and the Western Zambezian Grasslands (White, 1977; WWF)

Decades ago Davy (1922) and White (1977) reviewed the flora of these open lands and called attention to a peculiar phenomenon that transgresses our traditional concept of a forest (as described above): both men recognized that the suffrutex taxa in the grasslands mirror closely related tree taxa in the surrounding forests and thus in floristic terms embody a kind of forest – but an underground one. What at first look seem to be mere dwarf shrubs turns out to hide massive wooden structures underground that are akin to the canopy structure of common trees. The aboveground characteristics, however, resemble very much those of their closely related tree partners in the forests, i.e. leaf, flower and fruit traits are very similar. The crucial difference is the stands height, as the suffrutices protrude only low shoots aboveground (most < 50 cm) which after disturbances regrow rapidly from buds hidden underground (White, 1977). White (1977) thus coined the term “underground forest” for this phenomenon and he denominated the “underground trees” as *geoxylic suffrutex*.

Since the studies of Davy 1922 and White 1977 for many decades neither ecological nor evolutionary studies had been conducted focusing on the suffrutices and only recently new interest into this subject has sparked. The fact is striking that most of the suffrutex species have

closely related tree species within the aboveground forests, from whom they either evolved or with whom they share a close ancestor (Davy, 1922; White, 1977). However, it is still unknown for most taxa how close the relation with the closest sister species is and still controversial what drivers led to the evolution of the suffrutex habit. Several ecological drivers have been suggested in the past, and aside from herbivory, water logging, edaphic conditions (Davy 1922, White 1977) and frost (Finckh et al., 2016) the main focus had lain on fire (Simon & Pennington, 2012; Maurin et al., 2014).

Recent genetic studies determined the emergence and radiation of suffrutices to have mainly happened within the Pliocene (< 5.3 mya) with a major evolutionary peak at approx. 2.3 mya (Maurin et al., 2014) which is coinciding with the spread of flammable grasslands across the tropics of the world, including Miombo (Beerling & Osborne, 2006). Admittedly, fire plays an important role in current suffrutex ecosystems, but it does not explain all aspects of the astonishing evolution of suffrutices in the Miombo region. There is for example a discrepancy about the onset of suffrutices and the needed fire frequency required to have an impact on their evolution (Finckh et al., 2016). Therefore a thorough discussion about the ecological drivers is needed.

In addition to the active, ecological factors driving the evolution we also need to consider the passive, responsive part: the genetic setup of the suffrutex taxa. Currently, it is not known what genetic features and preadaptations might have facilitated the evolution of suffrutices. Furthermore, there are many knowledge gaps regarding the genetic biogeography of the Miombo region. We concordantly see that many tree-suffrutex species pairs must be closely related due to the aboveground trait accordance; disregarding the growth height, it is striking. However, in many cases the extent of relatedness is still unclear as we do not know if a pair represents two distinct species or rather one species with different phenotypes, depending on the environmental pressure. Maurin et al. (2014) dated the emergence of the suffrutex using molecular clocks on genetic differences between trees and their suffrutex partners; however, they did not elaborate further on the genetic divergences. Hence, until now there aren't any detailed genetic studies of the relationship between tree and suffrutex taxa from the Miombo region.

Modern molecular biological methods allow answering such questions. Population genetics as well as phylogenetic analyses of different scale and genetic resolution are nowadays widely applied, e.g. to investigate relatedness within families, genera, species or populations (Hirschhorn et al., 2002, Avise, 2012). Common methods use nuclear or chloroplast sequence data (CBOL et al.,2009; Schoch et al., 2012)or microsatellite data (SSR, Kalia et al., 2011) or a combination of them (Avise, 2012) to identify genetic patterns. Particularly the tools based on fragment analysis (SSR) are able to detect evolutionary processes like hybridization, gene drift, bottlenecks or isolation by distance, even at population or even individual scale (Jarne & Lagoda, 1996; Avise, 2012). They are therefore very suitable to shed light on the genetic setup of the tree – suffrutex pairs and how it came to be.

Thus, this study aims at investigating the evolution of underground trees further. To do so, I chose the woody Myrtaceae *Syzygium guineense* (Willd.) DC s.l. as a model taxon to analyse the genetic patterns between suffrutex and tree partners: The genus *Syzygium* is represented with some species in Africa (*S. cordatum*, *S. owariense*, *S. pondoense*) and *Syzygium guineense* is probably the most widespread of them as it occurs in rainforests in western and central Africa

(*guineense* = from Guinea), in the Ethiopian highlands and throughout the Miombo belt. Along with the large distribution goes a polymorphic appearance which is adapted to the respective habitat diversity. There is not only the large tree in the forests (*Syzygium guineense* ssp. *guineense*) and the suffrutex (*Syzygium guineense* ssp. *huillense*) in the grasslands; there are also intermediate forms and different tree phenotypes. Due to these features, the easy identification, and its ubiquitous distribution, *Syzygium guineense* (Willd.) DC s.l. is a suitable model taxon to investigate the genetic patterns of a tree-suffrutex species pair within the Miombo biome and to link genetic methods with biogeographic data.

Using *Syzygium guineense* as a model species, the following questions shall be investigated:

- Are tree - suffrutex pairs separate species, subspecies, or rather varieties of the same species with a high phenotypic variability?
- Do the tree and suffrutex partners, standing within eyeshot, show closer genetic relatedness to each other (despite deviant life forms) or are rather the phenotypes more closely related among themselves (over the huge distances of the whole Miombo region)? Is there still current gene flow between the different phenotypes?
- What genetic and environmental factors might have led to the evolution of the suffrutex habit?

Thus, the objective of this study is not only to better understand the genetic background of the tree-suffrutex pairs in the Miombo region. I want to discuss possible genetic processes (e.g. hybridization) and environmental filters (e.g. climate) that could explain the genetic patterns, and I want to put the findings into a greater ecological context in order to elucidate the drivers and facilitators of the suffrutescent habit. This approach aims to link “genetics” with “ecology” in order to show how the interplay of genes and environment shapes the vegetation of the Miombo landscapes.

# II. Material and Methods

## II.1 Field work

### II.1.1 Study region

The Miombo Region of southern Africa covers large parts of Angola, the south of DRC and Zambia, and large parts of Tanzania, Mozambique, Zimbabwe and north-eastern RSA. As the western part of this region, the area under study focuses on the biomes and WWF ecoregions "Angolan Miombo Woodlands" and "Western Zambezian Grasslands" (White, 1983; WWF: scientific codes AT0701 and AT0724). Both are characterized by a mosaic of more or less open woodlands and grasslands, though the Western Zambezian Grasslands are situated on the plains of the Zambezi graben whereas the Angolan part of the Miombo covers the undulating landscape of the Angolan plateau. Nevertheless, in both ecoregions trees are growing at favorable, somewhat elevated sites and the suffrutex-grasslands spread over lower slopes and valleys. Naturally this means that true forests are predominant over open grasslands in Angolan Miombo, whereas it is the other way round in the Zambezian Grasslands (WWF, pers. obs.). Some samples in this study have also been taken at sites outside the main Miombo region, e. g. at the Okavango Delta.

Angola's Miombo landscape, particularly in Central Angola where most of the research was conducted, is often dominated by a matrix of open forests on hilltops and ridges with nutrient poor ferralitic soils and secondary grasslands in sandy valleys and on the footslopes (see fig. 1a).The drainage lines are accompanied by linear peat bogs (Frost, 1996; Oldeland et al., 2013, pp. 43 – 44). The study regions in the Angolan Miombo Woodlands range from approx. 1200 m to 1800 m a.s.l.. The altitudinal difference between drainage lines and hill tops can reach more than 100 m (pers. obs.).

Pronounced precipitation seasonality is another feature of this biome. The overall annual precipitation varies between 800 – 1400 mm, but is limited to the rainy season from October to April; there is no precipitation in the dry season (Huntley, 1974; Hijmans, et al., 2005). This seasonality leads to temperature ranges between ca. 27°C/15°C (day/night) during the wet season and extreme amplitudes of up to 40°C/-6°C in the dry season. In places where cold air accumulates, e.g. in valleys and topographic depressions, this amplitude is particularly high; hills and topslopes are somewhat milder (Oldeland et al., 2013, pp. 47 – 50; SASSCAL WeatherNet, 2015: Cusseque station; Finckh et al., 2016).

The forests are mainly home to Caesalpinoid tree genera and species, particularly *Brachystegia* spp., *Cryptosepalum* spp., *Julbernardia paniculata* and *Isoberlinia angolensis*, but also to trees from other families like *Combretum* spp. (Combretaceae), *Uapaca kirkiana* (Euphorbiaceae) or *Syzygium guineense* (Myrtaceae). They form a closed forest of 10 m - 15 m height with little or no herbaceous understorey but extensive grass layer where the canopy is disturbed. The sloping rather open ecotones between forests and grasslands are colonized by the same or related species but these are much lower and sparser, giving space to grasses like *Loudetia* spp., *Hyparrhenia* spp. or *Monocymbium ceresiiforme* and dwarf shrubs like *Cryptosepalum*

*maravianse*, *Combretum platypetalum* or *Syzygium guineense* ssp. *huillense*, the latter ones closely related to forest trees (Huntley, 1974; Frost, 1996; Oldeland et al., 2013, pp. 59 – 63).

The Western Zambezian Grasslands ecoregion of the Zambezi basin is located on deep Kalahari sand deposits with no marked topography in the Zambesi graben; it is very flat, situated at around 1000 m a.s.l. (see fig. 1b). Both forests and grasslands grow on these sandy soils, though the grasslands in the lowest parts are seasonally flooded by swelling tributaries of the Zambezi from the middle of the rainy season on. These sites however are rather vegetated by Cyperaceae than by suffrutices (Coetzee & Werger, 1978; pers. obs.). The grasslands also provided forage for vast numbers of wildebeest, buffalos and other ungulates that migrated through this region (WWF) and still do so in some of the national parks. This ecoregions climate is of similar seasonality as the Angolan Miombo Woodlands; after a hot wet season from Nov. to Apr. follows a cool dry one (May – July) and a hot dry one (Aug. –Oct.). Annual precipitation concentrates in the wet season and accounts 800 to 1000 mm (Cole, 1963; Hijmans et al., 2005). The temperature ranges from 37°C/16°C (day/night) in wet and 30°C/6°C in the dry season (SASSCAL WeatherNet, 2015: Kalabo station), there is however nothing known about microclimatic conditions. Due to climatic disposition and human need of hunting and agriculture both ecoregions are prone to and affected by frequent fires in the dry season.

Some plant species occurring in the Angolan Miombo also occur in the Zambezian Grasslands; the forests for example are constituted by *Pterocarpus angolensis*, *Syzygium guineense* or *Baikiaea plurijuga*. Other tree species like *Dalbergia melanoxylon*, *Securidata longepedunculata* or *Croton megalobotrys* are not found in the Angolan Miombo ecoregion but in the Zambezian Grasslands (GBIF, Bingham et al., 2016: *Dalbergia melanoxylon*, *Securidata longepedunculata* & *Croton megalobotrys*). The transition from woodland to grassland is mostly slow, forming an ecotone of scattered smaller trees like *Erythrophleum africanum* or *Burkea africana* that give way to dominating grasses like *Monocymbium ceresiiforme*, *Loudetia simplex* or other wiry species like *Andropogon* spp. or *Eragrostis* spp.. The grasslands are furthermore home to a high diversity of dwarf shrubs which sprout and flower before the grasses do. Again, these are species related to those in the forests. Among them are for example *Parinari capensis* (tree partner*: P. curatellifolia*), *Syzygium guineense* ssp. *huillense* (tree partner*: S. g.* ssp. *guineense*), *Lannea edulis* (tree partner*: L. discolor*) or *Anisophyllea quangensis* (tree partner*: A. boehmii*) (Cole, 1963; Bingham et al., 2016: *Lannea edulis* & *Anisophyllea quangensis*).

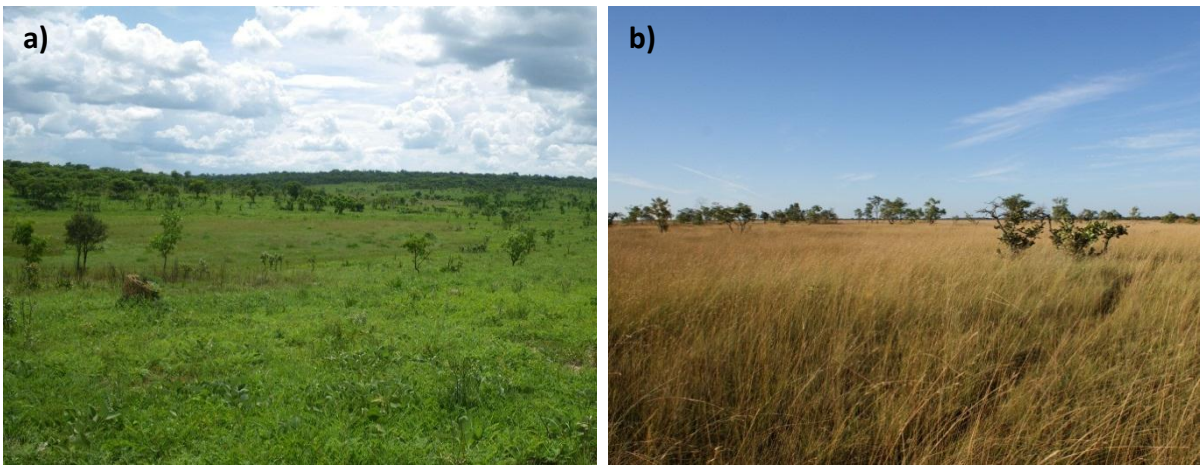



**Figure 1.** (a) Typical landscape of the Angolan Miombo woodlands on the Bíe Plateau in Central Angola. Photo taken in the beginning of the wet season when grasses hadn't sprouted yect. (b) The Western Zambezian Grasslands ecoregion of Cameia NP in Moxíco, Eastern Angola. In this case the photo taken at the beginning of the dry season when grasses are dominant, covering the suffrutices. Photos: a) P. Zigelski 2014; b) P. Zigelski 2016

### II.1.2 Sampling

The sampling was conducted from November 2014 to January 2015 in Namibia, Botswana, Zambia and Angola at sites with Miombo or Zambezian woodland, adjoining river systems and the ecotones in between. These form the natural habitats for the *Syzygium* phenotypes of interest. Additionally, some phenotypes were collected along the Angolan Escarpment in mountainous habitats. All locations of the sampled populations as well as probable locations of provided herbarium specimens are shown in fig. 2 (next page).

Collection took mostly place at sites that were easily accessible (near roads) and where enough individuals of a phenotype were present to provide a free sampling choice. Naturally the distances between the individuals in the sampled area varied between the populations. This problem affected especially sampling of the forest populations where the density and the success of finding *Syzygium* trees were lower compared to the suffrutices in the plains.

From each sampled population leaves and specimens of at least 10 individuals were taken, if possible. Per population individuals were collected partly close to each other (within 20 m), and partly in greater distance (up to 1.5 km between first and last plant of a sampling) in order to compare genetic differences. The distance between populations of the same kind (phenotype) was chosen so, that insect pollination is improbable (in one case 0.5 km, else > 5 km), so that there should be no direct gene flow in between.

Genetic material was collected in form of leaves that appeared intact and showed no signs of disease or infection by eukaryotic fungi, to avoid DNA contamination later on. Leaves from suffrutex and shrub individuals were taken from several sites plus the center of each plant due to possible (but unknown) underground connections, whereas tree samples were collected from a reachable branch. Additionally, a specimen was taken from each sampling site.

GPS position for each individual, some being so close to each other that they shared coordinates, were taken in order to relate genetic and geographic similarity later on. Since I initially assumed only three phenotypes (*S. guineense* ssp. *huillense*, *S. guineense* ssp. *guineense* and *S. cordatum*), the populations were named with the appropriate code (Sgh, Sgg and Sc). Only later did I recognize and classify the full spectrum of phenotype diversity (see II.1.2). Therefore some population names might be confusing (e.g. Sgg1 is actually a SGB); however I always either provided the attributed phenotype or a lower case letter to indicate the actual phenotype (Sgg1 >> Sgg1b). A list of all populations used in this study is supplemented in tab. A2 in the appendix and locations of populations sampled by me are described in the following chapter.

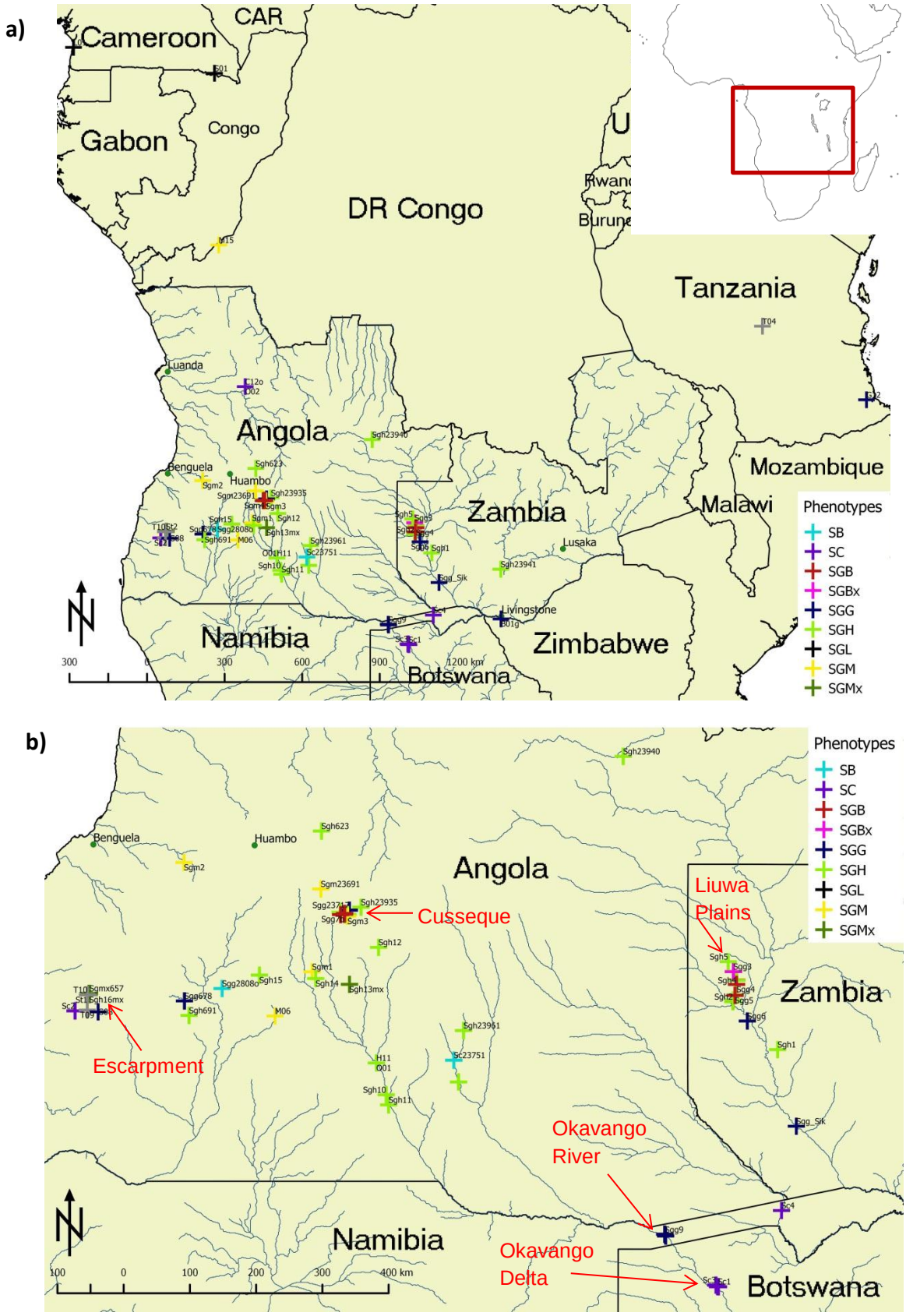



**Figure 2.** Overview of sample locations used in this study. Figure 2a shows the whole range including specimen from herbaria, and 2b shows locations of populations sampled by me or colleagues. The phenotype of the respective population is indicated by colored crosses. Important sites mentioned troughout the study are named for orientation.

### II.1.3 Habitat description

In total 16 open grassland habitats with suffrutex or shrub populations, 4 open and 2 dense Miombo woodland habitats as well as 5 Zambezian Graslands forest sites with trees, 5 river or floodplain habitats and two mountainous habitats were extensively sampled. A short description of the population habitats and their characteristics, as observed during sample collections, are given below.

Sgh1: Zambia, south of Mongu; sandy dambos grassland in a wide, shallow river valley. Dry treeless bottom, slight slopes with protruding islets of dwarf shrubs (e.g. *Parinari capensis*, *Ochna arenaria*), fringes with initially very open woodland getting denser with distance to valley bottom.

Sgh2: Zambia, north of Kalabo; wide, sandy, lightly grass covered plain with scarce distribution of low single or grouped trees (e.g. *Burkea africana*), in between islets of dwarf shrubs (dambos).

Sgh3: Zambia, Liuwa Plains NP; transition zone from Miombo woodland to sandy grass plain sharing characteristics of both habitats. Collection in line with woodland fringe, tree height approx. 8 m. Minimal elevation between woodland and plain (5 m).

Sgh4: Zambia, Liuwa Plains NP; sandy grass covered dambos plain some hundred meters away from woodland, separated from Sgh3 by dry river run with no dwarf shrub growth.

Sgh5: Zambia, Liuwa Plains NP; sandy, plain grassland with no trees in vicinity.

Sgh6: Zambia, Liuwa Plains NP; along the fringes of Miombo woodland and in between single trees (up to 10 m) on the sandy plain.

Sgh7: Angola, Cusseque catchment; east of Rio Cusseque inflowing river system, transition from sandy valley slope to peaty bottom, in line with river. Low vegetation (< 1 m), mostly islets of dwarf shrubs and grasses in between, those getting dominant and higher on the peat closer to the river. Slope to the Miombo woodland (trees up to 10 m high) much steeper and elevation difference greater (50 m from valley bottom to hill range top) than in Liuwa.

Sgh8: Angola, Cusseque catchment; east of Rio Cusseque inflowing river system, transition from open Miombo (trees up to 10 m) to tree and shrub scattered slope (vegetation height < 5 m) on ferralitic soils. *Brachystegia* shrubs dominate.

Sgh9: Angola, Cusseque catchment; west of Rio Cusseque inflowing river system, samples taken from the ferralitic Miombo woodland fringes (scattered trees lower than 10 m but mostly shrubs < 1.5 m) and Cassamba heath (*Cryptosepalum maraviense* and *C. pseudotaxus*) as well as from the lower sandy part of the slope near the river where vegetation consists of abundant dwarf shrub patches separated by tall grasses (1 m) or Cassamba.

Sgh10: Angola, lower Cubango; narrow sandy strip between peaty grassland (approx. 1 m high) and open Miombo woodland (10 m) with scattered dwarf shrubs, shrubs and smaller trees. Small river system joining the close by Cubango. Elevation difference low (10-20 m) between woodland and peatland.

Sgh11: Angola, lower Cubango; wide sandy plain, possibly former connected to floodplains of Cubango river. Grazed grass vegetation (< 0.5 m) with elevated patches of mostly *Syzygium* dwarf shrubs. Plain uphill surrounded by Miombo woodland, downhill merging into agricultural fields.

Sgh12: Angola, between Menongue and Chitembo; wide and shallow sandy river valley between Miombo covered hill ranges. Slopes dominated by *Parinari*, *Burkea* and other dwarf

shrubs and low, isolated trees, but no *Syzygium* (vegetation height in average < 1 m). Those could only be found along a ridge parallel to the river, being the closest dwarf shrubs to the water.

Sgh13: Angola, between Menongue and Cuchi; shallow, rather ferralitic river valley between Miombo (> 10 m tree height). Collection took place at woodland fringe and on Cassamba covered slope with some scattered small trees or large shrubs and moderate grass cover. Dwarf shrubs form loosely distributed patches.

Sgh14: Angola, between Menongue and Cuchi; wide peaty marshland covered by tall grasses (< 1.5 m). Short, steep elevation to the surrounding ferralitic Miombo (< 10 m tree height) and settlements, with patches of dwarf shrubs and shrubs (< 1 m).

Sgh15: Angola, between Dongo and Matala; flooded peaty marshland along a river, covered with tall grasses (< 1.5 m) and elevated islets of *Syzygium* dwarf shrubs. Sharp transition to open woodland (approx. 5 m tree height) with sandy and ferralitic soils.

Sgh16: Angola, Tundavala escarpment; withered granite substrate with shallow sandy cover. Scattered blocks of various sizes on lightly sloping plains that are locally interrupted by suddenly opening crevices and chasms and sharply ascending faults of intact, vegetated granite. Vegetation low (< 0.5 m), mostly grasses and dwarf shrubs, except for wind protected sites were sturdy shrubs and trees achieve greater heights (< 10 m).

Sgg1: Zambia, Liuwa Plains NP; slightly elevated, arboreous band on sandy plain with a rather patchy tree coverage and clearings, trees up to 15 m high (e.g. *Diplorhynchos condylocarpon*, *Pterocarpus angolensis*), little understorey and little grass cover, thin layer of litter under canopies and in clearings also occurrence of dwarf shrubs.

Sgg2: Zambia, Liuwa Plains NP; Miombo forest with high trees (up to 15 m) and closed canopy in center getting less dense and patchy towards the fringes. Layer of litter increasing according to density and height of trees, analogously with decreasing understorey and ground cover. No dwarf shrubs occur in closed forest. Soil is sandy with a thin organically enriched top horizon.

Sgg3: Zambia, Liuwa Plains NP; open Miombo woodland with a mosaic of denser forest patches (trees up to 15 m), grass and (dwarf) shrub covered clearings plus transitions between them. Soil is sandy, more or less covered with litter and top horizon enriched with organic materials, depending on amount of biomass on it.

Sgg4: Zambia, Liuwa Plains NP; greater band of dense forest, canopy closed. Single trees up to 15 m, majority around 10 m high, little understorey or ground coverage, except where canopy opens. Ground littered, soil sandy and in top horizon darkly enriched with organic material.

Sgg5: Zambia, Liuwa Plains NP; fringe of open Miombo woodland with fluid change to the dambo plains: gradient from closely grouped high trees (up to 15 m) dominated by *Syzygium* over loosely clustered or single trees up to 7 m height to shrub and suffrutex scattered open grass plain. Sandy soil with organically enriched top horizon, covered with thin layer of litter under shrubs and trees.

Sgg6: Zambia, Barotse Plains opposite of Mongu; seasonally flooded plain covered with grasses and sedges and intervened by water filled channels and holes. On higher, firmer ground groups of trees and shrubs up to 8 m high. No dwarf shrubs on the floodplain. Sandy, yellowish soils.

Sgg7: Angola, Cusseque catchment; dense Miombo forest with a canopy height up to 12 m. Lichen covered trunks, some understorey and ground cover as well as moss and litter. Scattered former agriculture fields (slash and burn) that had recovered to some extent, young trees and new sprouts shooting from old trunks reaching heights up to 4 m. Reddish clayey sand soils with

darker top horizon and increasing clay content in deeper horizons (see section “soil profiles” for more information).

Sgg8: Angola, Cusseque catchment; Miombo woodland on a slope towards the pan at the head of a tributary of river Cusseque. Mostly closed canopy with maximum height of 10 m. Some old former slash and burn fields that are now covered with young trees, new shoots (both up to 5 m high) as well as grasses, shrubs and ferns. Little understorey and ground cover but with layer of litter and high occurrence of lichens. Soil is reddish clayey sand with organic enriched top layer and strongly red sandy clay horizon in the deep.

Sgg9: Namibia, banks of the Okavango in the Caprivi Strip; summary of sites were single trees were sampled. Steep banks of the Okavango River that are forested with high trees up to 12 m and smaller ones (5 m) directly at the water. At some sites a belt of sedges and reed was also present. Sandy soils that are fixed through the roots.

Sgm1: Angola, between Menongue and Cuchi; loose Miombo woodland consisting of mostly young trees below 5 m height.

Sgm2: Angola, between Cubal and Alto Catumbela; secondary forest (*Eucalyptus* plantation) with high trees (up to 25 m) standing close to each other. Little understorey represented by some smaller shrubs and trees (below 4 m) and low herbs. Ground covered with litter. Reddish ferralitic soil. Wood is in silvicultural use.

Sgm3: Angola, Cusseque catchment; dense Miombo woodland on ridge, canopy nearly closed and about 8-10 m high. Lush understorey and ground covered by herbaceous layer. Some scattered former agricultural fields were canopy height is reduced but vegetation is denser compared to surroundings. Soil is constituted by ferralitic clayey sand with organically enriched top horizon and red withered clay in the deep.

Sgm4: Angola, Cusseque catchment; similar to Sgm3 a Miombo forest segment on the slope and top of a ridge. Dense woodland turning more open as it approaches the slope. Canopy height varies from 10 m to below 5 m with increasing understorey and herbaceous layer with decreasing tree height and density. Agricultural fields in recent utilization. Soil as at Sgm3 comprise enriched top layer followed by ferralitic clayey sand to sandy clay in the deep.

St1: Angola, Tundavala escarpment; stony precipice of large granite blocks covered with lichens and featuring sturdy vegetation in their protecting cracks and clefts. Aside from herbs and grasses also low shrubs and trees (below 5 m height) occur which cling to the barely withered soil and organic matter in the crevices.

St2: Angola, Tundavala escarpment; similar to St1 but further away from the precipice of the escarpment. Large granite blocks and hills in a rolling landscape covered by grasses and small shrubs. Blocks comprise alternate vegetation with herbs, shrubs and trees settled in the clefts. Soil rather clayey sand, barely withered with large amount of scattered rocks.

Sc1: Botswana, Okavango panhandle wetlands east of Seronga; Grass covered floodplain with gullies, channels and ponds filled with water and surrounded by sedges and papyrus. Termitaria constitute elevated islands in those wet surroundings, vegetated by trees and shrubs up to 15 m high and about 50 m in diameter. Wetlands are grazed by cattle. Soil is clayey sandy with darker fossile A-horizons and former sedimentation strati in the deep.

Sc2: Angola, stream bank below escarpment; small river that has dug its bed deep into the wooded slope of the foot of escarpment. Being next to a village, the wood comprises mango and

baobab and other plants of agricultural interest, aside of the typical species of Angola mountainous biome. Soil is clayey ferralitic with rocks of various sizes washed down by the stream in rainy season.

Sc3: Botswana, Okavango panhandle wetlands west of Seronga; similar to Sc1 situated in the floodplain but with bigger, hall-like tree islets mostly formed by *Syzygium*.

### II.1.4 Phenotype description

The phenotype categories used in this study correspond mostly to subspecies, or to species closely related to *Syzygium guineense* as described in literature (Exell et al., 1966; Coates Palgrave, 2002). For phenotypes without literature reference arbitrary names were chosen, for example *Syzygium "tundavalense"* since it was collected at Tundavala. The following description focuses on traits that differ between the phenotypes and are characteristic for them. To keep assigned codes for samples short, the first letter of the (sub)species epithet was used to indicate the phenotype, e.g. H01 is a sample of SGH type. Only exception is SB, there the letter of assignment is "O". As some samples had to be reassigned to another phenotype, this is indicated by a lower case letter at the end of the code (B01g). A complete list of samples with phenotype and location information is provided in the appendix (tab. A1)

**SGH:** Termed after *Syzygium guineense ssp. huillense*. It is the typical dwarf shrub growing in sandy soils, with little aboveground biomass (fig. 3). Only low shoots of averagely 20 - 50 cm, sometimes up to 100 cm height protrude from the ground. Its leaves are emerald, thick and stiff; small, rounded and almost circular at the branch base but becoming larger, oblongated and acutely pointed towards the terminal end. There is a short petiole and the leaf base is tapering. The terminal inflorescences are cymose corymb-shaped with white flowers; later fruits turn deep purple and are in form similar to rose hips. Fresh (re-)sprouts are light green and turn greyish brown later. The bark is lightly scaled. Underground in 5 to 70 cm depth grow horizontally thick wooden organs that origin from one central, vertical stem or sometimes several smaller ones. Near the surface they are enlarged and club-like, and from there the aboveground shoots origin. Also, some fine roots and some thicker, smooth and linear supporting roots sprout from the lower side of the clubs. Therefore one individual plant forms an approximately circular patch in grasslands of low height but diameters of up to 5 m. It is yet uncertain if this horizontal underground biomass is a shoot or root and how deep the central stem extents. In some analyses of this study SGH had been divided into SGH-Mio (from Miombo biome) and SGH-WZG (from Western Zambezian Grasslands biome). Populations assigned to this phenotype are Sgh1 to Sgh7, part of Sgh9, Sgh10 to Sgh12, Sgh14 and Sgh15; Sgh23935, Sgh23940, Sgh23941, Sgh623 and Sgh691.

**Figure 3**: SGH at Liuwa Plains NP, Zambia. Patch of approx. 4 m in diameter. Photo: P. Zigelski 2014

**SGB:** Termed after *Syzygium guineense ssp. barotsense*. It is a tall tree of up to 30 m height, found in forests not directly affected by water (rivers and streams) but rather in certain distance to it. Often multi-stemmed, it has a light, greyish bark with a brownish tinge on younger branches but the bark becomes scaly and somewhat peeling with age. The canopy is domed and hollow in the center, the shiny, dark green leaves contrasting with bare patches caused by dead, hanging branches in the outside hull (fig. 4). The leaves are oval, attenuately to caudately pointed with a tapering base and a long petiole. They are thin, of silky and smooth texture. The white flowers are on a slack, panicle-like inflorescence and the fruits in shape similar to those of SGH but somewhat smaller and more reddish. Inflorescences are not exclusively terminal. This phenotype is evident in populations Sgg1b, Sgg2b, Sgg4b and in parts of Sgg3b, Sgg5b, Sgg7 and Sgg8.

**Figure 1:** SGB at Liuwa Plains NP, Zambia, a grove around our campsite. Photo:P. Zigelski 2014

**SGBx**: A potential *Syzygium guineense ssp. huillense x barotsense* hybrid living in sympatry and with characteristics of both SGH and SGB. It is a multi-stemmed large shrub or small tree of 2 m – 5 m height with a bark similar to SGB but with young branches of slightly more reddish color (fig. 5). The shiny leaves are thicker, bigger and more acutely pointed, rather egg-shaped in comparison to SGB. The white flowers are arranged in a compact conical corymb and the fruits in shape and size similar to those of SGH. It is unknown if this phenotype also develops geoxylic structures. It occurs in Populations Sgg3bx and Sgg5bx.

**Figure 5:** A SGBx shrub in Liuwa Plains NP, Zambia. It grows in the woodland – savanna matrix in more open places. Photo: P. Zigelski 2014

**SGG:** The nominate form *Syzygium guineense ssp. guineense*. Found in gallery forests and on river banks, sometimes even standing in the water when shallow (fig. 6). It is a tree of medium height, mostly below 10 m, with very light and smooth, greyish bark. Younger sprouts have a yellowish-creamy tinge. The foliage is dark green and shiny, with long petioled, oval leaves of tapering base and attenuate to caudate apex. Similar to SGB the leaves are thinner and

**Figure 6:** SGG at Okavango River bank in the Caprivi Strip, Namibia. Photo: P. Zigelski 2014

softer than SGH. Slack, conical inflorescences of white flowers are found at the terminal branch ends; the ripe fruits are similar in size and form to SGH. Its populations are Sgg6, Sgg9 and Sgg678 and several single individuals from across the sampling range.

**SGM:** Named after *Syzygium guineense ssp. macrocarpum*. Found in primary and secondary Miombo woodlands and in adjacent ecotones towards the grasslands. In the first habitat it grows as small or medium sized tree (up to 8 m, fig. 7), often grouped, in the latter it appears as shrub or dwarf shrub patch of 0.5 m to 2 m height (fig. 8). Since this is an intermediate appearance between forest SGM and grassland SGH this sub-phenotype will be termed **SGMx** further on. Some populations that are labelled with Sgh, like Sgh8, part of Sgh9mx, Sgh13mx, Sgh16mx or Sgmx657 belong into this category. The bark on stem and older branches is greyish brown and shallowly fissured, younger branchlets are smoother, reddish brown with thickened nods. Young sprouts including leaves are wine red. Mature leaves are of bright emerald color. They are bigger than SGH leaves (up to 25 cm length) but similar in shape, with an acuminated apex and a long petiole. Less shiny than SGG or SGB, the leaves are somewhat smoother and less stiff than SGH, too. Inflorescences are slackly dome-shaped, often with two small leaves at its base. The ripe fruits are globular with a diameter of about 2 cm. SGMx develops geoxylic structure equal to SGH when growing in ecotones and forest fringes, however it is uncertain if the big SGM individuals in dense woodland do this alike. SGM populations are Sgm1 to Sgm4.

**SC:** Named after *Syzygium cordatum*, a closely related species. Strongly water associated, it grows at or even in rivers or streams. It is growing as a tree of variable size, from rather shrub-like to tall appearances of up to 15 m (fig. 9). Its bark is light grey and shallowly structured with a brown tinge on younger branchlets. Contrary to all other phenotypes the leaves have

**Figure 7:** SGM in open woodland, Huíla province, Angola. Photo: P. Zigelski 2014

**Figure 8:** SGMx on mid slope near Cusseque, Angola. Photo: P. Zigelski 2014

**Figure 9:** SC growing as tree in the Okavango Delta Panhandle near Seronga, Botswana. Photo: P. Zigelski 2014

almost no petiole and are sessile or subsessile. They are dull bright green with a more or less cordate base and a rounded or obtuse apex, to the touch similar like SGG. The flowers are mostly white but can have a rose hue. The inflorescences are compactly arranged in a corymb. Ripe fruits are bright magenta purple and inverted egg-shaped. Populations of this phenotype are Sc1 to Sc3 and several single individuals from across the sampling range.

**ST:** Arbitrary name *"Syzygium tundavalense",* since first collection was in Tundavala on the escarpment and no clear reference in literature could be found for this phenotype. Actually, specimen from Tundavala had been described in literature and herbaria as *S. g. ssp. guineense*, *S. g. ssp. afromontane* or *S. g. ssp. littorale*, but either way this arbitrary name was chosen to set the phenotype apart from the different looking SGG and because the description in Exell 1966 for *S. g. ssp. littorale* did not match ST. It grows as gnarled tree of up to 5 m height in crevices and between basalt boulders, its roots clawing into fissures and its crouching canopy and trunk hung with lichens (fig. 10). The leaves are the smallest of all phenotypes, narrowly oblongate with a long petiole, a tapering base and attenuated apex. They are dark, dull olive green, thick and somewhat angled along the rachis, the margins slightly rolled, therefore giving the leaf a stiff stability. The inflorescences are small and compact, a rounded corymb of flowers with white stamina and contrasting deep red ovaries (all other phenotypes having green hued ovaries). When ripe, the fruits are flatly globular shaped, approx. 1 cm – 1.5 cm in diameter and rouge to red. Two populations, St1 and St2, belong to this phenotype as do some specimen from the herbarium.

**Figure 10:** ST at Tundavala, Angolan Escarpment. Photo: P. Zigelski 2014

**SB:** Named for *Syzygium benguellense*. This phenotype was not deliberately sampled, but some individuals turned later out to belong rather to this phenotype which resembles SC quite much (fig. 11). According to Exell 1966 it often occurs in association with other *Syzygium sp.* in wet environments like streams where it grows as a tree up to 10 m height. Its leaves are soft, bicolored, narrow and elongated, with no pronounced but rounded tip. The petiole is very short or non-existent and the leaf base rounded. Inflorescences, terminally developed, are creamy white and form a compact umbel. Fruits are deep purple when ripe and egg-shaped.

**Figure 11:** SB in a nameless river in central Huíla province, Angola. Photo: P.Zigelski 2015

## II.1.5 Geoxyle excavations and soil profiles

In order to get an impression of the underground xylic biomass and to assess whether phenotypes other than the suffrutex develop geoxylic structures when exposed to similar environmental pressures, I excavated three individuals . One (suffrutex) was situated at the foot slope, the next at the middle of the slope (shrub) and one tree at top slope in the forest fringe, all near Cusseque (see fig. 12). In addition to that I dug a soil profile of 80 cm depth at each excavation site. This revealed a ferralitic arenosol for the forest site at 1553 m altitude, a ferralitic plinthisol at mid-slope (1540 m a.s.l.) and a sandy gleysol at the footslope (1518 m a.s.l.), at transition to the bog.


**Figure 12:** Geoxyle excavations on a slope in Cusseque. Three individuals were dug out; a tree form (SGM) at the forest fringe, 220 m further down a shrub form (SGMx) on mid-slope and 260 m further down at footslope a suffrutex (SGH). Position on the slope as well as soil profiles are shown next to the excavated underground structures.

Satellite image: Google Earth. Photos: P. Zigelski 2014

## II.2 Plant functional traits (PFT)

In order to evaluate morphological distances among phenotypes I conducted a PCoA with Euclidean distance measure and a one-way PERMANOVA, based on categories of plant functional traits (Cornelissen et al., 2003). Furthermore the Kruskal Wallis test was used to test for significant divergence in leaf area, leaf shape ratio and leaf petiole ratio among phenotypes. All traits used in these analyses that were performed in PAST v3 are listed in tab. 1.

**Table 1:** Traits and their respective bins used for plant functional trait analyses.

| **Traits** | **Bins** |
|---|---|
| Life form | tree; shrub; suffrutex |
| Height [cm] | < 50; 51 – 200; 201 – 1000; > 1000 |
| Leaf area [cm²] | < 50; 51 – 100; 101 – 150; 151 – 200; > 200 |
| Form of base | cordate; tapering |
| Form of apex | no; blunt; pointed |
| Leaf shape ratio (length:width) | < 1.00; 1.01 – 2.00; 2.01 – 3.00; > 3.01 |
| Leaf petiole ratio (length leaf:petiole) | < 5.00 ; 5.01 – 10.00; 10.01 – 15.00; 15.01 – 20.00; > 20.01 |
| Leaf texture | leathery/stiff; smooth/silky |
| Geoxyle | yes; no, unknown |

For this purpose all collected specimen as well as at least six leaves of each sampled population were measured in regard to their petiole length, leaf length, width and area. Other traits like approx. height and geoxylicity were recorded in the field or learned by personal communication with locals and colleagues who had been there. Measurements were grouped according to phenotype assignment. Herbarium specimens were excluded from these analyses due to insufficient information.

## II.3 Genetic analyses

### II.3.1 Sample preparation

The collected samples were first frozen for at least 3 days in order to kill any insects or other beings that could damage the plant tissue. Afterwards the material was examined with a microscope (47-50-52 9901 Zeiss, Oberkochen, Germany) and clean leaf pieces (with no damages or fungi) of about 1 $cm^2$ were cut out and processed further by powdering the material with a grinding mill (MM400, Retsch, Haan, Germany)

### II.3.2 DNA extraction

The standard extraction of DNA was performed according to Dumolin et al., (1995) but with slight modifications. The method of Dumolin itself is a modification of Doyle & Doyle (1987). The extraction buffer was provided with 50 mM DTT and 40 µg RNAse per sample. Lysis incubation was shortened to 30 min at 55°C and isopropanol incubation to 2 h at -20°C. DNA extraction from specimens of the Herbarium Hamburgense (HBG) of Universität Hamburg as well as from ISCED, Lubango, Angola (LUB) was conducted with the EZNA® HP Plant DNA Kit from Omega (Norcross, GA, USA) and followed the provided instructions. This was done with regard to the old age of the specimens, some being older than 100 years.

After DNA isolation, the yield was resuspended in 40 µL 1 x TE buffer and incubated at room temperature over night or for at 37 °C in a shaker at least one hour (BioSan, Riga, Latvia; Environmental Shaker-Incubator ES-20). The DNA concentration and purity was determined with a NanoPhotometer™ (IMPLEN, Munich, Germany) at 280 nm and working solutions of 20 ng DNA/µL were prepared.

### II.3.3 Sequence analysis

#### II.3.3.1 PCR

##### II.3.3.1a Internal transcribed spacer (ITS)

The internal transcribed spacer (ITS) is part of a nuclear gene operon coding for the ribosomal rRNA units 18S, 5.8S and 28S. The conserved sequences of these 3 units are separated by 2 highly variable spacer regions, ITS-1 and ITS-2. Both are commonly used for phylogenetic analysis (Baldwin, 1992). In my study I amplified the entire genome region via PCR with primers ITS-18F (5'-GTCCACTGAACCTTATCATTTAGAGG-3') (Käss & Wink, 1997, modified by Beyra-Matos & Lavin, 1999) and ITS-4R (5'-TCCTCCGCTTATTGATATGC-3') (White et al., 1990), yielding a fragment of 720 bp with both spacers and the 5.8S rRNA. They were ordered at Sigma-Aldrich (www.sigmaaldrich.com). Reaction conditions were as follows; using the my Budget HotStart DNA Polymerase Kit from BioBudget (Krefeld, Germany):

2 mM $MgCl_2$, 1x B2 Buffer, 0.2 mM each dNTP, 1 % DMSO (Carl Roth GmbH + Co. KG, Karlsruhe), 0.3 µM each primer, 0.25 U Hot-Start Taq, 10 ng DNA template in a 10 µL reaction volume.

Amplification was performed on a TProfessional Thermocycler (Biometra, Göttingen, Germany) with following conditions: Initial denaturation at 95°C for 15 min, for 39 cycles denaturation at 95°C for 30 sec, annealing at 50°C for 30 sec and elongation at 72°C for 1 min, followed by final elongationat 72°C for 10 min. PCR products were kept at 4°C or -20°C till further processing.

### II.3.3.1b *trnK/matK*

The chloroplast *trnK* region is a plastid gene coding for the tRNA of Lysin. The *matK* gene, coding for maturase K, is inserted inversely into its intron (Hausner et al., 2006). The *trnK* intron is highly variable and therefore usable for phylogenetic studies (Jeffrey et al., 1988; Dong et al. 2012). Primers were designed based on the chloroplast genome of *Syzygium cumini* available on Genbank (GQ870669.3) to amplify the region including spacers, yielding fragments of about 1346 bp. They were ordered at Sigma-Aldrich. Sequence of the forward primer matK-2F Syz is 5'-CAATACTCGTGGAGAAAGAAACG-3' and for the reverse one *trnK*-2R 5'-CAACGGTAGAGT ACTCGGCTTTTA-3' (Demesure et al., 1995). Reaction and cycling conditions were as follows, using the my Budget HotStart DNA Polymerase Kit from BioBudget (Krefeld, Germany):

2.5 mM $MgCl_2$, 1x B2 Buffer, 0.2 mM each dNTP, 3 % Trehalose (Carl Roth GmbH + Co. KG, Karlsruhe), 0.3 µM each primer, 0.25 U Hot-Start Taq, 10 ng DNA template in a 10 µL reaction volume.

Amplification was performed with following conditions: Initial denaturation at 95° for15 min, for 5 cycles denaturation at 95°C for 30 sec, annealing at 53°C for 1 min and elongation at 72°C for 2.5 min, followed by 30 cycles of denaturation at 95°C for 30 sec, annealing at 52°C for 1 min and elongation at 72°C for 2.5 min, with a final extension at 72°C for 10 min.

Since amplification for some samples failed, PCR was conducted with another primer pair: matK-4F Syz (5'-GACTATTCCAATACTCGTGG-3') and *trnK* 3914f (5'-GGGGTTGCTAACTCAACGG-3') (Gadek et al., 1996) with a concentration of 0.8 µM of each primer and 5 ng template DNA under same PCR conditions. For challenging samples a nested two-step PCR was conducted with matK-2F Syz/ *trnK*-2R as first primer pair and matK 4F Syz/*trnK* 3914f as second primer pair in the re-amplification. PCR products were kept at 4°C or -20°C till further processing.

### II.3.3.2 Gel electrophoresis

The success of PCR was checked by gel electrophoresis. Therefore, 1 µL PCR product was mixed with 2 µL loading buffer containing 0.25% bromophenol-blue and 30% glycerin and loaded on a 1 % agarose gel (0.5 x TBE) together with 2 µL of GeneRuler 1 kb™ size standard from Thermo Fisher Scientific (Waltham, MA, USA). A voltage of 70 V (Power Pac 300; Bio-Rad Laboratories, Karlsruhe, Germany) was applied and the gel ran for 30 – 40 min in an electrophoresis chamber (Mini-Sub® Cell GT; Bio-Rad Laboratories, Karlsruhe, Germany). Bands were visualized on a UV transilluminator (VWR Genoplex, Darmstadt, Germany) after incubation of each 5 – 10 min in a staining (75 µL 1% ethidium bromide in 250 mL deionized $H_2O$) and destaining bath (deionized $H_2O$), respectively.

### II.3.3.3 Purification

PCR products were then purified according to Werle et al. (1994) with Exonuclease I (Exo I) and thermosensitive Alkaline Phosphatase (FastAP™) in a combined clean up (Thermo Fischer Scientific). Thus, 5 µL PCR product, 0.5 µL Exo I and 1 µL FastAP™ was mixed on ice and incubated for 15 min at 37°, followed by 15 min at 85°C. The results were checked with gel electrophoresis as described above (II.3.3.2).

### II.3.3.4 Sequencing PCR

Subsequently, the purified products were amplified with fluorescent labeled ddNTP's using the BigDye™ Terminator Kit from Applied Biosystems with slightly modified manufacturer's protocol. This allows determination of DNA sequence by Sanger sequencing (Sanger et al., 1977). Each sample was sequenced forward and reverse, using the appropriate primer. Per 10 µL reaction mix were supplied:

2.6x Big Dye Sequencing Buffer, 0.1 µM one primer, 1 % DMSO (Carl Roth GmbH + Co. KG, Karlsruhe), 1x Big Dye Terminator Ready Reaction, diluted purified DNA template (dependent on amplification success; ranging from 1 to 1:5).

Reaction conditions in TProfessional Thermocycler (Biometra, Munich, Germany) were as follows: Initial denaturation at 96°C for 4 min followed by 40 cycles of denaturation at 96°C for 30 sec, annealing at 50°C for 15 sec and elongation at 60°C for 4 min. Light sensitive products were kept in the dark and frozen immediately when not processed further the same day.

### II.3.3.5 Precipitation and sequencing

To precipitate the products of sequencing PCR (Applied Biosystems, Sequencing chemistry guide, 2002), their total volume (10 µL) was provided with 1 µL of 125 mM EDTA, 1 µL of 3 M Na-Acetate and 25 µL of 99.8 % pure EtOH and thoroughly mixed. Reaction was incubated in the dark for 20 min at room temperature and subsequently centrifuged for 1 h at 13,000 rpm and 4 °. Supernatant was then carefully discarded and DNA pellet washed with 70 µL ice cold 76 % EtOH, followed by another centrifugation at 13,000 rpm and 4 °C for 1 h. Supernatant was again discarded and DNA pellet let dry until all alcohol evaporated. Tubes were then kept at 4 °C in the dark till further processing or pellet was directly redissolved in 20 µL Super-DI™ formamide (MCLAB, San Francisco, CA, USA) loading agent and incubated for at least 30 min at 37 °C. Samples were then ready for sequencing. The prepared samples were analyzed in a 3500 Genetic Analyzer (Applied Biosystems, Foster City, CA, USA; Hitachi, Tokyo, Japan) under the following conditions: with 8 capillaries of 50 mm length, POP7™ polymer, the dye set Z and the installation standard GeneScan™ Installation Kit DS-33. While running oven temperature was set to 60 °C, running voltage was to 8.5 kV, prerun voltage to 15 kV and injection voltage to 1.6 kV. Running time accounted for 5760 sec, pre-running time for 180 sec, injection time for 8 sec and there was 480 sec of data delay.

### II.3.3.6 Manual editing of sequences, alignment and phylogenetic tree construction

In total, 101 ITS and 104 *trnK* sequences were successfully produced (see tab. A1). *Syzygium cumini* served in both cases as an outgroup, its sequences were obtained from NCBI, Genbank: KF186456.1 (ITS) and GQ870669.3 (*trnK*). Thereafter ITS and *trnK* sequences of forward strand and reverse complement of reverse strand of each sample were aligned and corrected under inclusion of both chromatograms; gaps were filled accordingly using the software BioEdit (Hall et al., 1999) and FinchTV (Geospiza Inc., Seattle, WA, USA). Optimized forward sequences of 1 -3 individuals per population were then aligned with the ClustalW algorithm in MEGA v6 (Tamura et al., 2013) with default settings and edited such that each sequence began and ended at the same position. In the case of ITS some ambiguous positions occurred in samples that let to low genetic resolution. By specifying the ambiguous base if possible (under consultation of both forward and reverse chromatograms) the resolution could be improved if the ambiguity was weak and

appeared only in one chromatogram, otherwise this site was excluded. If both signals in the chromatogram were equally strong and lower than the average the position was kept ambiguous. Phylogenetic trees based on these alignments of ITS and *trnK* sequences were computed with Maximum Likelihood using the Tamura-3-Parameter substitution model with Gamma distribution and 5 discrete Gamma categories as well as 1000 bootstrap replicates, as this model described the data best according to the DNA models test in MEGA.

### II.3.4 Haplotype network analysis

Based on the *trnK* alignment a haplotype network analysis was with the software PopART (Leigh & Bryant, 2015). Each sample was assigned into one of these groups according to their respective phenotypes: SGB, SGBx, SGH, SC, SGM, SGMx, ST or SGG. Some samples from the herbarium were furthermore assigned to the categories *S.g.littorale*, *S.g. staudtii* and *S.benguellense*. Specimen of the taxa *Syzygium guineense* ssp. *angustifolia* and *Syzygium guineense* ssp. *latifolia* were put into the group ST and SGM, respectively. With these settings a medium joining network with ε = 0 was drawn (Bandelt et al., 1999). Inbuilt statistics of Tajimas D (Tajima, 1989), AMOVA (Excoffier et al., 1992), nucleotide diversity (Nei, 1972) and number of segregating and parsimony informative sites were used. In this analysis 105 aligned sequences were included (see Tab. A1).

### II.3.5 Microsatellite analysis

#### II.3.5.1 Fragment amplification and size determination

Short sequence repeats (SSR) or microsatellites are a variable number of iterations of mono- to hexanucelotides, leading to a fragment of varying length within a genome region (locus). The high variability in the number of repeats is a result of local mutations and mismatch errors of the DNA polymerase during strand replication. Due to this high polymorphism even between individuals of the same species, SSR's are commonly used to determine genetic divergence based on fragment length differences at one or more loci (Zietkiewicz et al., 1994). Microsatellite loci from related species *Syzygium sayeri*, 8 loci, (Hillyer et al., 2007) and *Syzygium paniculatum*, 14 loci, (Thurlby et al., 2011) were tested for applicability in *Syzygium guineense* variants. Multiplex PCR of loci, grouped according to sufficiently different product fragment sizes, was conducted with unlabeled primers. PCR conditions were as follows:

2 mM MgCl2, 1x B Buffer (my-Budget Taq DNA Polymerase Kit, Bio-Budget), 0.2 mM each dNTP, 0.3 µM each primer, 0.25 U Taq (my-Budget Taq DNA Polymerase Kit, Bio-Budget), 5 – 10 ng DNA template in a 10 µL reaction volume.

Amplification was performed with following conditions: Initial denaturation at 95°C for 4 min, followed by 35 cycles of denaturation at 94°C for 30 sec, annealing at 50°C for 30 sec, elongation at 72°C for 1 min and a final elongation at 72°C for 10 min.

Seven successfully amplified loci were then chosen for further analysis; those were SAY-1E4, SAY-3F10, SAY-6C10, SAY-6NE and SP38BGT, SP75BGT, SP85.1BGT (see tab. 2). These loci were amplified in two multiplex PCR's and two duplicates per sample with the same PCR conditions as above, but with Hex and 6-Fam labelled primers. At least two individuals of each population were analyzed plus all herbarium specimens available. For challenging samples a second PCR was conducted with the product of the first PCR as template ("reamplification"). The final products were diluted 1:30 and analyzed with the ABI to determine fragment sizes under the following conditions (devices were the same as in II.3.3.5 but with dye set G5). While running oven

**Table 2**: Seven microsatellite loci used for population genetic screening in this study, with primer sequences, repeat motifs, fragment sizes as well as multiplex combination and reference.

| Locus | Primer (5' - 3') | Motif | Fragment size (bp) | Multiplex | Source |
|---|---|---|---|---|---|
| SAY1E4 | F: 6fam-GCCGATGAATCAGGTCTTGAG | $(TGG)_8$ | 115 - 135 | A | Hillyer et al., |
| | R: GGCAAAGGGTACGCCTTCATC | | | | |
| SAY3F10 | F: hex-GCGTCCGAAGCTGGCCACCTC | $(TGC)_5$ …$(GGCGAC)_4$ …$(GAC)_5$ | 170 - 190 | A | Hillyer et al., |
| | R: CTGCATCTCCAAATCGCGAGC | | | | |
| SAY6C10 | F: hex-GAGGAGCAAACTACATGCTGC | $(GA)_{27}$ | 105 - 135 | A | Hillyer et al., |
| | R: GCATGGTCCCTGTGTGCAAGG | | | | |
| SAY6NE | F: hex-GTGAGTCAGTTGGGCTGTCTC | $(TC)_{18}$ | 135 - 145 | B | Hillyer et al., |
| | R: TCAAAGACTCAAAGTGGAAGC | | | | |
| SP38BGT | F: hex-GTCTCCTTCATGCTCCCTAT | $(AG)_8$ | 300 - 310 | B | Thurlby et al., |
| | R: AGTGCCTGTGCTTCACCCTC | | | | |
| SP75BGT | F: 6fam-GTCTTCTCAGAGTAGATATGGA | $(AG)_3$ GGG $(AG)_{17}$ | 115 - 135 | B | Thurlby et al., |
| | R: GAGGCACGAGAGGAGTTTC | | | | |
| SP85.1BGT | F: 6fam-AGGAGTGGATTCGGAGT | $(AG)_6$ A $(AG)_6$ | 235 - 250 | B | Thurlby et al., |
| | R: ACAGTGGAGCAAAGAAGCAC | | | | |

temperature was set to 60 °C, running voltage to 19.5 kV, prerun voltage to 15 kV and injection voltage to 1.6 kV. Running time accounted for 1330 sec, pre-running time for 180 sec injection time for 8 sec and there was 1 sec of data delay. 115 individuals were successfully genotyped and included in the simple sequence (SSR) analyses (see tab. A1 in the appendix).

### II.3.5.2 Data analysis

Subsequently, raw data was revised and edited with GeneMapper v5 software (Applied Biosystems, Inc., Foster City, CA, USA). For each marker (locus) an automatic binning and allele calling was used and if failed it was manually adjusted. In the case of split peaks, the higher peak was called or, if both were of the same height, the allele was kept undetermined and marked with a "?" (Applied Biosystems, Fact sheet microsatellite analysis, 2004). Stutter artifacts were dismissed. If allele calling was uncertain for one or both duplicates, the PCR and fragment separation was repeated. Only those samples were used in further analyses where all seven loci were successfully amplified and fragment sizes (of the alleles) were confirmed by at least two duplicates.

The samples and all detected alleles were arranged in a presence/absence matrix where the rows code for each sample which alleles of each locus are present (1) or absent (0). This table was the basis for Principle Coordinates Analysis (PCoA), one-way PERMANOVA test and Hierarchical Clustering with neighbor joining algorithm in PAST v3.06 (Hammer et al., 2001). Furthermore, the 0-1 matrix served the software Alleles In Space v1.0 (Miller 2005) to conduct Mantel tests (Smouse et al., 1986) and Spatial Autocorrelation (Sokal & Oden, 1978) in order to test for genetic and spatial dependencies.

Using the package *polysat* in R (Clark & Jasieniuk, 2011) I furthermore constructed a genotype matrix that I subsequently used in SPAGeDi v1.5a (Hardy & Vekemans, 2002). This software calculates multiple genetic statistics like fixation index $F_{ST}$ (Wright, 1965), gene diversity (Nei, 1978), allelic richness (Hardy & Vekemans, 2002), inbreeding coefficient (Hardy & Vekemans, 1999), selfing rate estimators (Hardy & Vekemans, 1999) and heterozygosity (Nei, 1978). Using package *pegas* and *adgenet* in R (Jombart, 2008), simple and nested AMOVA tests were conducted with grouping by phenotype, habitat and sampling region.

Finally, population genetic analysis was conducted with Bayesian approach as implemented in STRUCTURE (Pritchard et al., 2000). The burnin period had 30,000 replicates followed by 100,000 MCMC replicates. Population information was given in form of assigning each sample to its respective phenotype. The fitting was calculated for K = 1 to K = 10 with 3 repeats per K. The no-admixture model was implemented, meaning that individuals are assumed to originate from one population/phenotype only (models with assumed admixture were run but led to the same results. They are therefore not covered here). Allele frequencies were assumed to be independent. The results were interpreted with the Structure Harvester (Earl & von Holdt, 2012) and the most probable K chosen for further analysis. Due to no significant genetic differences between the suffrutices from Miombo and Western Zambezian Graslands (see III.4.2) I pooled both groups and treated them as one for subsequent analysis. Some individuals of ambiguous phenotype assignment that showed strongly deviating genotypes within their phenotypic group were reassigned to the other phenotype, for example, in the case of B01 (SGB) being reassigned to the SGG phenotype. Reassigned individuals changed the ID code, a lower case letter after the number indicates the new phenotype (B01 -> B01g). All statistics and analysis had therefore been

redone with the new phenotypic assignment. Visualization of the results was enhanced using Microsoft Excel, PAST and package *rgl* in Rstudio v0.98.501 (R v3.2.2).

## II.4 Factor Analysis for Mixed Data (FAMD) and Hierarchical Clustering on Principle Components (HCPC)

A combined profile of morpho-genetic-ecological traits was created for each population in complement to the separated genetic and morphometric characterization. Populations where two haplotypes occurred (e.g. Sgg1b) were duplicated and implemented with the respective haplotype (Sgg1b and Sgg1b_i). Climatic variables such as mean annual temperature or precipitation were derived from WorldClim database via QGIS 2.10 Pisa (Hijmans et al., 2005). Other parameters like height or soil/substrate were estimated by personal communication with locals and colleagues who know the sites and field observation. Herbarium specimens were excluded from this analysis due to unknown exact sampling location and insufficient data. Factors included in this analysis are both categorical and continuous and are listed in tab. 3. Using package *FactoMineR* in Rstudio v0.98.501 (R v3.2.2) I performed a Factor Analysis of Mixed Data (FAMD) with the phenotype as response in order to elucidate by which factors it was influenced (Pagès, 2004; Lê et al., 2008). Beyond that I conducted the FAMD again with excluded phenotypes to see how populations correlated to factors in an ordination space. Additionally a Hierarchical Clustering on Principle Components (HCPC) was conducted on the same dataset (Husson et al., 2010). This method ordinates samples according to their respective pairwise distances calculated from the FAMD. Also, a Pearson test for correlation from package *corrgram* was used to assess and visualize factor correlations.

**Table 3:** Factors and their categories or ranges used for FAMD in this study.

| Factors | Type | Range |
|---|---|---|
| Latitude [°] | continuous | -18.83 - -11.48 |
| Longitude [°] | continuous | 13.20 - 25.86 |
| Mean annual temperature [°C] | continuous | 15.5 – 24.1 |
| Mean temperature range [°C] | continuous | 17.3 – 28.9 |
| Annual precipitation [mm] | continuous | 459 – 1301 |
| Precipitation seasonality | continuous | 87 – 117 |
| Soil/substrate | categorical | Sand; ferralitic; bank; granite |
| *trnK* Haplotype | categorical | HT1 – HT7 |
| SSR: attribution to Cluster 1 | continuous | 0 – 1 |
| SSR: attribution to Cluster 2 | continuous | 0 – 1 |
| SSR: attribution to Cluster 3 | continuous | 0 – 1 |
| Leaf area [cm²] | continuous | 28.41 – 219.23 |
| Leaf shape ratio (length:width) | continuous | 1.50 – 3.03 |
| Leaf petiole ratio (length leaf:petiole) | continuous | 4.64 – 68.08 |
| Leaf texture | categorical | Stiff/leathery; smooth/silky |
| Geoxyle | categorical | Yes; no; unknown |
| Height [cm] | categorical | < 50; 51 – 200; 201 – 1000; > 1000 |
| Life form | categorical | Suffrutex; shrub, tree |

The entire R script is supplemented in the appendix (VI.3).

# III. Results

## III.1 PFT analysis

The analysis of morphological and functional traits of each population revealed a pattern that clearly separated the phenotypes from each other and simultaneously showed which phenotypes resemble one another (see fig. 13). The analyzed traits had been: life form, height, leaf area, leaf shape and texture characteristics and geoxylicity. Altogether was the difference between phenotypes significant according to one-way PERMANOVA (sum-of-squares$_{total}$ = 261.3, sum of squares$_{within\text{-}group}$ =98.49, F = 7.714, $p_{same}$ < 0.001). The pairwise comparison between phenotypes revealed where the significant differences lay exactly, as can be seen in table 4. The phenotypes grouped generally speaking into two (three) clusters. One contained the suffrutices from Angola and Zambia and the SGMx; they are on the left side of the plot. Opposed to them was the group of remaining tree phenotypes, whereby SB, SGG and SGB were very close and overlapping; ST, SGM and SGBx were set apart into the upper part of the plot and SC into the lower right corner. SGBx and ST were not displayed as polygons but as lines, due to low n. All in all we see a height gradient from left to right along the first axis and a leaf shape/petiole length gradient from bottom to top along the second axis.

**Table 4:** Results of the PERMANOVA, showing differences between phenotypes based on morphological features. Significant differences of p < 0.05 are highlighted in red. P values are Bonferroni-corrected.

| | **$SGH_{WZG}$** | **SGM** | **SGMx** | **SGB** | **SGBx** | **SGG** | **SC** | **ST** | **SB** |
|---|---|---|---|---|---|---|---|---|---|
| **$SGH_{Mio}$** | 1.000 | 0.005 | 0.014 | 0.005 | 0.734 | 0.005 | 0.036 | 0.729 | 0.153 |
| **$SGH_{WZG}$** | | 0.045 | 0.036 | 0.009 | 0.959 | 0.032 | 0.086 | 1.000 | 0.342 |
| **SGM** | | | 0.329 | 0.054 | 1.000 | 0.320 | 0.333 | 1.000 | 0.887 |
| **SGMx** | | | | 0.027 | 1.000 | 0.315 | 0.414 | 1.000 | 0.711 |
| **SGB** | | | | | 1.000 | 1.000 | 0.081 | 1.000 | 0.230 |
| **SGBx** | | | | | | 1.000 | 1.000 | 1.000 | 1.000 |
| **SGG** | | | | | | | 0.315 | 1.000 | 1.000 |
| **SC** | | | | | | | | 1.000 | 1.000 |
| **ST** | | | | | | | | | 1.000 |

As could be expected, the two SGH groups did not differ significantly and neither did SGM from the intermediate SGMx and SGB from SGBx. The latter, together with ST and SB, differed from none significantly; however one has to be careful since these phenotypes are shared by only two/three populations. This might be insufficient for significant calculations. As mentioned before, SGMx is an intermediate phenotype of SGH-Mio and SGM. In this analysis it differed significantly from SGH but not from SGM. Moreover there were morphological differences between the suffrutices and tree phenotypes SGB, SGG and SC, in decreasing significance.

Regarding the NJ tree in figure 14 we see that its basal group contains the populations of ST (Angolan Escarpment) and Sgg678 from Bicuar NP, Angola (approx. 170 km apart). In the higher branch it splits into two main groups, one containing all SC's, SB's, SGB's and SGG's (except Sgg678), the other the remaining populations of phenotypes SGH (from both biomes), SGM, SGMx and SGBx. There is an overall low bootstrap support, particularly in the more basal

branches, but also some significant values for terminal clades (e. g. Sc2 and Sc4). Basal to these groups is a clade. The next higher clade on the right side comprised all SC's plus the reclassified SB's Sc23751o and Sgg2809o (from the same location as Sgg2808o) on one branch and the remaining SGG's and all SGB's on the other.

The following branches were constituted by almost one phenotype each. SGM dispatched at the lower node except for Sgm 23691, with clustered together with the SGBx's. On the two next higher plains clustered SGMx, except for Sgh13 which clustered together with the SGH populations from the Miombo and the Western Zambezian Grasslands. The Cusseque populations Sgh8 and Sgh9 were slightly differentiated from Escarpment populations Sgh16 and Sgmx657. The terminal clade was formed by suffrutices from both Angola and Zambia mixed, though one branch contained more Angolan, the other more Zambian populations.

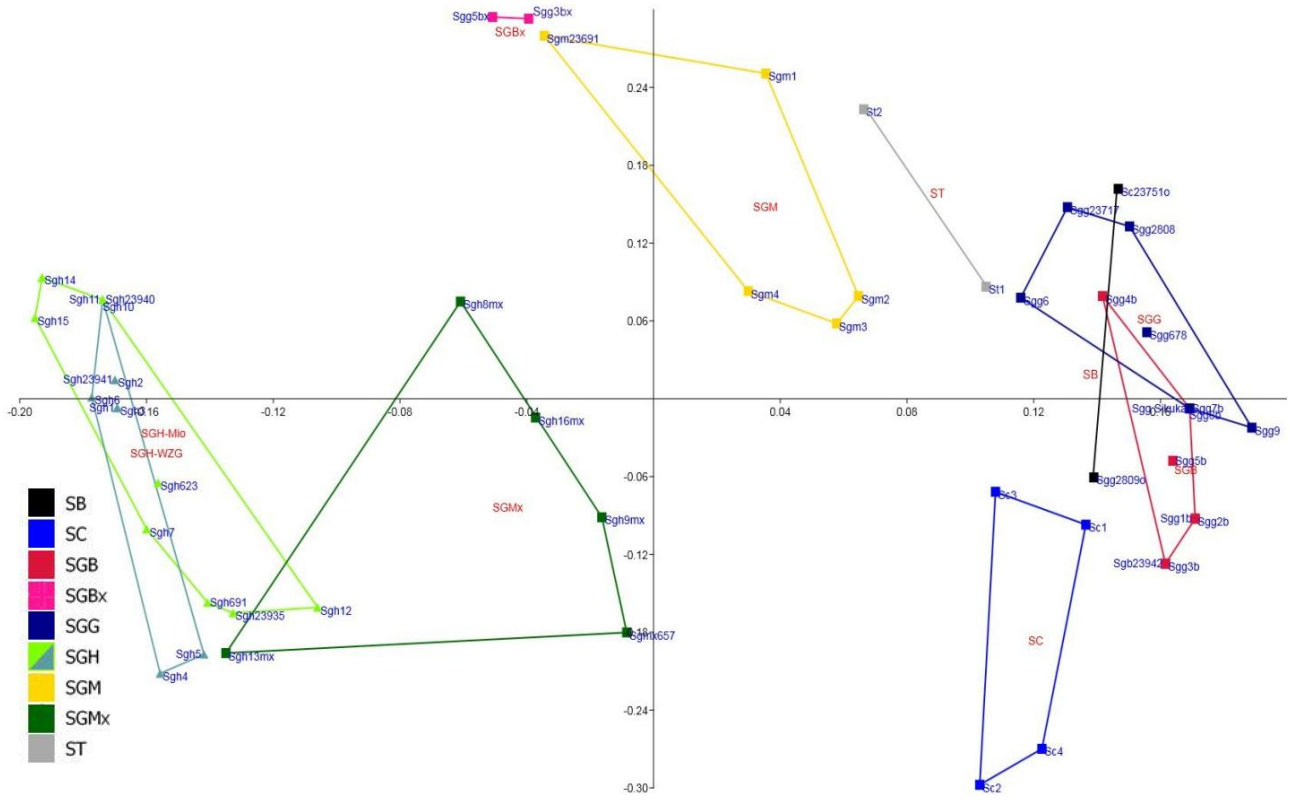



**Figure 13** PCoA of morphological and functional traits of *Syzygium* populations, variance explained by axis 1:35.35%; by axis 2: 11.27%. Phenotypes are indicated by different colors.

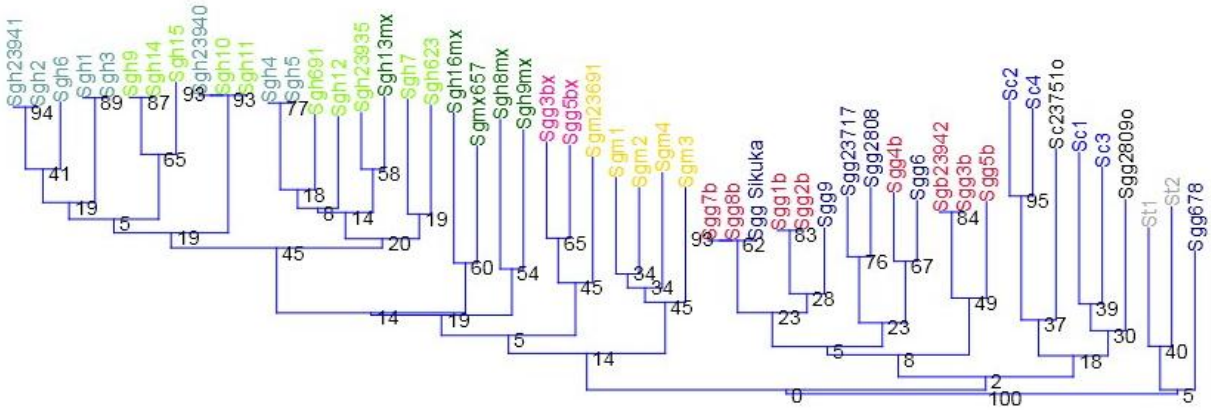


**Figure 14:** Neighbor-joining clustering of *Syzygium* populations. Colors indicate phenotypes as in fig. 13; numbers next to branches indicate bootstrap support in 1000 replicates.

## III.2 Phylogenetic relationship

### III.2.1 ITS

The Maximum Likelihood tree constructed upon the ITS alignment is evident in fig. 15. The analysis involved 101 nucleotide sequences of *Syzygium* specimen and one outgroup, *S. cumini* there were a total of 655 positions in the alignment. The phylogram was characterized by many unresolved clades as well as branches with low bootstrap support for many basic and terminal branches. This was due to ambiguous basepairs at heterozygous sites in some individuals. Bootstrap values range from 45 to 95.

At the base which was rooted with *Syzygium cumini*, the tree divided into two branches each with low 61% bootstrap support. The first mentioned clade (a) comprised a polytomy with mostly SGM and SGMx specimen from Central and West Angola (Benguela, Bíe, Huíla provinces), together with SGH individuals from near Kuito, Bíe in Central Angola (12.1% of all samples). Only Mx02 and Mx03 from Tundavala, Angola, are further separated from the remaining members of this group, however with only 66% support. The other branch (b)-(c)-(d)-(e) shows a big polytomy with all residual samples and is quite mixed in spatial and phenotypical regard. The clusters (c) (6.6% of all samples) and (d) (3.3% of all samples) which separate from this group have very low bootstrap support (46% and 45%, respectively) that does not allow any conclusions and only (b) had a significant support of 95% and contained SGG individuals from the Okavango River (5.5% of all samples). The last group (e) (71.4% of all samples) was a vast conglomeration of the remaining individuals of all phenotypes and with a wide geographic and environmental distribution.

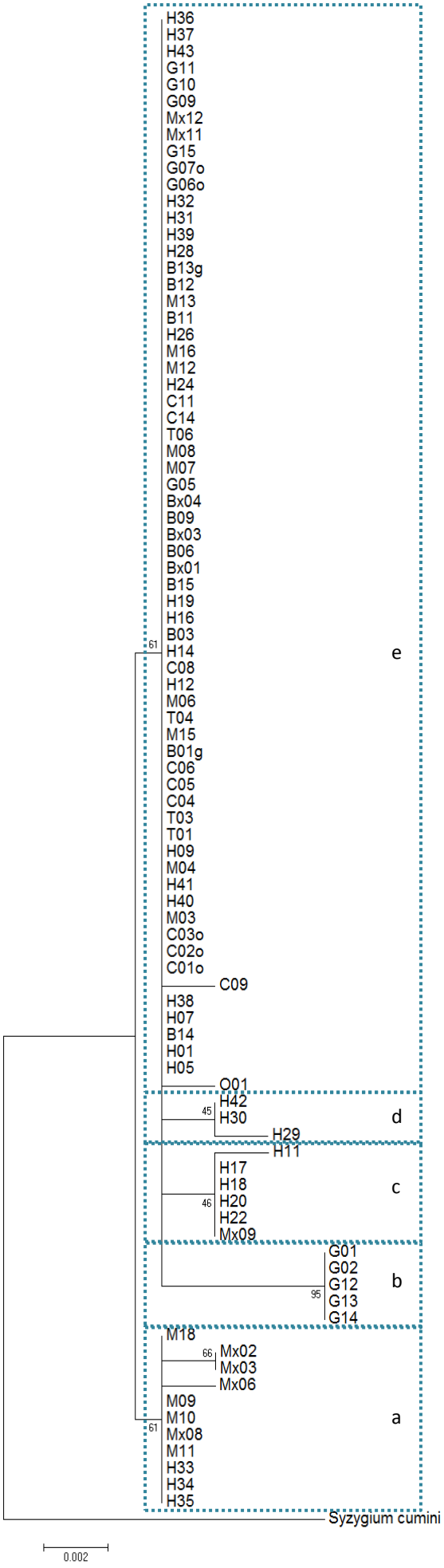



**Figure 15.** Molecular phylogenetic analysis by Maximum Likelihood method based on ITS sequence. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. Not all sequences used in the analysis are shown in order to improve readability; all leftout samples belonged to group (e). Numbers at bifurcations indicate bootstrap support in 1000 replications and rectangles indicate groups mentioned in the text.

### III.2.2 *trnK*

For this analysis, 105 *trnK* sequences of 1147 positions were used in the alignment. Compared to the previous ITS phylogram, the *trnK* phylogenetic tree was more structured, though this ML tree was characterized by polytomies with low bootstrap supports that does not exceed 67% (fig 16). Again, the tree is rooted with *Syzygium cumini* as an outgroup.

The phylogram branched at its base into a polytomy with three groups and two single ST specimen from Angolan escarpment (T11, population St1) and from Tanzanian mountains (T04; 1.9 % of all samples). The group (a) represents a polytomy (61 % support, 9.5% of all samples) of individuals from the foot (C09-C11, C14; population Sc2) and the top (T01-03, T06, T12-13; populations St1, St2) of the Angolan Escarpment. Yet another polytomy ((b), 57 % support, 7.6% off all samples) contained eight SGB individuals (B02, B07, B09, B15-16; Sgg1b, Sgg4b, Sgg5b) from Liuwa, Zambia, SB from Cuando Cubango province, Angola (C01o-02o; Sc23751o) and SC (C08; Sc4) from the Caprivi Strip, Namibia. These are characterized and set apart by an indel effect in their sequences that led to the deletion of a palindromic 20 basepair fragment. These two minor groups (a) and (b) plus the single individuals encompassed 19% of all samples.

The remaining branch (c)-(d)-(e)-(f) (62 % support) included most samples (80%) and was not resolved aside from three sub-branchings (d)-(e)-(f). At its base, group (c) gathered individuals from almost all phenotypes except ST and SGBx in a polytomy. Even herbarium specimens from other *Syzygium* varieties/species cluster here, e. g. *S. g. littorale* (SGL Herbar) or *S. g. staudtii* (SGS Herbar. Furthermore, the geographical range of this genetic type is enormous: From DR Congo and Cameroon in the North to Southern Angola as southern boundary, and from the Escarpment in the West to the Barotse Plains, Zambia in the East. Only the extent of ITS's unresolved central group is greater, see previous chapter.

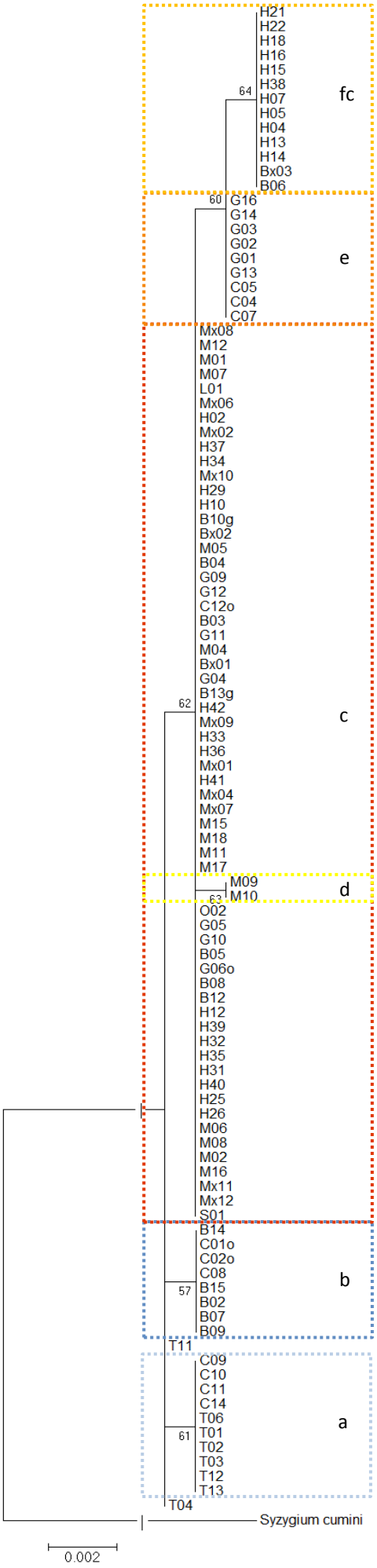



**Figure 16:** Molecular phylogenetic analysis by Maximum Likelihood method on *trnK* sequence alignment. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site numbers at bifurcations indicate bootstrap support. Rectangles highlight genetic clades. Their color hints at the haplotypes in the next chapter.

Within this central group (c) a minor polytomy (d) (63 % support, 1.9% of all samples) of only two SGM individuals from the same population (M09-10; Sgm2) sampled in Benguela province, Angola, was separated. Another group with 22 individuals was structured in two unresolved groups. The basal polytomy (e) (60 % support, 8.6% of all samples) consisted of SC and SGG populations (Sc1 (C07), Sc3 (C06) and Sgg9 (G01-03, G13-14) from the Okavango Delta and River, as well as one herbarium specimen from Lake Lutamba, Tanzania (G12). Finally there is terminal, unresolved branch (f) (64 % support, 12.4% of all samples) that constituted mostly of SGH individuals from the Western Zambezian Grassland biome. However, also the SGBx individuals clustered here. They inhabit similar and/or adjacent habitats as do the SGH's and SGB's.

## III.3 Haplotype network analysis

Since ITS follows another inheritance[1] as it is part of a nuclear gene, it cannot be used in haplotype network analyses. Therefore the following chapter will cover the *trnK* network only. The *trnK* intron is a haploid chloroplast genome region and hence linearly inherited. Altogether the *trnK* haplotype network analysis mirrors the results of the Maximum Likelihood phylogram (see III.2.2).

The haplotype network analysis of *Syzygium guineense* s.l. phenotypes revealed seven haplotypes (table 5), each differing only by one mutation from its neighbor and separated strongly from *S. cumini*. As to be seen in figure 17, the seven haplotypes are mainly linearly connected with only one reticulation, abifurcation at the left branch. On the one hand it is evident that almost all haplotypes were shared by more than one phenotype. Only some ST and SGM individuals constituted a private haplotype. On the other hand most phenotypes were not restricted to one haplotype but shared several. Exceptionally from the latter, all individuals of SGMx and SGBx share only one haplotype, HT1, as well as the herbarium specimens SGlittorale and SGstaudtii.

**Table 5**: Overview of identified haplotypes in *Syzygium* phenotypes and their abundance.

| Haplotype | Description | % of samples | Phenotypes |
|---|---|---|---|
| HT1 | Central | 57.1 | SGG, SGH, SGB, SGM, SGMx, SC, SB |
| HT2 | Okavango | 8.6 | SGG, SC |
| HT3 | Western Zambezian Grasslands | 12.4 | SGH, SGBx |
| HT4 | SGM - Benguela | 1.9 | SGM |
| HT5 | Escarpment a | 9.5 | SC, ST |
| HT6 | Escarpment b | 1.9 | ST |
| HT7 | Indel | 7.6 | SGB, SC, SB |
| (HT8) | *S. cumini* - NCBI | 1.0 | Outgroup |

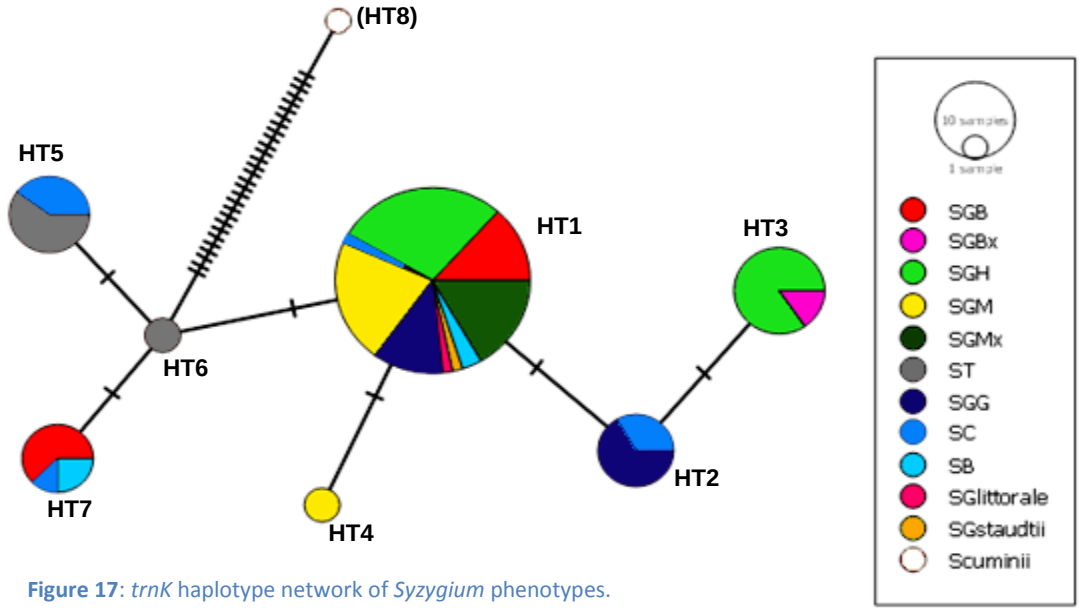



**Figure 17**: *trnK* haplotype network of *Syzygium* phenotypes.

1 Normally a diploid inheritance, but in this case, as will be revealed later on, it is even a tetraploid inheritance.

The central haplotype HT1 constituted the highest phenotypic diversity as it was shared by almost all phenotypes, except SGBx and ST. Also, many environmentally diverging populations from rivers, grasslands, woodlands or the escarpment were united in this genetic type. In addition this haplotype is widely distributed over southern, central and eastern Angola and western Zambia; even herbarium specimens from DRC or Cameroon share HT1 (see fig. 18).

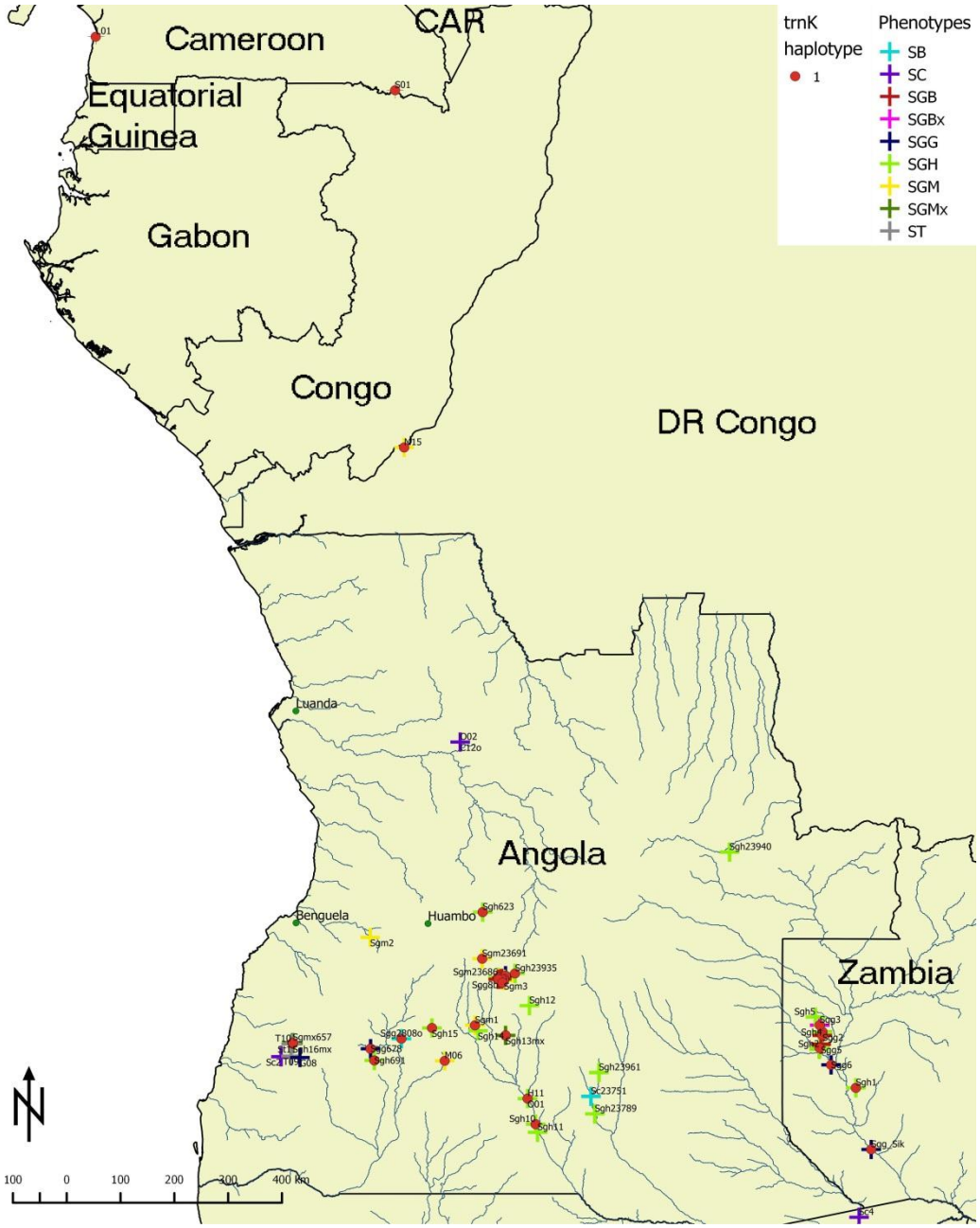



**Figure 18**: Map of haplotype distribution; distribution of trnK-HT1. Phenotypes are coded with colored crosses

Found almost exclusively at the Okavango River and Delta in Namibia and Botswana, HT2 is a haplotype shared by trees of the SC and SGG phenotype (fig. 19). Their environments are comparable and the populations are geographically close (Sgg9, Sc1, Sc3). However, this haplotype occurred also in one herbarium specimen from Lake Lutamba, Tanzania (fig. 20). However, information about the environment there is lacking.

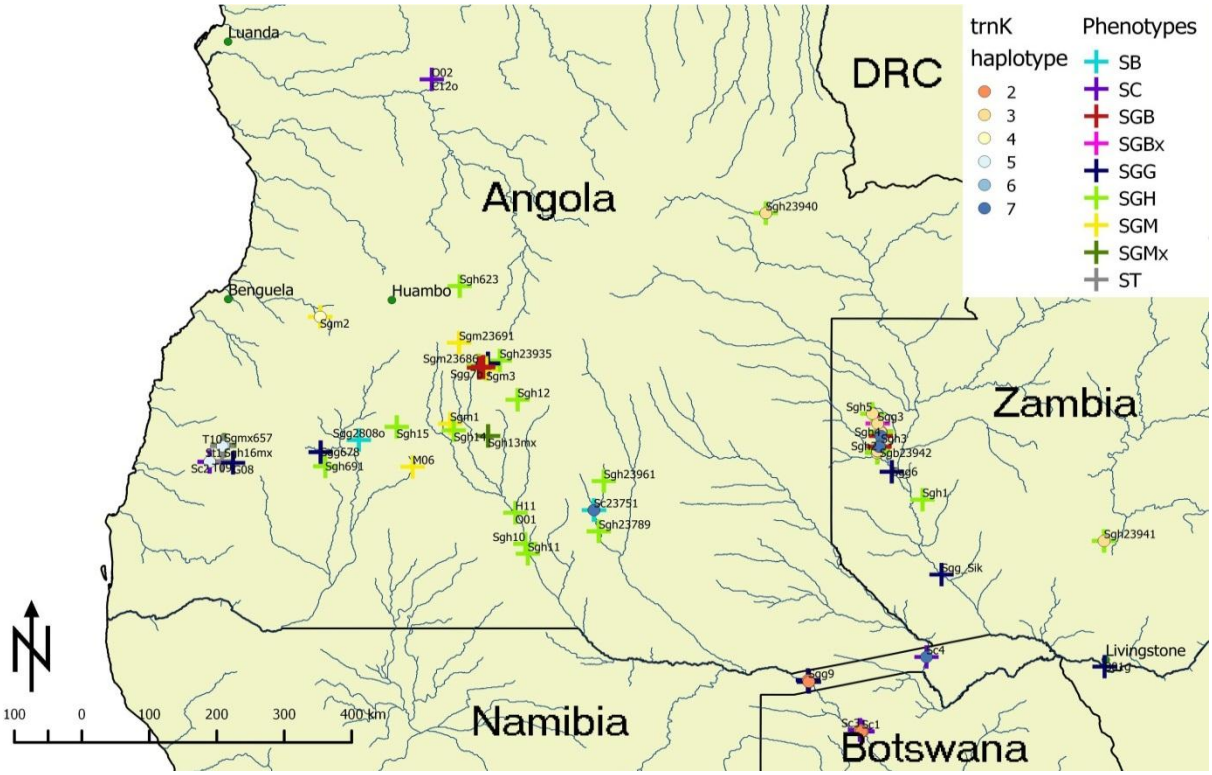



**Figure 19:** Distribution of haplotypes HT2 to HT7; western range Angola, Namibia, Botswana and Zambia. Colors indicate pheno- and haplotypes.


**Figure 20:** Distribution of *trnK* haplotypes HT2 to-HT7; eastern range Tanzania. Colors indicate pheno- and haplotypes.

Haplotype HT3 was found mostly in suffrutices (SGH phenotype) from the Western Zambezian Grasslands biome (fig. 19). Populations from Liuwa and Kafue National Park in Zambia as well as from Moxico province in eastern Angola, all belonging to this biome, share this genetic structure. But additionally another phenotype exhibited this haplotype, found until now only at Liuwa Plains National Park: SGBx. Being of intermediate appearance between SGH and SGB, it is closer related to SGH. Interestingly, this suffrutex dominated haplotype is not linked directly to the

other haplotype found in Angolan suffrutices. Instead it is bridged by the Okavango haplotype HT2, constituted by tree phenotypes. Reasons for this observation will be discussed later on.

Directly connected to the central haplotype is one that occurs only in a single population (Sgm2) of SGM phenotype from Benguela province in central Angola (HT4, see fig. 19).The remaining SGM's shared HT1. The next haplotype, HT5, occurred in individuals from the Angolan escarpment near Lubango. It was found in three populations from the Escarpment's hot and humid base (Sc2, SC phenotype in river) as well as from its colder and windy top (St1 and St2, ST phenotype between boulders and in crevices). Though differing strongly in terms of morphology and environment, they shared the same haplotype. Interestingly, the third phenotype occurring in immediate neighborhood atop the escarpment (SGMx), was characterized by a different haplotype, namely HT1.

Linking HT1, HT5 and HT7 is HT6, a haplotype that was found in only two samples with ST phenotype. Strikingly those two samples are from mountainous habitats from the opposite sites of the southern African continent: one from Angolan Escarpment (St1), the other from Tanzania (herbarium specimen T04, see fig. 20). This haplotype is furthermore the bridge and therefore closest relative to the outgroup *S. cumini*.

HT7, a haplotype found in some individuals from SC, SB and SGB phenotype, is a special case. In the SGB phenotype it occurred only in some individuals of the affected populations (Sgg1b, Sgg5b) in the Liuwa Plains in Zambia. The other individuals from Sgg1b and Sgg5b shared HT1. In SC, it was shared by one individual from Namibia and one SB population in southern Angola shared it, too. HT7 is characterized by an indel event where a 20 bp fragment had been deleted. This fragment contains a palindromic sequence, leading to the formation of secondary structure that could have been cut out (pers. comm. Barbara Rudolph). Although all samples in the study contain this palindromic sequence, it had only been cut in some individuals. Nevertheless, this indel haplotype was not formed exclusively by the deletion event since the exclusion of the 20 bp sequence from analysis lead to the same result, meaning that these individuals exhibit additional genetic divergences.

Looking at the statistics with excluded outgroup we see that the impression of low genetic divergence is confirmed. Within 1147 bp only 6 sites were segregating and parsimony informative which makes 0.05%. This is also highlighted by the low nucleotide diversity of $\pi = 1.15$. Tajimas D statistics however amounted to a significant strong D = 2459.3. The AMOVA statistic explained the variance within and among the phenotypes (tab. 6). We see though an almost even amount of genetic variation within phenotypes as well as between them, the latter having slightly higher values (53.6% to 46.3%, respectively). The fixation index $\Phi_{ST}$ emphasizes this variation as it accounts for $\Phi_{ST} = 0.464$.

**Table 6**: Results of the AMOVA using haploid chloroplast sequences.

| AMOVA | | df | Sum of squares | Variation [$\sigma^2$] | Variation [%] | Fixation [$\Phi_{ST}$] | Significance [p] |
|---|---|---|---|---|---|---|---|
| **Variance** | Among | 10 | 73.940 | 0.640 | 46.37 | 0.464 | < 0.001 |
| | Between | 93 | 68.819 | 0.740 | 53.63 | | |
| | Total | 103 | 142.760 | 1.380 | 100 | | |

## III.4 Microsatellite analysis

### III.4.1 General overview, heterozygosity, inbreeding and genetic fixation

Overall 115 individuals were included in this analysis: 37 SGH, 14 SGM, 12 SGMx, 10 SGB, 4 SGBx, 12 SGG, 8 SC, 10 ST and 8 SB. The analysis of seven SSR loci revealed that most examined *Syzygium* phenotypes and populations are assumed tetraploid as up to four alleles per locus were found. The exception could be phenotypes SGB and ST, where no individual had more than two alleles at one locus. Figure 21 displays average allele number per locus and SGB and ST had significantly lower values than the other phenotypes ($X^2 = 59.66$, $p < 0.001$). Allelic richness (AR) and effective allele number (NAe) were also reduced for ST, but not for SGB ($AR_{ST} = 2.09$, $AR_{total} = 2.67$; $NAe_{ST} = 2.94$, $NAe_{total} = 5.24$). Furthermore a high degree of polymorphic alleles was detected at all loci. Only one monomorphic allele from locus sp75BGT was present in all samples. Table 7 gives further information about allele counts, genetic diversity and heterozygosity at each locus. Loci with highest allele counts had also the highest heterozygosity and vice versa; the genetic diversity however did not correspond linearly to allele counts.

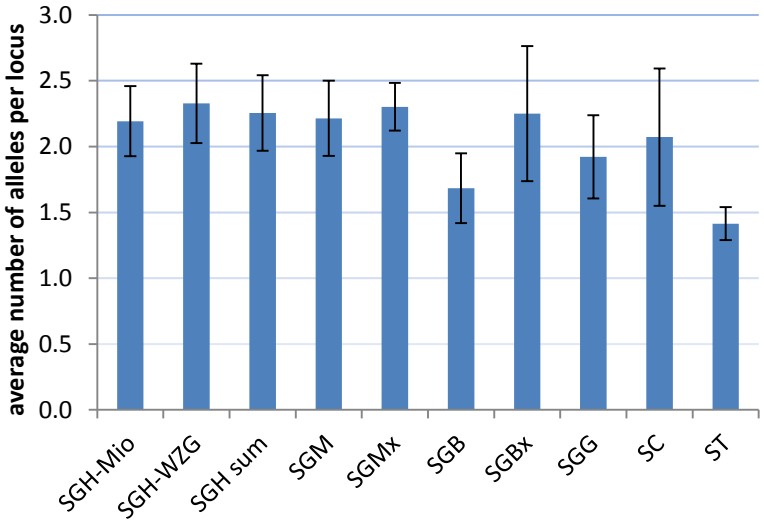



**Figure 21**: Mean number of alleles per locus shown for all phenotypes. Values for SGH are shown separately for the biomes as well as combined. SGB and ST have significantly lowered values compared to the rest which show c. 2 alleles per locus.

**Table 7:** Overview of microsatellite locus characteristics. $N_A$ = number of alleles, $H_e$ = gene diversity corrected for sample size, $H_o$ = heterozygosity.

| Locus | $N_A$ | $H_e$ | $H_o$ |
|---|---|---|---|
| **SAY-1E4** | 7 | 0.422 | 0.377 |
| **SAY-3F10** | 19 | 0.859 | 0.922 |
| **SAY-6C10** | 15 | 0.916 | 0.530 |
| **SAY-6NE** | 7 | 0.589 | 0.757 |
| **SP38BGT** | 7 | 0.746 | 0.868 |
| **SP75BGT** | 12 | 0.505 | 0.557 |
| **SP85.1BGT** | 17 | 0.868 | 0.870 |
| **Total/mean** | **84** | **0.701** | **0.697** |

Turning now to the statistical characterization of the phenotypes, there is a contrast to the findings of the haplotype network analysis (see III.3). In the former analysis the fixation index and $F_{ST}$ analogon $\Phi_{ST}$ was used ($\Phi_{ST} = 0.464$), but the genetic fixation according to microsatellites was rather low between phenotypes. As evident from table 8 we see that most pairwise $F_{ST}$ values were less than 0.1 (mean: 0.06) and only ST showed slightly higher fixation against SGB, SGBx, SC and SB.

Other statistics calculated on phenotype basis are shown in table 9. As mentioned before, though allelic richness was more or less even throughout the phenotypes, ST was poorer in allelic richness (AR), as well as in gene diversity ($H_e$) and heterozygosity ($H_o$). For the latter statistic we see that SGH, SGM, SGMx and SGBx displayed the highest heterozygosity ($H_o > 0.7$), though the rest was close behind. Heterozygosity did not

correspond directly to gene diversity (the same as above) where the highest values were achieved by SGB and SB ($H_e$ > 0.7).

**Table 8:** Comparison of pairwise $F_{ST}$ values between phenotypes. $F_{ST}$=genetic fixation index, ranging from 0 (=no fixation, apomixes to 1 (=total fixation, no genetic exchange). The highest values (>0.1) are marked red.

| $F_{ST}$ | SGM | SGMx | SGB | SGBx | SGG | SC | ST | SB |
|---|---|---|---|---|---|---|---|---|
| **SGH** | 0.021 | 0.019 | 0.039 | 0.016 | 0.02 | 0.021 | 0.053 | 0.020 |
| **SGM** | | 0.016 | 0.051 | 0.033 | 0.04 | 0.052 | 0.085 | 0.036 |
| **SGMx** | | | 0.065 | 0.049 | 0.05 | 0.057 | 0.080 | 0.046 |
| **SGB** | | | | 0.048 | 0.07 | 0.092 | 0.101 | 0.045 |
| **SGBx** | | | | | 0.05 | 0.079 | 0.113 | 0.032 |
| **SGG** | | | | | | 0.058 | 0.085 | 0.065 |
| **SC** | | | | | | | 0.124 | 0.062 |
| **ST** | | | | | | | | 0.124 |

**Table 9**: Population genetic statistics of *Syzygium* phenotypes, calculated from microsatellite data. $N_{Amax}$=max. number of alleles at one locus, AR=allelic richness, $H_e$ = gene diversity corrected for sample size, $H_o$ = obs. Heterozygosity, $F_I$ = inbreeding coefficient, $S_p$ = selfing rate estimate based on allele phenology (occurrence). Significant values are highlighted in red.

| | $N_{Amax}$ | AR | $H_e$ | $H_o$ | $F_I$ | $S_p$ |
|---|---|---|---|---|---|---|
| **SGH** | 4 | 2.54 | 0.675 | 0.776 | -0.023 | 0.210 |
| **SGM** | 4 | 2.56 | 0.676 | 0.776 | -0.021 | 0.259 |
| **SGMx** | 4 | 2.54 | 0.652 | 0.786 | -0.107 | 0.000 |
| **SGB** | 3 | 2.67 | 0.709 | 0.586 | 0.182 | 0.000 |
| **SGBx** | 4 | 2.47 | 0.658 | 0.786 | -0.201 | 0.545 |
| **SGG** | 4 | 2.45 | 0.651 | 0.595 | 0.287 | 0.439 |
| **SC** | 4 | 2.53 | 0.677 | 0.696 | 0.086 | 0.638 |
| **ST** | 2 | 2.09 | 0.469 | 0.414 | 0.123 | 0.000 |
| **SB** | 4 | 2.65 | 0.735 | 0.666 | 0.316 | 0.000 |
| **MEAN** | | **2.67** | **0.7007** | **0.697** | **0.119** | - |

Inbreeding coefficients were low for all phenotypes; the highest and statistically significant values belonged to SGB, SGG and SB. However, only in the SGG's case selfing was estimated to influence the genetic structure. Within SGB and SB selfing according to the model could not be identified. SGMx had a significantly negative inbreeding coefficient, in all other phenotypes the $F_i$ value ranged between -0.2 and 0.1 but was not significant. The allelic structures of SGMx, SGB, ST and SB are not estimated to be influenced by selfing, according to the selfing rate $S_p$. For SGH and SGM it plays a minor role (0.21 and 0.26, respectively) but for SGBx, SGG and SC a more important one (see table 9). In total we get the impression of genetically rich (high $H_e$ values), highly heterozygote (high $H_o$ values) set of individuals in all phenotypes, but to a lesser degree in ST. There is a slight trend to inbreeding in SGB, SGG, ST and SB (higher $F_i$ values), but even more so overall low genetic barriers and high genetic exchange within (low $F_{ST}$ values).

### III.4.2 PCoA and NJ-clustering

By plotting the genetic distances of all samples into a PCoA, the first impression one gets is chaotic since many polygons intersect. However, there is differentiation evident in their arrangement (see fig. 22, next page). Strikingly, despite the phenotypic and taxonomic differences, all individuals are rather close to each other, implying a close genetic relatedness. The first five axes of the PCoA cover in total only 32,70 % variance (8.44 %, 7.29 %, 6.39 %, 5.67 % and 4.92 %, respectively), and the first ten axis still only 51.49 %. Additionally, individuals of the same population do not necessarily cluster close together; some ordinate quite far apart from each other, though in most cases members of the same population stay close.

As for the general arrangement regarding the first two axes, all suffrutices (SGH-WZG and SGH-Mio) are found on the right side of the plot, whereas the shrub/tree phenotypes SGM and SGMx are on the other side in the bottom left quarter. The top left quarter is occupied by overlapping ST and SGB individuals. The SGG and SGBx phenotypes are cornered between these three greater groups in a central position. All mentioned phenotype polygons are not clearly separated from each other; the fringes of the greater groups rather overlap. The SC polygon is an exception; it stretches almost across all other polygons except SGM and SGMx, in spite of supposedly belonging to another species. Another important aspect is that the intermediate phenotype SGBx is situated between the trees and the suffrutices, linking them not only phenotypically but also genetically.

The two suffrutex groups from the Miombo (SGH-Mio) and Western Zambezian Grassland (SGH-WZG) biome overlap area-wide and form one unit, other than in the sequence analysis where they were more separated. This is also true when regarding the arrangement with first/third axis and second/third axis (see figure 23, next page). Their respective partner phenotypes, SGM and SGB, oppose the suffrutex polygon and overlap in some part; however they occupy different quarters themselves. SGB individuals B02 and B03, both from Sgg1b, were grouped apart, extending the polygon outside and more into SGM territory. While SGB and the suffrutices have no direct “genetic contact” in the plot, they are linked via SGBx.

SGM and SGMx have been placed covering each other. Both polygons connect to the suffrutices through some individuals in the lower part of the plot, however said individuals are geographically not close. Individuals from Tundavala population Sgmx657 (Mx11, Mx12), Huíla population Sgm1 (M07, M08) and from Benguela population Sgm2 (M09, M10) pierce the suffrutex groups. Surprisingly, the other Tundavala population Sgh16mx is clustered apart, but this is less evident regarding the other dimensions (axis 1-3 and 2-3). In all three first dimensions are the polygons of SGM and SGMx overlapping entirely; and in the dimension of second/third axis they are even almost entirely covered the suffrutices.

In the upper right quarter many tree phenotypes intersect and they do so in all first three dimensions. Half covered by the SGB polygon is the ST group which, if not in habitat, differs not much in leaf morphology and climatic variables. Its polygon intersects with SGB, SC and SB (and touches SGM). Herbarium specimens that looked much like ST - and had therefore been placed into this phenotype - correspond well to this group, even having been originally named *Syzygium guineense ssp angustifolia*, *S. g. littorale* or *S. g. afromontanum*.

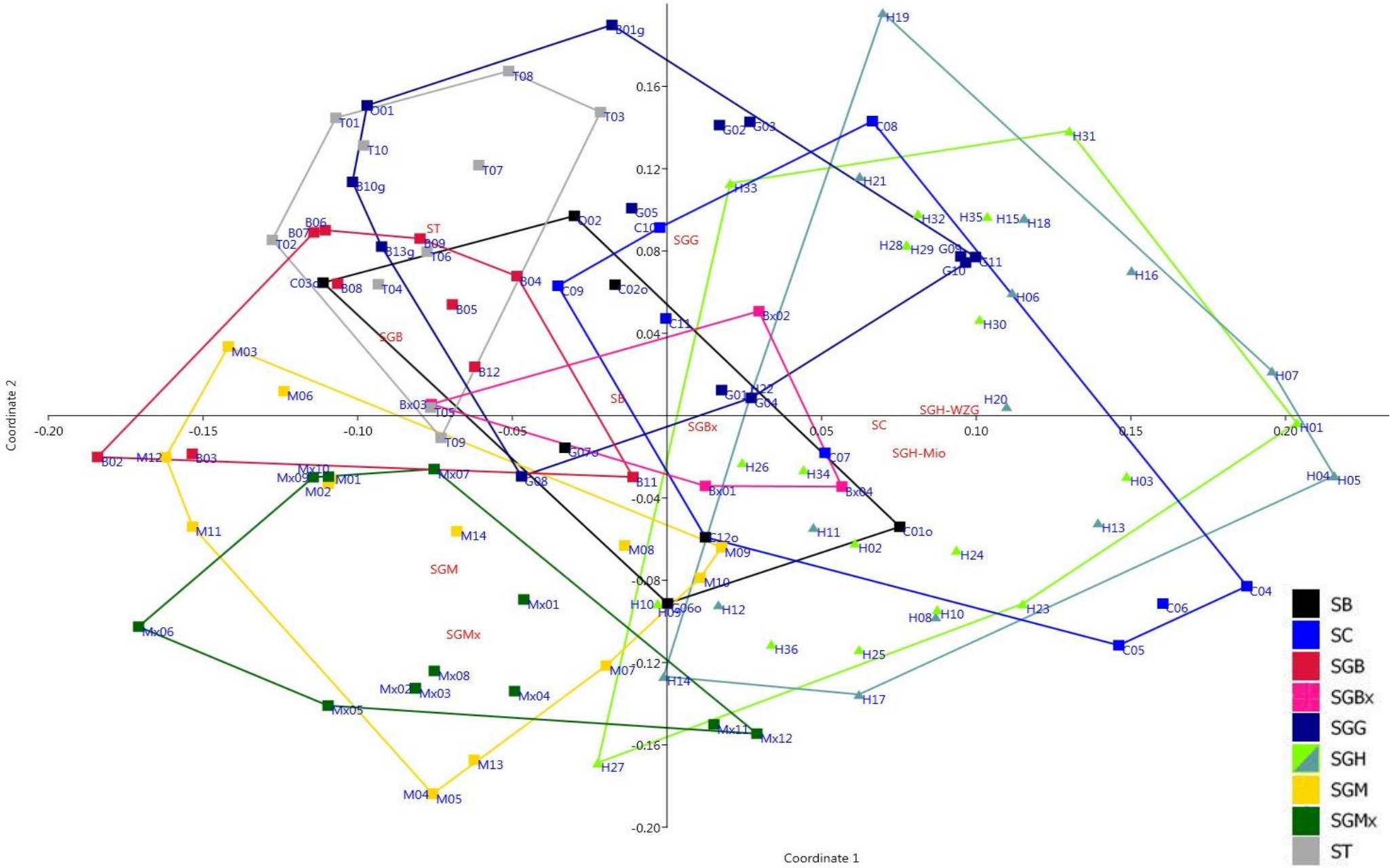



**Figure 22**: Principle Coordinate Analysis of full data set, polygons colored according to phenotype. 8.44 %, 7.29 % first two axes. Suffrutex samples are indicated as triangles, all other phenotypes as squares.

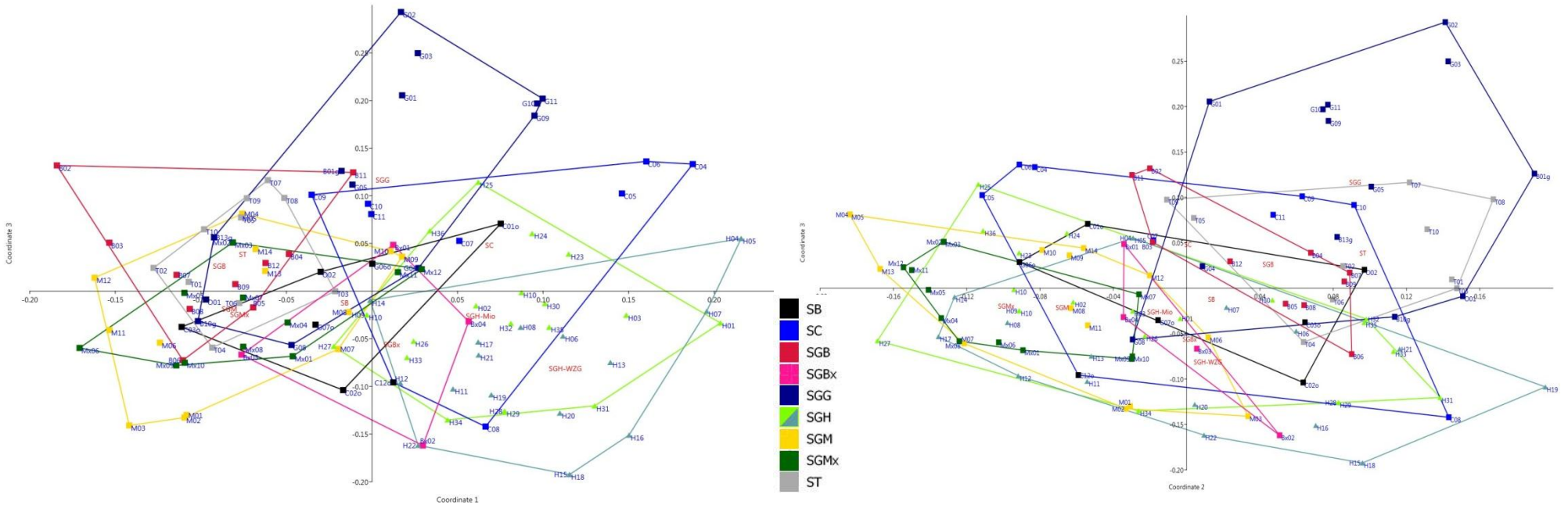
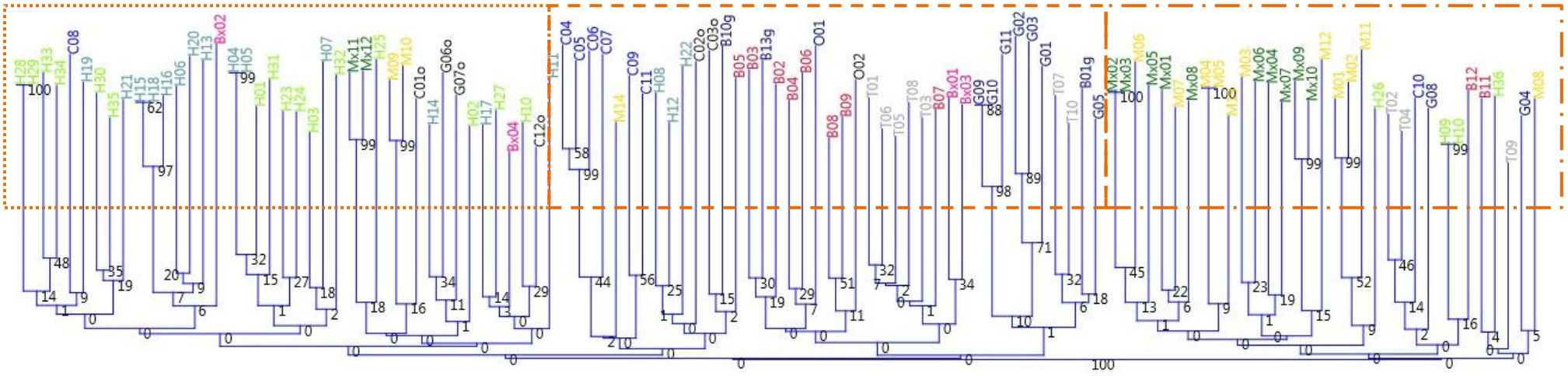



**Figure 23** (above): On the left: PCoA with axes 1 and 3 (Variance 8.44 % and 6.39 %, respectively). On the right: PCoA with axes 2 and 3 (Variance 7.29 % and 6.39 %, respectively). The colors indicate the same phenotype attribution as before.

**Figure 24** (below): Neighbor Joining Clustering Analysis on *Syzygium* microsatellite data. Bootstrap support at 1000 replicates is given as numbers at the bifurcations. The rectangles mark the greater genetic groups that are described in the text. The colors indicate the same phenotype attribution as before.



The individuals of "hybrid" phenotype SGBx fill the gap between the habitat-sharing SGB and SGH-WZG type. The SGBx polygon intersects with both SGH-WZG or SGB and links them, whereas most individuals cluster together with the suffrutices. Another group consists of SB individuals which were placed centrally. Some individuals that were previously assigned[2] to SGG or SC belong –after thorough review - to this group, together with resembling herbarium specimen. Therefore this polygon intersects with SGG and SC and together these three form a group of water associated phenotypes. Though the ecology of these samples is comparable, their genetic patterns and morphological appearance divergent.

This is especially true with the uncertain SGG assignment of herbarium specimen G08; it could as well be a SB according to the placement, in all dimensions. Apart from that the SGG polygon occupies the central upper part of the plot, overlapping with much of ST and SC. Population Sgg678 (G09, G10, G11) from near the Angolan Bicuar National Park protrudes into the SC and suffrutex polygons, and so does the Escarpment population Sc2 (C09, C10, C11) into the SGG polygon. SC individuals are surprisingly grouped intermingled within the suffrutices, though they are ecologically quite distinct. Within the SC group we also see a clear separation between Angolan (C09, C10, C11), Botswanan (C04, C05, C06 and C07) and Namibian (C08) individuals.

The analysis by hierarchical clustering is another method to assess similarities and differences and was conducted on the same data set. The dendrogram does not mirror the results of the PCoA exactly but rather, at first glance, distinguishes three greater clusters (see figure 24, previous page). We see again that phenotypes are not clearly separated and individuals of the same population do not necessarily cluster together. As the reliability of the dendrogram is compromised due to the low bootstrap support I will not go into detail. However, the tripartite pattern we see here also emerged in further genetic analysis and will be covered in the next chapters.

The first greater cluster encompasses c. 80% of individuals of SGM and SGMx type, though not all, as some Central and Western Angolan samples are clustered at the other side of the tree. In return, 11% of Miombo suffrutices are grouped to this cluster instead of the other. There are additionally single SGG, SGB, SC and ST individuals present which make out 24% of this branch.

The next greater cluster is formed to 90 % by tree individuals. Most SGB's, SGG's SC's and ST's are found here as well as half of the SGBx's and SB's. Three Zambezian suffrutices and one SGM are also assigned to this group. The remaining cluster is characterized to 76% by suffrutices and suffrutex associated SGBx. Within this branch are also half of the SB's (Angolan) and a single SC from Namibia and a SGM population from Benguela, AngolaThe significance of genetic differences between phenotypes, based on microsatellite data, can be evaluated by tests like PERMANOVA. The overall significance with 9999 permutations is very strong ($SS_{total}$: 872.4, $SS_{within}$: 683, F: 3.236, $p < 0.001$). In table 10 we see the detailed pairwise results of one-way PERMANOVA with Bonferroni-corrected p-values.

[2] Assigned to SGG/SC in the pre-screening due to morphological and ecological traits, e. g.leaf shape or growing in rivers. Though, leaf characteristics were often as not confusing as different shapes occurred on the same individual, which might have contributed to the reassignment campaign. In the final analyses these samples were treated as SB.

**Table 10:** Results of pairwise PERMANOVA test on genetic divergences between *Syzygium* phenotypes. Significant differences ($p < 0.05$) are highlighted red. P values are Bonferroni-corrected.

| | $SGH_{WZG}$ | SB | SGM | SGMx | ST | SGG | SC | SGB | SGBx |
|---|---|---|---|---|---|---|---|---|---|
| $SGH_{Mio}$ | 1.000 | 0.414 | 0.005 | 0.005 | 0.005 | 0.005 | 0.005 | 0.005 | 0.189 |
| $SGH_{WZG}$ | | 0.144 | 0.005 | 0.005 | 0.005 | 0.005 | 0.005 | 0.005 | 0.063 |
| SB | | | 0.734 | 0.176 | 0.009 | 0.126 | 0.095 | 1.000 | 0.774 |
| SGM | | | | 1.000 | 0.005 | 0.005 | 0.005 | 0.005 | 0.810 |
| SGMx | | | | | 0.005 | 0.005 | 0.005 | 0.005 | 0.122 |
| ST | | | | | | 0.005 | 0.014 | 0.005 | 0.059 |
| SGG | | | | | | | 0.005 | 0.054 | 0.063 |
| SC | | | | | | | | 0.005 | 0.135 |
| SGB | | | | | | | | | 0.081 |

The SGBx phenotype, having a very central and therefore overlapping position according to PCoA (see fig. 22), has no significant divergences. It differs barely significantly from SGH-WZG, ST and SGG, but it should be considered that its n is quite low (n=4). Apart from that SB is least genetically separated; it differs only from ST significantly. Again, SB takes a central, overlapping position in the PCoA. There is no significant genetic separation between the two suffrutex groups from the Miombo and the Western Zambezian Grasslands as well as none between SGM and SGMx. There is also none between SGB and SB, and between SGB and SGG. All other phenotypes show significant genetic delineations from each other.

### III.4.3 STRUCTURE

Under the assumption of independent allele frequencies and the no mixed heritage model in the multi-locus analysis with STRUCTURE 2.3.4 the following results were obtained. Modelling the highest probability by Structure Harvester achieved an optimal K = 3, meaning that genetically all individuals are best represented in three genetic clusters. Its mean likelihood is -2938.40 ± 0.70 which is the best for all tested two to ten K. Furthermore, the first and second derivation as an estimate for the rate of change, do also show highest values in their respective class. Delta K, as an indicator when representability starts to drop, shows the overall highest value for K = 3 (tab. 11 and fig 25).

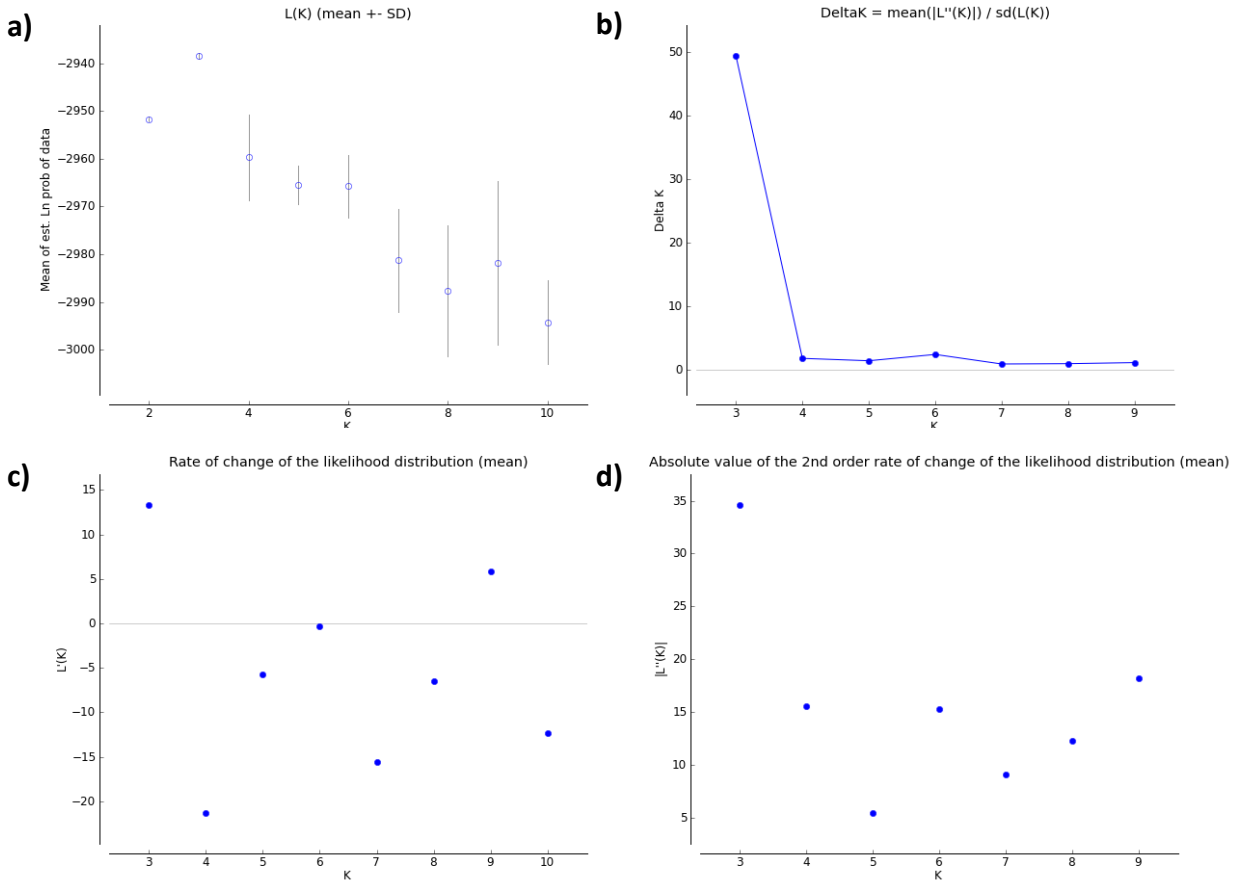



**Figure 25**: Structure harvester output, showing likelihood of K clusters (a), rate of likelihood change between K clusters (b) and first (c) and second (d) derivative of the likelihood function, indicating when probability starts to drop. All statistics point to K=3 as the best representative of the microsatellite data.

**Table 11:** Evanno table with probability results for K in 2 to 10, each K was replicated 3 times. LogLikelihood and characteristics based upon it are shown and indicate K=3 as the best representative number.

| K | Mean LnP(K) | Stdev LnP(K) | Ln'(K) | \|Ln''(K)\| | Delta K |
|---|---|---|---|---|---|
| **2** | -2951.700 | 0.529 | NA | NA | NA |
| **3** | **-2938.400** | **0.700** | **13.300** | **34.600** | **49.429** |
| **4** | -2959.700 | 9.014 | -21.300 | 15.567 | 1.727 |
| **5** | -2965.433 | 4.020 | -5.733 | 5.433 | 1.351 |
| **6** | -2965.733 | 6.494 | -0.300 | 15.233 | 2.346 |
| **7** | -2981.267 | 10.801 | -15.533 | 9.100 | 0.843 |
| **8** | -2987.700 | 13.727 | -6.433 | 12.233 | 0.891 |
| **9** | -2981.900 | 17.176 | 5.800 | 18.167 | 1.058 |
| **10** | -2994.267 | 8.749 | -12.367 | NA | NA |

The overall structure of the clusters is shaped by closeness and mixing. Estimated heterozygosity within the clusters as calculated by STRUCTURE is comparably similar; it accounts for 0.755 in Cluster 1 and 0.756 in Cluster 3. Cluster 2 shows a slightly higher value with 0.800. These values range near the previously calculated overall heterozygosity (see III.4.1). The divergences of allele frequencies among clusters, serving as an estimate of net nucleotide distance or, more precisely, the mean amount of pairwise differences between alleles from different clusters is low between all clusters. It is highest between Clusters 1 and 2 (0.032), followed by the divergence between Clusters 2 and 3 (0.022) and Clusters 1 and 3 (0.016), as calculated by STRUCTURE. As to the inner structure of the clusters, it can be said that some phenotypes are clearly grouped into one cluster, whereas others show a mixed heritage between clusters or even a phenotypically incoherent distribution, leading to a high variation within phenotypes. Details are given below in tab. 12 and fig. 26.

**Table 12:** Mean affiliation of *Syzygium* phenotypes to each of the three genetic clusters identified.

| | Cluster 1 | Cluster 2 | Cluster 3 |
|---|---|---|---|
| **SGM** | 0.032 | 0.966 | 0.002 |
| **SGMx** | 0.05 | 0.948 | 0.002 |
| **SGB** | 0.321 | 0.604 | 0.075 |
| **SB** | 0.379 | 0.576 | 0.045 |
| **SGBx** | 0.576 | 0.411 | 0.013 |
| **SGH** | 0.715 | 0.242 | 0.043 |
| **SGG** | 0.193 | 0.115 | 0.692 |
| **ST** | 0.585 | 0.178 | 0.237 |
| **SC** | 0.722 | 0.006 | 0.272 |

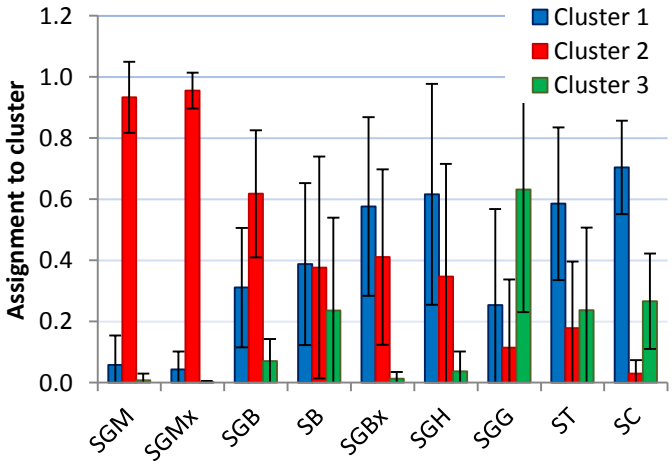



**Figure 26:** Phenotype assignments to the three genetic clusters. Standard deviation shown as error bars.

The rows in tab. 12 are deliberately ordered in a fashion to highlight the phenotype's clusters assignments. It can clearly be made out that SGM and SGMx were most strongly associated (97% and 95%, respectively) with Cluster 2 and with none over. As a second pattern we see a group of phenotypes (SGH, SGB, SGBx, SB) that were assigned to both Cluster 1 and 2, varying between 24% and 72%. Phenotype SC is divided between Cluster 1 (72%) and 3 (27%) and ST had a share of all clusters with highest attribution to Cluster 1 (58%), followed by 3 (24%) and 2 (18%). SGG had the highest assignment to Cluster 3 with 69% on average but strongly varying .Overall, by means of Pearsons correlation analysis, it was evident that assignment to Cluster 1 or 2 is negatively correlated so that there was a mutual exclusion (r = -0.83). Attribution to Cluster 2 or 3 on the other hand was negatively correlated to a minor degree of r = -0.42, and between Cluster 1 and 3 the value decreased to r = -0.16 (see III.5, fig. 38.)

All phenotypes except SGM and SGMx exhibited a certain degree of variation within their cluster assignment, as is evident from the error bars in fig. 26. To clarify this pattern it is necessary to look at the individual level. This is on the one hand done by putting the genotypes into a geographic context as can be contemplated in fig. 27-29. Assignment to Cluster 1, 2 or 3 is shown separately and there is little large scale geographic pattern evident. One aspect might be that Clusters 1 and 2 occupy rather space in Central Angola/West Zambia whereas Cluster 3 occurred more in samples at the southern and western fringes of the sampling range.

**Figure 27:** Distribution of SSR Cluster 1 within the sampling range. The deeper the blue hue, the stronger the association to Cluster 1.

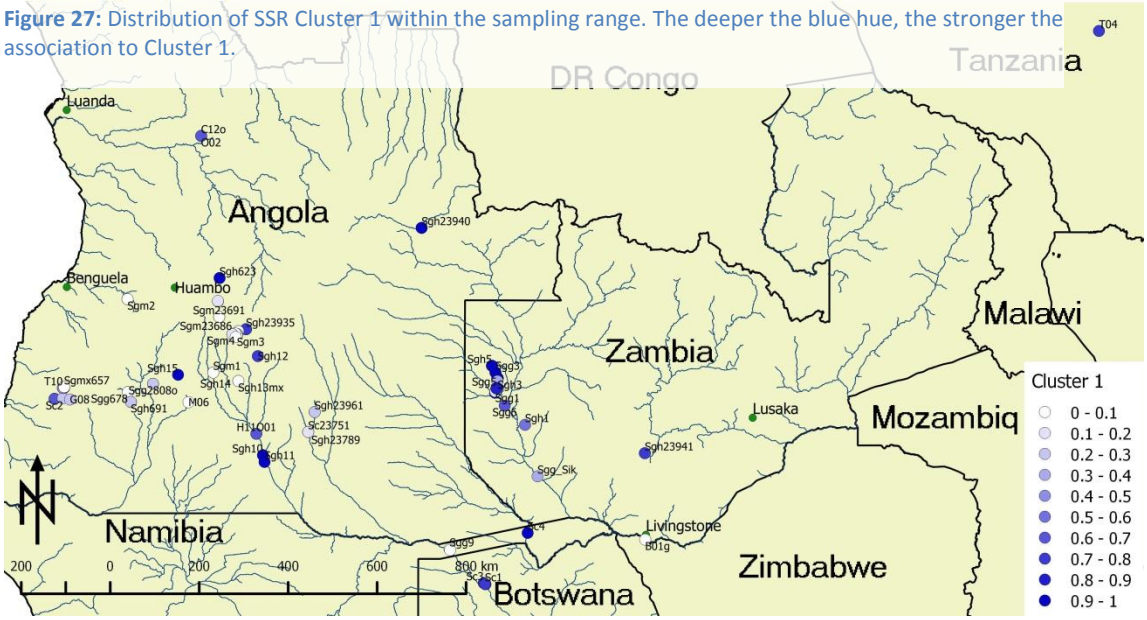



**Figure 28:** Distribution of SSR Cluster 2 within the sampling range. The deeper the red hue, the stronger the association to Cluster 2.

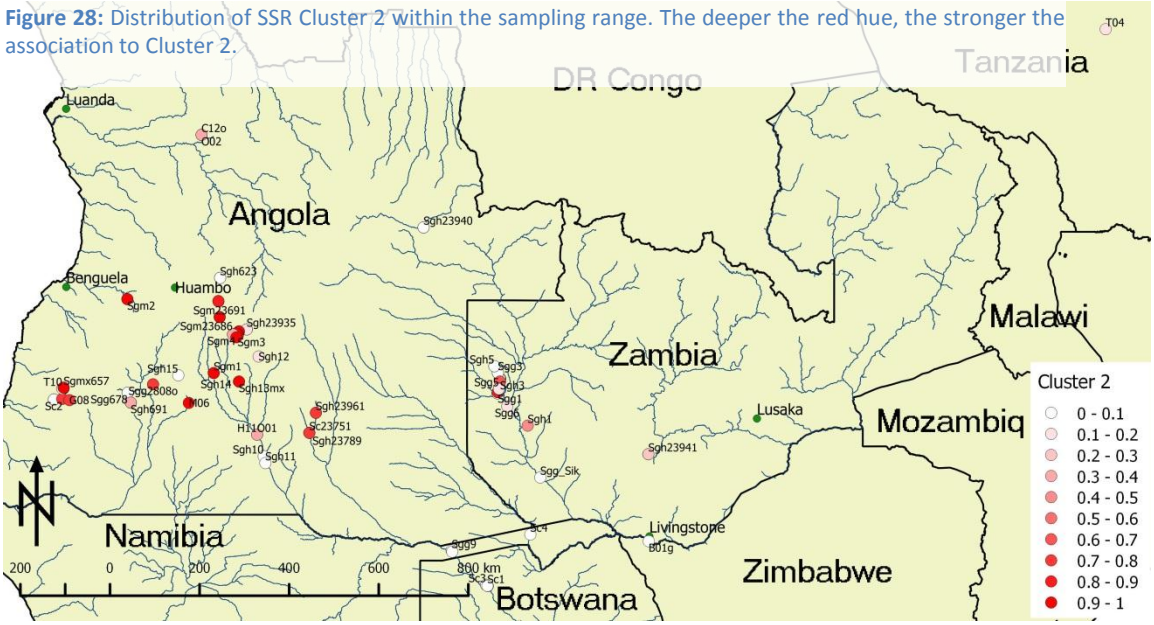



**Figure 29** Distribution of SSR Cluster 3 within the sampling range. The deeper the green hue, the stronger the association to Cluster 3.

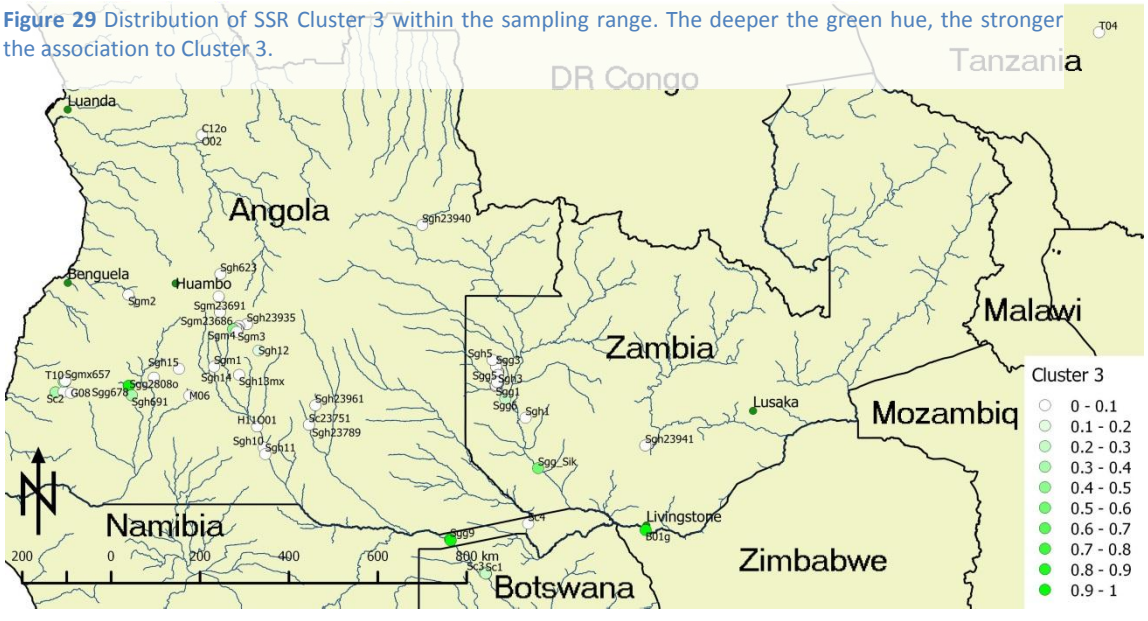


**Table 13:** Association to genetic SSR clusters of each analysed individual. Deeper hues indicate stronger association. Figures according to calculation of STRUCTURE with the no-admixture and independent allele assumption.

| | ID | Cluster 1 | Cluster 2 | Cluster 3 |
|---|---|---|---|---|
| *SGH* | H01 | 0.976 | 0.001 | 0.023 |
| | H02 | 0.211 | 0.776 | 0.013 |
| | H03 | 0.777 | 0.193 | 0.029 |
| | H04 | 0.902 | 0.007 | 0.091 |
| | H05 | 0.902 | 0.007 | 0.091 |
| | H06 | 0.985 | 0.001 | 0.014 |
| | H07 | 0.829 | 0.008 | 0.162 |
| | H08 | 0.762 | 0.230 | 0.007 |
| | H09 | 0.222 | 0.773 | 0.005 |
| | H10 | 0.222 | 0.773 | 0.005 |
| | H11 | 0.644 | 0.345 | 0.011 |
| | H12 | 0.513 | 0.483 | 0.003 |
| | H13 | 0.356 | 0.640 | 0.004 |
| | H14 | 0.969 | 0.017 | 0.015 |
| | H15 | 0.137 | 0.858 | 0.005 |
| | H16 | 0.999 | 0.001 | 0.001 |
| | H17 | 0.998 | 0.001 | 0.002 |
| | H18 | 0.559 | 0.438 | 0.003 |
| | H19 | 0.999 | 0.001 | 0.001 |
| | H20 | 0.974 | 0.001 | 0.025 |
| | H21 | 0.979 | 0.011 | 0.011 |
| | H22 | 0.921 | 0.033 | 0.046 |
| | H23 | 0.801 | 0.195 | 0.004 |
| | H24 | 0.323 | 0.665 | 0.012 |
| | H25 | 0.723 | 0.238 | 0.040 |
| | H26 | 0.217 | 0.719 | 0.064 |
| | H27 | 0.733 | 0.250 | 0.016 |
| | H28 | 0.350 | 0.569 | 0.080 |
| | H29 | 0.906 | 0.060 | 0.035 |
| | H30 | 0.906 | 0.059 | 0.035 |
| | H31 | 0.745 | 0.130 | 0.124 |
| | H32 | 0.989 | 0.006 | 0.005 |
| | H33 | 0.906 | 0.000 | 0.093 |
| | H34 | 0.870 | 0.022 | 0.108 |
| | H35 | 0.963 | 0.015 | 0.022 |
| | H36 | 0.915 | 0.026 | 0.059 |
| | H37 | 0.273 | 0.382 | 0.344 |
| *SGM* | M01 | 0.012 | 0.988 | 0.000 |
| | M02 | 0.006 | 0.994 | 0.000 |
| | M03 | 0.060 | 0.940 | 0.001 |
| | M04 | 0.014 | 0.984 | 0.001 |
| | M05 | 0.015 | 0.984 | 0.001 |
| | M06 | 0.001 | 0.999 | 0.000 |
| | M07 | 0.137 | 0.863 | 0.000 |
| | M08 | 0.047 | 0.947 | 0.006 |
| | M09 | 0.032 | 0.966 | 0.003 |
| | M10 | 0.006 | 0.993 | 0.001 |
| | M11 | 0.001 | 0.997 | 0.003 |
| | M12 | 0.080 | 0.909 | 0.012 |
| | M13 | 0.004 | 0.996 | 0.000 |
| | M14 | 0.101 | 0.899 | 0.000 |
| *SGMx* | Mx01 | 0.209 | 0.790 | 0.001 |
| | Mx02 | 0.001 | 0.999 | 0.000 |
| | Mx03 | 0.001 | 0.999 | 0.000 |
| | Mx04 | 0.003 | 0.997 | 0.000 |
| | Mx05 | 0.003 | 0.997 | 0.000 |
| | Mx06 | 0.006 | 0.994 | 0.000 |
| | Mx07 | 0.021 | 0.965 | 0.014 |
| | Mx08 | 0.060 | 0.939 | 0.000 |
| | Mx09 | 0.050 | 0.947 | 0.003 |
| | Mx10 | 0.064 | 0.934 | 0.002 |
| | Mx11 | 0.040 | 0.959 | 0.001 |
| | Mx12 | 0.096 | 0.901 | 0.002 |

| | ID | Cluster 1 | Cluster 2 | Cluster 3 |
|---|---|---|---|---|
| *SGB* | B02 | 0.018 | 0.825 | 0.158 |
| | B03 | 0.216 | 0.713 | 0.070 |
| | B04 | 0.533 | 0.369 | 0.098 |
| | B05 | 0.115 | 0.853 | 0.032 |
| | B06 | 0.501 | 0.490 | 0.008 |
| | B07 | 0.573 | 0.373 | 0.054 |
| | B08 | 0.357 | 0.626 | 0.017 |
| | B09 | 0.294 | 0.663 | 0.043 |
| | B11 | 0.405 | 0.368 | 0.226 |
| | B12 | 0.200 | 0.758 | 0.042 |
| *SGBx* | Bx01 | 0.455 | 0.498 | 0.047 |
| | Bx02 | 0.980 | 0.019 | 0.001 |
| | Bx03 | 0.296 | 0.703 | 0.001 |
| | Bx04 | 0.574 | 0.422 | 0.004 |
| *SGG* | B01g | 0.024 | 0.000 | 0.976 |
| | B10g | 0.805 | 0.093 | 0.102 |
| | B13g | 0.172 | 0.351 | 0.477 |
| | G01 | 0.000 | 0.001 | 0.999 |
| | G02 | 0.000 | 0.000 | 1.000 |
| | G03 | 0.000 | 0.000 | 1.000 |
| | G04 | 0.644 | 0.108 | 0.248 |
| | G05 | 0.356 | 0.080 | 0.564 |
| | G08 | 0.249 | 0.745 | 0.006 |
| | G09 | 0.041 | 0.000 | 0.959 |
| | G10 | 0.013 | 0.000 | 0.987 |
| | G11 | 0.009 | 0.000 | 0.991 |
| *SC* | C04 | 0.756 | 0.001 | 0.244 |
| | C05 | 0.766 | 0.001 | 0.233 |
| | C06 | 0.713 | 0.001 | 0.286 |
| | C07 | 0.771 | 0.009 | 0.220 |
| | C08 | 0.992 | 0.002 | 0.007 |
| | C09 | 0.493 | 0.019 | 0.487 |
| | C10 | 0.601 | 0.007 | 0.391 |
| | C11 | 0.685 | 0.004 | 0.311 |
| *ST* | T01 | 0.589 | 0.077 | 0.333 |
| | T02 | 0.787 | 0.119 | 0.094 |
| | T03 | 0.931 | 0.006 | 0.062 |
| | T04 | 0.800 | 0.171 | 0.029 |
| | T05 | 0.578 | 0.311 | 0.111 |
| | T06 | 0.626 | 0.309 | 0.065 |
| | T07 | 0.120 | 0.003 | 0.877 |
| | T08 | 0.700 | 0.021 | 0.279 |
| | T09 | 0.257 | 0.706 | 0.038 |
| | T10 | 0.465 | 0.058 | 0.477 |
| *SB* | C01o | 0.081 | 0.871 | 0.048 |
| | C02o | 0.694 | 0.246 | 0.059 |
| | C03o | 0.156 | 0.800 | 0.044 |
| | C12o | 0.644 | 0.355 | 0.001 |
| | G06o | 0.154 | 0.832 | 0.014 |
| | G07o | 0.271 | 0.725 | 0.003 |
| | O01 | 0.881 | 0.065 | 0.054 |
| | O02 | 0.153 | 0.712 | 0.135 |

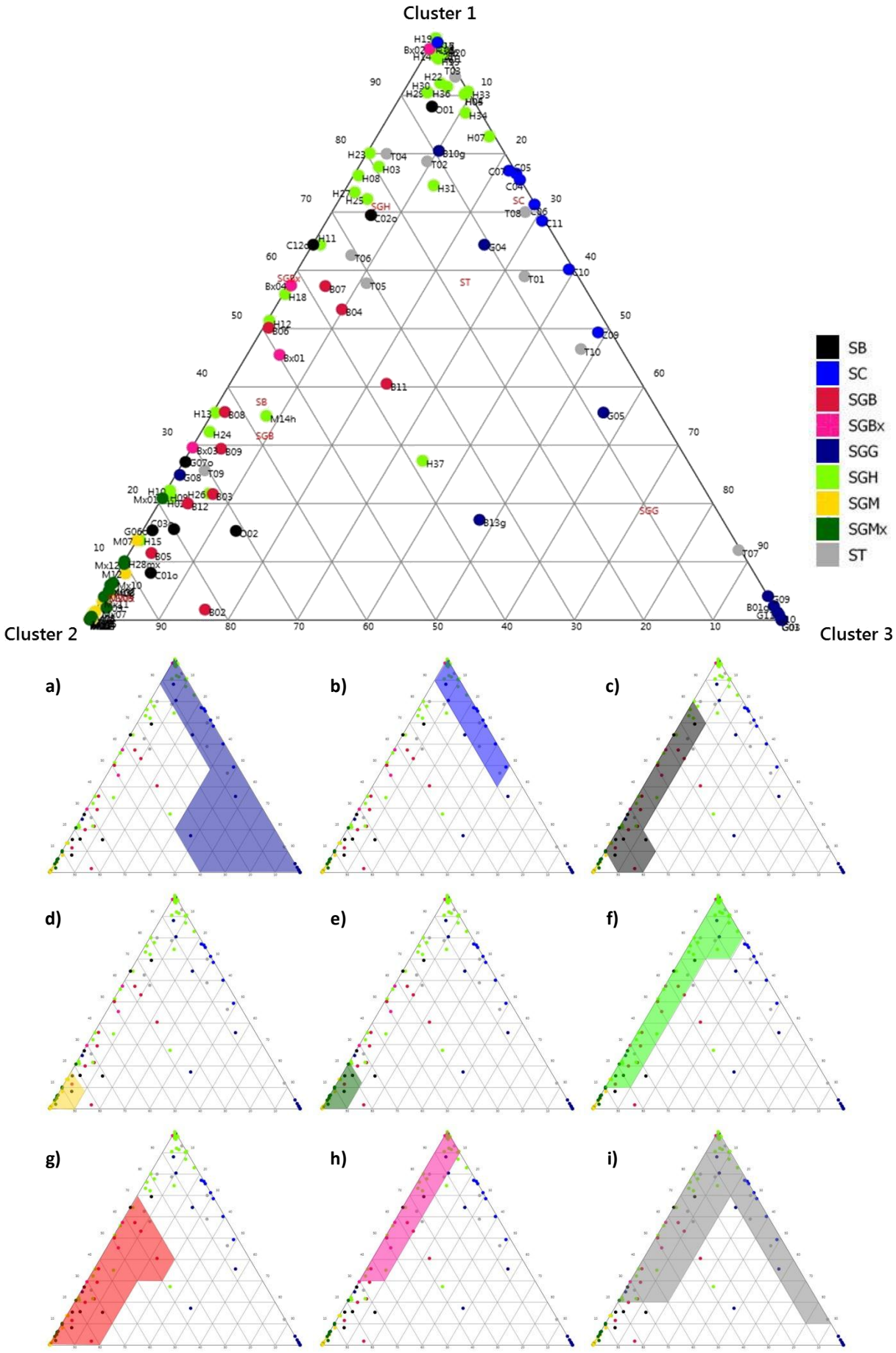



**Figure 30:** Top: Triangle plot of general cluster assignments. Each corner represents 100% assignment to the respective cluster; each edge indicates assignments between two clusters from 0% - 100%.
The smaller triangles below show the occupied space for each phenotype within the genetic triangle. a) SGG, b) SC, c) SB, d) SGM, e) SGMx, f) SGH, g) SGB, h) SGBx, i) ST. Strong outlier in SGH (in the triangle middle) is not included into the polygon.

On the other hand one should consider which individual pattern was formed not only in geographic but also genetic space (see tab. 13 and fig. 30). Starting with SGM and SGMx, the same strong association to Cluster 2 was observed though there were minor deviations (up to 21%) into Cluster 1 for individuals M07 (Huíla province) and Mx01 (Escarpment) which were not shared by the population. Both SGM and SGMx therefore occupy the "gold-green corner" of Cluster 2 (see fig. 30d/30e).

Phenotype SGH associated with emphasis Cluster 1, but also to some extent to Cluster 2; those individuals closer to Cluster 2 are from locations like Kafue Plains and Liuwa Plains, Zambia, and Cuando-Cubango province in Southern Angola, but not consistently. The samples H18 and H19 for example are from the same population Sgh4, Liuwa Plains, but while H18 is divided half-half, H19 belongs entirely to Cluster 1. Reasons for this uneven pattern will be discussed later on. Strikingly there is one deviating SGH which was placed almost central in the triangle due to even distribution between all clusters. However, its phenotype and population companions were grouped well within a pattern that is visible as a green suffrutex edge in the triangle plot (fig. 30f).

Phenotype SC was mainly attributed to Cluster 1 but also related to 22% - 50% to Cluster 3, with one exception, namely C08 from Namibian Caprivi Strip. This was the only SC with pure assignment to Cluster 1.While SC occupied the upper part of the Cluster 1 – Cluster 3 edge (see fig. 30), SGG did so with the lower part. Some SGG individuals were assigned to more than 95% to this corner; they belong to population Sgg9 from the Namibian Okavango River, Sgg678 from a small river north of Bicuar NP, Angola, and one herbarium sample from Victoria Falls, Zambia (B01g). Deviating individuals B10g, B13g were set apart which can be explained ecologically (see IV.1). Herbarium specimen G08 had been grouped in the far off corner. Disregarding the two latter individuals, a "blue SGG/SC edge" is evident (fig. 30a/30b).

The SGB individuals occupied the area around Cluster 2, not exceeding 70% in direction to Cluster 1 and 30% to Cluster 3. They overlapped with almost all other phenotypes except SC and constituted a consistent red area (fig. 30g), there were no strong outliers within this group though B02 (Liuwa Plains) and B11 (Cusseque) extended the area a bit to Cluster 3.

Within the same area and a bit further to Cluster 1, highlighted in fig. 28c, were the SB individuals. They are the third strongly water associated tree phenotype in this study; however they were genetically different as they clustered on the other leg of the triangle. Their only connection point is at Cluster 1 corner. According to table 14 there is fluctuation between 7% and 88% regarding assignment to Cluster 1 or 2.

Similarly, SGBx individuals were scattered along the edge of Cluster 1 and 2 (fig. 30h). Here, Bx02 deviated the most as it showed perfect association to Cluster 1. Lastly, ST exhibited the most confounding results since some individuals tended strongly towards Cluster 1, some to 2, some to 3 and some were in between (but never evenly between all three clusters, see fig. 30i). T03 was strongest associated to Cluster 1 (93%), T07 to Cluster 2 (88%) and herbarium T09 to Cluster 3 (71%) and they are all not more than 30 km apart from each other along the Angolan Escarpment. T04 from Tanzania showed similar assignments as the self-sampled T01-T03, T05-T06 and T08 from the Angolan Escarpment, despite the spatial distance.

The Structure software also gives opportunity to visualize cluster affiliation of all individuals in the analysis simultaneously, as shown in fig. 31 below. In the first figure are the cluster assignments sorted by phenotype. This figure corresponds directly to the previous tab. 13, each

line represents one individual. More important shall be fig. 32 where the individuals, again each represent by one line, are ordered according to their cluster assignments.

Evidently, each cluster had only few representable samples of perfect assignment; most samples were gradually divided between two clusters and very few between all three clusters (one each left and right from the central green area). Moreover, it is clear that Cluster 1 and 2 were shared by equal amounts and made out most of the genetic structure, whereas Cluster 3 was represented sparsely in both pure and admixed form.

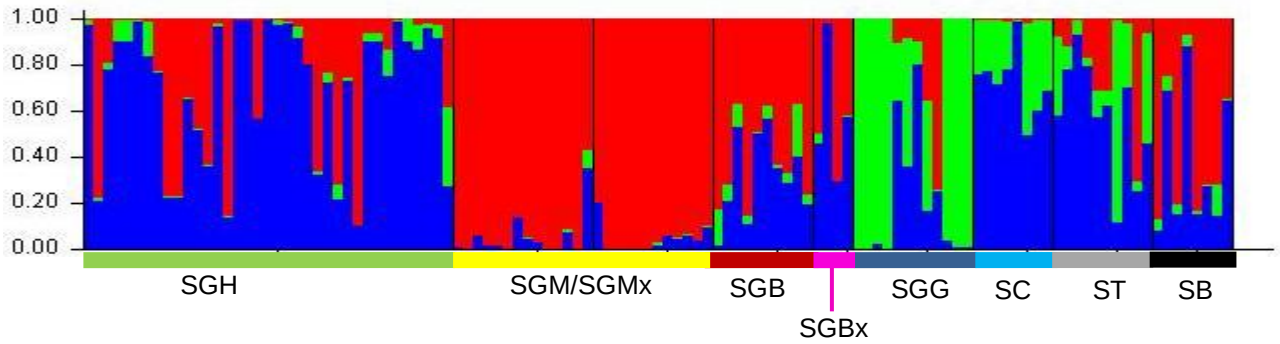



**Figure 31:** Genetic structure, represented by assignments to identified genetic clusters. Ordered by phenotype according to table 13. Each line represents one individual.

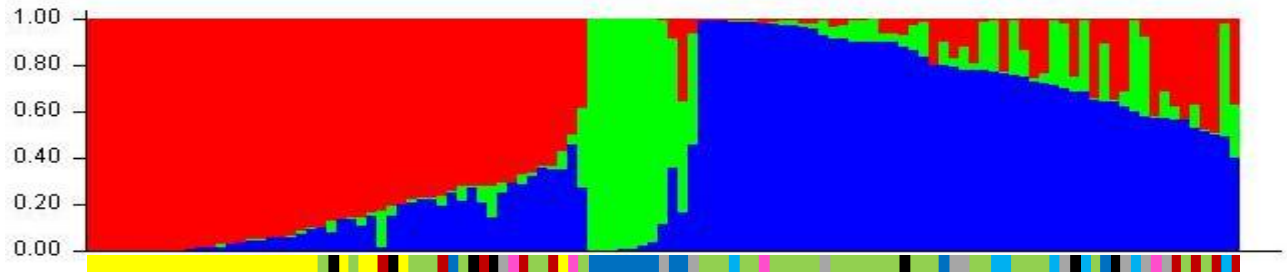



**Figure 32:** Genetic structure sorted by Q, the cluster membership coefficient. Each line represents one individual, colors indicate phenotype just as in fig. 31.

### III.4.4 Alleles In Space: linking geographic and genetic patterns

The software Alleles in Space offers possibilities to investigate the relationship between genetic and geographical patterns. It does so with various tools for SSR data, some of which had been used and the outcome is described in the following. Due to the tetraploidy of the studied *Syzygium* phenotypes, which is not supported by the software, it had been necessary to convert the data to haploid dominant format. This could result in loss of information.

#### III.4.4.1 AIS – Mantel test

The Mantel test estimates the correlation of genetic and geographical distance by calculating and assessing pairwise distances. Fig. 33 shows the graphical output in form of a distogram without making a difference to which phenotype the sample belongs. The self-sampled specimens were numerous and ranged between 0 km and approx. 1,200 km; therefore the data cloud is dense in this range. Other more spatially distant data points belonged to specimen from herbaria and represent a sampling bias. Nonetheless, the trend is the same over the whole range, there is no significant correlation between spatial and genetic distance (r = 0.035). The probability of observing a correlation greater than or equal to the calculated one was insignificant: $p(r_{obs} \geq r_{calc}) = 0.184$.

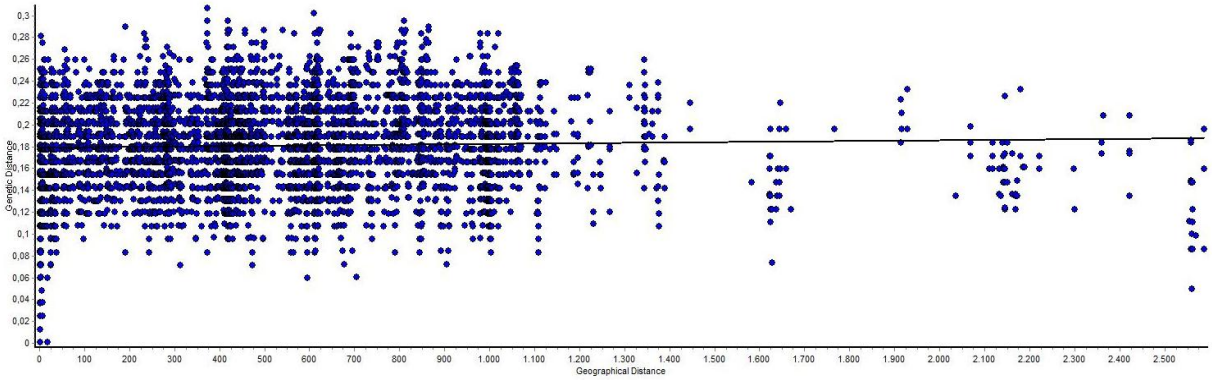



**Figure 33:** Graphical results of Mantel test for geographico-genetic correlations of all samples, all phenotypes. Spatial distance given in [km]. Correlation between genetic and geographical distance amounts to insignificant r = 0.035.

However due to the inclusion of all samples it was not possible to detect phenotypically characteristic patterns. Hence the test was redone with SGH only (fig. 34) and SGM+SGMx only (fig. 35) as major representatives of the suffrutex – tree – conundrum. For both SGH and SGM did the genetic divergence increase weakly (but significantly) with spatial distance (correlation: $r_{SGH} = 0.136/p_{SGH} = 0.006$; $r_{SGM} = 0.192/p_{SGM} = 0.017$).

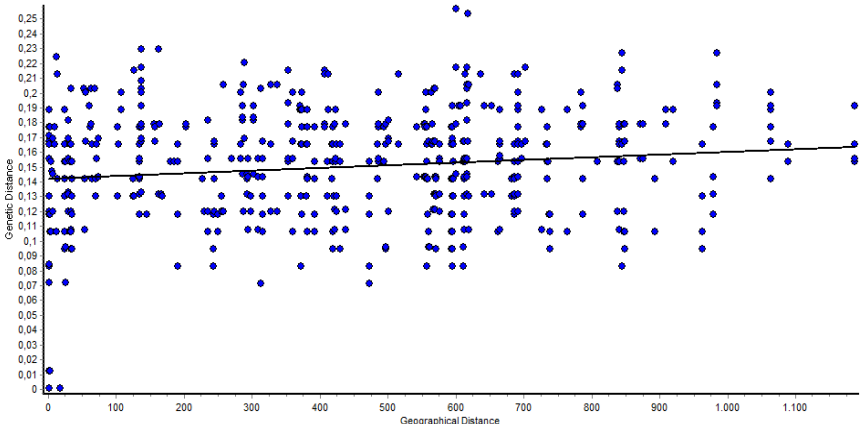



**Figure 34:** Distogram from Mantel test performed on SGH samples only. Spatial distance in [km]. Correlation between genetic and geographical distance amounts to significant r = 0.136.

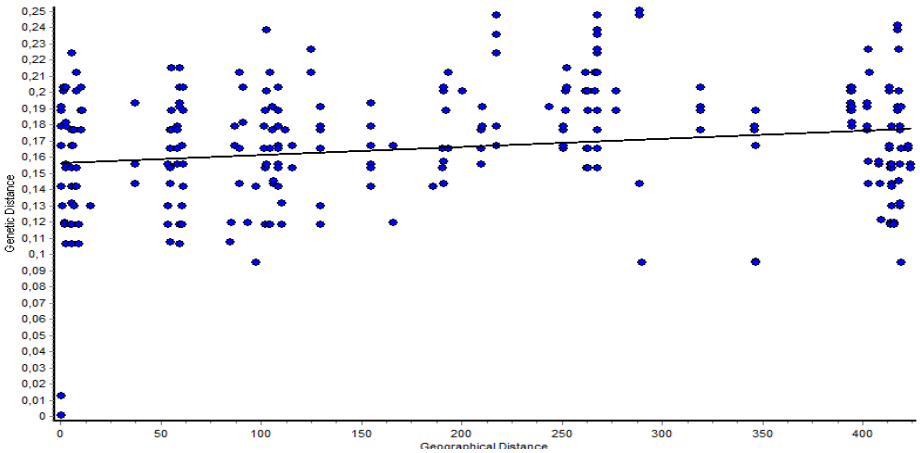



**Figure 35:** Distogram from Mantel test performed on SGM and SGMx samples only. Spatial distance in [km]. Correlation between genetic and geographical distance amounts to significant r = 0.192.

### III.4.4.2 AIS – Spatial Autocorrelation

The Spatial Autocorrelation (SAA) is a measure for evaluating dependencies among observations at different distance classes. As can be seen in figure 36, ten classes were defined for analysis. The overall averaged pairwise genetic distance was $A_y$ = 0.180 and the covariance among all observations accounted for V = 0.018, which was not significant ($p(V \geq V_{obs})$ = 0.149). Similar to the Mantel test results above, there was no evidence that genetic distance increases with spatial distance; though the graph was slightly drooping towards the end, this could also be accounted for the lower n in this distance classes.

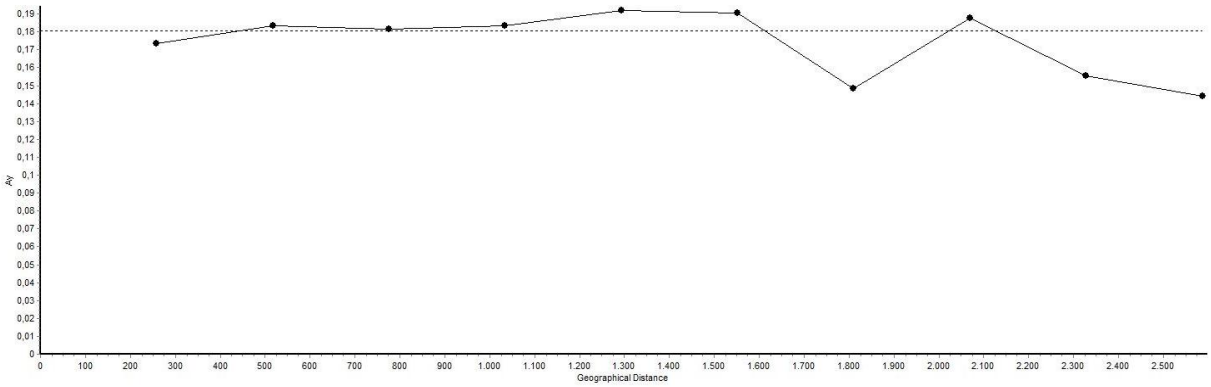



**Figure 36**: Spatial Autocorrelation graph, showing the correlation between geographic distance [km] and averaged pairwise genetic distance within 10 distance classes (points). The graph runs more or less constant, indicating no spatial autocorrelation.

In order to illustrate the pattern within suffrutices and SGM/SGMx the test was redone with the according data subset, just as above for the Mantel test. Figures are not shown here as they resembled the overall spatial autocorrelation from fig. 36 very much, but are available in the appendix (VI.4; fig. A1-A2). The averaged pairwise genetic distance in SGH accounted for $A_y$ = 0.149 and the covariance for V = 0.008 with an insignificance of p = 0.509. SGM's figures were slightly higher and significant with $A_y$ = 0.169, V = 0.017 and p = 0.036.

## III.5 FAMD/HCPC

The FAMD allows for a deeper understanding of correlations since it takes both categorical and continuous parameters into account. In this study it had been used to detect which factors influence and characterize the appearance of each single *Syzygium* phenotype. Variables under investigation had been of genetic nature like haplotype and SSR cluster of STRUCTURE analysis, of climatic nature such as precipitation or temperature, or of morphological nature like leaf area growth height. The outcome of the analysis is shown in fig. 37. Parameters and factor levels that influenced the phenotype significantly are displayed as barplots that indicate positive or negative correlation as well as its strength (v value > 0, v value < 0 and distance from 0, respectively).

**SC** was characterized mostly by low precipitation and low latitude, a circumstance that is correlated (see fig. 38). Furthermore it was neither associated with SSR Cluster 2 nor the central *trnK* haplotype HT1 (see III.3), but moreover with the Okavango haplotype (HT2) and SSR Cluster 1. Its non-existent petioles led to a positive correlation with high leaf-to-petiole length ratios. The most influencing factor was the preferred river bank substrate.

**SGB** was defined as being no suffrutex but tree of more than 10 m height with accordingly no geoxylic structures. Its leaf texture was not leathery but silky, the leaves rather small with a long petiole (low leaf-petiole ratio). Moreover, SGB was associated with higher longitudes (most populations were sampled in Zambia, the eastern part of the sampling range). Some SGB populations also shared the indel haplotype HT7.

Phenotype **SGBx** showed no negative associations but was positively connected to the Western Zambezian Grassland haplotype HT3. Secondly, the shrub appearance and the higher leaf area were characteristic, as well as the lack of knowledge regarding the underground whereabouts.

**SB** had only two positive attributes of lesser strength. For one this was the river bank as preferred substrate and for the other the leaf-petiole length ratio. SB is tagged by long narrow leaves with short petiole.

**SGG** on the other hand was associated to a lesser degree to low latitudes, and negatively correlated to SSR Cluster 2. It was connected to river banks as habitat and to a growth height of 2 m to 10 m. However, it was most strongly correlated to the attribution into Cluster 3 from SSR analysis (see III.4.3).

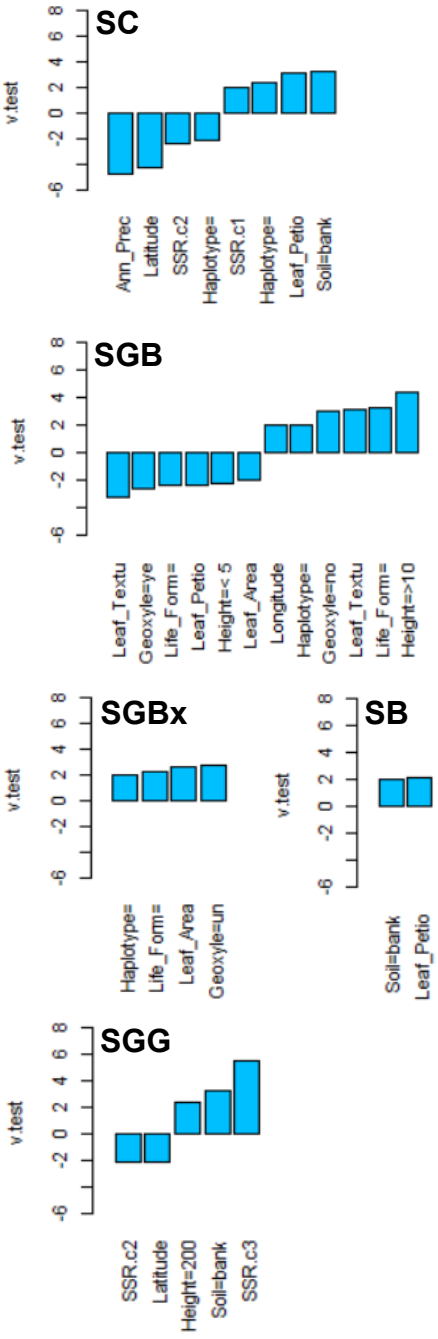



**Figure 37**: Impact factors based on FAMD for each phenotype. Positive and negative associations and the strength of their influence upon each phenotype are shown. Continued on next page. Abbreviated parameters: Annual precipitation (Ann_Prec), leaf-petiole-ratio (Leaf_Petio), leaf texture (Leaf_Textu), genetic clusters 1 -3 from STRUCTURE analysis: SSR.c1 – .c3. Categories (e. g. Haplotype=…) are explained in the text if not evident from figure.

**SGH** is the typical representative of geoxylicity in this analysis without biomic separation into Miombo and Western Zambezian Grassland populations. This is evident since it was most strongly correlated to being a geoxylic suffrutex and no tree and being of low height below 50 cm. Furthermore, according to the graph, it grows in sandy soils, not ferralitic or riverine ones. Its leaves are of leathery texture. Genetically it is positively associated with Western Zambezian Grassland haplotype HT3 and SSR Cluster 1, and negatively correlated to SSR Cluster 2.

The opposite was evident for **SGM**. It is genetically positively associated with SSR Cluster 2 and negatively with Cluster 1. Moreover, it grows in ferralitic soils, not in sandy ones and was characterized as a tree of 2 m to 10 m height with large, soft leaves. It occurred in the northern sampling range (lower latitude) where precipitation seasonality is mild.

SGM sister phenotype **SGMx** had similar associations in genetic aspects, however it was correlated to being a geoxylic shrub of 50 cm to 2 m height instead of a tree. Furthermore, it shared central haplotype HT1. SGMx is associated with lower temperature and mild precipitation seasonality, as well as with low longitude (western sampling range).

In said western range also occurred phenotype **ST**. This phenotype, or maybe its habitat, was characterized by low temperature and temperature range throughout the year, in addition to low precipitation seasonality. Its individuals display small leaves, grow on granite substrate and are 2 m to 10 m high. Genetically they were connected to one of the Escarpment haplotypes (HT5).

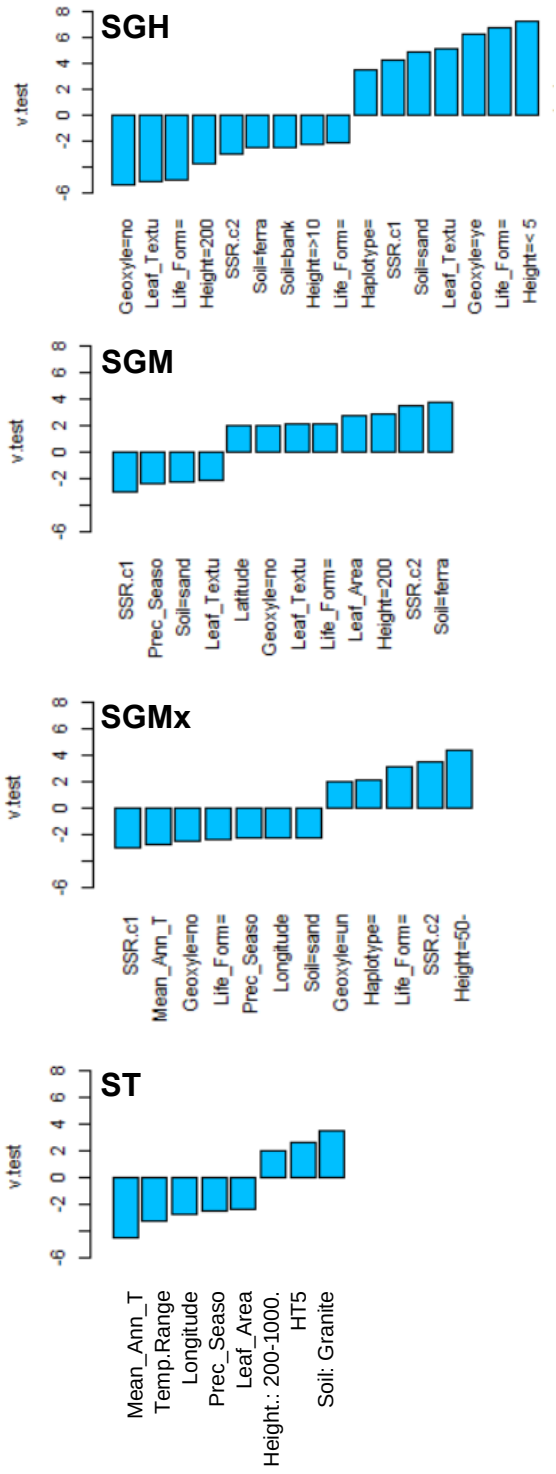



**Figure 37** (continued): Impact factors based on FAMD for each phenotype. Positive and negative associations and the strength of their influence upon each phenotype are shown. Abbreviated parameters: precipitation seasonality (Prec_Seaso), mean annual temperature (Mean_Ann_T), temperature range throughout year (Temp_Range). Categories (e. g. Haplotype=…) are explained in the text if not evident from figure.

As mentioned previously, certain factors are correlated and therefore it is not surprising if they occurred together in the FAMD. The strength and nature of correlation between continuous variables is shown in fig. 38. Climatic and spatial variables displayed the highest correlations which is neither new nor a revelation. All other factors were weakly to moderately correlated (-0.45 < r < 0.42) except for SSR Cluster 1 and Cluster 2, but this had been already treated in III.4.3. Those were however the factors with the strongest correlation apart from climatic/spatial variables. SSR Cluster 2 showed a moderate positive correlation to latitude (r = 0.38) and to annual precipitation (r = 0.42). Cluster 3 on the other hand showed a negative relation to both factors (r = -0.39 for latitude and r = -0.45 for precipitation).

A visualization of all factors and the ordination among them in the context of the FAMD is also given in fig. 39. This enables us to see patterns and correlations not only for the continuous variables as shown above, but also for the categorical ones. The first two dimensions of the FAMD explained 18.08 % and 17.11 %, respectively (the third dimension accounted for 11.90%, the fourth for 9.34%).

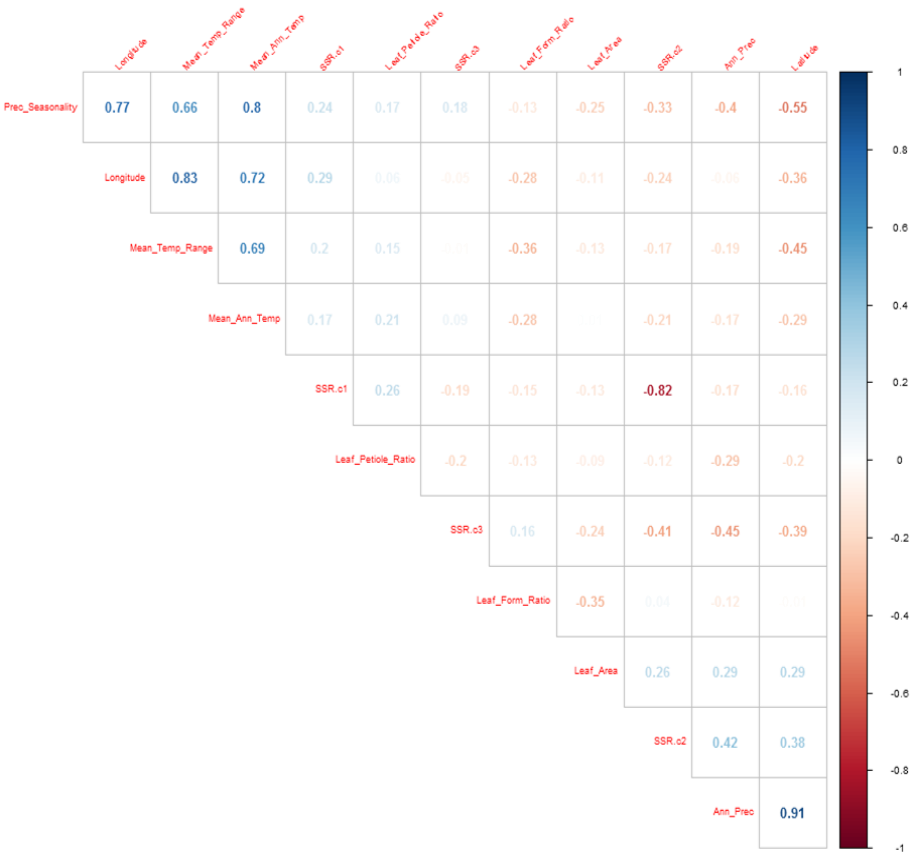



**Figure 38:** Correlation matrix plot of continuous parameters used in the FAMD. Correlations coefficients are calculated with Pearson statistic; the stronger the correlation, the darker the hue. Red colors indicate negative correlation, blue colors a positive one. Abbreviated parameters: Prec. = precipitation, Temp. = temperature, Ann. = annual.

While the continuous variables are shown as arrows radiating from the center, the length representing their strength of correlation, the categorical variables are displayed at their levels, ordinated into space. We see an overall ordination that mirrors the phenotype associations from fig. 37 as factors/factor levels that were grouped close to the phenotypes (SGH, SGM, …) also correlated with them.

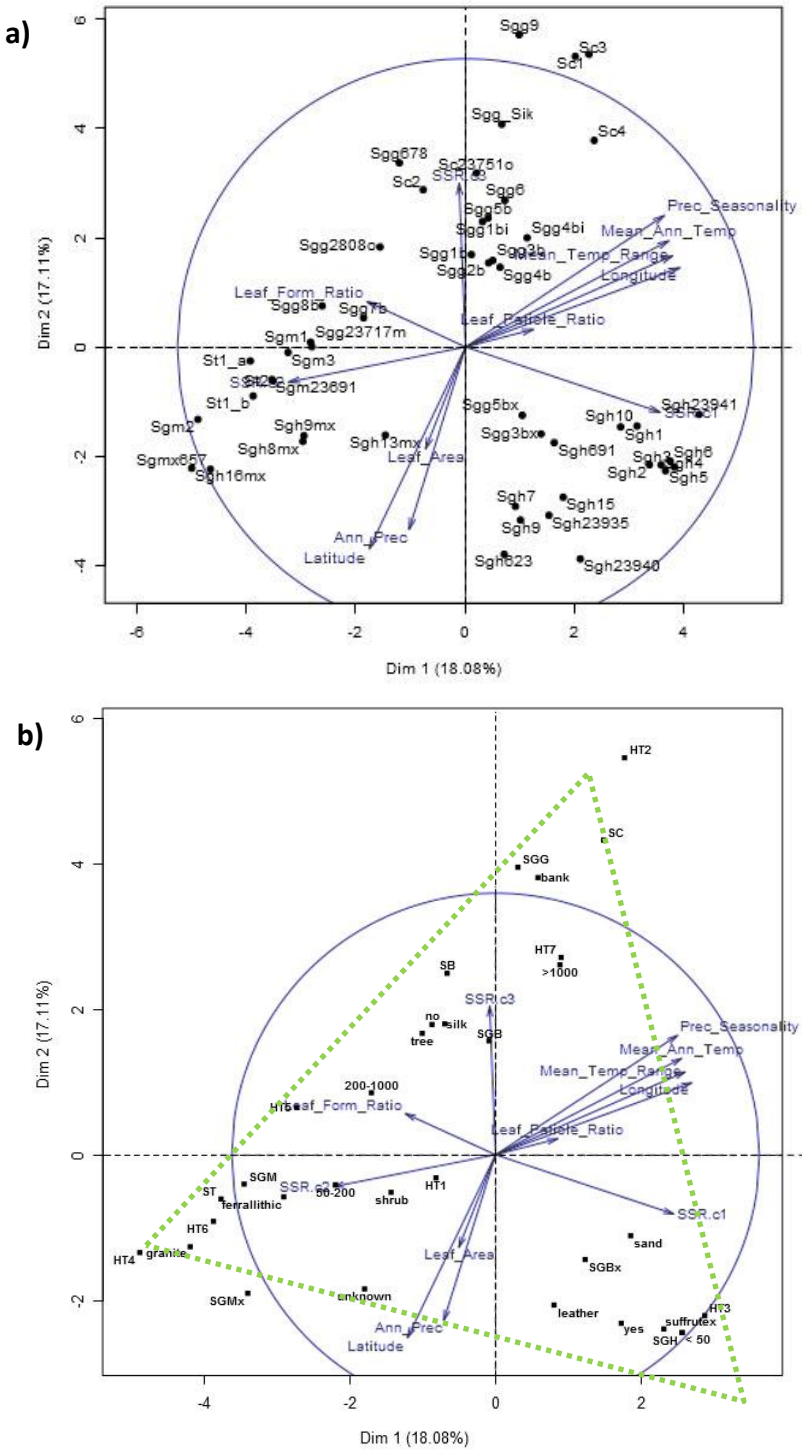



**Figure 39:** a) FAMD ordination of ecologically and genetically characterized populations plus quantitative variables. Some population names are corrected with the actual phenotype assignment if there was a mismatch, e.g. Sgh16mx which was sampled as a SGH but is correctly a SGMx population; b) FAMD ordination of categories from qualitative variables plus quantitative variables. "yes", "no" and "unknown" are the geoxyle categories. Other abbreviations are just as in fig. 37 and 38.

By overlaying the quantitative and qualitative factor map we see how the populations grouped in space and which continuous factors were associated with whom (fig. 39a). The same was done overlaying the quantitative parameters with the factor levels of categorical variables (fig. 39b).

In general, populations clustered according to their phenotypes (in the case of mismatching population names and phenotype assignments, the actual phenotype assignment is provided as lower case letter in the population names end, e.g. Sgh16mx). However, this plot does not echo exactly the ordination from the PFT analysis (see III.1) as much more factors were included here. Hence we have to account for the clustering trend of ecologically or genetically similar populations. As an example, please regard the left site, where Sgh16mx, Sgmx657, St1 and St2 had gathered. Those populations are all from the Angolan Escarpment, but genetically and morphologically distinct (see fig. 13).

Furthermore the same separation between suffrutices and trees is evident as all SGH populations (plus the SGBx's) clustered in the lower right part. The tree populations grouped in the upper and left part and are separated by a noticeable, central gap from the suffrutices. The tree groups themselves formed a band of gradual transition from SGMx to ST, SGM, SB, SGB and SGG/SC (left to right).

Strikingly, the populations ordinated somewhat in a triangle and so did the associated factor levels (see fig. 39b). In the previous chapter I reported about microsatellite results and there was also a triangular pattern evident. This triangle here corresponds very well with the genetic triangle as the three arrows indicating the correlation of SSR clusters point quite accurate into the corners of the triangle. Moreover, the corners here represent suffrutex traits (“Cluster 1”, right), SGM traits (“Cluster 2”, left) and wet habitat traits (“Cluster 3”, top). This pattern allows us to identify key factors associated with the genetic clustering in figure SSR-triangle-plot (fig. 30).

Before suggesting any clusters based on mere observation we should consider the results of the Hierarchical Clustering (HCPC, see fig. 40a). This analysis revealed seven groups that corresponded mostly to the phenotypes. Starting from the left: Group 1 contained the SGBx populations, group 2 all suffrutices and group 3 all SGG’ and SC’s on two branches (SGG left, SC right). Group 4 is characterized by SGB populations which are further divided into populations with haplotype HT7 (right) and those without (left). The fifth group comprised SGM populations on its left branching and Angolan SGG’s and SB’s on the right. Finally, clusters six and seven were attributed by the ST’s and SGMx’s, respectively. In the latter case the populations from Tundavala had split up from the remaining Angolan SGMx’s.

However, the full HCPC data set already contains phenotype assignments of the populations, which influences the clustering algorithm. Excluding the phenotype information from analysis led to another clustering pattern; yet the ordination on the factor map stayed the same and the dimensional variation improved slightly (20.36 % and 18.65 % for first and second axis). Three clusters were apparent in fig. 40b, with the first one (blue) containing all suffrutex populations and the SGBx. The second cluster comprised all Zambian SGB’s, all SC’s and SGG’s plus one SB. The final cluster was represented by all SGMx’s, SGM’s and ST’s, as well as the “dry” SGG’s from Cusseque and one SB. This trilateral clustering again mirrors the genetic structure from SSR analysis (and was therefore colored accordingly).

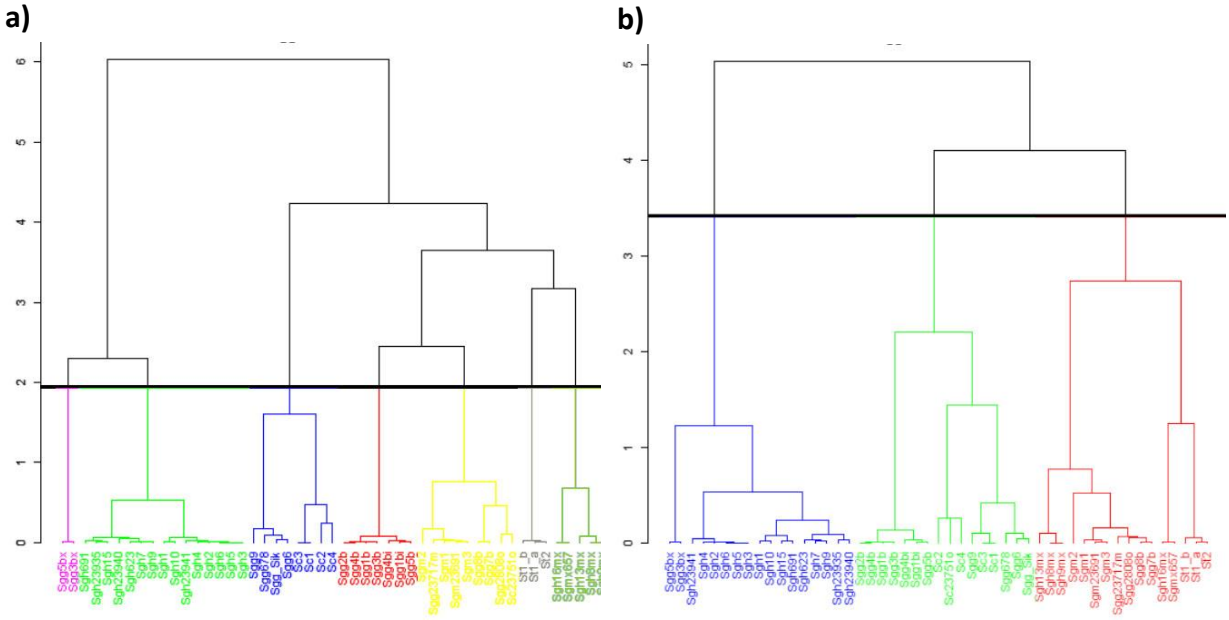



**Figure 40:** Dendrograms of Hierarchical Clustering on Principle Components (HCPC). a) Phenotype assignments implemented in HCPC calculation, leading to a clustering mirroring the phenotypes (indicated by the colors used before); b) calculation without a priori phenotype information, leading to a clustering pattern that resembles the SSR genetic pattern.

# IV. Discussion

Throughout the analysis and evaluation of the results one fact stays very prominent: The high phenotypic variety isn't mirrored in according genetic structure. Furthermore some of the analyzed phenotypes are genetically distinguishable and others are not, and reasons therefor will be discussed in this chapter. I will start with the morphological aspects and then turn to their genetic context, discussing explanations for the found patterns. Finally I want to draw a scenario that illustrates how the *Syzygium* suffrutices in the Miombo region may have evolved and developed the found genetic and ecological characteristics.

## IV.1 Morphological comparison of *Syzygium* phenotypes

The analysis of plant functional traits and therefore phenotypic characteristics confirmed that the defined *Syzygium* phenotypes are distinguishable based on the chosen morphological traits. However, some phenotypes are quite similar to each other, e. g. both suffrutex groups from the WZG and Angolan Miombo are basically identical in appearance despite the great spatial distance; the two slightly diverging populations (Sgh4 and Sgh5) do so due to missing values for some traits. This indicates comparable environmental pressures and probably one common ancestor for both groups, as the respective tree partners from Miombo and WZG are quite different. In the Miombo, SGH's partner is SGM which is morphologically more similar to the suffrutex as SGB in the WZG.

However, the suffrutices in both ecoregions are linked to their tree partners via an intermediate phenotype, SGMx in the Miombo and SGBx in the WZG. Whereas SGMx is clearly identifiable as a phenotypically adapted SGM, SGBx is represented by too few samples to be certain, but morphologically it appears to be a hybrid between SGB and SGH (see IV.2). Thus, assuming SGBx is a hybrid, there already had to be a suffrutex present in order to hybridize with SGB. This implies that SGBx cannot be the evolutionary intermediate step from SGB to SGH. For other possible explanations regarding SGBx see IV.2.

Regardless of the origin of the intermediary phenotypes, we find morphological gradients over short spatial distances between SGM-SGMx-SGH and SGB-SGBx-SGH, respectively. Both gradients are shaped by the steep environmental gradient between forests and grasslands (Swenson et al., 2011). They stretch from forests to grasslands, the only difference is the topography; SGM-SGMx-SGH grows in an undulating landscape, SGB-SGBx-SGH on rather flat plains. Therefore tree growth is possible within the forests and maybe at their fringes, but the increasingly adverse environmental conditions (e.g. frost, fire, water logged soils, herbivores; Beerling & Osborne, 2006) filter the individuals gradually and with increasing strength towards the open grassland (Lebrija-Trejos et al., 2010). Hence individuals tend to be lower and to show increasing adaptation by relocating underground (Havström et al., 1993).

On the other side of the morphological spectrum are SGB, SGG and SB clustered closely together and overlapping. Their similar appearance is the reason for the uncertain phenotype assignment of some samples from these groups, as well as for SC. After the first genetic

screening some of those individuals had been reassigned to new phenotypes due to absolutely not matching genotypes (e.g. B01g or G06o). However, some individuals with unclear genotypes and phenotypes remained in their originally allocated group, what may have caused distortion in the genetic analyses (e.g. G08, a herbarium specimen, see IV.2.2.3).

Particularly SGG includes several individuals that deviate from their "genetic mean" as identified by STRUCTURE analysis. Typically, when associated with riverine habitats, they exhibited a high affiliation to SSR Cluster 3 and no trace of Cluster 2. There are however populations from Central Angolan Miombo forests, for example near Cusseque, showing mixed genetic and phenotypic characteristics of SGG and SGB. Appropriately, those individuals with higher Cluster 3 affiliations and SGG appearance had been assigned to SGG ("dry SGG", not riverine), and those with higher Cluster 2 affiliation and SGB appearance to SGB. Perhaps the better way to deal with this conundrum would be to define another phenotype/entity which is very similar to SGG and SGB but occupies a different habitat (dense forests) and shows a varying threefold genetic pattern between the three clusters. This mixed genetic pattern then indicates that indeed the founders of these dense forest populations were of different phenotypes which intermingled and matched their appearance over time (Harbaugh et al., 2009).

Overall, the morphological analysis of *Syzygium* phenotypes presented in this study is merely a start to quantify distinguishable characteristics. For more reliable data in the future it will be necessary to include more traits, for example of fruits and flowers which have not been included here as they weren't available for all phenotypes. Furthermore, a larger sample number per phenotype is required, as well as measurement in vivo, in the field on living individuals. The measurement on dried herbarium specimens can lead to biased results as this pre-selected material does not necessarily represent the actual traits (though it should).

Another aspect is the confounding taxonomic diversity of *Syzygium guineense* subspecies and varieties in literature. In the Conspectus Florae Angolensis (Exell et al., 1966), the most reliable reference, six *Syzygium guineense* subspecies are described plus five congeneric species *S. huillense* (suffrutex), *S. cordatum*, *S. benguellense*, *S. rowlandii* and *S. owariense*. The two latter ones are distributed farther north towards the rainforest biomes and were probably not encountered in the sampling of this study. *Syzygium cordatum* and *S. benguellense* however were sampled and genotyped, and they show a close genetic similarity which does not support their status of an own species (Mallet, 1995, Abbott & Comes, 2007).

Strikingly, the suffrutex is assigned by Exell et al., (1966) to an own species, *Syzygium huillense*, and not marked as a subspecies of *S. guineense*. Of those, six are described: ssp. *littorale*, ssp. *macrocarpum*, ssp. *barotsense*, ssp. *afromontanum*, ssp. *urophyllum* and ssp. *guineense*. *S. g.* ssp. *urophyllum* is said to be found in Northern Angola and is therefore not covered in my study as well. Phenotypes that corresponded most to descriptions were named accordingly in this study; however, intermediate traits complicated the labelling.

The variety of not well delimitated *Syzygium* taxa proposed for Angola seems to produce many misidentifications. A thorough screening of the *Syzygium* specimens in the herbaria LUB at ISCED, Lubango, and HBG at University of Hamburg showed that many identical looking specimens were given different names by different collectors; and some different looking specimens were given the same name (see fig. 41). Because of this unclear, confusing labelling, I used "*S. tundavalense"* as I found three different names for the ST phenotype, all collected near Tundavala. If we take the phenotypically (and genetically) similar T04 from Tanzania into account, we even have four names: *S. g.* ssp. *littorale*, *S. g.* ssp. *afromontanum*, *S. g.* ssp. *guineense* and *S. g.* ssp. *angustifolia* (an epithet which is nowhere mentioned in literature). With emphasis on morphological accuracy, *S. g.* ssp *littorale* would be the best fitting referenced label for "*S. tundavalense*", particularly as specimen from Humpata near Tundavala had been labelled *S. g.* ssp *littorale* before, in 1961 (fig. 41a). But considering the habitat (mangroves vs. mountains), the best fit would definitively be *S. g.* ssp. *afromontanum*. However, the oldest annotation is usually valid in taxonomic naming.


**Figure 41:** Herbarium specimen from ISCED Lubango (LUB) of *Syzygium guineense* ssp. *littorale* from Tundavala, Angola (a); Specimen of *S. g.* ssp. *littorale/ afromontanum/gerrardii* from Tundavala, Angola (b); *Syzygium sp.* from Tundavala, Angola (c); self-sampled "*S. tundavalense*" from Tundavala, Angola (d). Photos: P. Zigelski 2014-2015, specimen photos taken with permission of F. Lages.

## IV.2 Genetics

### IV.2.1 The unresolved group

Sequence analysis by means of phylogenetic trees and haplotype network analysis revealed that there is little genetic differentiation between some phenotypes. In both gene sequence markers there is a large group of phenotypically mixed individuals that could not be resolved further, this is particularly true for ITS. This low resolution in ITS is caused by heterozygous individuals who display ambiguous positions in their sequences. As ITS is part of a nuclear gene and as at least some *Syzygium* phenotypes are tetraploid, this makes the correct inference of each haploid sequence via Sanger-sequencing as I conducted it impossible. Affected sites had therefore been excluded from analysis. As *trnK* already is haploid, *trnK* sequence analysis wasn't affected by this problem (Birky, 1978). Therefore less parsimony informative sites are needed for genetic resolution, but the central unresolved group persisted nonetheless. This group contains individuals from a wide geographical area, ranging from the Okavango Delta to Central African Rainforests and from Angola's coastline to Tanzania.

The results from microsatellite analyses are more informative but display a similar pattern as above. According to STRUCTURE, the majority of individuals had no clear assignment into one of the three genetic clusters but was mostly divided between two of these genetic groups. There was also one phenotypically mixed cluster evident in the PCoA which contained ST, SGB and SB (regarding the first 3 dimensions), which correspond to the unclear individuals (those that deviated genetically from the other members of their phenotype) in STRUCTURE (see III.4.3; tab. 13, fig. 30). In summary, regarding all genetic analyses combined, I can identify a group of individuals with an unclear genetic pattern; most of them belong to SB and SGB. ST is differentiated on the chloroplast, but not on the nuclear level which might be due to the reduced informative power of ITS.

Both SGB and SB are on the one hand groups with some uncertain phenotype assignments and reclassifications due to ambiguous genetic patterns. ST on the other hand is phenotypically clearly defined but genetically heterogeneous. Ecologically, according to FAMD (see III.5), it is closest to SGM/SGMx but still somewhat different, an own ecotype which is as yet not genetically profiled – at least not on the nuclear genome (see III.2.1). The differentiated chloroplast haplotype however could well be derived from the other side of the continent – the herbarium specimen T04 from Tanzania and sample T11 from Tundavala, Angola (not in SSR analysis) both share the same haplotype HT6, and the other escarpment haplotype HT5 is derived from it (see III.3). But how exactly HT6 spread from east to west, or if it was the other way round or just long distance dispersal, is unknown.

Returning to SGB and SB it can be stated that they are hard to characterize. SGB is morphologically very distinct due to its growth height, but genetically difficult to grasp. It appears to be diploid and is therefore ad hoc set apart, but nonetheless genetically close to SB and SGG (PERMANOVA, III.4.2, tab. 10) and to SGH, SGBx and SB (low fixation index, meaning genetic exchange). The inbreeding coefficient of SGB, SGG, SB and ST however was higher than average, indicating a reduced outcrossing. One possible explanation could be a recent and rapid abortion of genetic exchange with other phenotypes, maybe because the others polyploidized (Parisod et al., 2010), though the mentioned values are only weak indicators. After polyploidization some populations differentiated into new phenotypes (for example SB), leaving SGB (and ST) with their

diploid genetic pattern behind (Abbott & Comes, 2007; Ramsay et al., 2008). Cases where populations of mixed ploidy coexist and don't or seldom cross are long known for diverse other plant species (Hardy et al., 2000; Husband & Sabara, 2004).

Though SB and SC are phenotypically and ecologically similar and hard to differentiate, they seem to originate from different populations. SC is discussed in the next section in detail; it is set apart from SB by its mixed SSR Cluster 1/3 ancestry, whereas SB shows a mixed 1/2 ancestry. Furthermore not all of them share the same *trnK* haplotypes, except HT1, which are rather spatially than phenotypically patterned. Also according to FAMD, SB is very close to SC. Taking the findings for these unsettled groups together, a possible explanation could be that ST and SGB differentiated from an unknown, diploid progenitor before it underwent polyploidization. SB could have differentiated from this now polyploid progenitor as it is illustrated for other plant species in Husband & Sabara (2004) and Soltis et al. (2004). The markers or the sample set used in this study might not suffice to resolve these phenotypes genetically further; they therefore appear to be not differentiated. Additional and more thorough analysis might detect the genetic difference in future. The remaining groups are presented in the following sections.

### IV.2.2 The wet group

#### IV.2.2.1 The tailwaters of Southern Africa

The groups that were genetically resolved represent certain ecotypes and they appeared throughout all means of genetic analyses of *Syzygium* phenotypes used in this study. They differed significantly from the unresolved groups mentioned before. For one, there is the genetically fixed "wet ecotype" occurring in rivers and wetlands (e.g. population Sgg9), found even at great distances– the herbarium specimen G12 was sampled in Lake Lutamba in Tanzania. Individuals characterized by this genotype appear as small to medium trees that are always water associated, growing at or even in rivers. They are furthermore correlated with smaller leaf area, soft leaves and lower annual precipitation. This might be caused by the fact that plants growing close to water bodies have access to sufficiently high groundwater levels and are therefore less dependent on regular rainfalls (Oldeland et al., 2013, pp 19-22).

This ecotype is distinctly present in the *trnK* analysis as haplotype HT2 (III.3, fig. 17) or clade (e) in the phylogram (III.2.2, fig. 16), as well supported clade (b) in the ITS phylogram (III.2.1, fig. 15) and in the SSR analysis as Cluster 3 (III.4.3, fig. 26, tab. 12/13). In the SSR analysis, the ecotype included SGG's with an almost perfect Cluster 3 affiliation as well as SGG's with genetic contributions from both Cluster 1 and 3.It is further clearly identifiable in the FAMD and HCPC analyses as an own triangle corner (see fig. 39) and clade (see fig. 40), respectively.

In the genetic analysis, we also see a case of a gradual increase of Cluster 3 influence along the Zambezi River: G04, sampled in the Barotse Plains, is 64%/25% divided between Cluster 1 and 3, respectively; G05 from near Sioma 170 km further downstream shows a ratio of 36%/56% and finally B01g from Victoria Falls more than 300 km further shows 2%/98%. The latter individual is hence genotypically identical with the samples from the lower Okavango River.

Interestingly, the drainage routes and directions of Okavango, Zambezi and other minor rivers in this region changed in the past (Moore & Larkin, 2001; Burrough et al., 2009). The Okavango, Chobe, Chambeshi, Kafue and Zambezi had been connected and were capturing each other over time, draining endorheically into former lake systems in the early and mid-Pleistocene (approx. 2.2 – 0.5 mya) (Burrough et al., 2009; Moore & Eckardt, 2012). The Okavango Delta was once part of a much larger lake system extending over 80.000 km², forming the second biggest lake in Africa, the Makgadikgadi paleo-lake (Burrough et al., 2009). Today, both are only relicts of this paleo-lake which emerged around 2 mya and drained c. 4000 ya (Moore & Eckardt, 2012). This is just one example of recent interconnectivity between river associated wetlands that enabled exchanges of diaspores and represent a once much greater habitat. It might therefore explain the pattern of genetic similarity between individuals of different, nowadays unconnected river systems in South-Central Africa (Moore & Larkin, 2001).

### IV.2.2.2 Hybrids in the Okavango Delta?

Besides the SGG's from the Okavango and Zambezi rivers there are also the *S. cordatum* populations Sc1 and Sc3 from the Okavango Delta which exhibit a mixed affiliation between Cluster 3 and 1. They are furthermore morphologically intermediate between SGG and the typical *S. cordatum* as described in Exell et al., (1966) and Coates Palgrave (2002). They state that *S. cordatum* has sessile leaves with cordate base and rounded apex which are sitting on tetragonal twigs. The SC from the Okavango Delta however have rather subsessile or short-petiolate leaves, with tapering base and blunt apex (fig. 42). The only SC in the sample set in this study which is fully in accordance with the description of *S.cordatum* is sample C08 from the Cuando River in the Caprivi Strip, Namibia. It is the only SC sample perfectly affiliated to Cluster 1. It is therefore possible that the SC-populations of the Delta are of hybrid origin, or at least a product of gene flow between *S. cordatum* and *S. guineense* ssp. *guineense*. This origin is further supported by the fact that the Okavango River SGG and the Okavango Delta SC share the same haplotype, HT2, which can indicate chloroplast capture and hybridization (Smith & Sytsma, 1999; Tsitrone et al., 2003, Stegemann et al., 2012). Such a hybrid has already been proposed as *S. intermedium* from Zimbabwe, occurring also in the Okavango Delta (Coates Palgrave, 2002, JSTOR Global Plants: *Syzygium intermedium*).


**Figure 42:** (a) Self-sampled SC from the Okavango Delta. Bases are definitively tapering and petioles are indicated. (b) *S. cordatum* from Zimbabwe. Consider the rounded tips, cordate base and the sessile leaves in the picture. (c) *S. cordatum* from Malanje, Northern Angola: a specimen from the ISCED Herbario, Lubango. Photo: P. Zigelski 2014.

Image source for (b):

http://www.zimbabweflora.co.zw/speciesdata/image-display.php?species_id=142310&image_id=14

### IV.2.2.3 The Okavango – escarpment connection

Considerable Cluster 3 affiliation had been mostly found in individuals of before mentioned tailwaters, but also in the Kunene catchment and, surprisingly, in some individuals of the Angolan Escarpment (T07, T10). If there actually had been a gene flow in the (recent) past, it could either be from the escarpment to the Okavango/Zambezi basin (enrichment of the Cluster 3 specific genetic genepool) or vice versa (the translocated Cluster 3 genepool lost contact to its original genepool and got gradually substituted by genetic characteristics of Cluster 2 and Cluster 1 (see fig. 31). However the latter seems more likely as a SGG population from the Kunene catchment (Sgg678) shows perfect Cluster 2 affiliations. From there it is only 160 km direct line to the escarpment, instead of the 900 km between the Escarpment and the lower Okavango.

Though both habitats are perannually wet and therefore somewhat decoupled from rainfall seasonality (SASSCAL WeatherNet, 2015: Tundavala station), it is unlikely that a water course-mediated gene flow took place as there are no direct river connections linking all three sites today – the Okavango is endorheic, draining to the south-east and the Kunene draining west into the Atlantic Ocean. There is however the possibility that frugivore birds or mammals dispersed the seeds as they are edible and readily collected (Howe & Smallwood, 1982; Schupp, 1999). Where the environment was suitably humid, the seeds could sprout and genetic exchange in form of outcrossing or hybridization even at long distances could have happened (Tsitrone et al., 2003).


**Figure 43:** Herbarium specimen G08 from ISCED herbario Lubango. The label names Cascata Huíla in Angola as sampling location.

The herbarium sample G08 from Cascata, Huíla, is morphologically and by specimen name a SGG (see fig. 43). However, its nuclear genetic pattern (which was verified through repeated analysis) is very untypical as it is mostly derived from Cluster 2 and it is the only SGG with this pattern (a chloroplast genotype is unfortunately not available). This indicates a misclassification by the collector (despite the morphological accordance) and a different phenotype, probably the genetically similar ST. The sampling site Cascata is a waterfall in a protected canyon south of Lubango and only 30 km away from Tundavala. It is therefore possible that G08 is in fact G08t, a ST growing in a protected and humid place, due to which it developed the shape and morphological features typical of a SGG (Havström et al., 1993; Niklas, 1996).

In former times however, just as the Okavango and the Zambezi have a history changing courses and captures, this also applies to the Okavango/Cubango and the Kunene River. Both Kunene as well as Cubango Rivers originate in the Angolan Highlands, their catchments lying adjacent and parallel with no marked topographic barrier. The catchments embrace the Cuvelai drainage system farther south, whose ephemeral waters trickle off in the Owambo Basin (Wellington, 1955; Hipondoka, 2005). This Basin however, with its world-famous

Etosha Pans constituting the lowest parts, had in geologically recent times not only been the destination of the Cuvelai, but also of the Kunene and the Cubango River. These three river systems fed the water budget of the pans, forming another Palaeo-lake: the Etosha Lake. This lake existed in wetter climate cycles until approx. 2 mya ago, it dried out successively after the inflow cessation of Kunene and Cubango around 4 mya, being sustained only by the ephemeral Cuvelai system (Hipondoka, 2005, Miller et al., 2010). This circumstance represents another recent linkage of river wetlands suitable for genetic exchange. In addition to the recent connectivity of the tailwaters are the adjacent catchments of Cubango, Cuvelai and Kunene, as mentioned before, placed in a nearly “barrier-free” region of parallel, southwards directing drainage lines. The close proximity could therefore allow for genetic exchange via animals (Howe & Smallwood, 1982).

The time frame is recent enough to leave an imprint in the genetic pattern of the wet *Syzygium* ecotype, however it is still only a hint. To resolve this sketchy picture it would be necessary to sample more populations from the catchments of the affected hydrological systems, as riverine specimens are underrepresented in this study. This would help to see the routes of gene flow clearer and to check if the genetic patterns correspond to river courses or rather to biotic vectors.

### IV.2.3 The suffrutex group

#### IV.2.3.1 The "southern suffrutex"

The second genetic entity is the suffrutex (SGH) which is much better sampled and had therefore been expected to reveal a clearer picture. It is, however, heterogeneous in nature as for each method used there were always some outliers (verified by repeated analyses). At least the same individuals' proved to be outliers (e.g. H17 (Sgh3), H18 (Sgh4), H29 (Sgh10), H30 (Sgh11) or H42 (Sgh1)), but so far no convincing connecting pattern among them has been identified. One notable linkage is a geographic pattern as the genetically deviating suffrutices are mostly, but not entirely, located towards the southern limit of the sampling range and phenotype distribution. The slight spatio-genetic correlation within SGH in the Mantel test results (see III.4.4) is also a hint to this direction (Manel et al., 2003).

Said individuals are spread over two not supported branches (c) and (d) within the ITS analyses and are also distinguished, but with low support, within *trnK* as these are the individuals that share HT1 and not HT3 (see III.2.1; III.3.1). SSR analysis however does not support a conclusive genetic differentiation from the other suffrutices. Nevertheless these southern suffrutices are somewhat different to the classic Miombo and Western Zambezian Grassland ones, representing a kind of intermediary. Though they are of habitat requirements and morphology similar to the Western Zambezian Grassland suffrutices (pers. obs.), they constitute a genetic variation of them, growing in flat, sandy plains that suffer from water excess or logging in rainy season and longer and more pronounced droughts as well as temperature amplitudes during the dry season (so called dambos: Perera, 1982; WWF AT0724).

This additional environmental pressure could be the reason for their genetic divergence and water logging is Whites (1977) favorite explanation for the suffrutex habit, together with poor edaphic conditions. Nowadays, however, there is not much support for this hypothesis, as the suffrutic growth form could only be advantageous, compared to trees, if they would exhibit additional adaptations to inundated soils. Otherwise, both trees and suffrutices would suffer the same lack of oxygen as their roots and underground stems cannot transpirate (Crawford, 1982). Typical and widespread suffrutex species from Central Angola, however, don't display any primary adaptation to waterlogged soils; they lack aerenchymatic tissue, adventitious roots or increased porosity (Sanguino, 2015). They are furthermore not growing directly in the inundated sites of the valley bottoms; these boggy places are restricted to Cyperaceae and other inundation tolerant species. The suffrutices are rather growing above the riverine bogs on the lower slopes where water leaks out only in times of high precipitation (pers. obs.; Oldeland et al., 2013, pp 43-44). The suffrutices are therefore not directly affected by water logging. Hence it is likely that the genetic pattern should be similar due to similar pressures and that the seasonal water logging does not leave such an imprint in the "southern suffrutices".

To conclude, the genetic divergence is slight and not exactly locatable in the SSR results, but a trend is present. The affiliation to SSR Cluster 1 may be decisive. To the by ITS identified southern suffrutex group belong ca. 63 % Cluster 1 individuals (e.g. H20, H29) and 37% mixed cluster individuals (e.g.H11, H18) which is only a weak correspondence. However this might be the result of a genetic regime in transition in this area, where the process of genetic divergence is just in progress (Fitzpatrick et al., 2009). Supporting this assumption is the fact that in the eastern ranges of the "southern suffrutex", especially in Liuwa Plains, Zambia, they intermingle with the other genetic WZG type, represented by *trnK* haplotype HT3. The intergenetic spacer *trnK* is

furthermore the only marker that distinguishes between the southern suffrutices plus those of the Miombo biome (undulating landscape) and those of the Western Zambezian Grassland biome (flat, seasonally water logged). The first shared HT1, the latter HT3.

If the south-eastern distribution range really is crucial for this genetic pattern, then not only topographically flat or grassy plains might be explanatory but also the climatic conditions as mentioned before. Especially the southernmost populations Sgh10 and Sgh11 received 2015 only approx. 428 mm precipitation with a maximum temperature of 39.8°C (SASSCAL WeatherNet: Nangweshi station), whereas the northernmost population Sgh623 received two third more rainfalls: 705 mm and maximum temperature of 32.4°C (SASSCAL WeatherNet: Huambo station). Also precipitation seasonality (e.g. 3 vs 4 completely dry months in 2015, other years comparable) and temperature range (3.4°C – 39.8°C vs 6.7°C – 32.4°C in 2015, other years comparable) are harsher for the southern populations, though microclimatic conditions are not well represented here. The divergent genetic pattern might therefore be a response to the more challenging environmental conditions at the southern fringe of the Miombo (examples in Antonovics, 1971 and Mopper et al., 1991).

#### IV.2.3.2 Is SGBx of hybrid origin?

Haplotype HT3 is furthermore shared by the SGBx individuals who occur sympatrically with WZG-suffrutices (SGH) and SGB in Liuwa Plains, Zambia. The SSR analyses revealed a close relatedness of SGBx towards the suffrutices and SGB to whom there was no significant genetic distance in the PERMANOVA (see III.4.2). All three phenotypes however show a mixed genetic pattern between Cluster 1 and 2, whereby SGH had the closest Cluster 1 affiliation, followed by SGBx and SGB (see III.4.3). Furthermore, the selfing rate was high in SGBx, indicating that gene flow between SGBx and other individuals and populations is restricted (III.4.3). It is however uncertain if this *Syzygium guineense* phenotype is self-compatible; a study conducted on *Syzygium guineense* in Cameroon (therefore probably a SGM) indicates a mixed auto – and allogamy (with emphasis on allogamy), whereby mostly bees are responsible for pollination (Fohouo, 2008). Together with the intermediate morphological appearance of SGBx all these findings point to occasional gene flow and exchange events and perhaps an origin by hybridization as it is known from other species (Smith & Sytsma, 1990; Tsitrone et al., 2003, Stegemann et al., 2012). In the latter case SGBx would have inherited the chloroplast genome of its suffrutex mother, SGH (mechanism reviewed in Birky, 1978).

An important step to proof this hypothesis would be to determine exactly the ploidy level of SGH, SGBx and SGB and to check their fertility. My study confirms that SGH is tetraploid and suggests that SGB is diploid, and in an ideal world SGBx would then be triploid and infertile or incompatible with di- and tetraploids (Gottlieb, 1973; Soltis & Soltis, 2009). However these findings are based on the maximum observed number of alleles at one locus. Only four individuals of SGBx had been investigated of which three showed a maximum of three alleles and one a maximum of four alleles. Thus, the case is at the moment inconclusive, therefore SGBx was counted as tetraploid. Anyway, there are also cases were hybrids inherit the ploidy level of either of their parents (Soltis & Soltis, 2009) which in the latter (tetraploid) case would be the suffrutex. Some studies indicate that captured chloroplasts and/or hybridization can have advantageous effects, rendering the individuals more resilient to environmental pressures despite the problem of incompatibility (Ramsey et al., 2007; Parisod et al., 2010; te Beest et al., 2011).

If no hybridization occurred, there would be alternative (but less probable) explanations for the genetic pattern and suffrutex haplotype in SGBx. For one, SGBx might be a former suffrutex which lost the geoxylic habit due to lessened environmental stresses and returned to the tree habit. Similar cases of facultative tree growth are known from other suffrutex species (*Salacia kraussii* (Celastraceae)*, Ochna macrocalyx* (Ochnaceae)*, Paropsia brazzeana* (Passifloraceae)) of the Miombo region, though the genetic background is unknown (White, 1977; Coates Palgrave, 2002).

#### IV.2.3.3 The emergence of suffrutices

The third subgroup of suffrutices can be located mostly in the northwestern sampling range, in Huíla, Bíe and the northern Cuando-Cubango province, Angola (e.g. H02 (Sgh23935), H09/10 (Sgh23961) or H24 (Sgh7)). Most of the samples collected there (67%) show to a varying degree a mixed ancestry between SSR Cluster 1 and 2 with up to 77 % of Cluster 2. This is in contrast with the "southern suffrutex" mainly belonging to Cluster 1. It is much more likely that some variable that had not been considered in this study shapes the fine scale genetic pattern, particularly since the same genetic patterns occur over great distances and multiple genetic patterns occur within one population. The still open genetic questions in this study are furthermore evident by the fact that the *trnK* haplotypes shared by suffrutices (HT1 and HT3) are not directly linked but bypassed by the Okavango haplotype HT2, shared by SC and SGG. Considering the suffrutex closeness in SSR results, it is likely that there are some yet undetected haplotypes linking HT1 and HT3 directly.

One conclusion from this confounding genetic pattern can be drawn, namely that within the studied area (southwestern Miombo) there weren't more than two independent emergences of the suffrutex habit. This assumption is based on the twofold ancestry pattern within the phenotype (Cluster 1 and 2). There are two possibilities for this to arise:

(1) The suffrutex SGH is derived from two genetically distinct progenitor tree populations, one representing Cluster 1 the other Cluster 2. This implies that there should be a tree phenotype strongly affiliated to Cluster 2 (which there is: SGM) and one affiliated to Cluster 1 (which is not identified so far, not with the same extent of affiliation). There is furthermore a stronger geographic patterning expected as individuals derived from Cluster 1 should correspond genetically to the region of their origin and vice versa for Cluster 2, provided there was no means of genetic exchange beforehand. We see however a genetic pattern that does not clearly correspond geographically, at least not in the studied region. As the Miombo stretches still far to the east and northeast, it could be possible to find the other progenitor there, provided that the other progenitor really is as strongly affiliated to Cluster 1 as SGM is to Cluster 2. Another possibility is that the "Cluster 1 progenitor" itself is also of mixed ancestry, in this case SGB seems very likely as it occurs sympatrically in the Western Zambezian Grasslands towards the east. Some SGB individuals could have undergone polyploidization and were therefore favored to adapt rapidly to new environments and evolve the suffrutex habit. Examples of advantageous polyploidization, radiation and habitat adaption in other plant species are given in Ramsey et al. (2007) or te Beest et al. (2011). And in this case SGBx would represent the intermediate form like in the study of Haghighi & Ascher (1988) on beans. Hence the south-west of the Miombo region would be a contact zone between the "Cluster 2 suffrutices" and the "Cluster 1 suffrutices", whereby the Cluster 1 suffrutices seem to be more dominant. They make out the higher percentage within the

phenotype. But, morphologically there is no obvious difference. Examples of population contact zones are given in Hardy et al. (2000) or Kolář et al. (2009).

(2) The suffrutex SGH is derived from one progenitor tree genotype. This would be SGM (and SGMx, but more to that later) which is most strongly affiliated to Cluster 2. In this case all suffrutices evolved from SGM/SGMx by an increasing "Cluster 1 genepool", whereby individuals show different degrees of this divergence which could hint at the time elapsed/spatial distance gained since differentiation, a basic genetic feature true for all organisms (Ayala et al., 1974). The most probable evolutive region is a landscape like the one on the Bíe Plateau, Angola, where steep ecological gradients between forests and grasslands over small distances promote local adaptations (review and examples in Antonovics, 1971 and Rainey & Travisano, 1998). SGM/SGMx grows on ferralitic soils in the forests and forest-grassland ecotones of the hill ridges and slopes. Facilitated by the phenotypic plasticity which is in turn promoted by polyploidy (reviewed in Chen, 2007; Leitch & Leitch, 2008 and Parisod et al., 2010), some populations could have "moved" from the ferralitic hills to the sandy valleys, and survived this through gradual phenotypic adaptation. Once reaching the valley bottoms the individuals from those populations were exposed to different abiotic (frost, fire; Higgins et al., 2000; Oldeland et al., 2013, pp. 55-57; Finckh et al., 2016) and biotic (herbivory, competition with C4 grasses; Belsky, 1994; Sankaran et al., 2005) pressures which led to a genetic fixation of the phenotypical change over time (Pigliucci et al., 2006) which we see in the genetic pattern as "Cluster 1 genepool". I will return to this point in the next section.

At the current state of knowledge, both assumptions are possible. It should be stated that a direct evolutionary link between SGM/SGMx and SGH is highly probable as the former is able to form geoxylic structures given the needed environmental pressure. However, if SGB, SGBx and SGH are also directly linked in the same fashion remains unclear. SGBx is probably not geoxylic, at least to the same extent as SGMx, as the sampled individuals did not exhibit closely clustered, multi-stemmed growth patterns.

Wherever the suffrutex first emerged in the sandy valleys, from there it spread rapidly into similar habitats following those valleys (Willson, 1993) being dispersed by birds or other animals (Howe & Smallwood, 1982; Schupp, 1999) along the same or to adjacent valleys.

### IV.2.4 The ecotone group

#### IV.2.4.1 SGM and SGMx are the same

Finally, there is the genetic group of SGM and SGMx individuals. Initially, they had been regarded as two distinct groups of small trees growing in open woodland (SGM) and low shrubs growing in woodland-grassland ecotones (SGMx). However, in all analyses they had behaved very similar, making no genetic difference between them. Sequence analysis in both ITS and *trnK* showed them together in the unresolved clades (e) (ITS) and (d)/HT1 (*trnK*), though in ITS a sub-clade (a) was prevalent with some separated individuals. Also, population Sgm2 from Benguela province was set apart from the others in both sequences ((a) ITS and clade (e)/HT4 *trnK;* see III.2.1.-III.3.1.; fig. 15-17) but not in the microsatellite data.

The strongest evidence however is given by SSR analysis. No other phenotype grouped so consistently with high association to its respective cluster (Cluster 2, see III.4.3). Therefore, I present proof that SGM and SGMx are one genuine genetic unit which is markedly different from the other phenotypes. The SGM/SGMx entity inhabits open forests, forest-grassland ecotones and the neighbouring grasslands, all on ferralitic and psammoferralitic soils, and is characterized by a challenging environment prone to disturbances by frost and fires (Oldeland et al., 2013, pp 55-57). This unstable environment promoted the persistence of individuals with increased phenotypic plasticity and heterozygosity, as in the concept of environmental filtering (Lebrija-Trejos, 2010; Parisod et al., 2010, te Beest et al., 2011). The suffrutex phenotypes are possibly captured in a state of permanent post-disturbance recovery.

#### IV.2.4.2 The suffrutex progenitor

The polyploidy, the uniform genetic pattern which is identical throughout a wide geographical range (from Angola up to DRC and Cameroon), combined with the ecological flexibility (SGM: small to medium tree in forests, SGMx: small geoxylic shrub in grasslands) suggests strongly, that *Syzygium guineense* ssp. *macrocarpum* is one or even the progenitor of the suffrutex habit. Several authors reported that polyploidy facilitates colonization of novel habitats by enhancing phenotypic plasticity and heterozygosity and by altering physiology and morphology, leading to features that enable polyploids to spread and settle in originally unfavorable environments (Schlichting, 1986; Leitch & Leitch, 2008; Parisod et al., 2010, te Beest et al., 2011). Parisod et al. (2010) also state that such a genetic disposition would be particularly useful in changing, unstable environments.

To elucidate the suffrutex evolution from the tree's point of view, let us imagine a SGM population growing in the forests on the ferralitic hill tops. It is polyploid and due to that phenotypically responsive to changing environments. Some individuals growing closer to the ecotone adapted their appearance and resource allocation according to the harder conditions and increasingly frequent disturbance events by frost or fire; they got smaller, more multi-stemmed and tended to relocate resources and buds belowground (pers. obs., see fig. 44) as they had to resprout frequently due to frosts and fires (Bond & Midgley, 2001; Burrows et al., 2010). Not only allocation below ground is a protective response, but also the multiple trunks. The loss of one or more stems is insofar no longer lethal but can be compensated by remaining stems or multi-stemmed resprouting, and such a feature is often evident in frequently disturbed habitats (Bellingham & Sparrow, 2009). All these adaptations have not necessarily fixed genetic basis (yet), they could also likely be triggered by an epigenetic response (Chen, 2007).

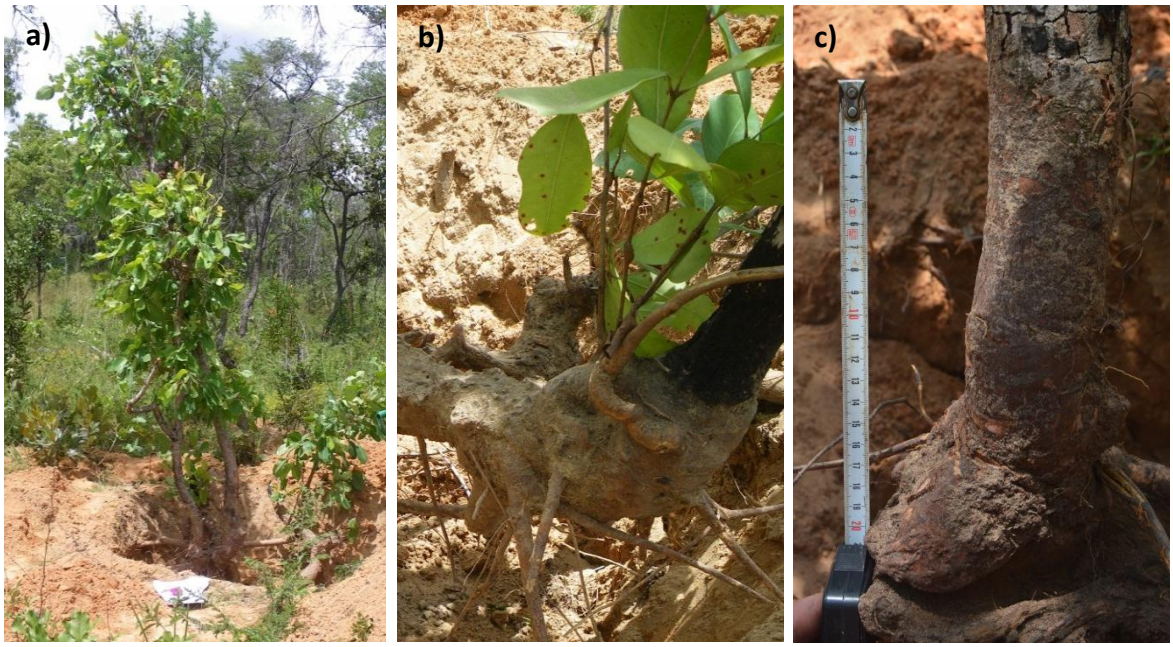



**Figure 44:** An excavated SGM in open woodland at Cusseque core site, Angola. Even though growing on the upper slope it displays thickened woody structures c. 20 cm belowground. In this case this could be a response to fire disturbances as signs of recent burning were evident. These charred stumps are surrounded by fresh resprouts from the belowground buds. Photos: a)+b): P. Zigelski 2014, c): M. Finckh 2014

Of these phenotypically adapted individuals again some were able to depart further from the protecting forests and into the grasslands. And again, they adapted appropriately by getting lower and relocating important stem parts into the ground. As the upper soil horizons are temperature buffered, even 10 cm depth is sufficient to evade short term frosts (Finckh et al., 2016), grass fires (Miranda et al., 1993) and of course larger herbivores (Bazzaz et al., 1987). A study on frost resilience of suffrutices and their tree partners indicates that the latter are more frost resilient than their suffrutex partners, which is at first glance surprising given that they grow much more exposed in the frost prone valleys. However, it makes sense since the suffrutices adapted another, better mechanism to avoid frosts by going underground and to resproute rapidly afterwards. Their tree partners cannot avoid frost – they are aboveground exposed to frost; therefore they depend on leaves which are more frost resilient (Sanguino, 2015).

If the second assumption of suffrutex evolution is true (see IV.2.3.3), then the stepwise movement into the grasslands (see fig. 45) probably happened several times without genetic accommodation, until an ecological barrier was hit that could not be overcome by phenotypic plasticity alone and that led to the “jump” from one genetic pattern to the other, hence evolving SGH (similar to West-Eberhard, 1989; Pigliucci et al., 2006). The SGMx of the mid-slopes and the SGH from the foot-slopes have most traits in common and they probably suffer the same damage through frost and fire. However, there is one important difference between them that might enhance the genetic shift: SGMx grows on ferralitic soils that are laterally drained and leached, not holding water and nutrients for long, and SGH grows on sandy soils where the lateral runoff pools (Mendelsohn & Weber, 2013) and which might get temporarily water logged (pers. obs.; Oldeland et al., 2013, p. 51-54). Both soils, the Ferralsols and the Arenosols, are therefore quite different.

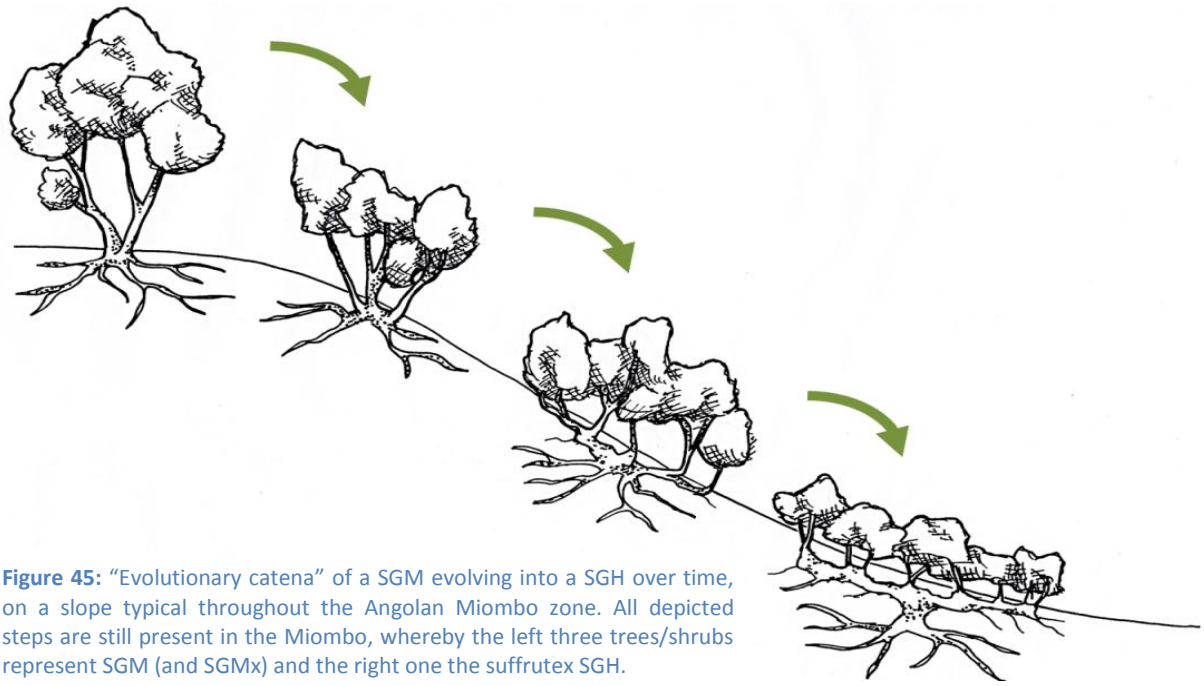


**Figure 45:** "Evolutionary catena" of a SGM evolving into a SGH over time, on a slope typical throughout the Angolan Miombo zone. All depicted steps are still present in the Miombo, whereby the left three trees/shrubs represent SGM (and SGMx) and the right one the suffrutex SGH.

If the first assumption as formulated in IV.2.3.3 is true, "Cluster 1 suffrutices" from the east, either derived from SGB or from a yet unknown tree population, had already been in place when SGM/SGMx moved into the grasslands. There they could have mixed with the suffrutex, forming the Cluster 1-Cluster 2 genetic pattern we see today. However, the geographical and genetic pattern is in my opinion to weak for that assumption.

There remains the question if SGM is not only the progenitor of the suffrutices but also of other phenotypes present today. All phenotypes except SC and greater part of SGG show affiliation to SGM's Cluster 2 to some extent. This hints that there had been gene flow and exchange in the past. A possible scenario could be that SGM populations were (and still are) widespread throughout tropical regions. This phenotype is strongly correlated with ferralitic soils, low precipitation seasonality and low temperature range throughout the year. These are features of the wet tropical zones in Central Africa and indicate an origin in these regions (Grant, 1981). Some SGM populations were isolated in refugia during unfavorable conditions (which will be discussed in the next section), where they changed genetically due to isolation or crossing with other phenotypes that took refuge there. This scenario is very similar to what happened during glacial/interglacial times across the world (Hilbert et al., 2007, Stewart et al., 2010, Keppel et al., 2012).

### IV.2.5 Comparison and similar cases

The genus *Syzygium* contains woody species growing mostly in riverine or rain forests; its distribution and diversity focus lies in South-East Asia and Australia but has also representatives throughout the old and new tropics (Chantaranothai & Jarnell, 1994; Biffin et al., 2010, Global Biodiversity Information Facility GBIF: *Syzygium*). The ploidy within the genus *Syzygium* is uneven; though most seem to be dipoid, there are also examples of tri- or tetraploid species with a general base chromosome number of n=11 (Mehra, 1972; Oginuma et al., 1992; 1993; Thurlby et al., 2012). The closest related species to *Syzygium guineense* are *S. pondoense* and *S. masukuense* (and *S. cordatum*) from Southern Africa as well as cultivated *S. cumini* and *S. jambos* from South-East Asia according to the limited phylogenetic species set of Biffin et al., (2006). Other Southern African *Syzygium* species like *S. legatii*, *S. gerarrdii* or *S. owariense* were not covered there as their distribution range is comparably small (Coates Palgrave, 2002).

However, these are all trees growing preferably in riverine forests (Exell et al., 1966, Coates Palgrave, 2002; Biffin et al., 2010), a habit that is also common to most *S. guineense* tree forms. The suffrutex form therefore represents a recent development away from the ancestral habitat, life form, and ecology. Another example of acquired traits within *Syzygium* that allow for a different habitat can be found in Australian *Syzygium* species. Most species there grow in rainforests but some, like *S. eucalyptoides* or *S. orbiculare,* also in savanna habitats (Burrows et al., 2010). The Myrtaceae in Australia exhibit a wide variety of epicormic structures which are buds located in the bark of the stems, being in this way able to sprout rapidly after disturbances like fires and frosts (Burrows et al., 2002; 2010). *Syzygium* species have such structures too, but indicating adaptation to wet and protected environments since the buds are located near the bark surface. In contrast, the Eucalypts are better adapted to dry habitats frequently disturbed by fires or frosts as their epicormic buds are hidden more deeply in the bark and are therefore thermically better protected (Burrows et al., 2002). All these findings hint to a recent colonization of and adaptation to drier habitats, as most *Syzygium* species display characteristics of sufficiently wet environments.

Within the Myrtaceae family are several examples of phenotypic plasticity and polyploidy known, one of the best studied is *Metrosideros polymorpha* from Hawaii. This species occurs on all Hawaiian Islands and is characterized by an enormous phenotypically continuous range from dwarf shrubs (without geoxylic structures) to trees with all stages and varying leave/flower traits in between (Cordell et al., 1998). This wide phenotypic plasticity is a response to the respective habitats it is growing in which encompass among others recent lava flows, bogs or mountainous forests (Joel et al., 1994). Several genetic studies (nDNA, cpDNA, SSR) on this complex indicated that the phenotypic variety is only slightly based on genetic differences but rather on local environmental responses (Wright et al., 2001; Percy et al., 2008; Harbaugh et al., 2009). Particularly the dwarf shrubs are local adaptions which happened on each island separately in recent times (< 3 mya, Percy et al., 2008).

Another example was studied in Southern Australia, where the prevalent *Eucalyptus globulus* displays dwarf growth under oceanic influence: At cliff tops facing the sea grow populations of low height (< 4 m) and only 100 – 200 m further inland are populations of the usual height of up to 60 m. Again, a genetic study with microsatellite markers was conducted to determine genetic relatedness of three such pairs from different sites. At all sites the dwarfs had originated locally

from the nearby tree populations as a response to the wind and spray from the sea; they were not closely related to the other dwarfs (Foster et al., 2007).

In both cases the authors found the opposite of my findings, namely that dwarf forms were directly derived from nearby trees. However, both cases describe a much smaller scaled evolutionary adaptation at single sites whereas the Miombo comprises a much bigger evolutionary arena. There had to be a fast large scale environmental change for so many suffrutex species to evolve (see next section). Furthermore, the above described cases represent the initial stage of speciation, with as yet unclear genetic fixation. Similar events of local morphologically differentiated sub populations could have been the starting point of the suffrutex differentiation (Abbott & Comes, 2007).

Interestingly, both *Metrosideros polymorpha* and *Eucalyptus globulus* are not polyploid (Fedorov, 1969; Bedi, 1989). Their phenoplasticity is based on other features. Nevertheless, aside from *Syzygium guineense* are several other species occurring in the Miombo that show polyploidy or from which it has been reported that they are polymorphic. For example, several members of Combretaceae (Brighton & Wickens, 1976), Rubiaceae (Kiehn, 1995) or Ochnaceae (Lewis, 1980) have undergone polyploidization. On the other hand, the particularly Miombo-typic Caesalpiniaceae or Mimosoideae don't seem to be polyploid (Atchison, 1951), and in the case of the tree-suffrutex pair *Parinari curatellifolia*/*P. capensis* the positions are swapped: the tree form is polyploid, the suffrutex not (Löve, 1982). These findings show that polyploidy alone is not sufficient to explain the adaption to new environments.

This statement is furthermore supported by the fact that polyploidy doesn't seem to play a major role in the Brazilian Cerrado (Forti-Martins et al., 2000). The Cerrado is the neotropical analogy to Africas Miombo (Ratter et al., 2006; WWF: NT0704); it is a savanna ecosystem with much congruence to the palaeotropical forest-grassland mosaic (see fig. 46), as both florae are affected by similar environments and both comprise geoxylic suffrutices (Oliviera et al., 2002; Simon et al., 2009; Pennington & Colin, 2014). Though there are some polyploid lineages within the suffrutices of the Cerrado, the authors themselves suggest that these are only few in numbers and that more species should be revised to allow for a more precise picture. The congruent development of geoxylic structures in suffrutex species of the Cerrado might not only be the product of similar environmental pressures, but also of relatedness between the African

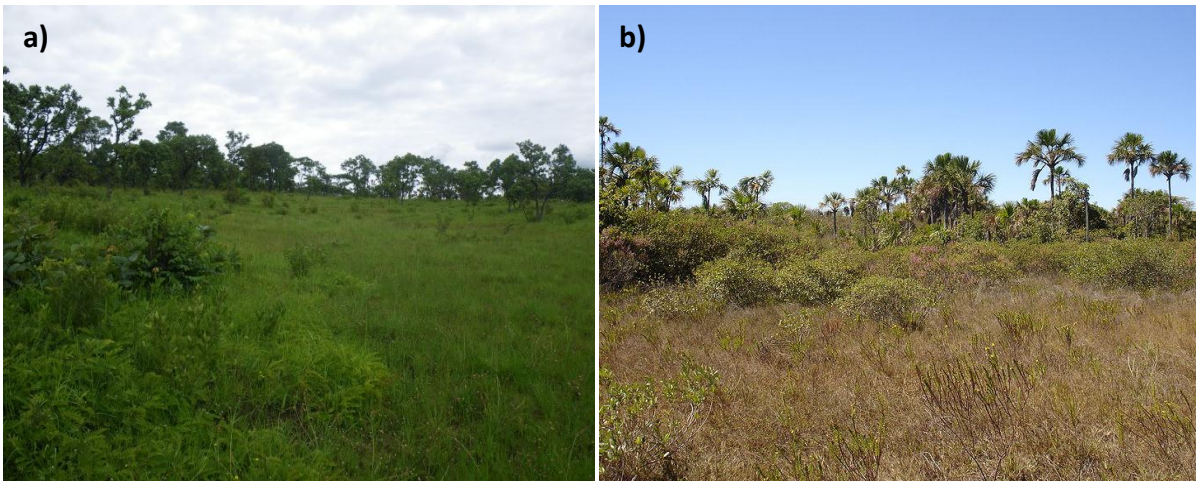



**Figure 46:** Comparison between typical Miombo as in the Cuchi catchment, Angola (a) and Cerrado (b) landscape. There is a mosaic of grasslands, forests ecotones between them in both ecoregions. Photo (a): P. Zigelski 2014, image source (b): http://www.infoescola.com/wp-content/uploads/2007/08/cerrado.jpg

and the Brazilian Flora. Due to their joined past as super-continent Gondwana, South America and Africa share a non negligible pool of families and genera (Givnish & Renner, 2004). Members of common genera, with therefore resembling genetic background, have evolved into geoxylic suffrutices on both continents, for example *Combretum mellifluum* and *Combretum platypetalum* or *Annona dioica* and *Annona stenophylla* ssp*. nana* (White, 1977; Ratter et al., 2003; Ratter et al.: Cerrado database).

However one should be careful when comparing Miombo and Cerrado – a one-to-one similarity is certainly not given, despite the physiogeographic similarities (see figure 46). Both ecoregions are characterized by a dry season, though the annual rainfall in the Cerrado is higher (1200 – 2000 mm) (Oliviera et al., 2002; WWF: NT 0704). Many studies suggest fire to be the key creator and shaper of the Cerrado landscape (Simon et al., 2009). However, there is little known about comparable microclimatic conditions as in the Miombo, though some authors ascribe the mosaic structure also to frost events (Silberbauer-Gottsberger et al., 1977; Brando & Durigan, 2005).

There is a general trend of floristic poverty of tropical African biomes compared to other tropical regions (Brennan, 1978). The species and endemic richness and diversity is higher there than in the Miombo, as more than 12,000 vascular plants are known of which 50 are endemic to the Cerrado (Oliviera et al., 2002; Mendonça et al., 2008; Simon et al., 2009; WWF: NT 0704). The Miombo on the other hand encompasses around 8,500 different plant species with only few true endemics (Frost, 1996; WWF: AT0701). However this number might be a considerable underestimation due to the understudied flora of Miombo core countries like Angola. The high biodiversity in the Cerrado indicates that one mechanism alone – polyploidy – cannot explain parallel evolution and radiation. But genetic pre-adaptation coupled with small-scale heterogeneous, uninhabited niches are strong drivers of speciation (Rainey & Travisano, 1998).


However, an example from the other side of the world stresses the importance of polyploidy to face adverse thermic environments. The alpine and arctic flora, best studied on the Norwegian Svalbard Islands, evolved several characteristics which are not unfamiliar to the suffrutices (see fig. 47). The woody plants there grow very low and form dwarf shrubs, huddling against rocks and crevicesin order to avoid cold temperatures and strong winds (Bliss, 1962; Billings, 1974; Rønning, 1996; Abbott & Brochmann, 2003). The alpine suffrutex species *Salix herbacea* is even able to penetrate the rocky ground to hide and protect its buds (Wijk, 1986). Most species are furthermore polyploid, and the degree of polyploidy increases with northern latitude and severity of arctic climate, indicating that polyploidy and suffrutication are advantageous in harsh climates (Stebbins, 1984; Brochmann et al., 2004).

**Figure 47:** Arctic growth forms of trees on Svalbard, Norway. Birch species *Betula nana* ssp. *tundrarum* (a) and dwarf willow *Salix herbacea* from the Alps (b).

Image sources:
http://svalbardflora.no/assets/images/species/Betula_nana_tundrarum/Betula_nana_tundrarum_3_full.jpg;
https://upload.wikimedia.org/wikipedia/commons/7/79/Salix_herbacea_a2.jpg

## IV.3 Evolution of the Miombo

### IV.3.1 The climatic turn: retreating forests and spreading grasslands

The suffrutex habit emerged in all affected plant families at a time frame mostly less than 5.3 mya, both in the Brazilian Cerrado as in the African Miombo, with a peak divergence from their tree sister taxa about 2.3 mya (Maurin et al., 2014). The transition of forest taxa to open grassland taxa certainly was promoted by basic mechanisms like avoidance of competition for space or scarce nutrients (Belsky 1994) and facilitated by the benefits of vacant ecological niches awaiting the evolving taxa (Bazzaz, 1986; Sakai et al., 2001; Kawato, 2002). However, the remarkable parallel evolution of the suffrutex habit across continents and plant families strongly indicates a common and sudden new environmental pressure (Pennington & Hughes, 2014) that strongly affected their evolutive course.

The emergence of suffrutices in Africa coincides indeed with a major climatic change: The Pliocene/Pleistocene boundary marks a global climatic break (Zachos, 2001; demenocal, 2004, Trauth et al., 2009). Before that, conditions were more humid, without a pronounced dry season (demenocal, 2004). Tropical Miombo-like forests the were growing even southwards in Namibia (Dupont, 2006; Morley, 2011) and in todays Miombo woodland region grew evergreen closed canopy forests at that time, as had been revealed by palynological and fossile studies (Axelrod & Raven, 1978; Scott, 1992, Maley, 1996; Morley, 2011). Many woody species were widely distributed in those environments, among them also the proto *Syzygium guineense* tree. As the epithet suggests, this species is associated with wet tropical Guinea in West African. Nevertheless *Syzygium guineense* s.l.still extends as far as northern Namibia and Botswana in suitable, wet habitats (Coates Palgrave, 2002; GBIF: *Syzygium*), which could be relicts from a once continuous distribution range.

With the Pleistocene advent around 2.5 mya ago (Trauth et al., 2009) climate became drier and more seasonal. With it changed the vegetation as the closed canopy forests could no longer be supported and retreated to the north, leaving thornbush savanna or Mopane woodlands where once had been Miombo and Miombo where once had been rainforest (Morley, 2011, Bonnefille, 2011). This forest retreat was furthermore intensified by parallel decreasing $CO_2$ levels below 500 ppm in the atmosphere (Cerling et al., 1998; Bond et al., 2003) as trees depend in dry hot climates on a sufficiently high $CO_2$ level (Saxe et al., 1998, Bond et al., 2003).

This new climate regime promoted the spread of open grassland habitats where forests were retreating to favorable extrazonal sites (Beerling & Osborne, 2006, Kgope et al., 2010). At that time C4-grasses began to fill the appearing niche as they were able to cope with the lower $CO_2$ and adverse dry season conditions through an adapted photosynthetic pathway (Bond & Keeley, 2005; Bond, 2008, Edwards et al., 2010). Edaphic conditions and annual rainfall were still sufficient to support closed canopy forests in most of today's Miombo region, but environmental filters arose that created the intercalated mosaic of forests and grasslands (Staver et al., 2010).

In flat areas with less pronounced topography fires fueled by the rapidly spreading grasses might have played and still play the role as abiotic herbivore (Bond & Keeley, 2005). On the one hand they prevent successful tree growth and proliferation if they occurr too frequent, trapping young tree populations in a "fire-trap" (Sankaran, 2004; Bond & Keeley, 2005). At least five years of growth or regeneration are needed to overcome the fire bottleneck; when trees reach a sufficient height, their fire prone canopy evades the flames on the ground if the fires are short-

lived (Bond & Keeley, 2005, Keeley et al., 2011). Natural fires kindled by thunderstorms are rare and it is questionable if they occur often enough to keep grasslands open and to drive the suffrutex evolution (Archibald et al., 2009; 2012). Human-made fires however are regular and frequent during the dry season and might therefore have played an important role in the last 200 – 400 kya, since humans learned to use fire (Archibald et al., 2012). Whether this time frame is sufficient enough to intensify the suffrutex evolution is unknown, particularly as the peak of humanmade fire frequency was reached only 4 – 40 kya ago (Archibald et al., 2012).

On the other hand grassland fires normally do not invade deep into forests as the prevalent, fueling, light demanding C4 grasses do not grow there densely enough (Keeley & Rundel, 2005; Archibald et al., 2009). The fires extinguish at the forest fringe under normal circumstances (pers. comm M. Finckh) but certain conditions such as fires in the late dry season or kindling during the hot and windy day time may lead to burning canopies and fire intrusion into the forests (pers. obs.). Though this might be an explanatory driver of the grassland-forest mosaic, there is also another factor more probable, as it had been acting on the landscape much earlier than humanmade fires (Trauth et al., 2009).

This aspect, particularly plausible in pronounced, undulating landscapes and highlands, is frost. With increased seasonality and marked dry seasons the temperature amplitude increases, too. During daytime temperatures of 40°C can be reached easily, and by contrast they can drop to -6°C or less early in the morning as the cold air accumulates especially in valleys (SASSCAL WeatherNet: Cusseque station; Finckh et al., 2016). The African Highveld for example is just one region where frosts prevent successful tree growth and where small scale sharp boundaries delineate tree-devoid grasslands and forests (Wakeling et al., 2012). The event of “frost currents” however is even probable in flat areas like the Liuwa Plains, since there are seasonally dry river runs that constitute slight depressions in the landscape. Even these marginal slopes could be sufficient to accumulate cold air, especially as the open grasslands are a prime generator of cold air (Sakai & Larcher, 2012). This is due to thermal radiation, as above the comparably uncovered ground less air moisture is generated through evaporation of plants and less is retained by them near the ground. Therefore temperatures can change more rapidly as water vapor is a key component of macro- and microclimate control (Maher et al., 2005). Many tropical woody plant families have not evolved physiological frost resistance and thus are prone to frosts, they do not survive these fluctuating conditions (Brando & Durigan, 2005), unless they develop strategies to avoid the frost (Sakai & Larcher, 2012). They are therefore limited to areas where frost events are rare, for example hill tops or other elevated sites, or they have to protect their structures and buds underground, to avoid the frost.

Thus both fire but particularly frost are environmental drivers that may lead to a patchy landscape of grasslands and forests, which is not surprising as both are thermic stresses to plants which can be coped with by similar means like thickened bark, tough leaves, protected buds or chemical compounds (Simon et al., 2009; Keeley et al., 2011). An adaption against frost can therefore concurrently be an preadaption against fires, and already frost-adapted species might be advantageous in an environment with increasing fire frequencies due to human proliferation and expansion. Human evolution had been coupled to the emergence of savannas and very soon human species learned to control and use fires as a tool to hunt game and for agriculture later on (Burton, 2011; Archibald et al., 2012). The increasing anthropogenic fires in the last millennia were most probably encountered by preadapted species that already evolved the suffrutex habit

and could therefore cope well with them. It is even imaginable that today's suffrutex diversity is a product of the combined frost/fire influence.

There might be other factors too that play an important role in maintaining this bimodal landscape. It has been suggested that seasonally inundated soils or biotic herbivory (Sankaran et al., 2005) also contribute to the typical Forest-open land mosaic in the Miombo belt. This might be the case in areas like Liuwa Plains in Western Zambia (the typical landscape of the Western Zambezian Grasslands biome), which is topographically flat, in large areas shallowly flooded during the rainy season and where vast numbers of large herbivores like zebra and blue wildebeest pass through every year (WWF: AT0724). However, this landscape is an exception to the typical Miombo mosaic. The Zambesi graben , a rift valley, is framed by the Highlands of Bíe to the West, the Lunda Ridge to the North and the Central African Plateau to the east; they are built by less weathered Precambrian bedrocks that support the undulating topography (Panagos et al., 2011). Thus, a different abiotic imprint, which is mirrored in flora and fauna and their genetic patterns, is therefore very likely for the plains with Zambezian Grasslands and is partially shown in this study.

### IV.3.3 The concept of changing niches

An alternative and summarizing way to look at the three genetically and phenotypically distinguishable ecotypes suffrutex, ecotone shrub and riverine tree is to consider what respective niches they occupy in the Miombo of today, compared to the niche of the proto-*Syzygium*. As proposed previously, the progenitor of the studied phenotypes grew/grows as a tree in evergreen tropical forests, all year round continuously supplied with water. In the changing environment, as the different phenotypes differentiated, this exact niche could only be kept when retreating north together with the closed forests.

But in order to stay and survive at the ancestral region, some populations had to retreat into year-round sufficiently wet extrazonal sites along rivers or lakes. These sites offer a hygric niche similar in water availability to the original forest niche as they are somewhat decoupled from seasonal precipitation patterns (Oldeland et al., 2013, pp. 19-22). Since these populations were then sufficiently water provided and protected from adverse conditions, they could maintain the tree lifeform. However, they had to adapt to strongly changing phreatic levels between flooding season and dry season. This group correspond to the wet ecotype we find today, the SGG phenotype.

Another way to persist in the ancestral region was to adapt to the drier climate and to change the niche. This could not happen without a compromise: As the protective, evergreen forest canopy gave way to frequently disturbed open woodland, its biota had to reply by becoming more resilient. *Syzygium* evolved into lower and more shrub-like, multi-stemmed trees with protogeoxylic structures. Seasonal precipitation is sufficient to support such a growth form, as it responds flexibly to frequent disturbances by resprouting rapidly from underground hidden buds. This niche is occupied by today's ecotone ecotype, represented by SGM.

Lastly, taking a step further, there is the niche of the open grasslands. Typically, grasses and herbs occupy niches where woody plants can not survive due to abiotic constraints like lack of nutrients, adverse thermic conditions, lack of water ressources or an adverse combination of the former. As adverse thermic conditions (either fire and/or frost) seem to be the exclusion factors

in the study area, avoidance strategies allow frost prone tropical woody plants to partly occupy this niche. Aboveground, the suffrutices are functional herbs, their aerial parts are often (facultatively) only annual due to frequent disturbances. Therefore, in order to stay in their ancestral distribution range, these populations had to completely change spatial niche and life form. Instead of building vertical structures they evolved horizontally spreading, near surface underground structures to compete for light. Above ground, they behave like herbs and they even occupy similar niches in the grasslands. What distinguishes them from true herbs are their geoxylic structures, a reminder of their tree origin and the key for their survival in a changed environment.

## IV.4 Conclusions and outlooks

Investigating genetic and ecological aspects of the Miombo is a start to understand the mechanisms shaping this biome. I aimed to shed light on some of these aspects using a model taxon and I hope this work will facilitate and contribute to future Miombo studies. I was able to provide some answers to the initially asked questions:

whether *Syzygium guineense* s.l. is a complex of separate (sub-) species or one species of high phenotypic plasticity:

it rather is a species complex with just diverging genepools and genetic clusters that contribute to the phenotypic diversity but do not entirely explain it;

whether the tree and suffrutex forms are closer related to their own kind over the whole studied area or rather to their closest neighbor despite the morphological discrepancy:

they are more related within their phenotypes over the whole region than to their neighbor of a different phenotype, there is no significant correlation between genetic and spatial distance;

and what genetic and environmental influences triggered and shaped the suffrutex habit:

polyploidy and with it phenotypic responsiveness as well as strong environmental drivers caused by a changing climate facilitated the evolution of the *Syzygium* suffrutex and other ecologically and genetically distinct phenotypes.

Summarizing, the phenotypical diversity of *Syzygium guineense* s.l. within the Zambezian Miombo region is a response of a species to a recently changing environment. This phenotypical response is becoming genetically fixed, a process that is still in process. To date three genetic types can be distinguished that correspond to ecological niches: the suffrutex of open grasslands, the shrub or small tree of frequently disturbed forest-grassland ecotones and the trees of humid and gallery forests and river banks. Hence the suffrutex in fact is a genetically different type; but if it is a distinct species could not be answered in this study. There are many different concepts of defining and distinguishing a biological species (Mallet, 1995; Abbott & Comes, 2007). Some of them, like the distinct niche or habit apply to the suffrutex, but it is unknown if interbreeding with the tree forms is still occuring, which, if not, would be the ultimate textbook proof.

However, genetic variation within these ecotypes is still high and not settled yet, and not all phenotypes yield plausible genetic patterns, an effect that might be owed to the few analysed individuals per population. Furthermore some phenotypes and locations are better represented in this study than others, leading to a bias and a still incomplete picture. It is also uncertain if this genetic pattern of relatedness is found in similar ways in other suffrutex species. However, it is likely as the suffrutices evolved more or less concertedly to the same environmental pressures. I therefore recommend broadening the analyses by investigating the genotypes of more individuals per population, by collecting more populations from yet unsampled places (e.g. North and East Angola, East of the Zambezi Basin) and by analyzing further suffrutex-tree species pairs from other families.

Besides from determining a more detailed genetic picture future studies should also focus on the environmental aspects causing these genetic patterns. Smallscale microclimatic measurements are until now available only for Cusseque in Central Angola. At least one typical Western Zambezian Grasslands location like Liuwa Plains, Zambia, should be microclimatically

analysed, plus one location at the southern fringes of the Zambezian suffrutex belt (population Sgh11). Spatial distribution patterns of suffrutices compared to their tree partners can indicate where and which environmental stresses act upon them and what influence these stresses have. Particularly the role of frost has been overlooked and neglected in favor of the fire hypothesis; therefore a closer look at this shaping element is necessary in order to understand the complex coherence. Furthermore the distribution pattern of species at a given site provides information about their ecology, and by involving more suffrutex species we can learn more about their diversity, competition or function within the underground forest.

Understanding the past and anticipating the future development of the Miombo is crucial for its survival and conservation. The Miombo is at this point of time one of the largest nearly intact ecosystems of the world (Leadley, 2010), but this landscape is nowadays under an immense anthropogenic pressure as increasing numbers of people live off and rely on the Miombo; livestock, agriculture and deforestation threaten to damage the Miombo irreversibly (Campbell et al., 1996; Abbot & Homewood, 1999; Lambin et al., 2003, Abdallah & Monela, 2007; Leadley, 2010; Hansen et al. 2013; Gasparri et al. 2016). However, this still natural biome provides important ecosystem services like carbon storage, climate regulation with respect to precipitation patterns and green house gases, biodiversity habitat or water provision (Leadley, 2010). Among others, four of the greatest rivers of Southern and Central Africa, Congo, Zambezi, Cuanza and Okavango, emanate in the Miombo region. It has been stated that the Miombo is one of the earth systems tipping points which means that further degradation of the biome would lead to a climatic and vegetational transition with possibly devastating social, economic, financial and environmental consequences (Leadley, 2010). By considering and heeding its importance we can and should elaborate strategies to conserve the Miombo in mutual agreement with the needs of local people.

The underground forests of the Miombo are inconspicuous as they hide mostly out of sight, but nevertheless they shouldn't be overlooked. Not all questions that I investigated in my work in order to reveal this under- and aboveground ecosystem could be answered in due detail. Moreover, all answers I found raised even more questions regarding its genetic and environmental setup. This study should therefore be regarded as a first step to elucidate the genetic profile of the Miombo woodlands and to find the environmental imprint within it; as genetics and environment are never separated, they interact and interleave permanently to produce organisms that are adapted best to cope with the conditions they were selected for.

# V. Literature


**Abbot, Joanne IO, and Katherine Homewood**. "A history of change: causes of miombo woodland decline in a protected area in Malawi." *Journal of Applied Ecology* 36.3 (1999): 422-433.

**Abbott, Richard J., and Christian Brochmann.** "History and evolution of the arctic flora: in the footsteps of Eric Hultén." *Molecular ecology* 12.2 (2003): 299-313.

**Abbott, R. J. and Comes, H. P.**, "Blowin' in the wind – the transition from ecotype to species". *New Phytologist*, 175 (2007): 197–200. doi:10.1111/j.1469-8137.2007.02127.x

**Abdallah, J. M., and G. G. Monela.** "Overview of Miombo woodlands in Tanzania." In *MITMIOMBO–management of indigenous tree species for ecosystem restoration and wood production in semi-arid Miombo woodlands in eastern Africa. Working Papers of the Finnish Forest Research Institute* 50 (2007): 9-23.

**Antonovics, Janis**. "The Effects of a Heterogeneous Environment on the Genetics of Natural Populations: The realization that environments differ has had a profound effect on our views of the origin and role of genetic variability in populations." *American Scientist* 59.5 (1971): 593-599.

**Archibald S, Roy DP, van Wilgen BW, Scholes RJ**. "What limits fire? An examination of drivers of burnt area in Southern Africa". *Global Change Biology 15* (2009)*: 613–630.*

**Archibald, Sally, A. Carla Staver, and Simon A. Levin**. "Evolution of human-driven fire regimes in Africa." *Proceedings of the National Academy of Sciences* 109.3 (2012): 847-852.

**Atchison, Earlene.** "Studies in the Leguminosae. VI. Chromosome Numbers Among Tropical Woody Species". *American Journal of Botany* 38.7 (1951): 538–546. (no polyploidy brachystegia etc)

**Axelrod, D. I., and P. H. Raven**. "Late Cretaceous and Tertiary vegetation history of Africa." *Biogeography and ecology of southern Africa*. Springer Netherlands, 1978. 77-130.

**Avise, John C.** "Molecular markers, natural history and evolution". *Springer Science & Business Media*, 2012.

**Ayala, Francisco J., et al.,** "Genetic differentiation during the speciation process in Drosophila." *Evolution* (1974): 576-592.

**Baldwin, Bruce G.** "Phylogenetic utility of the internal transcribed spacers of nuclear ribosomal DNA in plants: an example from the Compositae." *Molecular phylogenetics and evolution* 1.1 (1992): 3-16.

**Bandelt, H., Forster, P., & Röhl, A**. "Median-joining networks for inferring intraspecific phylogenies". *Molecular Biology and Evolution*, 16(1) (1999), 37–48.

**Bazzaz, F. A.** "Life history of colonizing plants: some demographic, genetic, and physiological features." *Ecology of biological invasions of North America and Hawaii*. Springer New York, 1986. 96-110.

**Bazzaz, Fakhri A., et al.,** "Allocating resources to reproduction and defense." *BioScience* 37.1 (1987): 58-67.

**Bedi, Y. S.** "Karyological studies in *Eucalyptus* L. Her." *Plant science research in India: present status and future challenges. Part* 1 (1989): 391-397.

**Beerling, David J., and Colin P. Osborne.** "The origin of the savanna biome." *Global change biology* 12.11 (2006): 2023-2031.

**Bellingham, Peter J., and Ashley D. Sparrow.** "Multi-stemmed trees in montane rain forests: their frequency and demography in relation to elevation, soil nutrients and disturbance." *Journal of Ecology* 97.3 (2009): 472-483.

**Belsky, A. Joy.** "Influences of trees on savanna productivity: tests of shade, nutrients, and tree-grass competition." *Ecology* 75.4 (1994): 922-932.

**Beyra-Matos, Angela, and Matt Lavin. "**Monograph of Pictetia (Leguminosae-Papilionoideae) and review of the Aeschynomeneae". *American Society of Plant Taxonomists*, 1999.

**Biffin, Ed, et al.,** "Molecular systematics of Syzygium and allied genera (Myrtaceae): evidence from the chloroplast genome." *Taxon* (2006): 79-94.

**Biffin, Ed, et al.,** "Evolution of exceptional species richness among lineages of fleshy-fruited Myrtaceae." *Annals of Botany* 106.1 (2010): 79-93.

**Billings, W. D.** "Adaptations and origins of alpine plants." *Arctic and alpine research* (1974): 129-142.

**Birky Jr, C. William.** "Transmission genetics of mitochondria and chloroplasts." *Annual review of genetics* 12.1 (1978): 471-512.

**Bliss, L. C.** "Adaptations of arctic and alpine plants to environmental conditions." *Arctic* 15.2 (1962): 117-144.

**Bond, William J., and Jeremy J. Midgley.** "Ecology of sprouting in woody plants: the persistence niche." *Trends in ecology & evolution* 16.1 (2001): 45-51.

**Bond, W. J., G. F. Midgley, and F. I. Woodward.** "The importance of low atmospheric CO2 and fire in promoting the spread of grasslands and savannas." *Global Change Biology* 9.7 (2003): 973-982.

**Bond, William J., and Jon E. Keeley.** "Fire as a global 'herbivore': the ecology and evolution of flammable ecosystems." *Trends in ecology & evolution* 20.7 (2005): 387-394.

**Bond, William J**. "What Limits Trees in $C_4$ Grasslands and Savannas?." *Annual Review of Ecology, Evolution, and Systematics* (2008): 641-659.

**Bonnefille, R.** "Rainforest responses to past climatic changes in tropical Africa." In *Tropical rainforest responses to climatic change*. Springer Berlin Heidelberg, 2011. 125-177.

**Brando, Paulo Monteiro, and Giselda Durigan.** "Changes in cerrado vegetation after disturbance by frost (São Paulo State, Brazil)." *Plant Ecology* 175.2 (2005): 205-215.

**Brenan, J. P. M**. "Some Aspects of the Phytogeography of Tropical Africa". *Annals of the Missouri Botanical Garden* 65.2 (1978): 437–478.

**Brighton, C. A., and G. E. Wickens.** "Some Chromosome Counts in the Genus Combretum (Combretaceae)". *Kew Bulletin* 31.1 (1976): 5–8.

**Brochmann, C., et al.,** "Polyploidy in arctic plants." *Biological Journal of the Linnean Society* 82.4 (2004): 521-536.

**Burrough, Sallie L., David SG Thomas, and Richard M. Bailey.** "Mega-Lake in the Kalahari: a Late Pleistocene record of the Palaeolake Makgadikgadi system." *Quaternary Science Reviews* 28.15 (2009): 1392-1411.

**Burrows, G. E.**, "Epicormic strand structure in *Angophora*, *Eucalyptus* and *Lophostemon* (Myrtaceae) – implications for fire resistance and recovery". *New Phytologist*, 153: 111–131 (2002). doi:10.1046/j.0028-646X.2001.00299.x

**Burrows GE, Hornby SK, Waters DA, Bellairs SM, Prior LD, Bowman DMJS.** "A wide diversity of epicormic structures is present in Myrtaceae species in the northern Australian savanna biome – implications for adaptation to fire". *Australian Journal of Botany* 58 (2010), 493–507.

**Burton, Frances D**. "Fire: The spark that ignited human evolution". *UNM Press*, 2011.

**Campbell, B, P. Frost, and N. Byron.** "Miombo woodlands and their use: overview and key issues". In *The Miombo in transition: woodlands and welfare in Africa*. Cifor,Bogor, 1996.

**CBOL Plant Working Group, et al.,** "A DNA barcode for land plants." *Proceedings of the National Academy of Sciences* 106.31 (2009): 12794-12797.

**Cerling, Thure E., James R. Ehleringer, and John M. Harris**. "Carbon dioxide starvation, the development of C4 ecosystems, and mammalian evolution." *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 353.1365 (1998): 159-171.

**Chantaranothai, P., and J. Parnell.** "Revision of Acmena, Cleistocalyx, Eugenia ss and Syzygium (Myrtaceae) in Thailand." *Thai Forest Bulletin [Botany]* (1994).

**Chen, Z. Jeffrey.** "Genetic and Epigenetic Mechanisms for Gene Expression and Phenotypic Variation in Plant Polyploids." *Annual review of plant biology* 58 (2007): 377–406. *PMC*. Web. 20 Feb. 2016.

**Clark, Lindsay V., and Marie Jasieniuk.** "POLYSAT: an R package for polyploid microsatellite analysis." *Molecular Ecology Resources* 11.3 (2011): 562-566.

**Coates Palgrave, Keith.** *Trees of southern Africa*. No. 3. ed. Struik, 2002. 833-837.

**Coetzee, B. J., and M. J. A. Werger.** "The Sudano-Zambezian region." In *The Biogegraphy and Ecology of Southern Africa*, The Hague, Junk (1978).

**Cole, Monica M.** "Vegetation and geomorphology in Northern Rhodesia: an aspect of the distribution of the savanna of central Africa." *The Geographical Journal* 129.3 (1963): 290-305.

**Cordell, S., et al.,** "Physiological and morphological variation in Metrosideros polymorpha, a dominant Hawaiian tree species, along an altitudinal gradient: the role of phenotypic plasticity." *Oecologia* 113.2 (1998): 188-196.

**Cornelissen, J. H. C., et al.,** "A handbook of protocols for standardised and easy measurement of plant functional traits worldwide." *Australian journal of Botany* 51.4 (2003): 335-380.

**Crawford, R. M. M**. "Physiological responses to flooding." *Physiological plant ecology II*. Springer Berlin Heidelberg, 1982. 453-477.

**Davy, Joseph Burtt.** "The suffrutescent habit as an adaptation to environment." *Journal of Ecology* 10.2 (1922): 211-219.

**Demenocal, Peter B.** "African climate change and faunal evolution during the Pliocene–Pleistocene." *Earth and Planetary Science Letters* 220.1-2 (2004): 3-24.

**Demesure, B., N. Sodzi, and R. J. Petit**. "A set of universal primers for amplification of polymorphic non-coding regions of mitochondrial and chloroplast DNA in plants." *Molecular ecology* 4.1 (1995): 129-134.

**Dong, Wenpan, et al.** "Highly variable chloroplast markers for evaluating plant phylogeny at low taxonomic levels and for DNA barcoding." *PLoS One* 7.4 (2012): e35071.

**Doyle, J. J. and J. L. Doyle.** "A rapid DNA isolation procedure for small quantities of fresh leaf tissue." *Phytochemical Bulletin* 19 (1987): 11-15

**Dumolin, S., B. Demesure, and R. J. Petit**. "Inheritance of chloroplast and mitochondrial genomes in pedunculate oak investigated with an efficient PCR method." *Theoretical and applied genetics* 91.8 (1995): 1253-1256.

**Dupont, L. M.** "Late Pliocene vegetation and climate in Namibia (southern Africa) derived from palynology of ODP Site 1082." *Geochemistry, Geophysics, Geosystems* 7.5 (2006).

**Earl, Dent A. and vonHoldt, Bridgett M.** "STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method." *Conservation Genetics Resources* vol. 4 (2) (2012), pp. 359-361 doi: 10.1007/s12686-011-9548-7

**Edwards, Erika J., et al.,** "The origins of C4 grasslands: integrating evolutionary and ecosystem science." *Science* 328.5978 (2010): 587-591.

**Swenson, Nathan G., Pedro Anglada-Cordero, and John A. Barone.** "Deterministic tropical tree community turnover: evidence from patterns of functional beta diversity along an elevational gradient." *Proceedings of the Royal Society of London B: Biological Sciences* 278.1707 (2011): 877-884.

**Excoffier, Laurent, Peter E. Smouse, and Joseph M. Quattro.** "Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data." *Genetics* 131.2 (1992): 479-491.

**Exell, Arthur Wallis, Francisco Ascenção Mendonça, and A. Fernandes.** "Conspectus florae angolensis". *Ministério do ultramar. Junta de investigaçoes colonais*, 1966.

**Fedorov, An A.** "Chromosome Numbers Of Flowering Plants", *Academy of Sciences of the USSR, VL Komarov Botanical Institute*. (1969).

**Finckh, Manfred, Rasmus Revermann, and Marcos PM Aidar**. "Climate refugees going underground–a response to Maurin et al.,(2014)." *New Phytologist* 209.3 (2016): 904-909.

**Fitzpatrick, B. M., J. A. Fordyce, and S. Gavrilets.** "Pattern, process and geographic modes of speciation." *Journal of evolutionary biology* 22.11 (2009): 2342-2347.

**Fohouo, F. N., Denis Djonwangwe, and D. Brückner.** "Foraging behaviour of the african honey bee (*Apis mellifera adansonii*) on *Annona senegalensis*, *Croton macrostachyus*, *Psorospermum febrifugum* and

*Syzygium guineense* var. *guineense* flowers at Ngaoundere (Cameroon)." *Pakistan journal of biological sciences: PJBS* 11.5 (2008): 719-725.

**Forni-Martins, Eliana Regina, and Fernando Roberto Martins.** "Chromosome studies on Brazilian cerrado plants." *Genetics and Molecular Biology* 23.4 (2000): 947-955.

**Foster, Susan A., et al.,** "Parallel evolution of dwarf ecotypes in the forest tree *Eucalyptus globulus." New Phytologist* (2007) 175: 370-380.

**Frost, P.** "The ecology of miombo woodlands." *The miombo in transition: woodlands and welfare in Africa* (1996): 11-57.

**Gadek, P. A., P. G. Wilson, and C. J. Quinn.** "Phylogenetic reconstruction in Myrtaceae using matK, with particular reference to the position of Psiloxylon and Heteropyxis." *Australian Systematic Botany* 9.3 (1996): 283-290.

**Gasparri, Nestor Ignacio, et al.,** "The emerging soybean production frontier in Southern Africa: Conservation challenges and the role of south-south telecouplings." *Conservation Letters* 9.1 (2016): 21-31.

**Gilchrist, Alan R., Henk Kooi, and Christopher Beaumont.** "Post-Gondwana geomorphic evolution of southwestern Africa: Implications for the controls on landscape development from observations and numerical experiments." *Journal of Geophysical Research: Solid Earth* 99.B6 (1994): 12211-12228.

**Givnish, Thomas J., and Susanne S. Renner.** "Tropical intercontinental disjunctions: Gondwana breakup, immigration from the boreotropics, and transoceanic dispersal." *International Journal of Plant Sciences* 165.S4 (2004): S1-S6.

**Gottlieb, Leslie D**. "Genetic differentiation, sympatric speciation, and the origin of a diploid species of Stephanomeria." *American Journal of Botany* (1973): 545-553.

**Grant, Verne.** "Plant speciation". New York: *Columbia University Press* xii, En 2nd edition, 1981.

**Haghighi, Kobra R., and Peter D. Ascher**. "Fertile, intermediate hybrids between Phaseolus vulgaris and P. acutifolius from congruity backcrossing." *Sexual Plant Reproduction* 1.1 (1988): 51-58.

**Hall, T.A.** "BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT." *Nucl. Acids. Symp. Ser*. 41 (1999): 95-98.

**Hammer, Ø., D. A. T. Harper, and P. D. Ryan.** "PAST: Paleontological Statistics Software Package for education and data analysis". *Palaeontolia Electronica* 4. (2001).

**Hannah, L., Lovejoy, T.** "Conservation, climate change, and tropical forests" *Tropical rainforest responses to climatic change*. Springer Berlin Heidelberg, 2011. 431-441.

**Hansen, Matthew C., et al.,** "High-resolution global maps of 21st-century forest cover change." *Science* 342.6160 (2013): 850-853.

**Harbaugh, Danica T., et al.,** "Genetic structure of the polymorphic Metrosideros (Myrtaceae) complex in the Hawaiian Islands using nuclear microsatellite data." *Plos One* 4.3 (2009): e4698.

**Hardy, Olivier J., and Xavier Vekemans**. "Isolation by distance in a continuous population: reconciliation between spatial autocorrelation analysis and population genetics models." *Heredity* 83.2 (1999): 145-154.

**Hardy, Olivier J., et al.,** "Ecological, morphological and allozymic differentiation between diploid and tetraploid knapweeds (Centaurea jacea) from a contact zone in the Belgian Ardennes." *New Phytologist* 146.2 (2000): 281-290.

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

**Hausner, Georg, et al.,** "Origin and evolution of the chloroplast trnK (matK) intron: a model for evolution of group II intron RNA structures." *Molecular Biology and Evolution* 23.2 (2006): 380-391.

**Havström, Mats, Terry V. Callaghan, and Sven Jonasson.** "Differential growth responses of Cassiope tetragona, an arctic dwarf-shrub, to environmental perturbations among three contrasting high-and subarctic sites." *Oikos* (1993): 389-402.

**Higgins, Steven I., William J. Bond, and Winston SW Trollope.** "Fire, resprouting and variability: a recipe for grass–tree coexistence in savanna." *Journal of Ecology* 88.2 (2000): 213-229.

**Hijmans, R. J., et al.,** "WorldClim—global climate data." *Very High Resolution Interpolated Climate Surfaces for Global Land Areas* (2005).

**Hilbert, David W., Andrew Graham, and Mike S. Hopkins.** "Glacial and interglacial refugia within a long-term rainforest refugium: the Wet Tropics Bioregion of NE Queensland, Australia." *Palaeogeography, Palaeoclimatology, Palaeoecology* 251.1 (2007): 104-118.

**Hillyer, Mia J., et al.,** "Isolation and characterization of eight polymorphic microsatellite loci in the rainforest canopy tree, Syzygium sayeri (Myrtaceae)." *Molecular Ecology Notes* 7.6 (2007): 1199-1201.

**Hipondoka, Martin HT.** "The development and evolution of Etosha Pan, Namibia." Unpublished PhD, *Institut für Geographie and Geologie, Bayerischen Julius-Maximilians-University of Würzburg* (2005).

**Hirschhorn, Joel N., et al.** "A comprehensive review of genetic association studies." *Genetics in Medicine* 4.2 (2002): 45-61.

**Howe, Henry F., and Judith Smallwood.** "Ecology of seed dispersal." *Annual review of ecology and systematics* 13 (1982): 201-228.

**Hughes, F. M. R.** "The ecology of African floodplain forests in semi-arid and arid zones: a review." *Journal of Biogeography* (1988): 127-140.

**Huntley, Brian J.** "Outlines of wildlife conservation in Angola." *JS Afr. Wildl. Managet. Assoc* 5 (1974): 157-166.

**Husband, Brian C., and Holly A. Sabara.** "Reproductive isolation between autotetraploids and their diploid progenitors in fireweed, Chamerion angustifolium (Onagraceae)." *New Phytologist* 161.3 (2004): 703-713.

**Husson, François, Julie Josse, and Jérôme Pages.** "Principal component methods – hierarchical clustering - partitional clustering: why would we need to choose for visualizing data?" *Technical Report–Agrocampus*.(2010)

**Jarne, Philippe, and Pierre JL Lagoda.** "Microsatellites, from molecules to populations and back." *Trends in ecology & evolution* 11.10 (1996): 424-429.

**Joel, Geeske, Gregory Aplet, and Peter M. Vitousek.** "Leaf morphology along environmental gradients in Hawaiian Metrosideros polymorpha." *Biotropica* (1994): 17-22.

**Jombart, Thibaut**. "adegenet: a R package for the multivariate analysis of genetic markers." *Bioinformatics* 24.11 (2008): 1403-1405.

**Kalia, Rajwant K., et al.** "Microsatellite markers: an overview of the recent progress in plants." *Euphytica* 177.3 (2011): 309-334.

**Käss, Ernst, and Michael Wink.** "Phylogenetic relationships in the Papilionoideae (family Leguminosae) based on nucleotide sequences of cpDNA (rbcL) and ncDNA (ITS 1 and 2)." *Molecular Phylogenetics and Evolution* 8.1 (1997): 65-88.

**Kawata, Masakado.** "Invasion of vacant niches and subsequent sympatric speciation." *Proceedings of the Royal Society of London B: Biological Sciences* 269.1486 (2002): 55-63.

**Keeley, Jon E., and Philip W. Rundel**. "Fire and the Miocene expansion of C4 grasslands." *Ecology Letters* 8.7 (2005): 683-690.

**Keeley, Jon E., et al.,** "Fire as an evolutionary pressure shaping plant traits." *Trends in plant science* 16.8 (2011): 406-411.

**Keppel, Gunnar, et al.,** "Refugia: identifying and understanding safe havens for biodiversity under climate change." *Global Ecology and Biogeography* 21.4 (2012): 393-404.

**Kgope, Barney S., William J. Bond, and Guy F. Midgley.** "Growth responses of African savanna trees implicate atmospheric [CO2] as a driver of past and current changes in savanna tree cover." *Austral Ecology* 35.4 (2010): 451-463.

**Kiehn, Michael.** "Chromosome Survey of the Rubiaceae". *Annals of the Missouri Botanical Garden* 82.3 (1995): 398–408.

**Kolář, Filip, et al.,** "Towards resolving the Knautia arvensis agg.(Dipsacaceae) puzzle: primary and secondary contact zones and ploidy segregation at landscape and microgeographic scales." *Annals of Botany* 103.6 (2009): 963-974.

**Lambin, Eric F., Helmut J. Geist, and Erika Lepers.** "Dynamics of land-use and land-cover change in tropical regions." *Annual review of environment and resources* 28.1 (2003): 205-241.

**Lê, Sébastien, Julie Josse, and François Husson.** "FactoMineR: an R package for multivariate analysis." *Journal of statistical software* 25.1 (2008): 1-18.

**Leadley, Paul.** "Biodiversity scenarios: projections of 21st century change in biodiversity, and associated ecosystem services: a technical report for the global biodiversity outlook" *UNEP/Earthprint* 3. No. 50. (2010).

**Lebrija-Trejos, E., Pérez-García, E. A., Meave, J. A., Bongers, F., & Poorter, L.** "Functional traits and environmental filtering drive community assembly in a species-rich tropical system." *Ecology*, *91* (2010), 386-398.

**Leigh, Jessica W., and David Bryant**. "popart: full-feature software for haplotype network construction." *Methods in Ecology and Evolution* 6.9 (2015): 1110-1116.

**Leitch, A. R., and I. J. Leitch.** "Genomic plasticity and the diversity of polyploid plants." *Science* 320.5875 (2008): 481-483.

**Lewis, Walter H.** "Polyploidy in angiosperms: dicotyledons". *Springer US*, 1980. (Ochnaceae)

**Löve, Áskell**. "IOPB Chromosome Number Reports LXXV". *Taxon* 31.2 (1982): 342–368. .(Parinari)

**Maceira, N. O., P. Jacquard, and R. Lumaret.** "Competition between diploid and derivative autotetraploid Dactylis glomerata L. from Galicia. Implications for the establishment of novel polyploid populations." *New Phytologist* 124.2 (1993): 321-328.

**Maher, Eliza L., Matthew J. Germino, and Niles J. Hasselquist.** "Interactive effects of tree and herb cover on survivorship, physiology, and microclimate of conifer seedlings at the alpine tree-line ecotone." *Canadian Journal of Forest Research* 35.3 (2005): 567-574.

**Maley, Jean.** "The African rain forest–main characteristics of changes in vegetation and climate from the Upper Cretaceous to the Quaternary." *Proceedings of the Royal Society of Edinburgh. Section B. Biological Sciences* 104 (1996): 31-73.

**Mallet, James**. "A species definition for the modern synthesis." *Trends in Ecology & Evolution* 10.7 (1995): 294-299.

**Manel, Stéphanie, Schwartz, M. K., Luikart, G., & Taberlet, P**. "Landscape genetics: combining landscape ecology and population genetics." *Trends in ecology & evolution* 18.4 (2003): 189-197.

**Marengo, J. A. et al.,**"Climate change in the Amazon Basin: Tipping points, changes in extremes,and impacts on natural and human systems" *Tropical rainforest responses to climatic change*. Springer Berlin Heidelberg, 2011. 259-278.

**Maurin, Olivier, et al.,** "Savanna fire and the origins of the 'underground forests' of Africa." *New Phytologist* 204.1 (2014): 201-214.

**Mehra, P. N.** "Cytogenetical evolution of hardwoods." *Nucleus* 15.1 (1972): 64-83.

**Mendelsohn, J. and B. Weber. „**An atlas and profile of Huambo, it's environment and people." *Development Workshop Angola*, Luanda (2013)

**Mendonça, R.C. de, Felfili, J.M., Walter, B.M.T., Silva Junior, M.C. da, Rezende, A.V., Filgueiras, T.S., Nogueira, P.E. & Fagg, C.W.** "Cerrado - Ecologia e Flora" ed. 2 vol. 2. Ch. 15. *Flora Vascular do Bioma Cerrado:checklist com 12,356 espécies*. EMBRAPA CERRADOS, Brasília, DF, Brazil (2008)

**Miller, M. P**. "Alleles In Space (AIS): computer software for the joint analysis of interindividual spatial and genetic information." *Journal of Heredity* 96.6 (2005): 722-724.

**Miller, R. McG, M. Pickford, and B. Senut.** "The geology, palaeontology and evolution of the Etosha Pan, Namibia: Implications for terminal Kalahari deposition." *South African Journal of Geology* 113.3 (2010): 307-334.

**Miranda, Antonio Carlos, et al.,** "Soil and air temperatures during prescribed cerated fires in Central Brazil." *Journal of tropical ecology* 9.03 (1993): 313-320.

**Moore, A. E., and P. A. Larkin.** "Drainage evolution in south-central Africa since the breakup of Gondwana." *South African Journal of Geology* 104.1 (2001): 47-68.

**Moore, A. E., and F. D. Eckardt**. "The evolution and ages of Makgadikgadi palaeo-lakes: Consilient evidence from Kalahari drainage evolution south-central Africa." *South African Journal of Geology* 115.3 (2012): 385-413.

**Mopper, Susan, et al.,** "Genetic differentiation and heterozygosity in pinyon pine associated with resistance to herbivory and environmental stress." *Evolution* (1991): 989-999.

**Morley, R. J.** "Cretaceous and Tertiary climate change and the past distribution of megathermal rainforests." *Tropical rainforest responses to climatic change*. Springer Berlin Heidelberg, 2011. 1-34.

**Mueller, Ulrich G., and L. LaReesa Wolfenbarger**. "AFLP genotyping and fingerprinting." *Trends in Ecology & Evolution* 14.10 (1999): 389-394.

**Nei, M.** "Genetic distance between populations." *American Naturalist* 106 (1972): 283-292.

**Nei, M.** "Estimation of average heterozygosity and genetic distance for small number of individuals." *Genetics* 89 (1978): 583-590.

**Niklas, Karl J.** "Differences between Acer saccharum leaves from open and wind-protected sites." *Annals of Botany* 78.1 (1996): 61-66.

**Oginuma, K., S. K. Y. Lum, Y. H. Lee & H. Tobe.** "Karyomorphology of some Myrtaceae from Singapore." *Gard. Bull. Singapore* 44 (1992): 135–139.

**Oginuma, K., A. Kato, H. Tobe, S. G. Mathenge & F. D. Juma.** "Chromosomes of some woody plants in Kenya". *Acta Phytotax. Geobot*. 44 (1993): 53–58.

**Oldeland, J., Erb, C., Finckh, M. & Jürgens, N. [Eds.]**: "Environmental Assessments in the Okavango Region". *Biodiversity & Ecology* 5 (2013). DOI: 10.7809/b-e.00244.

**Oliveira, Paulo S., and Robert J. Marquis.** "The cerrados of Brazil: ecology and natural history of a neotropical savanna". *Columbia University Press*, 2002.

**Pagès, Jérome.** "Analyse factorielle de données mixtes." *Revue de statistique appliquée* 52.4 (2004): 93-111.

**Palm, Cheryl Ann, et al.,** eds. "Slash-and-burn agriculture: the search for alternatives". *Columbia University Press*, 2013.

**Palmer, Jeffrey D., et al.,** "Chloroplast DNA variation and plant phylogeny." *Annals of the Missouri Botanical Garden* (1988): 1180-1206.

**Panagos P., Jones A., Bosco C., Senthil Kumar P.S**. "Esquisse structurale Provisoire de l'Afrique" European digital archive on soil maps (EuDASM): Preserving important soil data for public free access (2011) *International Journal of Digital Earth*, 4 (5), pp. 434-443.

**Parisod, C., Holderegger, R. and Brochmann, C.** "Evolutionary consequences of autopolyploidy." *New Phytologist*, 186 (2010): 5–17. doi:10.1111/j.1469-8137.2009.03142.x

**Pennington, R. Toby, and Colin E. Hughes**. "The remarkable congruence of New and Old World savanna origins." *New Phytologist* 204.1 (2014): 4-6.

**Percy, Diana M., et al.,** "Progressive island colonization and ancient origin of Hawaiian Metrosideros (Myrtaceae)." *Proceedings of the Royal Society of London B: Biological Sciences* 275.1642 (2008): 1479-1490.

**Perera, N. P.** "Ecology of wetlands (Dambos) of Zambia, and their evaluation for agriculture--a model for the mangement for wetlands in sub-humid eastern and southern Africa." *International journal of ecology and environmental sciences* (1982).

**Pigliucci, Massimo, Courtney J. Murren, and Carl D. Schlichting.** "Phenotypic plasticity and evolution by genetic assimilation." *Journal of Experimental Biology* 209.12 (2006): 2362-2367.

**Pritchard, Jonathan K., Matthew Stephens, and Peter Donnelly**. "Inference of population structure using multilocus genotype data." *Genetics* 155.2 (2000): 945-959.

**Rainey, Paul B., and Michael Travisano.** "Adaptive radiation in a heterogeneous environment." *Nature* 394.6688 (1998): 69-72.

**Ramsey, Justin, Alexander Robertson, and Brian Husband.** "Rapid adaptive divergence in New World Achillea, an autopolyploid complex of ecological races." *Evolution* 62.3 (2008): 639-653.

**Ratter, J. A., Bridgewater, S., Ribeiro, J. F., Fonsêca-Filho, J., Rodrigues da Silva, M., Milliken, W., Pullan, M., Pott, A., Oliveira-Filho, A., Durigan, G. & R.T. Pennington**. *Analysis of the floristic composition of the Brazilian cerrado vegetation IV: Presentation of a Revised Data-Base of 367 Areas*. (2003)

**Ratter, James A., Samuel Bridgewater, and J. Felipe Ribeiro**. "Biodiversity patterns of the woody vegetation of the Brazilian Cerrados." In *Neotropical savannas and seasonally dry forests: plant diversity, biogeography, and conservation*. CRC Press, 2006.

**Ratnam, Jayashree, et al.,** "When is a 'forest'a savanna, and why does it matter?." *Global Ecology and Biogeography* 20.5 (2011): 653-660.

**Rice, A., Glick, L., Abadi, S., Einhorn, M., Kopelman, N. M., Salman-Minkov, A., Mayzel, J., Chay, O. and Mayrose, I.** "The Chromosome Counts Database (CCDB) – a community resource of plant chromosome numbers." *New Phytologist (2014)*. doi: 10.1111/nph.13191

**Rønning, Olaf I.** "The flora of Svalbard." (1996).

**Sakai, Ann K., et al.,** "The population biology of invasive species." *Annual Review of Ecology and Systematics* (2001): 305-332.

**Sakai, Akira, and Walter Larcher**. *Frost survival of plants: responses and adaptation to freezing stress*. Vol. 62. Springer Science & Business Media, 2012.

**Sanger, Frederick, Steven Nicklen, and Alan R. Coulson.** "DNA sequencing with chain-terminating inhibitors." *Proceedings of the National Academy of Sciences* 74.12 (1977): 5463-5467.

**Sanguino, Graciela**: "Wood anatomy and adaptation strategies of suffrutescent shrubs in south-central Angola*".* MSc. Thesis*,Department: Biodiversity, Ecology & Evolution, Universität Hamburg*, (2015)

**Sankaran M, Ratnam J, Hanan NP**. *"*Tree–grass coexistence in savannas revisited – insights from an examination of assumptions and mechanisms invoked in existing models*." Ecology Letters* 7 (2004)*:* 480–490*.*

**Sankaran, Mahesh, et al.,** "Determinants of woody cover in African savannas." *Nature* 438.7069 (2005): 846-849.

**Saxe, Henrik, David S. Ellsworth, and James Heath.** "Tree and forest functioning in an enriched CO2 atmosphere." *New Phytologist* 139.3 (1998): 395-436.

**Schlichting, Carl D**. "The evolution of phenotypic plasticity in plants." *Annual review of ecology and systematics* (1986): 667-693.

**Schoch, Conrad L., et al.** "Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for Fungi." *Proceedings of the National Academy of Sciences* 109.16 (2012): 6241-6246.

**Schupp, Eugene W.** "Quantity, quality and the effectiveness of seed dispersal by animals." *Frugivory and seed dispersal: ecological and evolutionary aspects*. Springer Netherlands, 1993. 15-29.

**Scott, L.** "Pollen evidence for vegetational and climatic change in southern Africa during the Neogene." in *Paleoclimate and evolution, with emphasis on human origins,* Vrba, Elisabeth S. Yale University Press, 1995.

**Shapcott, A**: "Vagile but Inbred: Patterns of Inbreeding and the Genetic Structure Within Populations of the Monsoon Rain Forest Tree Syzygium nervosum (Myrtaceae) in Northern Australia". *Journal of Tropical Ecology*, Vol. 14, No. 5 (Sep., 1998), pp. 595-614

**Silberbauer-Gottsberger, I., W. Morawetz, and G. Gottsberger**. "Frost Damage of Cerrado Plants in Botucatu, Brazil, as Related to the Geographical Distribution of the Species". *Biotropica* 9.4 (1977): 253–261.

**Silvertown, Jonathan.** "Plant coexistence and the niche." *Trends in Ecology & Evolution* 19.11 (2004): 605-611.

**Simon, Marcelo F., Grether, R., de Queiroz, L. P., Skema, C., Pennington, R. T., & Hughes, C. E.** "Recent assembly of the Cerrado, a neotropical plant diversity hotspot, by in situ evolution of adaptations to fire." *Proceedings of the National Academy of Sciences* 106.48 (2009): 20359-20364.

**Simon, Marcelo F., and Toby Pennington**. "Evidence for adaptation to fire regimes in the tropical savannas of the Brazilian Cerrado." *International Journal of Plant Sciences* 173.6 (2012): 711-723.

**Smith, Randy L., and Kenneth J. Sytsma.** "Evolution of Populus nigra (sect. Aigeiros): introgressive hybridization and the chloroplast contribution of Populus alba (sect. Populus)." *American journal of botany* (1990): 1176-1187.

**Smouse, Peter E., Jeffrey C. Long, and Robert R. Sokal**. "Multiple regression and correlation extensions of the Mantel test of matrix correspondence." *Systematic zoology* 35.4 (1986): 627-632.

**Sokal, Robert R., and Neal L. Oden**. "Spatial autocorrelation in biology: 1. Methodology." *Biological journal of the Linnean Society* 10.2 (1978): 199-228

**Soltis, Douglas E., Pamela S. Soltis, and Jennifer A. Tate**. "Advances in the study of polyploidy since plant speciation." *New phytologist* 161.1 (2004): 173-191.

**Soltis, Pamela S., and Douglas E. Soltis**. "The role of hybridization in plant speciation." *Annual review of plant biology* 60 (2009): 561-588.

**Staver, A. Carla, Sally Archibald, and Simon A. Levin.** "The global extent and determinants of savanna and forest as alternative biome states." *Science* 334.6053 (2011): 230-232.

**Stebbins, G. L.** "Polyploidy and the distribution of the arctic-alpine flora: new evidence and a new approach." *Bot. Helvetica* 94.1 (1984): 1-13.

**Stegemann, Sandra, et al.,** "Horizontal transfer of chloroplast genomes between plant species." *Proceedings of the National Academy of Sciences* 109.7 (2012): 2434-2438.

**Stewart, John R., et al.,** "Refugia revisited: individualistic responses of species in space and time." *Proceedings of the Royal Society of London B: Biological Sciences* 277.1682 (2010): 661-671.

**Tajima, Fumio**. "Statistical method for testing the neutral mutation hypothesis by DNA polymorphism." *Genetics* 123.3 (1989): 585-595.

**Tamura K.** "Estimation of the number of nucleotide substitutions when there are strong transition-transversion and G + C-content biases." *Molecular Biology and Evolution* 9 (1992):678-687.

**Tamura K, Stecher G, Peterson D, Filipski A, and Kumar S** "MEGA6: Molecular Evolutionary Genetics Analysis version 6.0" *Molecular Biology and Evolution*: 30 (2013), 2725-2729.

**te Beest, Mariska, et al.,** "The more the better? The role of polyploidy in facilitating plant invasions." *Annals of Botany* (2011): mcr277.

**Thomas Hovestadt, Paul Yao and K. Eduard Linsenmair**: "Seed Dispersal Mechanisms and the Vegetation of Forest Islands in a West African Forest-Savanna Mosaic (Comoé National Park, Ivory Coast)" *Plant Ecology*,Vol. 144, No. 1 (1999), pp. 1-25

**Thurlby, Katie AG, et al.,** "Development of microsatellite loci for Syzygium paniculatum (Myrtaceae), a rare polyembryonic rainforest tree." *Conservation Genetics Resources* 3.2 (2011): 205-208.

**Thurlby, Katie AG, et al.,** "Reproductive bet-hedging in a rare yet widespread rainforest tree, Syzygium paniculatum (Myrtaceae)." *Austral Ecology* 37.8 (2012): 936-944.

**Trauth, Martin H., Juan C. Larrasoaña, and Manfred Mudelsee.** "Trends, rhythms and events in Plio-Pleistocene African climate." *Quaternary Science Reviews* 28.5 (2009): 399-411.

**Tsitrone, Anne, Mark Kirkpatrick, and Donald A. Levin.** "A model for chloroplast capture." *Evolution* 57.8 (2003): 1776-1782.

**Wakeling, Julia L., Michael D. Cramer, and William J. Bond**. "The savanna-grassland 'treeline': why don't savanna trees occur in upland grasslands?." *Journal of Ecology* 100.2 (2012): 381-391.

**Wellington, John H.** Southern Africa: a geographical study. *Physical Geography Vol. 1*. Cambridge: University Press, 1955.

**Werle, E., et al.** "Convenient single-step, one tube purification of PCR products for direct sequencing." *Nucleic Acids Research* 22.20 (1994): 4354.

**West-Eberhard, Mary Jane**. "Phenotypic plasticity and the origins of diversity." *Annual review of Ecology and Systematics* (1989): 249-278.

**White, Frank.** "The underground forests of Africa: a preliminary review." *Gard. Bull. Singapore* 29 (1977): 57-71.

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

**White TJ, Bruns T, Lee S, Taylor J.** "Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics." In: *PCR Protocols: a guide to methods and applications*. (Innis MA, Gelfand DH, Sninsky JJ, White TJ, eds). Academic Press ,(1990), New York, USA: 315–322.

**Wijk, S.** "Performance of Salix herbacea in an alpine snow-bed gradient." The Journal of Ecology (1986): 675-684.

**Willson, M. F.** "Dispersal mode, seed shadows, and colonization patterns." *Vegetatio* 107.1 (1993): 261-280.

**Wright, Sewall.** "The interpretation of population structure by F-statistics with special regard to systems of mating." *Evolution* (1965): 395-420.

**Wright, Shane D., et al.,** "Stepping stones to Hawaii: a trans-equatorial dispersal pathway for Metrosideros (Myrtaceae) inferred from nrDNA (ITS+ ETS)." *Journal of Biogeography* 28.6 (2001): 769-774.

**Zachos, James, et al.,** "Trends, rhythms, and aberrations in global climate 65 Ma to present." *Science* 292.5517 (2001): 686-693.

**Zietkiewicz, Ewa, Antoni Rafalski, and Damian Labuda.** "Genome fingerprinting by simple sequence repeat (SSR)-anchored polymerase chain reaction amplification." *Genomics* 20.2 (1994): 176-183.

**Internet ressources:**

**Bingham, M.G., Willemen, A., Wursten, B.T., Ballings, P. and Hyde, M.A. (2016)**. *Flora of Zambia:* http://www.zambiaflora.com/speciesdata/, retrieved 9. Mar. 2016 and 22. Jun. 2016

**Applied Biosystems:** Fact sheet Microsatellite analysis on the 3130 Series Systems (2004). http://www3.appliedbiosystems.com/cms/groups/mcb_marketing/documents/generaldocuments/cms_040581.pdf ; accessed 3. Jul. 2016

**Applied Biosystems:** 3730/3730xl DNA Analyzers, Sequencing chemistry guide (2002). http://www3.appliedbiosystems.com/cms/groups/mcb_support/documents/generaldocuments/cms_041258.pdf ; accessed 3. Jul. 2016

**JSTOR Global Plants. "JSTOR Global Plants." ITHAKA,** http://plants.jstor.org/compilation/syzygium.intermedium **;** accessed on 7. Apr. 2016

**Global Biodiversity Information facility GBIF**: www.gbif.org ; accessed 9. Mar. 2016

**Global Forest Change**: http://earthenginepartners.appspot.com/science-2013-global-forest ; accessed 28. Jun. 2016

**Ratter et al.,:** Database on Cerrado Plant Biodiversity: http://cerrado.rbge.org.uk/cerrado/index.php ; accessed 2. Jun. 2016

**SASSCAL WeatherNet (2015),** www.sasscalweathernet.org ; accessed 9. Mar. 2016

**SvalbardFlora**, Betula nana: http://svalbardflora.no/index.php?id=382 ; accessed 12. Feb. 2016

**SvalbardFlora,** Salix Polaris: http://svalbardflora.no/index.php?id=606 ; accessed 12. Feb. 2016

**WWF ecoregions** Angolan Miombo Woodland: http://www.worldwildlife.org/ecoregions/at0701 ; accessed 9. Mar. 2016

**WWF ecoregions** Western Zambezian Grasslands: http://www.worldwildlife.org/ecoregions/at0724 ; accessed 9. Mar. 2016

**WWF ecoregions** Cerrado: http://www.worldwildlife.org/ecoregions/nt0704 ; accessed 29. Jun. 2016

# VI. Appendix

## VI.1 List of all samples

**Table A1:** List of all genotyped samples. Next to collection date and location also coordinates are given, though exact sampling site of herbarium specimen is unknown and therefore printed in italic. Phenotypes are indicated by color (blue: SGG, red: SGB, pink: SGBx, lilac: SB, turquoise: SC, green: SGH, yellow: SGM and SGMx) and mode of genotyping by crosses in the respective column.

| Code | Sampling date | Latitude | Longi-tude | Alti-tude | Taxon | Popul-ation | Location | SSR | *trnK* | ITS | Comment |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **B01g** | 1976 | *-17.91600* | *25.85200* | *900* | *Syzygium guineense ssp. barotsense* | | Zambia, Southern Province, Victoria Falls | x | | x | herbarium specimen HBG |
| **B02** | 25.11.2014 | -14.88863 | 22.65252 | 1026 | *Syzygium guineense ssp. barotsense* | Sgg1b | Zambia, Western Province, Liuwa Plains NP | x | x | | |
| **B03** | 25.11.2014 | -14.88773 | 22.65240 | 1026 | *Syzygium guineense ssp. barotsense* | Sgg1b | Zambia, Western Province, Liuwa Plains NP | x | x | x | |
| **B04** | 26.11.2014 | -14.61143 | 22.70230 | 1034 | *Syzygium guineense ssp. barotsense* | Sgg2b | Zambia, Western Province, Liuwa Plains NP | x | x | | |
| **B05** | 26.11.2014 | -14.60945 | 22.70262 | 1035 | *Syzygium guineense ssp. barotsense* | Sgg2b | Zambia, Western Province, Liuwa Plains NP | x | x | | |
| **B06** | 26.11.2014 | -14.49405 | 22.65407 | 1039 | *Syzygium guineense ssp. barotsense* | Sgg3b | Zambia, Western Province, Liuwa Plains NP | x | x | x | |
| **B07** | 27.11.2014 | -14.67135 | 22.69927 | 1048 | *Syzygium guineense ssp. barotsense* | Sgg4b | Zambia, Western Province, Liuwa Plains NP | x | x | | |
| **B08** | 27.11.2014 | -14.67175 | 22.69883 | 1037 | *Syzygium guineense ssp. barotsense* | Sgg4b | Zambia, Western Province, Liuwa Plains NP | x | x | | |
| **B09** | 28.11.2014 | -14.82513 | 22.67742 | 1027 | *Syzygium guineense ssp. barotsense* | Single | Zambia, Western Province, Liuwa Plains NP | x | x | x | |
| **B10g** | 12.12.2014 | -13.69258 | 17.06932 | 1574 | *Syzygium guineense cf ssp. guineense* | Sgg7 | Angola, Bíe, Cusseque | x | x | | |
| **B11** | 12.12.2014 | -13.68900 | 17.06873 | 1586 | *Syzygium guineense cf ssp. barotsense* | Sgg7 | Angola, Bíe, Cusseque | x | | x | |
| **B12** | 16.12.2014 | -13.69912 | 17.01387 | 1582 | *Syzygium guineense cf ssp. barotsense* | Sgg8 | Angola, Bíe, Cusseque | x | x | x | |
| **B13g** | 16.12.2014 | -13.69980 | 17.01403 | 1573 | *Syzygium guineense cf ssp. guineense* | Sgg8 | Angola, Bíe, Cusseque | x | | x | |
| **B14** | 26.11.2014 | -14.49308 | 22.65550 | 1036 | *Syzygium guineense ssp. barotsense* | Sgg3b | Zambia, Western Province, Liuwa Plains NP | | x | x | |
| **B15** | 06.10.2013 | -14.81695 | 22.68402 | 1036 | *Syzygium guineense ssp. barotsense* | | Zambia, Western Province, Liuwa Plains NP | | x | x | collected by M. Finckh |
| **B16** | 06.10.2013 | -14.81695 | 22.68402 | 1036 | *Syzygium guineense ssp. barotsense* | | Zambia, Western Province, Liuwa Plains NP | | x | x | collected by M. Finckh |
| **Bx01** | 26.11.2014 | -14.49407 | 22.65460 | 1038 | *Syzygium guineense cf ssp. huillense x barotsense* | Sgg3bx | Zambia, Western Province, Liuwa Plains NP | x | x | x | |
| **Bx02** | 26.11.2014 | -14.49255 | 22.65472 | 1038 | *Syzygium guineense cf ssp. huillense x barotsense* | Sgg3bx | Zambia, Western Province, Liuwa Plains NP | x | | | |
| **Bx03** | 28.11.2014 | -14.82533 | 22.67737 | 1032 | *Syzygium guineense cf ssp. huillense x barotsense* | Sgg5bx | Zambia, Western Province, Liuwa Plains NP | x | x | x | |
| **Bx04** | 28.11.2014 | -14.82335 | 22.67732 | 1029 | *Syzygium guineense cf ssp. huillense x barotsense* | Sgg5bx | Zambia, Western Province, Liuwa Plains NP | x | | x | |
| **C01o** | 06.09.2013 | -15.72271 | 18.64008 | 1182 | *Syzygium cf cordatum* | Sc23751o | Angola, Cuando Cubango, Longa tributary | x | x | x | collected by M. Finckh |
| **C02o** | 06.09.2013 | -15.72271 | 18.64008 | 1182 | *Syzygium cf cordatum* | Sc23751o | Angola, Cuando Cubango, Longa tributary | x | x | x | collected by M. Finckh |
| **C03o** | 06.09.2013 | -15.72271 | 18.64008 | 1182 | *Syzygium cf cordatum* | Sc23751o | Angola, Cuando Cubango, Longa tributary | x | | x | collected by M. Finckh |
| **C04** | 03.01.2015 | -18.83187 | 22.43997 | 972 | *Syzygium cf cordatum* | Single | Botswana, Northwest, Okavango Delta Panhandle | x | x | x | |
| **C05** | 09.01.2015 | -18.80765 | 22.40385 | 976 | *Syzygium cf cordatum* | Sc3 | Botswana, Northwest, Okavango Delta Panhandle | x | | x | |
| **C06** | 09.01.2015 | -18.80792 | 22.40405 | 979 | *Syzygium cf cordatum* | Sc3 | Botswana, Northwest, Okavango Delta Panhandle | x | x | x | |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **C07** | 18.11.2014 | -18.82767 | 22.43198 | 970 | *Syzygium cf cordatum* | Sc1 | Botswana, Northwest, Okavango Delta Panhandle | x | x | | |
| **C08** | 24.11.2014 | -17.79052 | 23.34500 | 961 | *Syzygium cordatum* | Sc4 | Namibia, Caprivi Strip, Cuando River | x | x | x | |
| **C09** | 04.12.2014 | -15.04012 | 13.20072 | 730 | *Syzygium cordatum* | Sc2 | Angola, Huíla, Escarpment foot | x | x | x | |
| **C10** | 04.12.2014 | -15.03920 | 13.20178 | 740 | *Syzygium cordatum* | Sc2 | Angola, Huíla, Escarpment foot | x | x | | |
| **C11** | 04.12.2014 | -15.03922 | 13.20188 | 737 | *Syzygium cordatum* | Sc2 | Angola, Huíla, Escarpment foot | x | x | x | |
| **C12o** | 1965 | *-9.54700* | *16.34500* | *1128* | *Syzygium cordatum* | | Angola, Malanje | x | | | herbarium specimen LUB |
| **C13** | 1994 | *-19.42000* | *23.37500* | *950* | *Syzygium cordatum* | | Botswana, Northwest, Okavango Delta, 5km N of Delta Camp | | x | | herbarium specimen HBG |
| **C14** | 04.12.2014 | -15.03843 | 13.20325 | 746 | *Syzygium cordatum* | Sc2 | Angola, Huíla, Escarpment foot | | x | x | |
| **G01** | 31.12.2014 | -18.11995 | 21.67357 | 997 | *Syzygium guineense ssp. guineense* | Single | Namibia, Caprivi Strip, Okavango River | x | x | x | |
| **G02** | 31.12.2014 | -18.11515 | 21.66927 | 995 | *Syzygium guineense ssp. guineense* | Single | Namibia, Caprivi Strip, Okavango River | x | x | x | |
| **G03** | 30.11.2014 | -18.13985 | 21.68160 | 993 | *Syzygium guineense ssp. guineense* | Single | Namibia, Caprivi Strip, Okavango River | x | x | x | |
| **G04** | 29.11.2014 | -15.17940 | 22.85377 | 1018 | *Syzygium guineense ssp. guineense* | Sgg6 | Zambia, Western Province, Barotse Floodplain | x | x | | |
| **G05** | 29.11.2014 | -16.63358 | 23.55755 | 994 | *Syzygium guineense ssp. guineense* | Single | Zambia, Western Province, Zambezi River | x | x | x | |
| **G06o** | 25.12.2014 | -14.72857 | 15.31283 | 1300 | *Syzygium guineense cf ssp. guineense* | Single | Angola, Huíla, E of Matala | x | x | x | |
| **G07o** | 25.12.2014 | -14.72857 | 15.31283 | 1300 | *Syzygium guineense cf ssp. guineense* | Single | Angola, Huíla, E of Matala | x | | x | |
| **G08** | 1966 | *-15.05300* | *13.53300* | *1800* | *Syzygium guineense ssp. guineense* | | Angola, Huíla, Cascata | x | | | herbarium specimen LUB |
| **G09** | 01.12.2015 | -14.90133 | 14.7725 | 1258 | *Syzygium guineense cf ssp guineense* | Sgg678 | Angola, Huíla, Bicuar NP | x | x | x | |
| **G10** | 01.12.2015 | -14.90133 | 14.7725 | 1258 | *Syzygium guineense cf ssp guineense* | Sgg678 | Angola, Huíla, Bicuar NP | x | x | x | |
| **G11** | 01.12.2015 | -14.90133 | 14.7725 | 1258 | *Syzygium guineense cf ssp guineense* | Sgg678 | Angola, Huíla, Bicuar NP | x | x | x | |
| **G12** | 1934 | *-10.03100* | *39.46400* | *150* | *Syzygium guineense ssp guineense* | | Tanzania, Lindi, Lake Lutamba | | x | x | herbarium specimen HBG |
| **G13** | 1910 | *-4.47200* | *15.30100* | *288* | *Syzygium guineense ssp guineense* | | Kongo, Kinshasa, 17km S of Leopoldville | | x | | herbarium specimen HBG |
| **G14** | 30.11.2014 | -18.13793 | 21.68080 | 993 | *Syzygium guineense ssp. guineense* | Single | Namibia, Caprivi Strip, Okavango River | | x | x | |
| **G15** | 1911 | *3.50000* | *12.41700* | *650* | *Syzygium guineense* | | Cameroon, East, Dja River | | | x | herbarium specimen HBG |
| **G16** | 30.11.2014 | -18.13792 | 21.68062 | 996 | *Syzygium guineense ssp. guineense* | Single | Namibia, Caprivi Strip, Okavango River | | x | | |
| **H01** | 20.09.2013 | -13.59665 | 17.30775 | 1468 | *Syzygium guineense ssp. huillense* | Sgh23935 | Angola, Bíe, Caluanda | x | | x | collected by M. Finckh |
| **H02** | 20.09.2013 | -13.59665 | 17.30775 | 1468 | *Syzygium guineense ssp. huillense* | Sgh23935 | Angola, Bíe, Caluanda | x | x | x | collected by M. Finckh |
| **H03** | 20.09.2013 | -13.59665 | 17.30775 | 1468 | *Syzygium guineense ssp. huillense* | Sgh23935 | Angola, Bíe, Caluanda | x | | | collected by M. Finckh |
| **H04** | 26.09.2013 | -11.48048 | 21.07199 | 1123 | *Syzygium guineense ssp. huillense* | Sgh23940 | Angola, Moxico, Cameia NP | x | x | | collected by M. Finckh |
| **H05** | 26.09.2013 | -11.48048 | 21.07199 | 1123 | *Syzygium guineense ssp. huillense* | Sgh23940 | Angola, Moxico, Cameia NP | x | x | x | collected by M. Finckh |
| **H06** | 26.09.2013 | -11.48048 | 21.07199 | 1123 | *Syzygium guineense ssp. huillense* | Sgh23940 | Angola, Moxico, Cameia NP | x | | | collected by M. Finckh |
| **H07** | 30.09.2013 | -16.15888 | 25.85778 | 1052 | *Syzygium guineense ssp. huillense* | Sgh23941 | Zambia, Lusaka, Kafue NP | x | x | x | collected by M. Finckh |
| **H08** | 30.09.2013 | -16.15888 | 25.85778 | 1052 | *Syzygium guineense ssp. huillense* | Sgh23941 | Zambia, Lusaka, Kafue NP | x | | | collected by M. Finckh |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **H09** | 10.07.2013 | -15.31176 | 18.77962 | 1248 | *Syzygium guineense ssp. huillense* | Sgh23961 | Angola, Cuando Cubango, Longa River | x | | x | collected by M. Finckh |
| **H10** | 10.07.2013 | -15.31176 | 18.77962 | 1248 | *Syzygium guineense ssp. huillense* | Sgh23961 | Angola, Cuando Cubango, Longa River | x | | | collected by M. Finckh |
| **H11** | 1899 | *-15.76100* | *17.529* | *1163* | *Syzygium guineense ssp huillense* | | Angola, Cuando Cubango, Kuebe-Cubango junction | x | x | x | herbarium specimen HBG |
| **H12** | 25.11.2014 | -15.57550 | 23.28942 | 1018 | *Syzygium guineense ssp. huillense* | Sgh1 | Zambia, Western Province | x | x | x | |
| **H13** | 25.11.2014 | -14.91278 | 22.64578 | 1018 | *Syzygium guineense ssp. huillense* | Sgh2 | Zambia, Western Province, Liuwa Plains NP | x | x | | |
| **H14** | 25.11.2014 | -14.91235 | 22.64715 | 1023 | *Syzygium guineense ssp. huillense* | Sgh2 | Zambia, Western Province, Liuwa Plains NP | x | | x | |
| **H15** | 25.11.2014 | -14.88750 | 22.6523 | 1027 | *Syzygium guineense ssp. huillense* | single | Zambia, Western Province, Liuwa Plains NP | x | x | | |
| **H16** | 26.11.2014 | -14.61315 | 22.70215 | 1032 | *Syzygium guineense ssp. huillense* | Sgh3 | Zambia, Western Province, Liuwa Plains NP | x | x | x | |
| **H17** | 26.11.2014 | -14.61535 | 22.70382 | 1035 | *Syzygium guineense ssp. huillense* | Sgh3 | Zambia, Western Province, Liuwa Plains NP | x | x | x | |
| **H18** | 26.11.2014 | -14.61592 | 22.69633 | 1032 | *Syzygium guineense ssp. huillense* | Sgh4 | Zambia, Western Province, Liuwa Plains NP | x | x | x | |
| **H19** | 26.11.2014 | -14.61510 | 22.69652 | 1026 | *Syzygium guineense ssp. huillense* | Sgh4 | Zambia, Western Province, Liuwa Plains NP | x | | x | |
| **H20** | 26.11.2014 | -14.35500 | 22.57905 | 1041 | *Syzygium guineense ssp. huillense* | Sgh5 | Zambia, Western Province, Liuwa Plains NP | x | | x | |
| **H21** | 26.11.2014 | -14.35443 | 22.57913 | 1042 | *Syzygium guineense ssp. huillense* | Sgh5 | Zambia, Western Province, Liuwa Plains NP | x | x | | |
| **H22** | 27.11.2014 | -14.81543 | 22.68427 | 1030 | *Syzygium guineense ssp. huillense* | Sgh6 | Zambia, Western Province, Liuwa Plains NP | x | x | x | |
| **H23** | 28.11.2014 | -14.82350 | 22.6773 | 1031 | *Syzygium guineense ssp. huillense* | single | Zambia, Western Province, Liuwa Plains NP | x | | | |
| **H24** | 07.12.2014 | -13.69815 | 17.11612 | 1520 | *Syzygium guineense ssp. huillense* | Sgh7 | Angola, Bíe, Cusseque | x | | x | |
| **H25** | 07.12.2014 | -13.69755 | 17.11777 | 1525 | *Syzygium guineense ssp. huillense* | Sgh7 | Angola, Bíe, Cusseque | x | x | | |
| **H26** | 08.12.2014 | -13.70080 | 17.07298 | 1541 | *Syzygium guineense ssp. huillense* | Sgh9 | Angola, Bíe, Cusseque | x | x | x | |
| **H27** | 10.12.2014 | -13.70292 | 17.07137 | 1527 | *Syzygium guineense ssp. huillense* | Sgh9 | Angola, Bíe, Cusseque | x | | | |
| **H28** | 18.12.2014 | -13.00915 | 16.70327 | 1649 | *Syzygium guineense ssp. huillense* | single | Angola, Bíe, Cachingues | x | | | |
| **H29** | 22.12.2014 | -16.20037 | 17.66852 | 1139 | *Syzygium guineense ssp. huillense* | Sgh10 | Angola, Cuando Cubango, S of Caiundo | x | x | x | |
| **H30** | 22.12.2014 | -16.33993 | 17.70382 | 1129 | *Syzygium guineense ssp. huillense* | Sgh11 | Angola, Cuando Cubango, S of Caiundo | x | | | |
| **H31** | 23.12.2014 | -14.15552 | 17.5602 | 1504 | *Syzygium guineense ssp. huillense* | Sgh12 | Angola, Cuando Cubango, N of Menongue | x | | x | |
| **H32** | 25.12.2014 | -14.54228 | 15.84942 | 1508 | *Syzygium guineense ssp. huillense* | Sgh15 | Angola, Huíla, E of Dongo | x | | x | |
| **H33** | 06.11.2015 | -12.52991 | 16.74041 | 1741 | *Syzygium guineense ssp huillense* | Sgh623 | Angola, Bíe, W of Kuito | x | x | x | |
| **H34** | 06.11.2015 | -12.52991 | 16.74041 | 1741 | *Syzygium guineense ssp huillense* | Sgh623 | Angola, Bíe, W of Kuito | x | x | x | |
| **H35** | 06.11.2015 | -12.52991 | 16.74041 | 1741 | *Syzygium guineense ssp huillense* | Sgh623 | Angola, Bíe, W of Kuito | x | x | x | |
| **H36** | 08.12.2015 | -15.10281 | 14.83834 | 1239 | *Syzygium guineense ssp huillense* | Sgh691 | Angola, Huíla, Bicuar NP | x | x | x | |
| **H37** | 08.12.2015 | -15.10281 | 14.83834 | 1239 | *Syzygium guineense ssp huillense* | Sgh691 | Angola, Huíla, Bicuar NP | x | x | x | |
| **H38** | 06.10.2013 | -14.81635 | 22.68413 | 1034 | *Syzygium guineense ssp. huillense* | Sgh23943 | Zambia, Western Province, Liuwa Plains NP | | x | x | collected by M. Finckh |
| **H39** | 22.12.2014 | -16.19953 | 17.66958 | 1140 | *Syzygium guineense ssp. huillense* | Sgh10 | Angola, Cuando Cubango, S of Caiundo | | x | x | |
| **H40** | 16.06.2013 | -13.66613 | 17.00935 | 1570 | *Syzygium guineense ssp. huillense* | Sgh23714 | Angola, Bíe, Cusseque | | x | x | collected by M. Finckh |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **H41** | 16.06.2013 | -13.66613 | 17.00935 | 1570 | *Syzygium guineense ssp. huillense* | Sgh23714 | Angola, Bíe, Cusseque | | x | x | collected by M. Finckh |
| **H42** | 25.11.2014 | -15.57525 | 23.28920 | 1022 | *Syzygium guineense ssp. huillense* | Sgh1 | Zambia, Western Province | | x | x | |
| **H43** | 08.12.2015 | -15.10281 | 14.83834 | 1239 | *Syzygium guineense ssp huillense* | Sgh691 | Angola, Huíla, Bicuar NP | | x | x | |
| **H44** | 06.10.2013 | -14.81635 | 22.68413 | 1034 | *Syzygium guineense ssp. huillense* | Sgh23943 | Zambia, Western Province, Liuwa Plains NP | | | x | collected by M. Finckh |
| **H45** | 06.10.2013 | -14.81635 | 22.68413 | 1034 | *Syzygium guineense ssp. huillense* | Sgh23943 | Zambia, Western Province, Liuwa Plains NP | | | x | collected by M. Finckh |
| **H46** | 06.10.2013 | -14.81635 | 22.68413 | 1034 | *Syzygium guineense ssp. huillense* | Sgh23943 | Zambia, Western Province, Liuwa Plains NP | | | x | collected by M. Finckh |
| **H47** | 06.10.2013 | -14.81635 | 22.68413 | 1034 | *Syzygium guineense ssp. huillense* | Sgh23943 | Zambia, Western Province, Liuwa Plains NP | | | x | collected by M. Finckh |
| **H48** | 06.10.2013 | -14.81635 | 22.68413 | 1034 | *Syzygium guineense ssp. huillense* | Sgh23943 | Zambia, Western Province, Liuwa Plains NP | | | x | collected by M. Finckh |
| **H49** | 22.12.2014 | -16.33923 | 17.70372 | 1126 | *Syzygium guineense ssp. huillense* | Sgh11 | Angola, Cuando Cubango, S of Caiundo | | | x | |
| **H50** | 25.12.2014 | -14.54208 | 15.84993 | 1507 | *Syzygium guineense ssp. huillense* | Sgh15 | Angola, Huíla, E of Dongo | | x | x | |
| **L01** | 1911 | *3.01600* | *9.95600* | *15* | *Syzygium guineense ssp littorale* | | Cameroon, South, Coast between Kribi/Lonji | | x | | herbarium specimen HBG |
| **M01** | 30.10.2012 | -13.34144 | 16.73154 | 1665 | *Syzygium guineense ssp. macrocarpum* | Sgm23691 | Angola, Bíe, Chitembo | x | x | | collected by M. Finckh |
| **M02** | 30.10.2012 | -13.34144 | 16.73154 | 1665 | *Syzygium guineense ssp. macrocarpum* | Sgm23691 | Angola, Bíe, Chitembo | x | x | | collected by M. Finckh |
| **M03** | 30.10.2012 | -13.34144 | 16.73154 | 1665 | *Syzygium guineense ssp. macrocarpum* | Sgm23691 | Angola, Bíe, Chitembo | x | | x | collected by M. Finckh |
| **M04** | 17.06.2013 | -13.63722 | 17.14019 | 1577 | *Syzygium guineense* | Sgg23717 | Angola, Bíe, Chitembo | x | x | x | collected by M. Finckh |
| **M05** | 17.06.2013 | -13.63722 | 17.14019 | 1577 | *Syzygium guineense* | Sgg23717 | Angola, Bíe, Chitembo | x | x | | collected by M. Finckh |
| **M06** | 1900 | *-15.11600* | *16.07700* | *1315* | *Syzygium guineense ssp latifolia* | | Angola, Cunene, Menempremp/Cassinga | x | x | x | herbarium specimen |
| **M07** | 01.12.2014 | -14.49627 | 16.60492 | 1543 | *Syzygium guineense ssp. macrocarpum* | Sgm1 | Angola, Cuando Cubango, between Capelongo & Cuchi | x | x | x | |
| **M08** | 01.12.2014 | -14.49603 | 16.60465 | 1551 | *Syzygium guineense ssp. macrocarpum* | Sgm1 | Angola, Cuando Cubango, between Capelongo & Cuchi | x | x | x | |
| **M09** | 06.12.2014 | -12.97030 | 14.76572 | 1258 | *Syzygium guineense ssp. macrocarpum* | Sgm2 | Angola, Benguela, Alto Catumbela | x | x | x | |
| **M10** | 06.12.2014 | -12.97015 | 14.76617 | 1267 | *Syzygium guineense ssp. macrocarpum* | Sgm2 | Angola, Benguela, Alto Catumbela | x | x | x | |
| **M11** | 07.12.2014 | -13.72255 | 17.11370 | 1562 | *Syzygium guineense ssp. macrocarpum* | Sgm3 | Angola, Bíe, Cusseque | x | x | x | |
| **M12** | 07.12.2014 | -13.72077 | 17.11425 | 1570 | *Syzygium guineense ssp. macrocarpum* | Sgm3 | Angola, Bíe, Cusseque | x | x | x | |
| **M13** | 12.12.2014 | -13.68257 | 17.07407 | 1580 | *Syzygium guineense cf ssp. macrocarpum* | Sgg7 | Angola, Bíe, Cusseque | x | | x | |
| **M14** | 18.12.2014 | -13.01533 | 16.70540 | 1694 | *Syzygium guineense ssp. macrocarpum* | single | Angola, Bíe, Cachingues | x | | x | |
| **M15** | 1910 | *-4.47200* | *15.30100* | *288* | *Syzygium guineense ssp. macrocarpum* | | Kongo, Kinshasa, 17km S of Leopoldville | | x | x | herbarium specimen HBG |
| **M16** | 08.12.2014 | -13.69907 | 17.06853 | 1572 | *Syzygium guineense ssp. macrocarpum* | Sgm4 | Angola, Bíe, Cusseque | | x | x | |
| **M17** | 08.12.2014 | -13.69810 | 17.06898 | 1571 | *Syzygium guineense ssp. macrocarpum* | Sgm4 | Angola, Bíe, Cusseque | | x | | |
| **M18** | 30.10.2012 | -13.66676 | 17.00826 | 1578 | *Syzygium guineense ssp. macrocarpum* | Sgm23686 | Angola, Bíe, Cusseque | | x | x | collected by M. Finckh |
| **M19** | 30.10.2012 | -13.66676 | 17.00826 | 1578 | *Syzygium guineense ssp. macrocarpum* | Sgm23686 | Angola, Bíe, Cusseque | | | x | collected by M. Finckh |
| **M20** | 30.10.2012 | -13.66676 | 17.00826 | 1578 | *Syzygium guineense ssp. macrocarpum* | Sgm23686 | Angola, Bíe, Cusseque | | | x | collected by M. Finckh |
| **Mx01** | 27.12.2014 | -14.81457 | 13.40637 | 2295 | *Syzygium guineense ssp. macrocarpum* | Sgh16mx | Angola, Huíla, Escarpment Tundavala | x | x | | |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Mx02** | 27.12.2014 | -14.80043 | 13.40087 | 2254 | *Syzygium guineense ssp. macrocarpum* | Sgh16mx | Angola, Huíla, Escarpment Tundavala | **x** | **x** | **x** | |
| **Mx03** | 27.12.2014 | -14.80052 | 13.40120 | 2249 | *Syzygium guineense ssp. macrocarpum* | Sgh16mx | Angola, Huíla, Escarpment Tundavala | **x** | | **x** | |
| **Mx04** | 07.12.2014 | -13.70070 | 17.11800 | 1542 | *Syzygium guineense ssp. macrocarpum* | Sgh8mx | Angola, Bíe, Cusseque | **x** | **x** | **x** | |
| **Mx05** | 07.12.2014 | -13.70052 | 17.11715 | 1538 | *Syzygium guineense ssp. macrocarpum* | Sgh8mx | Angola, Bíe, Cusseque | **x** | | | |
| **Mx06** | 08.12.2014 | -13.69963 | 17.07183 | 1550 | *Syzygium guineense ssp. macrocarpum* | Sgh9mx | Angola, Bíe, Cusseque | **x** | **x** | **x** | |
| **Mx07** | 10.12.2014 | -13.70245 | 17.07117 | 1536 | *Syzygium guineense ssp. macrocarpum* | Sgh9mx | Angola, Bíe, Cusseque | **x** | **x** | | |
| **Mx08** | 19.12.2014 | -13.76015 | 17.08635 | 1519 | *Syzygium guineense ssp. macrocarpum* | Single | Angola, Bíe, Cusseque | **x** | **x** | **x** | |
| **Mx09** | 25.12.2014 | -14.67065 | 17.14155 | 1392 | *Syzygium guineense ssp. huillense* | Sgh13mx | Angola, Cuando Cubango, W of Menongue | **x** | **x** | **x** | |
| **Mx10** | 25.12.2014 | -14.67057 | 17.14250 | 1383 | *Syzygium guineense ssp. huillense* | Sgh13mx | Angola, Cuando Cubango, W of Menongue | **x** | **x** | | |
| **Mx11** | 18.11.2015 | -14.79592 | 13.40861 | 2245 | *Syzygium guineense ssp macrocarpum* | Sgmx657 | Angola, Huíla, Escarpment Tundavala | **x** | **x** | **x** | |
| **Mx12** | 18.11.2015 | -14.79592 | 13.40861 | 2245 | *Syzygium guineense ssp macrocarpum* | Sgmx657 | Angola, Huíla, Escarpment Tundavala | **x** | **x** | **x** | |
| **Mx13** | 16.12.2014 | 13.69977 | 17.01330 | 1589 | *Syzygium guineense ssp. macrocarpum* | Single | Angola, Bíe, Cusseque | | **x** | | |
| **O01** | 1899 | *-15.76100* | *17.52900* | *1163* | *Syzygium benguellense* | | Angola, Cuando Cubango, Kuebe-Cubango junction | **x** | **x** | **x** | herbarium specimen HBG |
| **O02** | 1964 | *-9.54700* | *16.34500* | *1128* | *Syzygium benguellense* | | Angola, Malanje | **x** | | | herbarium specimen LUB |
| **S01** | 1911 | *2.05500* | *15.20300* | *400* | *Syzygium guineense ssp staudtii* | | Cameroon, East, Woodland near Molundu | | **x** | | herbarium specimen HBG |
| **T01** | 27.12.2014 | -14.80738 | 13.40215 | 2283 | *"Syzygium tundavalense"* | St2 | Angola, Huíla, Escarpment Tundavala | **x** | **x** | **x** | |
| **T02** | 27.12.2014 | -14.80693 | 13.40182 | 2284 | *"Syzygium tundavalense"* | St2 | Angola, Huíla, Escarpment Tundavala | **x** | **x** | | |
| **T03** | 27.12.2014 | -14.80455 | 13.40203 | 2280 | *"Syzygium tundavalense"* | St2 | Angola, Huíla, Escarpment Tundavala | **x** | **x** | **x** | |
| **T04** | 1907 | | | | *Syzygium guineense ssp angustifolia* | | Deutsch-Ostafrika (Tanzania) | **x** | **x** | **x** | herbarium specimen HBG |
| **T05** | 03.12.2014 | -14.81618 | 13.38075 | 2246 | *"Syzygium tundavalense"* | St1 | Angola, Huíla, Escarpment Tundavala | **x** | | | |
| **T06** | 03.12.2014 | -14.81618 | 13.38075 | 2246 | *"Syzygium tundavalense"* | St1 | Angola, Huíla, Escarpment Tundavala | **x** | | | |
| **T07** | 03.12.2014 | -14.81688 | 13.38145 | 2246 | *"Syzygium tundavalense"* | St1 | Angola, Huíla, Escarpment Tundavala | **x** | **x** | **x** | |
| **T08** | 03.12.2014 | -14.81688 | 13.38145 | 2246 | *"Syzygium tundavalense"* | St1 | Angola, Huíla, Escarpment Tundavala | **x** | | | |
| **T09** | 1961 | *-15.02800* | *13.37700* | *1890* | *Syzygium guineense ssp. littorale* | | Angola, Huíla, Humpata | **x** | | | herbarium specimen LUB |
| **T10** | 1971 | *-14.81600* | *13.38000* | *2250* | *Syzygium guineense ssp. afromontanum* | | Angola, Huíla, Escarpment Tundavala | **x** | | | herbarium specimen LUB |
| **T11** | 03.12.2014 | -14.81618 | 13.38075 | 2246 | *"Syzygium tundavalense"* | St1 | Angola, Huíla, Escarpment Tundavala | | **x** | | |
| **T12** | 27.12.2014 | -14.80738 | 13.40215 | 2283 | *"Syzygium tundavalense"* | St2 | Angola, Huíla, Escarpment Tundavala | | **x** | | |
| **T13** | 27.12.2014 | -14.80455 | 13.40203 | 2280 | *"Syzygium tundavalense"* | St2 | Angola, Huíla, Escarpment Tundavala | | **x** | | |

## VI.2 List of populations

**Table A2:** Locations and characteristics of self-sampled and provided populations and single samples used in this study (coordinates of herbarium samples approx.).

| Population | Latitude | Longitude | Altitude | Life form | Habitat | |
|---|---|---|---|---|---|---|
| Sgh1 | -15.57550 | 23.28942 | 1018 | suffrutex | Zambia, dambos grassland | *Entire populations, self-sampled* |
| Sgh2 | -14.91278 | 22.64578 | 1022 | suffrutex | Zambia, Liuwa plains grassland | |
| Sgh3 | -14.61315 | 22.70215 | 1032 | suffrutex | Zambia, Liuwa plains grassland/ecotone | |
| Sgh4 | -14.61592 | 22.69633 | 1032 | suffrutex | Zambia, Liuwa plains grassland | |
| Sgh5 | -14.35500 | 22.57905 | 1041 | suffrutex | Zambia, Liuwa plains grassland | |
| Sgh6 | -14.81617 | 22.68355 | 1028 | suffrutex | Zambia, Liuwa plains grassland | |
| Sgh7 | -13.69815 | 17.11612 | 1520 | suffrutex | Angola, Cusseque grassland | |
| Sgh8 | -13.70070 | 17.11800 | 1525 | shrub | Angola, Cusseque grassland/forest ecotone | |
| Sgh9 | -13.69963 | 17.07183 | 1542 | suffrutex/shrub | Angola, Cusseque grassland/forest ecotone | |
| Sgh10 | -16.20037 | 17.66852 | 1140 | suffrutex | Angola, Cuando-Cubango forest/bog ecotone | |
| Sgh11 | -16.33923 | 17.70372 | 1126 | suffrutex | Angola, Cuando-Cubango grassland | |
| Sgh12 | -14.15552 | 17.56020 | 1504 | suffrutex | Angola, Cuando-Cubango grassland near river bank | |
| Sgh13 | -14.67065 | 17.14155 | 1392 | suffrutex/shrub | Angola, Cuando-Cubango grassland/forest ecotone | |
| Sgh14 | -14.58978 | 16.65658 | 1508 | suffrutex | Angola, Cuando-Cubango grassland/bog ecotone | |
| Sgh15 | -14.54228 | 15.84942 | 1508 | suffrutex/shrub | Angola, Huíla grassland/bog ecotone | |
| Sgh16 | -14.80052 | 13.40120 | 2295 | shrub | Angola, mountainous escarpment top near Tundavala | |
| Sgg1 | -14.88863 | 22.65252 | 1026 | tree | Zambia, Liuwa plains forest | |
| Sgg2 | -14.61143 | 22.70230 | 1034 | tree | Zambia, Liuwa plains forest | |
| Sgg3 | -14.49308 | 22.65550 | 1036 | tree/shrub | Zambia, Liuwa plains forest | |
| Sgg4 | -14.67135 | 22.69927 | 1048 | tree | Zambia, Liuwa plains forest | |
| Sgg5 | -14.82533 | 22.67737 | 1032 | tree/shrub | Zambia, Liuwa plains forest/ecotone | |
| Sgg6 | -15.17988 | 22.85437 | 1018 | tree | Zambia, Barotse plains | |
| Sgg7 | -13.69258 | 17.06932 | 1574 | tree | Angola, Cusseque dense Miombo forest | |
| Sgg8 | 13.69977 | 17.01330 | 1586 | tree | Angola, Cusseque dense Miombo forest | |
| Sgg9 | -18.11515 | 21.66927 | 993 | tree | Namibia, Okavango river bank | |
| Sgm1 | -14.49627 | 16.60492 | 1543 | tree | Angola, Huíla open Miombo forest | |
| Sgm2 | -12.97030 | 14.76572 | 1258 | tree | Angola, Benguela open Miombo forest | |
| Sgm3 | -13.72255 | 17.11370 | 1562 | tree | Angola, Cusseque open Miombo forest | |
| Sgm4 | -13.69907 | 17.06853 | 1572 | tree | Angola, Cusseque open Miombo forest | |
| St1 | -14.81618 | 13.38075 | 2246 | tree | Angola, mountainous escarpment top near Tundavala | |
| St2 | -14.80738 | 13.40215 | 2283 | tree | Angola, mountainous escarpment top near Tundavala | |
| Sc1 | -18.82767 | 22.43198 | 970 | tree | Botswana, Okavango Delta tree islet | |
| Sc2 | -15.04012 | 13.20072 | 730 | tree | Angola, escarpment bottom/riverbed | |
| Sc3 | -18.80765 | 22.40385 | 976 | tree | Botswana, Okavango Delta tree islet | |

| Population | Latitude | Longitude | Altitude | Life form | Habitat | |
|---|---|---|---|---|---|---|
| Sgh23935 | -13.59665 | 17.30775 | 1463 | suffrutex | Angola, grassland northeast of Cusseque | *Entire populations, sampled by M. Finckh* |
| Sgh23940 | -11.48048 | 21.07199 | 1123 | suffrutex | Angola, grassland in Moxico province | |
| Sgh23941 | -16.15888 | 25.85778 | 1053 | suffrutex | Zambia, grassland in Kafue Plains NP | |
| Sgh623 | -12.52991 | 16.74041 | 1741 | suffrutex | Angola, grassland and agricultural fields near Kuito | |
| Sgh691 | -15.10281 | 14.83834 | 1239 | suffrutex | Angola, grassland in Bicuar NP | |
| Sgm23691 | -13.34144 | 16.73154 | 1665 | tree | Angola, open Miombo woodland near Chitembo | |
| Sgmx657 | -14.79592 | 13.40861 | 2245 | shrub | Angola, mountainous escarpment top near Tundavala | |
| Sgg678 | -14.90133 | 14.77250 | 1258 | tree | Angola, river bank near Bicuar NP | |
| Sgg23717 | -13.63722 | 17.14019 | 1577 | tree | Angola, dense Miombo forest north of Cusseque | |
| Sc23751 | -15.72271 | 18.64008 | 1182 | tree | Angola, riverbed in Cuando-Cubango province | |
| C08 | -17.79052 | 23.34500 | 961 | shrub | Namibia, riverbank near Kongola | *Single samples* |
| Mx08 | -13.76015 | 17.08635 | 1519 | shrub | Angola, forest/grassland ecotone south of Cusseque | |
| M14 | -13.01533 | 16.70540 | 1694 | tree | Angola, open Miombo woodland in Cuchi catchment | |
| H28 | -13.00915 | 16.70327 | 1649 | suffrutex | Angola, bog fringe in Cuchi catchment | |
| G05 | -16.63358 | 23.55755 | 994 | tree | Zambia, Zambezi River at Sikuka | |
| G06o | -14.72857 | 15.31283 | 1300 | tree | Angola, riverbed in Huíla province | |

| Specimen | Latitude | Longitude | Taxon | Location | |
|---|---|---|---|---|---|
| B01g | -17.916 | 25.852 | *Syzygium guineense ssp barotsense* | Zambia, Victoria Falls | *Samples from herbarium specimen* |
| C12o | -9.547 | 16.345 | *Syzygium cordatum* | Angola, Malanje | |
| C13 | -19.420 | 23.375 | *Syzygium cordatum* | Botswana, Okavango Delta, 5km N Delta Camp | |
| G08 | -15.053 | 13.533 | *Syzygium guineense ssp. guineense* | Angola, Cascata do Huíla | |
| G12 | -10.031 | 39.464 | *Syzygium guineense ssp guineense* | Tanzania, Lindi, Lake Lutamba | |
| G13 | -4.472 | 15.301 | *Syzygium guineense ssp guineense* | DR Congo, Kimuenza, 17km S of Leopoldville | |
| H11 | -15.761 | 17.529 | *Syzygium guineense ssp huillense* | Angola, Kuebe mouth into Kubango | |
| L01 | 3.016 | 9.956 | *Syzygium guineense ssp littorale* | Cameroon, Coast between Kribi/Londgji | |
| M06 | -15.116 | 16.077 | *Syzygium guineense ssp latifolia* | Angola, Menempremp (Cassinga), Kubango | |
| M15 | -4.472 | 15.301 | *Syzygium guineense ssp macrocarpum* | DR Congo, Kimuenza, 17km S of Leopoldville | |
| O01 | -15.761 | 17.529 | *Syzygium benguellense* | Angola, Kuebe mouth into Kubango | |
| O02 | -9.547 | 16.345 | *Syzygium benguellense* | Angola, Malanje | |
| S01 | 2.055 | 15.203 | *Syzygium guineense ssp staudtii* | Cameroon, dense Woodlands of Moloundou | |
| T04 | -7.272 | 37.591 | *Syzygium guineense ssp angustifolia* | Tanzania, Mountain ridge in Deutsch-Ostafrika | |
| T09 | -15.028 | 13.377 | *Syzygium guineense ssp littorale* | Angola, Huíla, Humpata | |
| T10 | -14.816 | 13.380 | *Syzygium guineense ssp. afromontanum* | Angola, Huíla, Tundavala | |

## VI.3 R Scripts

**###Supplementary R Script, Master Thesis. Paulina Zigelski: TREES GOING UNDERGROUND, HOW ENVIRONMENT AND GENETICS SHAPE AFRICAS ZAMBEZIAN MIOMBO REGION. Universität Hamburg (2016)####**

**###Evaluating SSR data set with package polysat, computing statistics considering the tetraploidy. Also conversion into SPAGeDi format for further analyses####**

```
setwd("C:/Paule/Syzygium/Auswertung/SSR/POLYSAT")library(polysat)
input <- read.Structure(infile="20160227 structure popinfo final rearranged.txt",
ploidy=4, missingin=-9, sep="\t", markernames=T, labels=T, popinfocol=1,
extracols=1, ploidyoutput="one")

###checking some general structures and statistics of the STRUCTURE genotype data
set####
summary(input)
viewGenotypes(input)
alleleDiversity(input)
genotypeDiversity(input, threshold=0.3)
newploids <- estimatePloidy(input)  ##ploidy estimation based on max. number of
alleles/locus
freq <- simpleFreq(input)  ##allele frequencies
Fst <- calcFst(freq)  ##Wright's F statistic
Fst

###calculating genetic distances based on allele frequencies (simple and
advanced)####
dist <- meandistance.matrix(freq, all.distances=T)
dist2 <- meandistance.matrix2(input, all.distances=T, freq=freq)

###writing tables for SPAGeDi use
write.table(dist$MeanMatrix, file="distances.txt", sep="\t")
write.freq.SPAGeDi(freq, file="dists.txt", usatnts=Usatnts(input), digits=3,
pops=row.names(freq))
all.dist <- read.table("allele freqs test.txt", header=F, sep="\t")
dists <- dist(all.dist, method="euclidean", diag=T, upper=T)
write.table(as.matrix(dists), file="all-distances.txt", sep="\t")

###Calculating AMOVA between different groupings of genetic, ecological and spatial
context####
setwd("C:/Paule/Syzygium/Auswertung/SSR/ADGENET")
library(pegas)
dist2 <- read.table(file="distances2.txt", sep="\t", row.names=1, header=T)
##matrix of genetic distances between all individuals
suppl <- read.table(file="pheno-habitat-region.txt", sep="\t", header =T)
##dataframe with phenotype, habitat and sampling location information of each
individual

####computing AMOVA with phenotypes/sampling regions/habitats as groups####
suppl2 <- na.omit(suppl)
```

```
attach(suppl2)
amo1b <- amova(dist2~phenotype, nperm=1000)
amo1b
amo2b <- amova(dist2~habitat, nperm=1000)
amo2b
amo3b <- amova(dist2~region, nperm=1000)
amo3b
amo4 <- amova(dist2~phenotype/habitat, nperm=1000)
amo4
amo4b <- amova(dist2~habitat/phenotype, nperm=1000)
amo4b
amo5 <- amova(dist2~phenotype/region, nperm=1000)
amo5
amo5b <- amova(dist2~region/phenotype, nperm=1000)
amo5b
 detach(suppl2)

####trying AMOVA with package vegan
library(vegan)
ado1 <- adonis(dist2~phenotype, data=suppl2, permutations=1000)
ado1
ado2 <- adonis(dist2~habitat, data=suppl2, permutations=1000)
ado2
ado3 <- adonis(dist2~region, data=suppl2, permutations=1000)
ado3
ado4 <- adonis(dist2~phenotype/habitat, data=suppl2, permutations=1000)
ado4
ado4b <- adonis(dist2~phenotype/habitat/region, data=suppl2, permutations=1000)
ado4b
ado4e <- adonis(dist2~habitat+phenotype, data=suppl2, permutations=1000)
ado4e
ado4d <- adonis(dist2~phenotype+habitat+region, data=suppl2, permutations=1000)
ado4d

###looking at distances between SGH and SGM####
dist3 <- read.table(file="distances3.txt", sep="\t", header=T, row.names=1)
suppl3 <- read.table(file="pheno-habitat-region3.txt", sep="\t", header =T,
row.names=1)
suppl3 <- suppl3[-43,]
ado5 <- adonis(dist3~region, data=suppl3, permutations=1000)
ado5

###looking at distances between SGH, SGB(zambian), SGBx####
dist4 <- read.table(file="distances4.txt", sep="\t", header=T, row.names=1)
suppl4 <- read.table(file="pheno-habitat-region4.txt", sep="\t", header =T,
row.names=1)
ado6 <- adonis(dist4~habitat+phenotype, data=suppl4, permutations=1000)
ado6
```

**###Performing a Factor Analysis on Mixed Data (FAMD) and a Hierarchical Clustering on Principle Components (HCPC) on Syzygium populations, characterized by morphological, ecological and genetic variables.####**

```
setwd("C:/Paule/Syzygium/Auswertung/FAMD")

###reading in raw data, consisting of a matrix of populations (rows) and their
scores for continuous and categorical parameters (columns)####
data <- read.table(file="20160211 FAMD final new.csv", sep=";", header=T,
na.strings="NA", row.names=1)
attach(data)
str(data)

library(FactoMineR)
library(plyr)
library(Cairo)

###renaming some values as they are listed as integer numbers but shall be treated
as categories####
data$Haplotype <- as.factor(Haplotype)
data$Haplotype <- revalue(data$Haplotype, c("1"="HT1", "3"="HT3", "2"="HT2",
"7"="HT7", "5"="HT5","6"="HT6", "4"="HT4"))

###checking for correlations between the parameters. which variables are redundant
and should be excluded. only for continuous parameters####
library(corrgram)
corrgram(data, diag.panel=NULL, lower.panel=NULL, upper.panel=panel.shade)

###trying another package
library(corrplot)
cor.dat <- cor(data.num, method="pearson")
cplot<-corrplot(cor.dat, method="number", type = "upper", diag=F, order="FPC",
tl.srt=45, tl.cex=.7, title="", tl.offset=0.4, mar=c(1,1,2,2))

###saving resulting plot as separate file
Cairo(1000, 1000, file="corrplot.png", type="png", bg="white")
corrplot(cor.dat, method="number", type = "upper", diag=F, order="FPC", tl.srt=45,
tl.cex=.7, title="", tl.offset=0.4, mar=c(1,1,2,2))
dev.off()

###_defining a FAMD, combination of MCA (multiple correspondence
analysis=categorical variables) and PCA (principle correspondence
analysis=continuous variables)####
famd <- FAMD(data, graph=T, ncp=5)      ##full data set
famd.nomorph <- FAMD(data[,-1], graph=T, ncp=7) ##without phenotype information
famd.gen <- FAMD(data[,c(1,4,11,12,13)], graph=T, ncp=7)   ##only genetic
information
famd.eco <- FAMD(data[,-c(4,11,12,13)], graph=T, ncp=7)    ##only ecological
information

###saving as separate image file
Cairo(700, 700, file="FAMD ind plots.png", type="png", bg="white")
```

```
plot(famd, choix="ind", cex=0.8)
dev.off()

###results of each FAMD variant for comparison. Later on only full model/model
without phenotype information will be used####
summary.FAMD(famd, nbelements=Inf, nbind=Inf, ncp=7, file="20160216 FAMDfull
results final")
summary.FAMD(famd.nomorph, nbelements=Inf, nbind=Inf, ncp=7, file="20160229
FAMDnomorph results final")
summary.FAMD(famd.gen, nbelements=Inf, nbind=Inf, ncp=7, file="20160216 FAMD.GEN
results final")
summary.FAMD(famd.eco, nbelements=Inf, nbind=Inf, ncp=7, file="20160216 FAMD.ECO
results final")

###interpreting the results####
res.dims1 <- dimdesc(famd, axes=1:3) ##which factors influence the first 3
dimensions
res.dims1
res.dims2 <- dimdesc(famd.nomorph, axes=1:3)
res.dims2

res.cats1 <- catdes(data, proba=0.05, num.var=1) ##which factors influence
phenotype
res.cats1
res.cats2 <- catdes(data[,-1], proba=0.05, num.var=1)
res.cats2

###saving as separate image file
Cairo(2500, 1000, file="FAMD pheno results.png", type="png", bg="white")
plot(res.cats)
par(mfrow=c(1,1))
dev.off()

Cairo(2500, 1000, file="FAMD pheno.nomorph results.png", type="png", bg="white")
plot(res.cats2)
par(mfrow=c(1,1))
dev.off()

###performing a HCPC on the FAMD results####
hcpc <- HCPC(famd, nb.clust=0, consol=T)
plot.HCPC(hcpc, axes=c(1,2), choice="tree", centers.plot=T, t.level="centers",
ind.names=F)

#####exclusion of phenotype assignment, to see if it still clusters this way
hcpc.nomorph <- HCPC(famd.nomorph, nb.clust=0, consol=T)
plot.HCPC(hcpc.nomorph, axes=c(1,2), choice="tree", draw.tree=T)
```

## VI.4 Supplementary figures

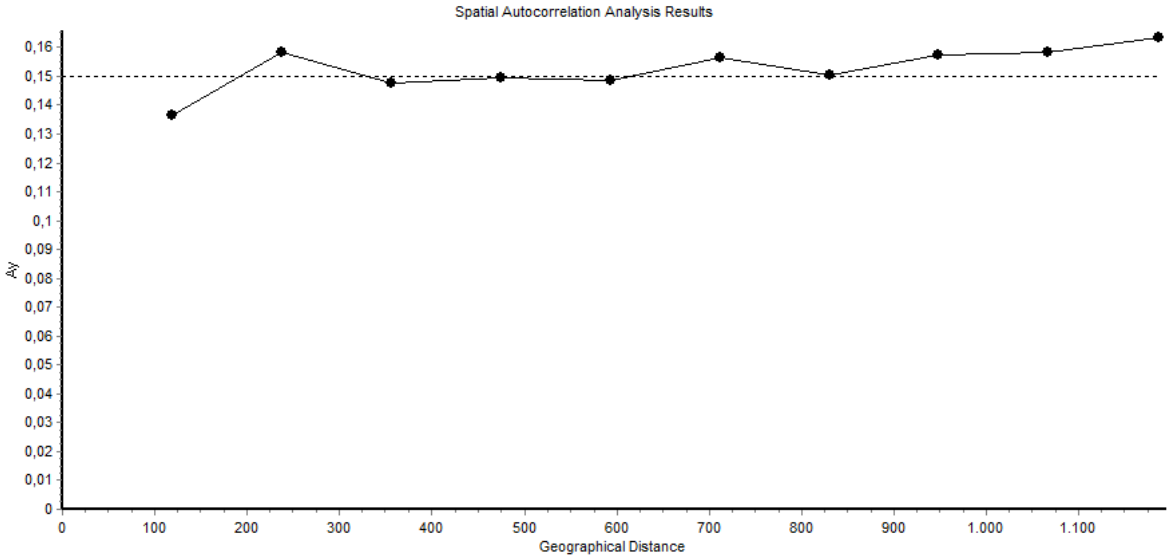



**Figure A1:** Spatial Autocorrelation Analysis of SGH individuals from the whole sampling range. The averaged pairwise genetic distance accounted for Ay = 0.149 and the covariance for V = 0.008 with an insignificance of p = 0.509

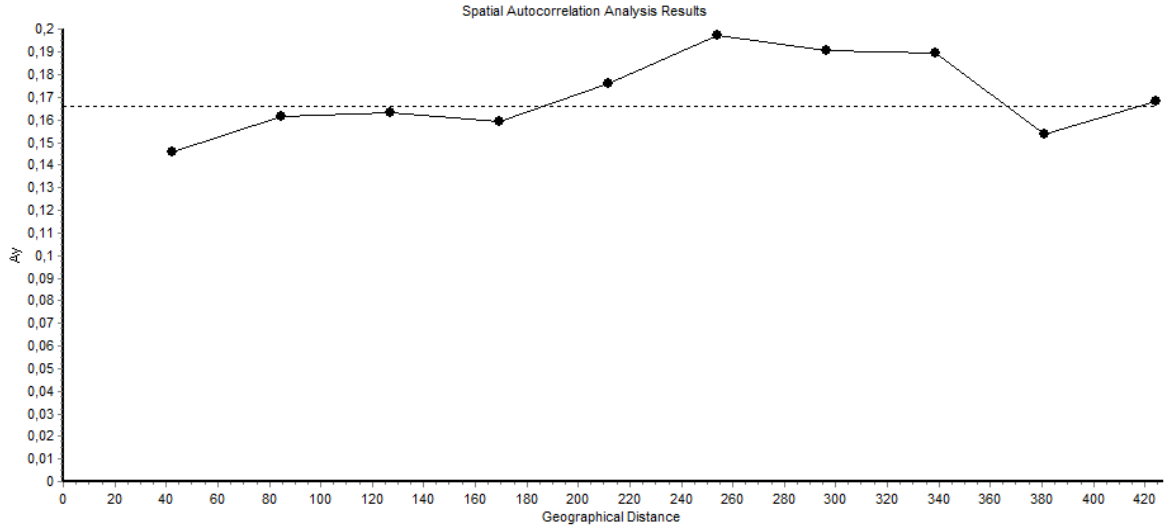



**Figure A2:** Spatial Autocorrelation Analysis of SGM and SGMx individuals from the whole sampling range. The averaged pairwise genetic distance accounted for Ay = 0.169 and the covariance for V = 0.017 with an significance of p = 0.036