Analysis of genetic diversity of two African leafy vegetables, spider plant (*Cleome gynandra*) and cowpea (*Vigna unguiculata*), reproductive characteristics and nutritional analysis of spider plant

Von der Naturwissenschaftlichen Fakultät der

Gottfried Wilhelm Leibniz Universität Hannover

zur Erlangung des Grades

Doktor der Gartenbauwissenschaften

(Dr. rer. hort.)

genehmigte Dissertation

von

Emmanuel Gordon Otunga Omondi, M. Sc.

2017

Referentin:	Prof. Dr. rer. hort. Traud Winkelmann
Korreferenten:	Prof. Dr. rer. nat. Thomas Debener
	Prof. Mary Abukutsa-Onyango
Tag der Promotion:	04.10.2017

Abstract

African leafy vegetables have been recognized for their important contribution in combating malnutrition and hunger in several communities in the sub-Saharan Africa and are also a source of income for the poor-resource farm families. The demand for spider plant (*Cleome gynandra*) and cowpea (*Vigna unguciulta* L. Walp) is unsatisfied due to the poorly adapted low yielding farmers' cultivars. Information on basic cytogenetic and reproductive properties, and the extent of genetic variability are prerequisites for genetic improvement. This study therefore, focused on the genetic diversity assessment and evaluation of the agronomic performance of spider plant and cowpea entries, analysis of the cytological basis and reproductive biology of spider plant, and analysis of nutritional content in spider plant. Thirty spider plant entries and fifteen cowpea entries from six African countries were used in this study. The entries consisted of advanced lines, gene bank entries from the WorldVeg, Arusha, Tanzania and farmers' cultivars collected directly from the farmers in Kenya

Amplified fragment length polymorphism (AFLP) and simple sequence repeat (SSR) markers were used. The SSR markers for spider plant analysis were designed from Illumina MiSeq sequence data from one of the entries, HTT-Select, while those of cowpea were obtained from previous publications. For the final genetic diversity analysis, combined matrix data for 499 polymorphic bands of the 11 AFLP and 9 SSR markers for spider plant and 544 AFLP and 18 SSR markers for cowpea were used giving genetic distances of 0.13 - 0.77 and 0.002 - 0.193 for spider plant and cowpea, respectively. The SSRs revealed heterozygosity values of 0.60 and 0.037 for spider plant and cowpeas respectively. Entries for both crops could not be fully resolved neither according to country of origin. Farmers' cultivars however grouped apart from the advanced lines and gene bank entries. An assignment of most of the entries into entry specific clades was possible for cowpea but not for spider plant. The agronomic traits evaluated in the field in Kenya showed significant differences for some of the traits for both crops. In spider plant, no or only low levels of self-incompatibility were determined from the seed set analysis and pollen tube growth observed on the stigma and in the style under a fluorescence microscope after self- and cross pollination in the greenhouse. A further analysis of pollen germinability *in vitro* showed at least 60% pollen germinability for all the entries. Ploidy determination and the genome sizes estimation of the spider plant entries by flow cytometry revealed that the entries were all diploid with a relatively small genome size of approximately 1.19 pg/1C. A chromosome count in root tips of one of the entries (GS-Sel) showed a chromosome number of 2n=34.

Glycosides of quercetin, kaempferol and isorhamnetin were identified as the main flavonoids in the leaves and flower using high performance liquid chromatography (HPLC) after methanol extraction. Glucosinolates were determined as desulfo-glucosinolates and separated using ultra high pressure liquid chromatography (UHPLC). The aliphatic glucosinolate 3-hydroxypropyl glucosinolate was the main glucosinolate in all the plant organs analysed. High mineral concentrations in the leaf tissue

were observed using an inductively coupled spectrophotometer (ICP-OES) especially for potassium, calcium, magnesium, phosphorus, iron, manganese and zinc.
Farmers' cultivars could be considered as a source of genetic variation for both spider plant and cowpea. The nutrient analysis of the spider plant could be used to promote spider plant consumption among the sub-Saharan population for improved health. Further single plant analyses are necessary to identify the variations between the genotypes.

Keywords: African leafy vegetables, *Cleome gynandra,* genetic diversity, minerals, secondary metabolites, *Vigna unguiculata*.

Zusammenfassung

Einheimischen afrikanischen Blattgemüsearten wird zunehmend Beachtung geschenkt, weil sie einen wichtigen Beitrag zur Vermeidung von Mangelernährung in vielen Regionen Afrikas südlich der Sahara leisten und eine wichtige Einkommensquelle für arme Kleinbauern darstellen. Die Nachfrage nach Spinnenpflanzen (*Cleome gynandra* L.) und Kuhbohnen (*Vigna unguiculata* L. Walp) als Blattgemüse kann oft nicht befriedigt werden, weil nur wenig entwickelte lokale Bauernrassen mit geringen Erträgen verwendet werden. Informationen zu zytogenetischen und Reproduktionseigenschaften und zum Ausmaß genetischer Variabilität sind wichtige Voraussetzungen für eine züchterische Verbesserung. In dieser Arbeit wurden deshalb Untersuchungen zur genetischen Diversität sowie zur agronomischen Leistung verschiedener *C. gynandra* und *V. unguiculata* Akzessionen sowie zu zytologischen und reproduktionsbiologischen Grundlagen von *C. gynandra* vorgenommen. Dazu fanden 30 *C. gynandra* und 15 *V. unguiculata* Akzessionen Verwendung, darunter entwickelte Linien, Genbankakzessionen vom World Vegetable Center (Arusha, Tansania) und lokale Rassen, die direkt bei Bauern in Kenia gesammelt worden waren.

Amplified Fragment Length Polymorphism (AFLP) und Simple Sequence Repeat (SSR) Marker wurden für die Bestimmung der genetischen Diversität verwendet. Die SSR Marker für *C. gynandra* wurden von Illumina MiSeq Sequenzdaten einer Akzession (HTT-Select) abgeleitet, während sie für *V. unguiculata* aus der Literatur übernommen wurden. Daten für 499 polymorphe Banden von 11 AFLP und 9 SSR Markern für *C. gynandra* und 544 AFLP sowie 18 SSR Markerbanden für *V. unguiculata* gingen in die Auswertung ein und resultierten in genetischen Distanzen von 0,13 – 0,77 für *C. gynandra* und 0,002 – 0,193 für *V. unguiculata*. Die SSR Marker ergaben Heterozygotiewerte von 0,60 für *C. gynandra* und 0,037 für *V. unguiculata.* Die Akzessionen konnten bei beiden Arten nicht vollständig getrennt werden, auch nicht nach ihren Herkunftsländern, lediglich die Bauernrassen bildeten eine von den entwickelten Linien und Genbankakzessionen getrennte Gruppe. Im Gegensatz zu *C. gynandra* fielen bei *V. unguiculata* die meisten Akzessionen in eigene Kladen. Für einige agronomische Eigenschaften, die in einem Feldversuch in Kenia erfasst wurden, wurden bei beiden Arten signifikante Unterschiede festgestellt.

Aus Beobachtungen zum Samenansatz und zum Pollenschlauchwachstum nach Selbstbestäubungen und Kreuzungen im Gewächshaus bei *C. gynandra* wurde abgeleitet, dass keine oder nur gering ausgeprägte Selbstinkompatibilität vorlag. Die In-vitro-Keimfähigkeit des Pollens lag bei allen Akzessionen über 60 %. Die Ploidiestufe und die Genomgröße wurden durchflusscytometrisch bestimmt und mit dem Errgebnis, dass alle *C. gynandra* Akzessionen diploid waren mit einem relativ geringen DNA-Gehalt von 1.19 pg/1C. In Wurzelspitzen einer Akzession (GS-Sel) wurde eine Chromosomenzahl von 2n=34 ermittelt.

Mittels High Performance Liquid Chromatography (HPLC) nach Methanolextraktion wurden Glycoside von Quercetin, Kämpferol und Isorhamnetin als Hauptflavonoide in Blüten und Blättern von *C. gynandra* identifiziert. Glucosinolate wurden als Desulfo-Glucosinolate nach Auftrennung über Ultra High Pressure Liquid Chromatography (UHPLC) bestimmt: das aliphatische 3-Hydroxypropyl Glucosinolat wurde als Hauptglucosinolat in allen Pflanzenorganen nachgewiesen. In Blättern wurden zudem hohe Mineralstoffgehalte, vor allem für Kalium, Kalzium, Phosphor, Eisen, Mangan und Zink, über Inductively Coupled Spectrophotometry (ICP-OES) ermittelt.
Bauernrassen können als wichtige Quelle genetischer Variabilität sowohl für *C. gynandra* als auch *V. unguiculata* angesehen werden. Die Analysen der wertgebenden Inhaltsstoffe bei *C. gynandra* bestätigen den ernährungsphysiologischen Wert dieser Pflanze für die Bevölkerung in Subsahara-Afrika.

Schlagwörter: Afrikanische Blattgemüse, *Cleome gynandra,* genetische Diversität, Mineralstoffe, Sekundärmetabolite, *Vigna unguiculata*.

Table of Contents

Abbreviations

AFLP	Amplified Fragment Length Polymorphisms
ALV	African leafy vegetables
ANOVA	Analysis of variance
AVRDC	Asian Vegetable Research and Development Center
BS	Bundle sheath
cpSSR	chloroplast simple sequence repeats
CTA	Technical Centre for Agricultural and Rural Cooperation (CTA)
CV	Coefficient of variation
CVA	Conical variates
DNA	Deoxyribonucleic acid
DPPH	1,1-diphenyl-2-picrylhydrazyl
DW	Dry weight
FAO	Food and Agriculture Organization
GA	Gibberellic acid
GS	Glucosinolate
HCDA	Horticultural Crops Development Authority
HORTINLEA	Horticultural Innovation and Learning for Improved Nutrition and Livelihood in East Africa
HPLC	High-performance liquid chromatography
IFAD	International Fund for Agricultural Development
IRD	Infrared
ISSR	Inter-simple sequence repeat
ITS	Internal transcribed spacers
JKUAT	Jomo Kenyatta University of Agriculture and Technology
ML	Maximum likelihood
MS	Mass spectroscopy
mtSSR	mitochondrial simple sequence repeats
NJ	Neighbor joining
NO	Nitric oxide
nuSSR	nuclear simple sequence repeats
PCA	Principal component analysis
PCoA	Principal coordinate analysis
PCR	Polymerase chain reaction
PIC	Polymorphism information content
RAPD	Random Amplification of Polymorphic DNA
RFLP	Restriction Fragment Length Polymorphism
RP	Resolving power
SD	Standard deviation
SNP	Single nucleotide polymorphisms
SSR	Simple sequence repeats
UHPLC-DAD	ultra HPLC-diode array detector
UV	Ultraviolet rays
UVB	Ultraviolet B rays
WFP	World Food Programme
WHO	World health organization

1. Introduction

1.1. African leafy vegetables

African leafy vegetables (ALVs) have been used in the sub-Saharan Africa (SSA) as food and traditional medicine for generations(Smith & Eyzaguirre, 2007). The most important are *Solanum* spp., *Amaranthus* spp., and *Corchorus* genera, spider plant (*Cleome gynandra*), *Brassica carinata* and cowpea (*Vigna unguciulata*), but others such as pumpkins and cucurbits are also included (Abukutsa-Onyango, 2007; Van Rensburg et al., 2004). They are indigenous to SSA and are well adapted to the climatic conditions in these regions. They have also received little research attention at least until recently but a resurgence of interest during the past years is evident due to the growing awareness about the health promoting and protecting property of the bioactive compounds found in the ALVs (Smith and Eyzaguirre, 2007). Besides, they are also rich in minerals and proteins(Abukutsa-Onyango, 2007; Odhav et al., 2007) and have started to be also studied for their secondary metabolites with antioxidant effect which are important for prevention of chronic diseases (Mibei et al., 2012; Neugart et al., 2017).

The ALVs were initially considered as poor man's food and generally collected from the wild in lean seasons of main crop failure and drought. In fact, the frequency of consumption dropped over the years since they were considered inferior in taste and nutrition compared to the exotic vegetables (Weinberger and Msuya, 2004). Cultivation of the crops is progressively increasing beyond kitchen garden to a large scale production for the larger markets in the urban and peri-urban areas (Ndenga et al., 2013; Gevorgyan et al., 2016; Vorster et al., 2008). Species preference is also influenced by gender, age, cultural background and geographical location (Van Rensburg et al., 2004; Kimiywe et al., 2007). The likelihood of adoption of the ALVs was reported to decreases with age (Krause et al., 2017). One reason for this could be that older persons have more knowledge about AIVs(Modi et al, 2006). Cultivation area of the ALVs has increased significantly, for instance, in Kenya the area under ALVs increased by 6% while the yields and value increased by 6% and 10%, respectively, in 2014(HCDA, 2014). The demand in Nairobi alone for instance was estimated at 3,600 - 4,500 tons of ALVs which translates to an income of about Ksh 45 million (Muhanji et al., 2011). These vegetables therefore, hold tremendous potential for improving food security, nutritional intake and economic welfare in sub-Saharan Africa where under-nutrition and under-employment is prevalent (FAO, IFAD, & WFP, 2015).

The production of the ALVs are still faced by a number of constraints which call for urgent research investment. These constraints include seed availability, variability in seed quality, lack of seed selection for uniformity of desired traits, plant pests and diseases, marketing, processing and storage (Abukutsa-Onyango, 2007; Smith and Eyzaguirre, 2007; Vorster et al., 2008).

1.2. Project cooperation

Horticultural innovation and learning for improved nutrition and livelihood in East Africa (HORTINLEA)

This study has been conducted as part of a cooperation in a multi-disciplinary project HORTINLEA (http://www.hortinlea.org) among several Kenyan and German institutions together with the World Vegetable Center, Eastern and Central Africa based in Arusha, Tanzania. The project was financed by the German Federal Ministry of Education and Research (BMBF) and the German Federal Ministry of Economic Cooperation and Development within the programme "GlobE- Global food security" and national research strategy BioEconomy 2030. The main focus of the project was to strengthen the production of African indigenous vegetables (AIVs), which has a potential to address pressing challenges of under-nutrition, poverty and sustainability among vulnerable people in rural and urban areas of Kenya. The species under investigation include spider plant (*Cleome gynandra*), cowpea (*Vigna unguciulata*), *Amaranthus spp*., Ethiopian kale (*Brasicca carinata*) and African nightshade (*Solanum scabrum*).

This dissertation focuses on spider plant (*Cleome gynandra*) and cowpea (*Vigna unguciulata*). The cowpea investigation has formed the master thesis of Max Menssen. Cowpea is an autogamous species and has also been well studied before compared to spider plant.

1.3. Spider plant (*Cleome gynandra* (L) Briq.)

1.3.1. Taxonomy and botanical description

Spider plant (*Cleome gynandra*L.) is also known as cat's whiskers, spider flower, spider wisp, African cabbage or *Gynandropsis gynandra* L. Briq. Other common names exist for different dialects in African countries for the communities that grow spider plant e.g. Kenya - mkabili, dek, Chinsaga; Malawi - Brede caya, pissat des chiens; Tanzania - ekeyo, eshogi, ekaboi; South Africa - Tamaleika, akaki, ziri; Germany - Senfkapper, Benzoinbaum, Fieberstrauch (Chweya and Mnzava, 1997). It is a herbaceous annual plant belonging to the botanical family Cleomaceaeas has been suggested by phylogenetic studies (Hall et al., 2004) and is now widely accepted as opposed to the earlier traditional classification in subfamily Cleomoideae of Capparaceae. The Cleomaceae is related to Brassicaceae family to which the model plant *Arabidopsis thaliana* and other brassica vegetables belong. *Cleome* is the largest genus in Cleomaceae according to Hall et al., (2002) with approximately 200-250 species and about 50 occurring in Africa.

Spider plant is erect and grows up to about 150 cm tall, is strongly branched, has a long tap root and a few secondary roots with root hairs. The stems and leaf petioles are thickly haired but glabrous in very few occasions, the stem and petiole colourations vary from green to purple**,** the leaf always has 3-7 leaflets with 5 leaflets being common (Chweya and Mnzava, 1997; Masuka et al., 2012; Wasonga et al., 2015) **(Figure 1A)**. Cleome flowers form a terminate raceme. The flowers are white but

sometimes tinged with purple **(Figure 1B)**. The flower may have long or short stamens, and long or short gynophores making Cleome flowers resemble a 'spider', hence its common name is 'spider flower'. Spider plant flowering stops further leaf formation, usually after about three to five weeks after planting (Chweya and Mnzava, 1997).The fruits are long, slender, green or yellow (when ripe)siliques with seeds inside **(Figure 1C)**. The fruits occur in zones at the inflorescence stalk, alternating with non-fruited sterile zones, hence it's other common name is 'cat's whiskers' (Chweya and Mnzava, 1997).The seeds are small, black, round resembling the shell of a snail with a rough surface **(Figure 1D)**.

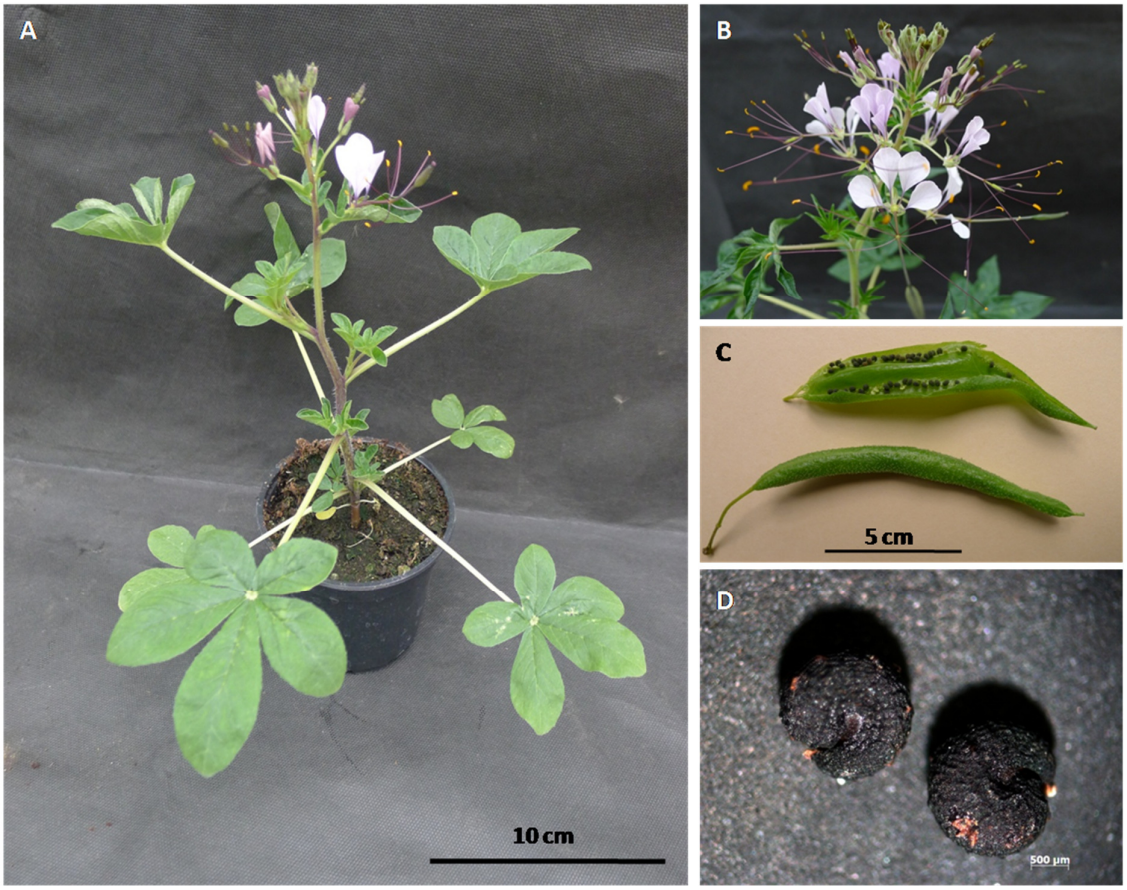


Figure 1:Spider plant morphology A- plant starting to flower after four weeks in the greenhouse; B-Inflorescence; C-Siliques; D-seeds.

The spider plant flower is usually hermaphroditic but occasionally staminate or pistillate **(Figure 2A),** a phenomenon that has also been observed in a related specie *Cleome viscosa* L. (Saroop and Kaul, 2015).The staminate or pistillate flowers usually don't set seed. Only flowers with elevated stigma set seed upon pollination by fertile pollen grains. In the field, pollination is predominantly carried out by bees. Buds with unelongated gynophores lack normal seeds in the ovary. Different developmental stages of the flower till anthesis are shown in **Figure 2B**.

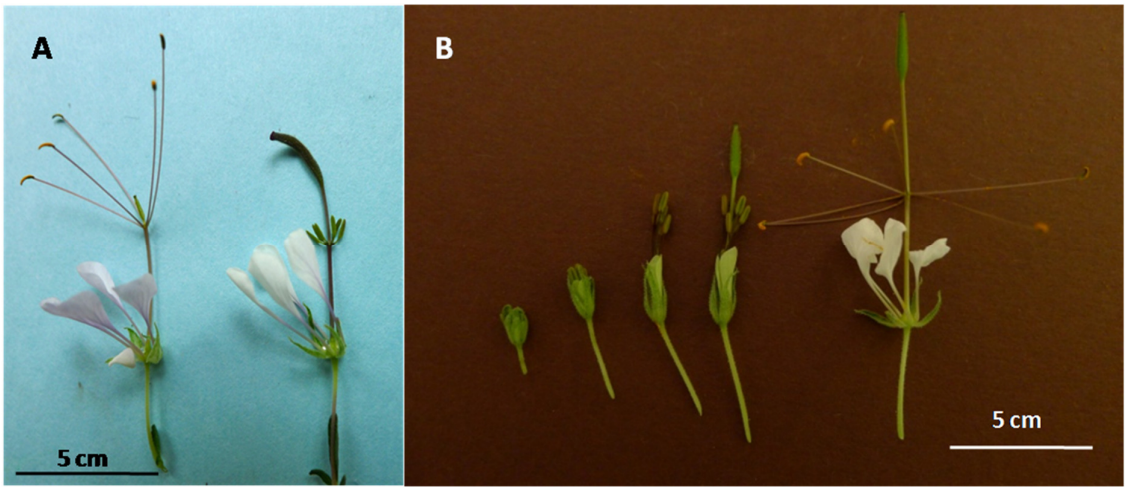


Figure 2: Staminate flower (left) and pistillate flower (right) (A); developmental stages of a spider plant flower (B).

Gibberellic acid (GA_3), dark treatment and temperatures ranging between 20° and 30° C may improve the germination of the spider plant seed(Motsa et al., 2015; Sowunmi and Afolayan, 2015a; Ochuodho and Modi, 2006). Pre-heating the seeds at 40° C for 1-5 days was found effective for breaking the dormancy after studying the effects of different treatments including pre-chilling, soaking, treatment with various concentrations of GA_3, KNO_3, leaching and pre-heating at various temperatures. Viable seeds take up to approximately 5 days to germinate (Ekpong et al., 2009).

Cleome gynandra is a C_4 plant as has been demonstrated since it possesses traits that are associated with this type of photosynthesis such as increased venation in the leaves, large bundle sheath (BS) cells, proliferation of the mitochondria and chloroplasts, and accumulation of transcripts and proteins needed for C_4 photosynthesis (Marshall et al., 2007).The genus *Cleome,* generally, provides a potential model for the study of the progression from C_3 to C_4 photosynthesis since some of the C_3 species have shown some of the traits associated with C4 photosynthesis (Marshall et al., 2007). This is enables spider plant to survive in drier and hot environment and explains it's spread in the sub-tropical regions.

Chweya& Mnzava, (1997) reviewed chromosome numbers of 2n=18, 20, 22, 32 and 34 for samples of the genus *Cleome* from Asia and South America. An ancestral chromosome base number of x=10 is proposed in the genus Cleome. *Cleome gynandra* is assumed to have resulted from the duplication of the chromosome base number to x=20 and a subsequent karyotype reduction to x=16-17 (Inda et al., 2008).

1.3.2. Distribution and cultivation.

Spider plant (*Cleome gynandra*) is believed to have originated in tropical Africa and Southeast Asia and later widely distributed in the warm tropical and subtropical regions where they were regarded as weeds but have now been domesticated (Chweya and Mnzava, 1997). In Africa, it is widespread mostly in Eastern (Kenya, Tanzania, Uganda, Rwanda) and Southern Africa (South Africa, Malawi,

Namibia and Zimbabwe) (Masuka et al., 2012; Wasonga et al., 2015)with some occurrences in West African (Kiebre et al., 2015).

Spider plant grows well within an altitude range of sea level to about 2400 meters with warm temperatures ranging from about 18° to 25° C. The plant can survive water shortage and hence can thrive well in areas of short rains, however prolonged drought stress may hasten flowering and senescing thereby reducing yield and quality (Chweya and Mnzava 1997). It grows fast with the seeds germinating within a week and harvesting may begin six weeks after sowing. The crop is adapted to many soil types but grows especially in well drained sandy loam or clay loam soils (Chweya and Mnzava, 1997). Spider plant is seed propagated. The seeds are usually sown directly into the soil by drilling holes or broadcasting on the soil surface. Shallow planting holes are recommended to reduce emergence failure and seed rot under the soil. Thinning after about three weeks is necessary for proper sharing of nutrient for excellent growth and yield. Nitrogenous fertilizers have been applied for good yield although well decomposed farm yard or compost manure are commonly used by farmers or in small kitchen gardens in the urban areas (Abukutsa-Onyango, 2007). The farm yard and compost manure give better yields compared to the fertilizers since they additionally improve the water holding capacity of the soil, cation exchange and soil structure. Twenty kilograms of farmyard manure per square meter is recommended or 200 g diammonium phosphate fertilizer per m^2 is recommended at planting (Chweya and Nzava, 1997; Mauyo et al., 2008).

Harvesting of spider plant is by topping, cutting at the ground level or uprooting the whole plant, but picking off individual leaves or leafy branches periodically until the plant senesces is the most common. The picking period may last for several weeks (Chweya and Nzava, 1997). In several cases the crop has been cultivated by intercropping with other main crops such as maize.

1.3.3. Uses and economic importance

Spider plant has been part of the traditional food systems in the sub-Saharan Africa for many generations and thus is considered to be a traditional leafy vegetable. Mainly leaves are eaten but sometimes also the whole shoots or even the flowers. Several studies have indicated that spider plant leaves are highly nutritious and rich in vitamins especially A and C, and mineral elements especially calcium and iron (Jinazali et al, 2017;van Jaarsveld et al., 2014; Odhav et al., 2007; Nesamvuni et al., 2001). Vitamin C is reported to be significantly higher in mature leaves than in younger leaves of spider plant (Ayua et al., 2016). Spider plant is usually eaten after brief cooking but also uncooked dried grinded leaves are incorporated in weaning food. Sometimes they are mixed with other vegetables or cooked by adding some milk to neutralize its bitter taste (Abukutsa-Onyango, 2007; Chweya and Mnzava, 1997). Spider plant is also forage to herbivores in the wild grazing land.

Besides being food, spider plant has also been used as traditional medicine over generations to treat different ailments among local communities. Dried leaves are used to make concoction for ailments and conditions (Chinsembu, 2016; Shanmugam et al., 2012; Chweya and Mnzava, 1997).

It contains glucosinolates and their hydrolysis products which are attractants to specialists pests on cruciferous plant and hence spider plant has been used as a trap crop (Zedler et al., 2016). Its volatile compounds are also repellants to cattle tick larvae and diamond back moths larvae attacking cabbages (Chweya and Mnzava, 1997; Nyalala and Grout, 2007; Nyalala et al., 2011). Spider plant also has a potential as a biofumigant against weeds during lawn establishment (Chongori et al., 2016).*Cleome gynandra* extracts is also used as an antifungal for the treatment of *Tinea capitis* (Imanirampa and Alele, 2016).

Small scale cultivation of the spider plant as intercrops for small markets for sale hence a source of income. Most of the cultivation is done by rural women who also allocate most of the land for the ALV cultivation compared to men (Dinssa et al., 2016; Ndengwa et al., 2013).

1.3.4. Pests and diseases

Some of the pests and diseases in spider plant have been reported from unpublished data. The fungal diseases in spider plant include powdery mildew (*Sphaerotheca fuliginea*, *Oidiopsis taurica*) and leaf spot (*Cercospora uramensis*). The crop is also attacked by Cabbage aphid (*Brevicoryne brassicae*) which causes stunted growth and wrinkling of the leaves and growing tips. The aphid also spreads virus diseases. The hurricane bug (*Bagrada* spp.) may similarly affect spider plant especially in dry periods (Chweya and Mnzava, 1997; Dinssa et al.,2016). This bug is however easily controlled by insecticides. Spider plant is also attacked by other insects such as flea beetles and other insects and worms harbored in the seed that consume the young seeds. The roots are also prone to root-knot nematodes (*Meloidogyne* spp.) attacks (Chweya and Mnzava, 1997; Dinssa et al.,2016). Young seeds are eaten by weaver birds (Chweya and Mnzava, 1997).

1.3.5. Breeding goals for spider plant

Generally, studies on various ALVs are now being undertaken to tap their health benefits and to improve these vegetables in terms of resilience to various stress conditions and to improve the yields. Inclusivity of vegetable producer target groups, commercial producers and small- holder farmers in breeding program has been identified as important. This kind of participatory plant breeding exploits both specifically and widely adapted genotypes as well as to target users. It further helps to maintain genetic diversity, stemmed from different areas preferring different genotypes, for sustainable production (Fufa et al. 2007).Concerns have also been raised about eliminating the traits such as disease resistance that make these vegetables so desirable in the first place (Cernansky, 2015).

According to Dinssa et al., (2016) specific breeding objectives would vary among the ALVs depending on environmental conditions and consumer demand and this might apply within the spider plant where different morphotypes occur. Some of the major breeding goals for spider plant include (i) yield since the existing farmers' cultivars show scanty leafing, (ii) time to maturity. Early maturity to first harvest and a long production cycle for repeated harvesting as important traits in traditional

vegetables (Adeniji and Aloyce, 2012; Keller, 2004; Weinberger and Msuya, 2004), (iii) Late flowering varieties of spider plant would be preferable for vegetable production since flowering halts leafing in spider plant, (iv) water use efficiency to enable production in water shortage conditions/ seasons, (v) pure line of uniform plant types in terms of colour of stems/leaf petioles to target different traditional preferences and ages with respect to bitterness, (vi) longer shelf-life. Leafy vegetables especially spider plant have short shelf-life that limits time available in marketing them(Dinssa et al., 2016; Gevorgyan et al., 2016) which calls for multidisciplinary approach between breeders and postharvest specialists to improve the shelf-life of the crop.

1.4. Cowpea (*Vigna unguciulata*)

Cowpea is a member of the family Fabaceae native to Africa (Padulosi and Ng, 1997). Two centers of primary domestication of cowpea, Zambezian region in East Africa and West Africa, were proposed by Ng and Marechal, (1985) and later supported by Baudoin and Marechal, (1985) who also suggested that South Eastern Asia could be a secondary center of diversity following the occurrences of cultigroups *Sesquipedalis* and *Biflora* in this region. Only a few samples of the wild might have been domesticated (Vaillancourt & Weeden, 1992) and the domesticated cowpea is characterized by large seeds and non-shattering pods compared to the wild cowpea (Kouam et al., 2012). Cowpea is mainly grown in tropical and sub-tropical areas including Africa, Asia, south America, Southern parts of Europe and the United States (Singh, 1997). It is a diploid plant with 2n = 2x = 22 chromosomes and a total nuclear DNA content of 1.27 pg/2C. The genome size of cowpea is stated with 620 Mbp (Chen et al., 2007). Cowpea mainly treated as highly self-pollinating plant owing to is flower structure (cleistogamous) and the simultaneous activities of pollen shedding and stigma receptivity (Ehlers and Hall, 1997) but outcrossing has also been observed in some subspecies (Kouam et al., 2012; Pasquet, 1999).

It is used as a vegetable crop as green or dry fodder. Dry seeds of cowpea are also used for food and contain 20-25% protein, 1.8% fat, 60.3% carbohydrate and are rich sources of iron and calcium (Timko & Singh, 2008). It's atmospheric nitrogen fixing ability is extremely valuable when it is cultivated with cereal crops in crop rotation system (Timko & Singh, 2008). Cowpea (*Vigna unguiculata*) is able to perform well even under low moisture conditions, moreover, its genome share a high degree of collinearity with other warm season legumes such as common bean (*Phaseolus vulgaris* L.) (Muñoz-Amatriaín et al., 2016; Vasconcelos et al., 2015).

The main breeding goals in cowpea are to develop lines with high grain yield potential, resistance to biotic stresses, tolerance to abiotic factors and adaptation to major production agro-ecologies (Boukar et al., 2016).

From over 15, 000 accessions of cultivated cowpea and more than 2000 wild relatives maintained at the International Institute of Tropical Agriculture (IITA) genetic resources center, about 20 breeding lines were released in approximately 10 countries between 2005 and 2015, many of the new varieties

having a combination of high grain yield and resistance to striga (*Striga hermonthica*) (Boukar et al., 2016).

1.5. Genetic diversity studies

Genetic diversity studies quantify the extent of genetic variation or relatedness within or between populations, species or individuals (Hughes et al., 2008). Genetic diversity studies have generated important information to facilitate efforts in germplasm conservation and guidance for better germplasm use during in crop improvement (Fu, 2015). Genetic diversity is key for providing genetic barriers against different biotic and abiotic stress factors that threaten survival of organisms in different environments (Hajjar et al., 2008; Hughes and Stachowicz, 2004; King and Lively, 2012).Genetic diversity is determined by factors such as domestication, selection, breeding among others (Christiansen et al., 2002; Rauf et al., 2010).

1.5.1. Genetic markers

The assessment of genetic diversity has applied the analysis of classical markers such as morphology, physiology, biochemical markers and the most recently used DNA or molecular marker (Govindaraj et al., 2015; Weising et al., 2005). The morphological markers are based on visually accessible traits, for instance, flower characteristics, growth habit and pigmentation, and seed traits. The morphological markers are cheap but on the other hand require large pieces of land for the experiments and are also influenced by environmental factors (Fu, 2015; Govindaraj et al., 2015). Biochemical markers are allelic variants of enzymes called isozymes that are detected by electrophoresis with specific staining (Govindaraj et al., 2015).DNA or molecular markers show variations at different positions of the chromosomes and they arise from deletions, duplications, inversions or insertions(Govindaraj et al., 2015; Hughes et al., 2008). They are broadly classified as PCR or non-PCR based. The restriction fragment length polymorphism (RFLP) is non-PCR based while amplified fragment length polymorphism (AFLP), single nucleotide sequence repeats (SSR) and random amplified sequence repeats (RAPD) require a PCR reaction and are much more informative than RFLP. The robust informative AFLP and SSR makers were used in this study.

1.5.1.1. Amplified fragment length polymorphism (AFLP)

Amplified fragment length polymorphism (Vos et al., 1995) is based on the selective PCR amplification of restricted fragments from genomic DNA. Three steps involved are digestion of the genomic DNA and ligation of nucleotide adapters, selective amplification of the digested fragments and finally the gel analysis of the amplified fragments **(Figure 3)**. Polymorphisms between genotypes shown by AFLPs may arise either from sequence variation in one or both restriction sites flanking a particular fragment, insertions or deletions within an amplified fragment or the differences in the nucleotide sequences immediately adjacent to the restriction sites (Weising et al., 2005).

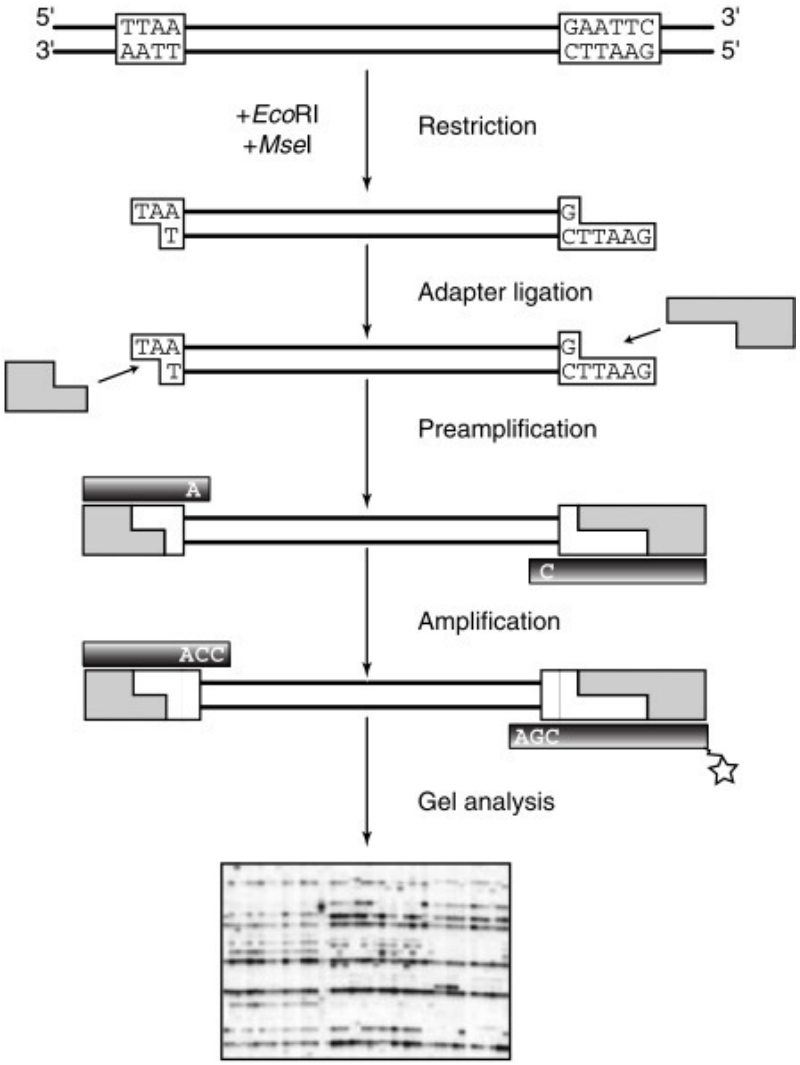


Figure 3: Amplified fragment length polymorphism mechanism (Weising et al., 2005)**.**
(i) Genomic DNA is digested with two restriction enzymes (here *EcoRI* and *MseI*) followed by specific or non-specific adapters ligation to fragments **(ii)** Two successive PCRs are then performed using specific or non-specific primer pairs and usually one of the primers is labeled by a radioisotope or a fluorochrome (indicated by a star). **(iii)** Amplification products obtained by the second, selective PCR are separated on sequencing gels.

The AFLP is a dominant marker, laborious due to the many steps to produce results, costly due to additional chemicals required and is more sensitive to the DNA quality and quantity. However, it possesses several advantages such as no prior sequencing information is required, it is able to simultaneously screen DNA regions that are randomly distributed over the whole genome and is reliable and reproducible(Semagn, et al., 2006; Mueller and Wolfenbarger, 1999).

1.5.1.2. Simple sequence repeats (SSR)

Simple sequence repeats (SSRs), also referred to as microsatellites, are tandem repeats of short DNA sequence motifs. They were first studied in humans (Hamada et al, 1982; Weber & May, 1989) but

have also now been found in other eukaryotes including plants and animals (Estoup et al., 1993; Morgante and Olivieri, 1993)and prokaryotes (Van Belkum et al., 1998; Field and Wills, 1996; Gur-Arie et al., 2000). They are may be classified based on their sizes (number of nucleotide per repeat unit), the nature of the repeat unit or their position in the genome (Miah et al., 2013). Based on the sizes they may be classified as mono-, di-, tri-, etc (Kalia et al., 2011).Based on the arrangement of the repeat units within the motif, they may be referred to as perfect, imperfect or compound according to Weber (1989) while Wang et al. (2009) coined the terms simple perfect, simple imperfect, compound perfect and compound imperfect. While perfect repeats are continuous repeats of single motifs, imperfect repeats are perfect repeats interrupted by non-repeat sequences. Compound repeats consists of two basic repeat motifs in different configurations. Based on the SSRs location, they may be referred to as nuclear (nuSSRs), mitochondrial (mtSSRs) or chloroplastic (cpSSRs) (Kalia et al., 2011). The illustrations of these classification is shown in **Table 1**.

Table 1: Classification of simple sequence repeats

Based on the nucleotides arrangement in the repeat motifs
Pure or perfect or simple perfect (CA)*n* Simple imperfect (AAC)*n* ACT (AAC)*n* + 1
Compound or simple compound (CA)*n* (GA)*n*
Interrupted or imperfect or compound imperfect (CCA)*n* TT (CGA)*n* + 1
Based on the number of nucleotides per repeat
Mononucleotide (A)*n*
Dinucleotide (CA)*n*
Trinucleotide (CGT)*n*
Tetranucleotide (CAGA)*n*
Pentanucleotide (AAATT)*n*
Hexanucleotide (CTTTAA)*n* (*n* = number of variables)
Based on location of SSRs in the genome
Nuclear (nuSSRs)
Chloroplastic (cpSSRs)
Mitochondrial (mtSSRs)

Adopted from Miah et al.(2013) with some modifications.

Simple sequence repeats are distributed throughout the genome and are highly polymorphic (Tautz and Renz, 1984). They are also inherited in a codominant manner and because most of the SSRs are resident in the noncoding DNA, they may be assumed to be selectively neutral (Weising et al, 2005). SSRs allow unambiguous allele assignment and have a high degree of reproducibility. They also require low amounts of template DNA that does not need to be of high quality. The major drawbacks of these markers is the time and cost of developing the species specific primers that demand sequence information.

Simple sequence repeats may be readily available from DNA sequences for a species under investigation in public databases or may be transferable from already developed microsatellites from a related species. They can also be developed by screening libraries of clones (McCouch et al., 1997). The mutations of the SSRs occur through strand slippages during DNA replication (Tautz and Renz, 1984). Slippage implicates mispairing of the newly replicated strand during the replication process

and most often involves gain or loss of a single repeat unit. As molecular markers, polymorphisms between the alleles in the tandem repeats are detected by PCR amplification of the nuclear or organellar microsatellite using flanking primer (sequences of the repeat-flanking regions the SSR). The PCR products are then separated on a denaturing polyacrylamide gels and visualized by autoradiography or stains such as ethidium bromide or silver. In this study, Infra-red dye (IRD 700 and IRD 800) were used the PCR products on denaturing polyacrylamide gels.

Further discussions of the working mechanisms, potentials and limits of the molecular markers are discussed here in **chapter 2.1**.

1.5.2. Evaluation of genetic diversity

The molecular marker data are finally evaluated using various statistical methods to translate it into a meaningful biological information such to assess the of genetic diversity or relatedness. A summary of the commonly used metrics to measure diversity among genotypes using molecular markers are defined in **Table 2**.

Table 2: Commonly used metrics for measuring genetic diversity.

Source of data	Metric of diversity	Definition
Molecular marker data	Allelic diversity (A)	The number of alleles per locus, averaged over all loci tested.
	Nucleotide diversity (π)	The average number of nucleotide differences per site between two random individuals selected from a population
	Allelic richness	Average number of alleles per locus
	Genotypic richness	The number of genotypes within a population
	Heterozygosity	The average proportion of loci that carry two different alleles at a single locus within an individual
	Mutational diversity and effective population size (θ)	A measure of nucleotide diversity that provides a combined measure of effective population size (Ne) and mutation rate (μ). Q is typically calculated using Watterson's estimator ($\theta = 4N_e\boldsymbol{\mu}$), which is equal to the expected number of segregating sites between two genotypes
	Percentage polymorphic loci	The percentage of loci that are polymorphic.
Continuous traits e.g morphological	Coefficient of genetic variance (CV)	Genetic variance in a trait (VG) corrected by the trait mean, calculated as ($VG^{0.5}/mean_{trait}$) x100%.
	Genetic variance (V_G)	The variance in a phenotypic trait among individuals due to genetic differences.
	Heritability	The ratio of the genetic variance to the total phenotypic variance in the population

(Fu, 2015; Hughes et al., 2008; Weising et al., 2005)

Allelic diversity (A) is sensitive to sample size since the larger the sample size, the higher the chances of detecting new (rare) alleles.

$$A_e = \frac{1}{\Sigma p_i^2}$$

Equation 1: Calculation of allelic diversity (A).
P_i is the frequency of each allele observed across all investigated loci

Allelic richness provides a more sensitive tool for detecting genetic bottlenecks with microsatellite markers than the more commonly used allelic evenness measures such as heterozygosity (Greenbaum et al., 2014)

Most of the statistical analysis are based on genetic distances and similarity indices (coefficients of similarity). The indices are calculated from band sharing data and are able to quantify the genetic variations between pairs of samples (pairwise similarity/ distance) which are afterward applied in multivariate analyses. Some of the commonly applied coefficients of similarity include:

Dice's coefficient

This is also commonly referred to as Nei and Li's coefficient(Nei and Li, 1979). It is calculated as:

$$S = \frac{2n_{ab}}{n_a + n_b}$$

Equation 2: Dice coefficient formula.
Where $\mathbf{n_a}$ and $\mathbf{n_b}$ represent the numbers of bands present in lanes a and b, respectively, and $\mathbf{n_{ab}}$ represents the number of bands shared by both lanes. **S** has a value between 0 and 1, where 0 means no bands in common, and 1 means patterns are identical.

Jaccard's coefficient

Jaccard's coefficient (Jaccard, 1901) is calculated as:

$$S_J = \frac{2n_{ab}}{n_a + n_b - n_{ab}}$$

Equation 3: Jaccard's coefficient.
Where $\mathbf{n_a}$ and $\mathbf{n_b}$ represent the numbers of bands present in lanes a and b, respectively, and $\mathbf{n_{ab}}$ represents the number of bands shared by both lanes. $\mathbf{S_J}$ has a value between 0 and 1, where 0 means no bands in common, and 1 means patterns are identical.

Simple matching coefficient

Simple matching coefficient is calculated as:

$$S_s = \frac{n_{ab} + n_{AB}}{N}$$

Equation 4: Simple matching coefficient.
where, $\mathbf{n_{ab}}$ represents the number of bands shared by both lanes, $\mathbf{n_{AB}}$ represents the total number of bands that are absent in both lanes a and b (but present in some other lanes), and N is the total number of bands (De Riek et al., 1999).

Generally, Dice's and Jaccard's are commonly used and have been reported to give the most consistent outcome (Maguire and Sedgley, 1997).

1.5.3. Ordination, clustering and dendrograms.

Visual techniques are employed to help interpret the genetic distances. They also help to better visualize the relationships among samples especially when a large number of samples are studied. These techniques simplify the data and also allow as many characters as possible to be used to differentiate the samples under study. One of the major applications is to help selecting a smaller core collection with preserved genetic diversity from a larger collection (Liu et al., 2015; El Bakkali et al., 2013; Shashidhara et al., 2003). Principal coordinate analysis (PCoA), principal component analysis (PCA) and conical variates (CVA) analysis are the most commonly used ordination methods. While the starting point for PCA is the data matrix derived from the presence or absence of all the bands, the data matrix for the PCO is derived from the distances (or similarities) between the samples (Weising et al., 2005).

Dendrograms can additionally show the phylogenetic (evolutionary history) relationships among samples in study. Three main methods for dendrogram construction include distance methods (Cluster analysis, or phenetic methods), parsimony methods and the maximum likelihood (ML) method. The distance based method starts from the pairwise distance data matrix from the raw data calculated using either the Dice or Jaccard algorithms. The distance method uses a step by step grouping according to the distances until the tree is completed. The resulting dendrogram reflects the phenetic similarities among the unit and is sometimes called a phenogram but does not necessarily reflect phylogenetic relationships. Neighbor-joining (NJ) (Saitou and Nei, 1987) and Unweighted Pair Group Method using Arithmetic average (UPGMA) distance matrix algorithms are frequently used in tree construction using the distance method. UPGMA assumes the same evolutionary rates along all the tree branches unlike NJ algorithm which assumes minimum evolution. Parsimony method (Saitou and Imanishi, 1989) is character based and does not involve distance calculations but rather reconstructs the phylogenetic patterns. It is used to further select and compare only those trees that explain the data set with the smallest number of changes.

Maximum likelihood (ML) method on the other hand aims at finding the tree with the largest probability that reflects the actual data set on the background of an evolutionary model. The estimation of the statistical support for the individual tree branches is done by various methods such as such as resampling with replacement (bootstrapping) (Felsenstein, 1985), or resampling without replacement (jackknifing and parsimony jackknifing) (Farris et al., 1996; Mort et al., 2000).

1.6. Objectives

Spider plant was identified as one of the major strategic ALV crops for breeding at AVRDC Eastern and Southern Africa after Amaranth (leafy type) and African eggplant (fruit type) during a stakeholder priority-setting exercise conducted in 2007/2008 (AVRDC, unpublished data) (Dinssa et al., 2016). The potential has not yet been fully exploited due to the little research attention it has received until recently. Seed system remain a challenge (Vorster et al., 2008). Seed is a key determinant of agricultural productivity in terms of quantity and quality of the output and quality seed has potential to double or even triple the yields of most crops, offering higher value and consumer appeal (CTA, 2014). Development of new varieties of will also broaden the consumers and farmers choice of selection (Cernansky, 2015).

Spider plant production has so far been based on poorly-adapted traditional farmers' cultivars (Dinssa et al., 2016). The farmers produced and saved their own seed from season to season and sold the surplus to other growers. Farmers in different localities grow local cultivars that differ from location to location which limit productivity due to different biotic and abiotic factors in these locations (Dinssa et al., 2016). Seed companies did not consider the production and marketing of traditional African vegetables seeds to be a profitable business and mostly concentrated on crops that sufficiently capture a large market size (Afari-Sefa et al., 2011). The production of seeds adapted to each locality is also expensive to seed companies. Insufficient characterization of indigenous vegetables at the morphological and especially at the molecular level causes severe difficulties for gene bank curators and breeders (Ngwediagi et al., 2009). Generally, the genetic diversity studies of various ALVs has been undertaken using molecular markers and morphological traits in the field. Most of the genetic diversity studies have covered cowpea, African nightshade and Amaranth in that order, whereas for spider plant, only a single study using random amplified polymorphic DNA (RAPD) molecular makers to assess genetic diversity of spider plant morphotypes from Western Kenya is reported (K'Opondo et al.,2009).

Information on basic physiology and reproduction, genetics and traits for adaptation to different agro-ecologies and various stresses are essential to help improve breeding efficiency. Characterization of germplasm collections for spider plant and cowpea will facilitate utilization by breeders for parental selection in breeding programs and allow for marker-trait associations, paving the way for using molecular breeding techniques such as marker-assisted selection. This information will also be important for germplasm conservation efforts because later on farmers are expected to opt for

improved varieties that are be able to produce better quality produce to the markets to replace poor-yielding local varieties. Analysis of the nutritional content of spider plant, especially the secondary metabolites, is necessary since little research based evidence on the profile and richness of secondary metabolites in the ALVs exists (Neugart et al., 2017).

To contribute to the overall aim of developing improved varieties and supply farmers with high-quality, affordable seed of spider plant and cowpea vegetable, the following objectives were set out:

1. To analyze the genetic diversity in selected spider plant entries and farmers' cultivars using morphological and molecular markers.
   i. Development and establishment of molecular markers (Amplified Fragment Length Polymorphism (AFLP) and Simple Sequence Repeat (SSR) markers) for genetic diversity studies
   ii. Analysis of genetic diversity within the spider plant and cowpea entries.
   iii. Evaluation of the spider plant and cowpea entries using morphological traits in the field.

2. To analyze the cytological basis and reproductive biology of spider plant.
   i. Determination of the ploidy levels and DNA content of the selected spider plant entries
   ii. Assessment of self-compatibility by *in vitro* pollen germinability, pollen tube development *in situ* stigma and seed set after pollination experiment.

3. To analyze the nutritional content of selected spider plant entries.
   i. Determination of mineral element concentrations in the leaves.
   ii. Identification of secondary metabolites (glucosinolates and flavonoids) in different organs of the spider plant.

The thesis is composed of

1. A review manuscript summarizing the use of molecular markers for genetic diversity studies in African leafy vegetables.

2. Mating biology, nuclear DNA content and genetic diversity in spider plant (*Cleome gynandra*) germplasm from various African countries.

3. Nutritional compound analysis and morphological characterization of spider plant (*Cleome gynandra*) germplasm from six African countries.

Further work involving assessment of genetic and morphological diversity in cowpea (*Vigna unguciulata*) germplasm is additionally presented in chapter 2 which was the outcome of a co-supervised M Sc. thesis of Max Menssen.

2. Publications and manuscripts

2.1. Molecular markers for genetic diversity studies in African leafy vegetables.

Emmanuel O. Omondi[1], Thomas Debener[2], Marcus Linde[2], Mary Abukutsa-Onyango[3], Fekadu F. Dinssa[4], Traud Winkelmann[1]

[1]Institute of Horticultural Production Systems, Woody Plant and Propagation Physiology, Leibniz Universität Hannover, Hannover, Germany

[2]Institute for Plant Genetics, Molecular Plant Breeding, Leibniz Universität Hannover, Hannover, Germany

[3]Department Horticulture, Jomo Kenyatta University of Agriculture & Technology (JKUAT), Juja, Kenya

[4]AVRDC - The World Vegetable Center, Eastern and Southern Africa, Arusha, Tanzania

Type of authorship	First author
Type of article	Review article
Contribution to the article	Wrote the paper
Contribution of other authors	Thomas Debener: Manuscript design and correction Marcus Linde: Manuscript correction Mary Abukutsa-Onyango: Manuscript correction Fekadu F. Dinssa: Manuscript correction. Traud Winkelmann: Revision and writing of parts of the manuscript.
Journal	Advances in Bioscience and Biotechnology
Date of publication	31 March 2016
DOI	10.4236/abb.2016.73017

Advances in Bioscience and Biotechnology, 2016, 7, 188-197
Published Online March 2016 in SciRes. http://www.scirp.org/journal/abb
http://dx.doi.org/10.4236/abb.2016.73017

Scientific Research Publishing

Molecular Markers for Genetic Diversity Studies in African Leafy Vegetables

Emmanuel O. Omondi[1], Thomas Debener[2], Marcus Linde[2], Mary Abukutsa-Onyango[3], Fekadu F. Dinssa[4], Traud Winkelmann[1]

[1]Institute of Horticultural Production Systems, Woody Plant and Propagation Physiology, Leibniz Universität Hannover, Hannover, Germany
[2]Institute for Plant Genetics, Molecular Plant Breeding, Leibniz Universität Hannover, Hannover, Germany
[3]Department Horticulture, Jomo Kenyatta University of Agriculture & Technology (JKUAT), Juja, Kenya
[4]AVRDC - The World Vegetable Center, Eastern and Southern Africa, Arusha, Tanzania
Email: traud.winkelmann@zier.uni-hannover.de

Received 22 February 2016; accepted 28 March 2016; published 31 March 2016

Abstract

African leafy vegetables are becoming important crops in tackling nutrition and food security in many parts of sub-Saharan Africa, since they provide important micronutrients and vitamins, and help resource-poor farm families bridge lean periods of food shortage. Genetic diversity studies are essential for crop improvement programmes as well as germplasm conservation efforts, and research on genetic diversity of these vegetables using molecular markers has been increasing over time. Diversity studies have evolved from the use of morphological and biochemical markers to molecular markers. Molecular markers provide valuable data, since they detect mostly selectively neutral variations at the DNA level. They are well established and their strengths and limitations have been described. New marker types are being developed from a combination of the strengths of the basic techniques to improve sensitivity, reproducibility, polymorphic information content, speed and cost. This review discusses the principles of some of the established molecular markers and their application to genetic diversity studies of African leafy vegetables with a main focus on the most common *Solanum*, *Amaranthus*, *Cleome* and *Vigna* species.

Keywords

AFLPs, Allozymes, *Amaranthus*, *Cleome*, ISSRs, Microsatellites, RAPDs, SNPs, *Solanum*, SSRs, *Vigna*

How to cite this paper: Omondi, E.O., Debener, T., Linde, M., Abukutsa-Onyango, M., Dinssa, F.F. and Winkelmann, T. (2016) Molecular Markers for Genetic Diversity Studies in African Leafy Vegetables. *Advances in Bioscience and Biotechnology*, **7**, 188-197. http://dx.doi.org/10.4236/abb.2016.73017

1. Introduction

African leafy vegetables include those native to Africa [1], as well as introduced vegetable crops that have been integrated into local food cultures and have become indigenized. They grow spontaneously in the wild, semi-domesticated or domesticated in home gardens. They are important emerging crops on the continent due to their contribution in nutrition and food security—they help people bridge food shortage seasons because of their earliness, robustness and easiness to cultivate. Moreover, from our experience, they are important sources of income for resource-poor people, mainly women, in rural and peri-urban areas in sub-Saharan Africa. In the past, these crops have received little research attention, but nowadays they are studied in various aspects, genetic diversity being one main research focus. In order to establish efficient and sustainable breeding programmes and to decide for sound conservation strategies, knowledge on the existing genetic diversity is essential. Genetic diversity can be detected by morphological markers, biochemical constituents (e.g. secondary metabolites), and/or macromolecules (proteins and deoxyribonucleic acids—DNA). Molecular markers detect mostly selectively neutral variations at the DNA level and are an indispensable tool for genetic diversity studies with a high level of precision and reproducibility.

Morphological traits have been used for diversity analyses in African leafy vegetables such as *Solanum* spp., *Vigna unguiculata*, *Amaranthus* spp., and *Cleome gynandra* [2]-[5]. The limitations of these traits result from the plasticity of certain traits and modifications caused by the environmental conditions. Moreover, morphological markers are limited in number to cover all genome regions of a plant species and are therefore less suitable to be used as markers. A correlation of morphological markers and molecular markers, however, can provide basic information to explain the genetics of phenotypic variation [6]. Other classical strategies for evaluation of genetic diversity such as isozymes have also been complemented by DNA markers due to their several drawbacks such as limited number of suitable loci in a given genome and the requirement of fresh tissue of high developmental stage as well as the subjection to variation due to the environment. Today, molecular markers are an indispensable tool for genetic diversity studies with a high level of precision and reproducibility. An ideal molecular marker should have the following qualities [7]:

1) Highly polymorphic and reproducible;
2) Adequate resolution of genetic differences e.g. high information content;
3) Simple, quick and inexpensive;
4) Needing small amounts of tissue and DNA;
5) Linkage to distinct phenotypes; and
6) Requiring no prior information about the genome of an organism.

Finding all these qualities in a single marker technique is nearly impossible. Hence, the selection of a marker or a combination of markers will depend on the type of study and the species. Well established DNA marker based techniques such as Restriction Fragment Length Polymorphisms (RFLPs), Random Amplified Polymorphic DNAs (RAPDs), Simple Sequence Repeats (SSRs), Inter-Simple Sequence Repeats (ISSRs), Amplified Fragment Length Polymorphisms (AFLPs) and Single Nucleotide Polymorphisms (SNPs) have been used in genetic diversity studies in African leafy vegetables. These markers differ in their genomic abundance, level of polymorphism, locus specificity, reproducibility, dominance or codominance, technical requirements and financial investment. In general, molecular markers are classified as either non-polymerase chain reaction (PCR) based markers, *i.e.* RFLPs, or the most commonly used PCR-based markers, *i.e.* RAPDs, AFLPs, SSRs, ISSRs and SNPs [8]. This review aims at discussing the principles of some of the established molecular markers and their application to genetic diversity studies of African leafy vegetables with a main focus on the most common *Solanum*, *Amaranthus*, *Cleome* and *Vigna* species.

2. Molecular Markers in Genetic Diversity Studies in African Leafy Vegetables

2.1. Allozymes

Allozymes are variants of enzymes which differ in one or few amino acids due to allelic differences of the encoding orthologous genes [7]. For allozyme analyses, proteins are extracted from plant tissues and separated by electrophoresis by their net charge, conformation and size. Mutations in the DNA may result in replacement of an amino acid, hence in modification of net electric charge and the overall shape (conformation) of the protein. These protein modifications affect the migration rate of the proteins in an electric field allowing allelic variation

detection by gel electrophoresis and subsequent enzyme-specific stains. These stains contain the respective substrate for the enzyme, co-factors and an oxidized salt (e.g. nitro-blue tetrazolium). Thus, the allozymes become visible as bands in the gel, and their numbers reflects the number of loci and alleles (homozygous or heterozygous), as well as in some cases the number of subunits of the protein that can be separated [7].

The allozyme technique is simple since it does not require DNA extraction or sequence information. However, in some species considerable effort for optimization for certain enzymes was noted. Allozymes can be applied at low costs, if no expensive enzyme staining reagents are required. Allozymes are codominant markers that are highly reproducible. Some limitations however, include a restricted number of suitable allozyme loci that can be used, the requirement of large amounts of fresh tissue and sometimes limited variation due to identical electrophoretic mobility even for distantly related germplasm [9].

Allozymes combined with RAPDs were used in analyses of genetic diversity in *Vigna luteola* and *V. marina* in Nigeria [10]. Seven out of thirteen tested allozyme loci were found to be polymorphic. Mainly variation among different populations was detected, whereas the genetic diversity within both species based on the allozyme analysis was low. Using RAPDs a higher resolution was obtained [10]. Other genetic diversity studies in the genus *Vigna* using allozymes include a collection of wild and cultivated cowpea (*V. unguiculata*) accessions of different parts of Africa [11]. The results of the study confirmed previous morphological classifications and allowed a clear separation of the different breeding systems (outcrossing and selfing) present in the species. A collection of wild cowpea populations from West Africa with focus on Ghana, Burkina Faso and Nigeria studied by Kouam *et al*. (2012) [12] using nine polymorphic allozyme loci revealed pronounced genetic differences between and low levels of genetic diversity within populations due to the prevailing inbred breeding system.

In *Amaranthus*, Di Renzo *et al*. (2000) [13] used seven enzyme systems to examine the genetic diversity of commercial cultivars and experimental strains representing seven different species from an amaranth breeding programme. Based on genetic diversity indices it was shown that 60% of the variation was interspecific while 40% was intraspecific.

2.2. Random Amplified Polymorphic DNAs (RAPDs)

RAPDs [14] involve the use of short random oligonucleotide primers to realize DNA polymorphisms produced by mutations at or length mutations between the primer sites in the genome. The key innovation in this technique was the use of an arbitrary primer that requires no prior sequence information of the genome. The RAPD primers, as all PCR primers, have to meet two criteria: a minimum GC content of 40% and the absence of a palindromic sequence (a base sequence that reads exactly the same from left to right and from right to left) as suggested by Williams *et al*. (1990) [14]. Polymorphisms in RAPD fragments are detected mainly as presence or absence of bands. Its simplicity and low cost are reasons for its wide application. However, the RAPD technique is highly sensitive to the reaction conditions, hence reproducibility between different laboratories is low [15].

Among and within six *Amaranthus* species from different regions of the Indo-Gangetic plains genetic diversity and relationships were analysed using RAPD markers [16]. High variation in the genetic diversity within the different amaranth species was found with similarity coefficients ranging between 0.16 and 0.97.

In cowpea, Nkongolo (2003) [17] characterised 38 Malawian landrace accessions using twenty RAPD primers and scored 143 DNA fragments. Their study revealed that the variation among accessions within geographic regions or groups accounted for 96% of the total molecular variance and that this was due to the uncontrolled gene flow among the populations. Ba *et al*. (2004) [18] characterised domesticated cowpea, and its wild progenitors represented by wild accessions obtained from West, East and Southern Africa using 202 RAPD marker bands. RAPD analysis showed wild accessions from East Africa, the proposed area of origin, to be more diverse within the species *V. unguiculata* var. *spontanea.* Moreover, the primitive cultivars expressed a higher genetic diversity than the more advanced ones. Zannou *et al*. (2008) [19] used RAPD markers to reveal genetic diversity in 70 cowpea accessions in Benin. Although being based on a rather low number of only 32 polymorphic bands, their study detected a high genetic diversity among the accessions, a confirmation of a previous more detailed study involving 120 primers by Mignouna *et al*. (1998) [20]. They suggested that the high variation may be an indication of different ancestry to the accessions studied. A combined approach using morphological traits, ISSR and RAPD markers by Ghalmi *et al*. (2010) [21] compared twenty cowpea landrace accessions in Algeria. The RAPD markers alone did not correlate with the morphological data to cluster the accessions, but when combined with the ISSR data a significant correlation was found. Malviya *et al*. (2012) [22]

used 18 sets of RAPD primers resulting in 148 polymorphic bands to analyse the genetic diversity among 10 Indian cultivars of cowpea. The detected relatively narrow genetic base was suggested to be due to a single domestication event occurring in the origin of this crop. The authors also concluded that seed conservation practices led to only limited exchange of germplasm throughout India and also did not integrate genotypes from foreign sources into local breeding programmes.

RAPDs and morphological markers were also used to study genetic variation and phylogenetic relationships of 12 accessions belonging to the section *Solanum* by Poczai *et al*. (2010) [5]. By 210 polymorphic RAPD markers and the morphological markers clustering of the accessions into similar groups was achieved, however, the morphological differences observed in some *S. scabrum* var. *scabrum* accessions were not correlated to the respective molecular data. In spider plant (*C. gynandra*), K'Opondo *et al*. (2009) [23] studied four morphotypes (defined by the coloration of stems and petioles) collected from small-scale farmers or wild grown plants in Kenya based on 31 polymorphic RAPD bands and could separate all four morphotypes.

2.3. Amplified Fragment Length Polymorphisms (AFLPs)

The AFLP technology [24] has received recognition as one of the most efficient markers currently available. It has the capacity to concurrently screen representative DNA regions distributed randomly throughout the genome. It combines the strength of RFLP with the flexibility of PCR-based technology by ligating adaptor cassettes with primer recognition sequences to the restricted DNA and selective PCR amplification of restriction fragments using a limited set of primers. AFLP generates fingerprints of any DNA and does not require prior knowledge of the DNA sequences. Reproducibility is also high in AFLP markers. Disadvantages include the need for pure, high molecular weight DNA, the need for dominant scoring of this marker type, and the possible non-homology of co-migrating fragments belonging to different loci [8]. In addition, due to the high number and different intensity of bands per primer combination, there is the need to adopt certain strict but subjectively determined criteria for acceptance of bands in the analysis. Detection of the AFLP fragments is done on denaturing polyacrylamide gels mostly using automated sequencers or automatic capillary sequencers, but silver stained gels can also be evaluated visually.

AFLPs have been used to study genetic diversity among several species of African leafy vegetables. In *Solanum* spp., Jacoby *et al*. (2003) [25] compared the genetic relationships of 14 genotypes mostly of the *S. nigrum* complex established by morphological markers to those obtained in AFLP analyses including 222 polymorphic markers. Both methods separated the different genotypes and clustered them into similar groups. Dehmer and Hammer (2004) [26] employed two AFLP primer combinations to characterise the genetic diversity present in a collection of 44 accessions of the *S. nigrum* L. complex in the Gatersleben genebank in Germany. They were able to classify taxonomically unknown material, and to correlate the clustering of the examined accessions with their geographic origin. Moreover, different levels of genetic diversity were detected in the four identified groups with higher infraspecific variation within *S. americanum* than interspecific variation of the remaining species. In order to gain new insight into the taxonomically difficult *S. nigrum* complex, Olet *et al*. (2011) [27] assessed the genetic relationships among 107 accessions (90 collected in Uganda and 17 from a genebank) representing eight *Solanum* species using AFLPs. Although out of 510 AFLP bands only 31 were polymorphic, Olet *et al*. (2011) [27] were able to conclude that the accessions analysed represented only five species and that most species must have been introduced to Uganda.

Cowpea genetic diversity has also been examined in several studies using AFLPs. Coulibaly *et al*. (2002) [28] evaluated genetic relationships within a total of 117 cowpea accessions, including both domesticated and wild forms from eastern and western Africa. They found AFLPs to be superior to allozymes that had been applied in earlier studies. Higher genetic diversity in wild than in cultivated accessions was observed and the higher diversity of wild materials from eastern Africa supported the ideas of the origin of the species in eastern Africa. Moreover, the AFLP data were in line with a proposed unique domestication event in northern Africa [28]. By three AFLP primer combinations resulting in 253 bands in total, the different subspecies of *V. unguiculata* could not be fully resolved due to pronounced intra-accession variability [29]. Fang *et al*. (2007) [30] also examined genetic relationships among 60 advanced breeding lines of cowpea from different breeding programmes and 27 landrace accessions of different origins by scoring 382 AFLP bands in total (207 polymorphic). Despite the diverse origin of the materials analysed very high genetic similarity was observed. Since the accessions clearly clustered according to the breeding programmes an exchange of materials especially with West African breeding

programmes was recommended by the authors.

Dendrograms based on AFLP (and ISSR) analyses of 30 accessions and cultivars of the *Amaranthus* species revealed a clear assignment of all taxa to different cluster groups, which was not possible by sequence comparisons due to low sequence variation in the ITS (internal transcribed spacer) region [31].

2.4. Simple Sequence Repeats (SSRs)

SSRs markers, also known as microsatellites or short tandem repeats, are very short (1 to 5) nucleotide motif repetitions occurring as interspersed repetitive elements in all eukaryotic genomes [32]. Schlotterer and Tautz (1992) [33] suggested strand slippage during DNA replication to cause the variation in the number of tandem repeat units, since the repeats allow matching via excision or addition of repeats. SSRs utilize either unlabelled primer pairs or one radioactive or fluorescent labelled primer. Unlabelled SSR-PCR products are analysed using polyacrylamide or sometimes agarose gels. The detection by lasers of automated sequencers has enabled an efficient and high-throughput application of fluorescent labelled microsatellite primers [34], but on the other hand this has made the SSR technique relatively costly compared to other markers [6]. SSR primer development and testing is also time-consuming especially in species whose primer sequences have to be designed newly or are inadequate for unstudied groups. These markers however have a number of advantages such as codominance of alleles, high reproducibility, a low amount of required template DNA that does not need to be of high quality, a high genomic abundance in eukaryotes and a random distribution throughout the genome [6].

SSRs have been used in recent times for most of the genetic diversity studies in African leafy vegetables: Van Biljon *et al*. (2010) [35] studied accessions of the *S. nigrum* complex and their progenies by SSR primers that had partly been transferred from other economically important *Solanum* species. They detected a close relationship among the accessions, hence confirmed the fact that the *S. nigrum* complex is taxonomically difficult to resolve.

By means of only five out of 27 polymorphic SSR markers Li *et al*. (2001) [36] distinguished 88 of 90 cowpea breeding lines. For the complete set of 27 primers, between two to seven alleles were detected per primer, but the primers differed strongly in their PIC (polymorphic information content) which ranged between 0.02 and 0.73 (average: 0.45) [36]. Using the same primer set, Diouf and Hilu (2005) [37] found SSR markers to be a more powerful tool than RAPDs to elucidate the relationships between cowpea local cultivars and breeding lines from Senegal. Later, Asare *et al*. (2010) [38] designed primers from sequence reads for studying the diversity of 141 cowpea accessions from nine geographical locations in Ghana. They found a PIC range of 0.07 to 0.66 with an average of 0.38, and an average heterozygosity of 0.19. The accessions clustered into five main clades, but only a loose correlation with the geographical origin was observed. Phylogenetic relationships and genetic diversity among 16 cowpea genotypes that were used in breeding programmes for resistance to *Striga gesnerioides* in Burkina Faso were investigated using 16 SSR primer combinations [39]. The *Striga* resistant cultivars were found to be very similar in their SSR profiles and only very few primer combinations revealed polymorphisms that could discriminate resistant from susceptible cultivars. SSRs were also used to assess the genetic diversity among 22 local cowpea cultivars and inbred lines collected in Senegal [40]. Forty-four polymorphic primer combinations deduced from expressed sequence tags showed a lower PIC range of 0.08 to 0.33 compared to the earlier studies and the cultivars clustered in groups which were characterised by certain morphological traits.

2.5. Inter-Simple Sequence Repeats (ISSRs)

ISSRs are DNA fragments about 100 - 3000 bp in length which are located between flanking microsatellite regions. The ISSR technique [41] involves amplification of DNA segments present between two identical microsatellite regions that are oppositely oriented to each other. By single microsatellite primers ISSRs of different sizes are amplified in a PCR reaction targeting multiple genomic loci. Thus, fragments of several loci are generated at once, separated by gel electrophoresis and scored for presence or absence. ISSRs are mostly dominant markers, though occasionally a few of them exhibit codominant inheritance. ISSRs are similar to AFLPs and RAPDs in that they do not require sequence data for primer construction [42]. They also require low quantities of DNA templates for PCR and are randomly distributed throughout the genomes. In addition, they are simple, quick and use of radioactivity is not necessary. Despite these advantages, ISSRs can have problems in reproducibility, and their dominant inheritance and homology of co-migrating amplification products result in similar problems like for RAPDs. According to Kojima *et al.* (1998) [43] ISSRs show high levels of polymorphism, but

this depends on the method of detection used. Polyacrylamide gel electrophoresis (PAGE) in combination with radioactively labelled primers was shown to be most sensitive, followed by PAGE with $AgNO_3$ staining and then agarose gels with detection by ethidium bromide staining.

ISSR in combination with RAPDs and SSRs were applied to assess the genetic diversity in 31 *Amaranthus* accessions [44]. All markers separated the accessions, but resulted in different cluster structures making further analyses necessary. Another study [45] used ISSRs in diversity analyses of 56 *Amaranthus* accessions belonging to three species and separated the *Amaranthus* species based on 11 ISSR primers with only few exceptions.

In the genus *Vigna* Ajibade *et al*. (2000) [46] applied ISSRs to investigate 62 taxa within the genus with a focus on the species *V. unguiculata*. Of the 19 primers tested, 15 were very effective and resulted in 63 DNA fragments. It was possible to distinguish the taxa at the species level and below, however, ISSRs were not suitable to clearly differentiate subgenera within *Vigna*. Ghalmi *et al*. (2010) [21] characterised 20 landrace accessions of cowpea in Algeria by 12 ISSR primers and found a correlation of the genetic data with the geographical distribution. Vila-Nova *et al*. (2014) [47] studied the genetic variability and weevil resistance among 27 cowpea cultivars. Ten ISSR primers uncovered a large genetic variability and sufficient polymorphisms to discriminate all the 27 cowpea cultivars. However, the genetic variability could not be related to the resistance of the cultivars tested.

Poczai and Hyvönen (2011) [48] applied ISSR markers combined with start codon targeted polymorphisms (SCoT) and chloroplast sequence data as well as morphological traits to analyse the genetic relationship among diploid, tetraploid and hexaploid *Solanum* species of the *S. nigrum* complex. They found out that all the accessions of the diploid species shared a cluster with all the polyploid species, and concluded that the polyploid species have originated in few combinations of genetically differentiated diploids.

2.6. Single Nucleotide Polymorphisms (SNPs)

SNPs are single-base pair positions in the genome of two or more individuals at which different sequences alternatives (alleles) occur in a population. SNPs are generally abundant, but their density differs considerably between different regions of a genome between genotypes in any species, and more so between species. For instance, Sachidanandam *et al*. (2001) [49] reported that the average density of SNPs in the human genome was estimated at about 1 in 1000 bp but is considerably larger in some genomic areas such as the noncoding human leukocyte antigen (HLA) regions [50]. In plant species analysed so far, approximately one SNP was usually present per 200 to 500 bp with large differences between different plant species. For instance, maize has 1 SNP per 60 - 120 bp [51]. The SNP technique combines two elements, namely the generation of allele-specific products and the analysis of the products. Detection methods for SNPs are either by direct hybridization techniques or those involving the generation and separation of allele-specific products [52]. The high costs for the development of the SNP markers by comparative sequencing of a large numbers of genotypes is a limitation for the use of SNP in plants with limited economic importance. SNPs stand out because of their total number per genome, their comparatively low mutation rates, their distribution across the genome and their relative ease of detection. Applicability in high-throughput genotyping methods such as DNA chips make SNPs striking as genetic markers. Automation with SNPs is also possible and is used for example for identification of genotypes and construction of high-density genetic maps [53]. Only recently the first application of SNPs in genetic diversity studies in African leafy vegetables appeared aiming at genotyping of a worldwide collection of landraces and African ancestral wild cowpeas using 1200 SNPs [54], which in contrast to earlier statements revealed the presence of two major gene pools (eastern and western Africa) and divergent domestication events in cultivated cowpeas in Africa. Further reduction in costs for next generation sequencing will allow low cost SNP detection and GBS (genotyping by sequencing) or related techniques to be used even in crops with minor economic importance in the near future.

3. Conclusion

The 33 studies (published between 1998 and 2015) examined in this review indicate that RAPDs have been used in the majority (25% of the studies) while SNPs were applied only in 3%, allozymes in 12%, AFLPs in 22%, and each of SSRs and ISSRs in 19% of the studies. RAPDs have been used particularly in the investigation of genetic relationship among accessions of a single species from different geographic areas and phylogenetic relationship among species. Allozymes, AFLPs, SSRs and ISSRs have been used at the interspecific and intraspe-

cific levels. Most of the studies covered *Vigna* (58%), *Solanum* (24%), and *Amaranthus* (15%), whereas *C. gynandra* has received little research attention (only 3% of all studies) when genetic diversity studies are concerned. Molecular tools will increasingly be important enabling genetic studies of these and other African leafy vegetables addressing questions regarding the evolutionary origin, centers of diversity, domestication, genetic structure of populations, characterization of germplasm and establishing markers for important agronomic traits. The marker techniques used for this will change with further advances in technical development. If the PCR based markers are compared, RAPDs have been used frequently due to the fact that they are cheap and do neither need sequence information nor high-quality DNA. However, ALFPs and SSRs and possibly SNPs are recommended in future studies due to their better reproducibility and higher information content. The results will considerably support germplasm collection and maintenance strategies and enable the development of improved breeding methods

Acknowledgements

The authors gratefully acknowledge funding within the project HORTINLEA by the German Federal Ministry of Education and Research (BMBF) and the German Federal Ministry of Economic Cooperation and Development (BMZ) within the framework of the program GlobE-Global Food Security.

References

[1] Schippers, R.R. (2000) African Indigenous Vegetables. An Overview of the Cultivated Species. Natural Resources Institute/ACP-EU Technical Centre for Agricultural and Rural Cooperation, Chatham, UK.

[2] Costa, E.M.R, Damasceno-Silva, K.J., Rocha, M.M., Medeiros, A.M. and Anunciação Filho, C.J. (2013) Genetic Divergence among African Cowpea Lines Based on Morpho-Agronomic Traits. *Genetics and Molecular Research*, **12**, 6773-6781. http://dx.doi.org/10.4238/2013.December.16.3

[3] Celine, V.A., Shankaran, S.S. and Seema, S. (2007) Characterization and Evaluation of Vegetable Amaranthus (*Amaranthus tricolor* L.) for High Yield, Quality and Resistance to *Rhizoctonia solani*. *Acta Horticulturae*, **752**, 447-452. http://dx.doi.org/10.17660/ActaHortic.2007.752.81

[4] Diouf, M., Kanté, A. and Mbegue, N.B. (2007) African Leafy Vegetable Agro-Morphological Characterization and Participatory Plant Breeding on Four Species. *Acta Horticulturae*, **752**, 175-178. http://dx.doi.org/10.17660/ActaHortic.2007.752.27

[5] Poczai, P., Mátyás, K., Taller, J. and Szabó, I. (2010) Study of the Origin of the Rarely Cultivated Edible *Solanum* Species: Morphological and Molecular Data. *Biologia Plantarum*, **54**, 543-546. http://dx.doi.org/10.1007/s10535-010-0096-x

[6] Agarwal, M., Neeta, S. and Harish, P. (2008) Advances in Molecular Marker Techniques and Their Applications in Plant Sciences. *Plant Cell Reports*, **27**, 617-631. http://dx.doi.org/10.1007/s00299-008-0507-z

[7] Weising, K., Nybom, H., Wolff, K. and Kahl, G. (2005) DNA Fingerprinting in Plants, Principles, Methods, and Applications. CRC Press. http://dx.doi.org/10.1201/9781420040043

[8] Kumar, P., Gupta, V.K., Misra, A.K., Modi, D.R. and Pandey, B.K. (2009) Potential of Molecular Markers in Plant Biotechnology. *Plant Omics Journal*, **2**, 141-162.

[9] Spooner, D., van Treuren, R. and de Vicente, M.C. (2005) Molecular Markers for Genebank Management. IPGRI Technical Bulletin No. 10. International Plant Genetic Resources Institute, Rome.

[10] Sonnante, G., Anne, S., Angela, M. and Domenico, P. (1997) Isozyme and RAPD Analysis of Genetic Diversity within and between *Vigna luteola* and *V. marina*. *Annals of Botany*, **80**, 741-746. http://dx.doi.org/10.1006/anbo.1997.0511

[11] Pasquet, R.S. (1999) Genetic Relationships among Subspecies of *Vigna unguiculata* (L.) Walp. Based on Allozyme Variation. *Theoretical and Applied Genetics*, **98**, 1104-1119. http://dx.doi.org/10.1007/s001220051174

[12] Kouam, E.B., Pasquet, R.S., Campagne, P., Tignegre, J.B., Thoen, K., Gaudin, R., Ouedraogo, J.T., Salifu, A.B., Muluvi, G.M. and Gepts, P. (2012) Genetic Structure and Mating System of Wild Cowpea Populations in West Africa. *BMC Plant Biology*, **12**, 113. http://dx.doi.org/10.1186/1471-2229-12-113

[13] Di Renzo, M.A., Bonamico, N.C., Ibanez, M.A., Samame, S.S. and Gesumaria, J.J. (2000) Isozyme Diversity in Seven Amaranth Species. *Phyton-International Journal of Experimental Botany*, **66**, 165-174.

[14] Williams, J.G.K., Kublelik, A.R., Livak, K.J., Rafalski, J.A. and Tingey, S.V. (1990) DNA Polymorphisms Amplified by Arbitrary Primers Are Useful as Genetic Markers. *Nucleic Acids Research*, **18**, 6531-6535. http://dx.doi.org/10.1093/nar/18.22.6531

[15] Jones, C.J., Edwards, K.J., Castaglione, S., Winfield, M.O.,Sala, F., Wiel, C., van de, Bredemeijer, G., Vosman, B., Matthes, M., Daly, A., Brettschneider, R., Bettini, P., Buiatti, M., Maestri, E., Malcevschi, A., Marmiroli, N., Aert, R., Volckaert, G., Rudea, J., Linacero, R., Vazquez, A. and Karp, A. (1997) Reproducibility Testing of RAPD, AFLP and SSR Markers in Plants by a Network of European Laboratories. *Molecular Breeding*, **3**, 381-390. http://dx.doi.org/10.1023/A:1009612517139

[16] Ray, T. and Roy, S.C. (2009) Genetic Diversity of *Amaranthus* Species from the Indo-Gangetic Plains Revealed by RAPD Analysis Leading to the Development of Ecotype-Specific SCAR Marker. *Journal of Heredity*, **100**, 338-347. http://dx.doi.org/10.1093/jhered/esn102

[17] Nkongolo, K.K. (2003) Genetic Characterization of Malawian Cowpea (*Vigna unguiculata*(L.)Walp) Landraces: Diversity and Gene Flow among Accessions. *Euphytica*, **129**, 219-228. http://dx.doi.org/10.1023/A:1021935431873

[18] Ba, F.S., Pasquet, R.S. and Gepts, P. (2004) Genetic Diversity in Cowpea [*Vigna unguiculata* (L.) Walp.] as Revealed by RAPD Markers. *Genetic Resources and Crop Evolution*, **51**, 539-550. http://dx.doi.org/10.1023/B:GRES.0000024158.83190.4e

[19] Zannou, A., Kossou, D.K., Ahanchédé, A., Zoundjihékpon, J., Agbicodo, E., Struik, P.C. and Sanni, A. (2008) Genetic Variability of Cultivated Cowpea in Benin Assessed by Random Amplified Polymorphic DNA. *African Journal of Biotechnology*, **7**, 4407-4414.

[20] Mignouna, H.D., Ng N.Q., Ikea, J. and Thottapilly, G. (1998) Genetic Diversity in Cowpea as Revealed by Random Amplified Polymorphic DNA. *Journal of Genetics and Breeding*, **52**, 151-159.

[21] Ghalmi, N., Malice, M., Jacquemin, J.M., Ounane, S.M., Mekliche, L. and Baudoin, J.P. (2010) Morphological and Molecular Diversity within Algerian Cowpea (*Vigna unguiculata* (L.) Walp.) Landraces. *Genetic Resources and Crop Evolution*, **57**, 371-386. http://dx.doi.org/10.1007/s10722-009-9476-5

[22] Malviya, N., Yadav, D., Sarangi, B.K. and Yadav, M.K. (2012) Analysis of Genetic Diversity in Cowpea (*Vigna unguiculata* L. Walp.) Cultivars with Random Amplified Polymorphic DNA Markers. *Plant Systematics and Evolution*, **29**, 523-526. http://dx.doi.org/10.1007/s00606-011-0545-9

[23] K'Opondo, F.B.O., van Rheenen, H.A. and Muasya, R.M. (2009) Assessment of Genetic Variation of Selected Spiderplant *(Cleome gynandra* L.) Morphotypes from Western Kenya. *African Journal of Biotechnology*, **8**, 4325-4332.

[24] Vos, P., Hogers, R., Bleeker, M., Reijans, M., van de Lee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J., Kuiper, M. and Zabeau, M. (1995) AFLP: A New Technique for DNA Fingerprinting. *Nucleic Acids Research*, **23**, 4407-4414. http://dx.doi.org/10.1093/nar/23.21.4407

[25] Jacoby, A., Labuschagne, M.T. and Viljoen, C.D. (2003) Genetic Relationships between Southern African *Solanum retroflexum* Dun. and Other Related Species Measured by Morphological and DNA Markers. *Euphytica*, **132**, 109-113. http://dx.doi.org/10.1023/A:1024657827796

[26] Dehmer, K.J. and Hammer, K. (2004) Taxonomic Status and Geographic Provenance of Germplasm Accessions in the *Solanum nigrum* L. Complex: AFLP Data. *Genetic Resources and Crop Evolution*, **51**, 551-558. http://dx.doi.org/10.1023/B:GRES.0000024163.86762.fc

[27] Olet, E.A., Lye, K.A. and Heun, M. (2011) Amplified Fragment Length Polymorphisms (AFLPs) Analysis of Species of *Solanum* Section *Solanum* (Solanaceae) from Uganda. *African Journal of Biotechnology*, **10**, 6387-6395.

[28] Coulibaly, S., Pasquet, R.S., Papa, R. and Gepts, P. (2002) AFLP Analysis of the Phenetic Organization and Genetic Diversity of *Vigna unguiculata* L. Walp. Reveals Extensive Gene Flow between Wild and Domesticated Types. *Theoretical and Applied Genetics*, **104**, 358-366. http://dx.doi.org/10.1007/s001220100740

[29] Gillaspie, A.G., Hopkins, M.S. and Dean, R.E. (2005) Determining Genetic Diversity between Lines of *Vigna unguiculata* subspecies by AFLP and SSR Markers. *Genetic Resources and Crop Evolution*, **52**, 245-247. http://dx.doi.org/10.1007/s10722-004-6693-9

[30] Fang, J.G., Chao, C.T., Roberts, P.A. and Ehlers, J.D. (2007) Genetic Diversity of Cowpea [*Vigna unguiculata* (L.) Walp.] in Four West African and USA Breeding Programs as Determined by AFLP Analysis. *Genetic Resources and Crop Evolution*, **54**, 1197-1209. http://dx.doi.org/10.1007/s10722-006-9101-9

[31] Xu, F. and Sun, M. (2001) Comparative Analysis of Phylogenetic Relationships of Grain Amaranths and Their Wild Relatives (*Amaranthus*; Amaranthaceae) Using Internal Transcribed Spacer, Amplified Fragment Length Polymorphism, and Double-Primer Fluorescent Inter-Simple Sequence Repeat Markers. *Molecular Phylogenetics and Evolution*, **21**, 372-387. http://dx.doi.org/10.1006/mpev.2001.1016

[32] Tautz, D. and Renz, M. (1984) Simple Sequences Are Ubiquitous Repetitive Components of Eukaryotic Genomes. *Nucleic Acids Research*, **12**, 4127-4138. http://dx.doi.org/10.1093/nar/12.10.4127

[33] Schlotterer, C. and Tautz, D. (1992) Slippage Synthesis of Simple Sequence DNA. *Nucleic Acids Research*, **20**, 2211-2215. http://dx.doi.org/10.1093/nar/20.2.211

[34] Wenz, H.M., Robertson, J.M., Menchen, S., Oaks, F., Demorest, D.M., Scheibler, D., Rosenblum, B.B., Wike, C., Gilbert, D.A. and Efcavitch, J.W. (1998) High-Precision Genotyping by Denaturing Capillary Electrophoresis. *Genome Research*, **3**, 69-80.

[35] Van Biljon, A., Labuschagne, M. and Koen, E. (2010) Microsatellite-Based Assessment of Five *Solanum nigrum* Complex Species and Their Progeny. *Acta Agriculturae Scandinavica*, *Section B*, **60**, 494-499. http://dx.doi.org/10.1080/09064710903233839

[36] Li, C.D., Fatokun, C.A., Ubi, B., Singh, B.B. and Scoles, G.J. (2001) Determining Genetic Similarities and Relationships among Cowpea Breeding Lines and Cultivars by Microsatellite Markers. *Crop Science*, **41**, 189-197. http://dx.doi.org/10.2135/cropsci2001.411189x

[37] Diouf, D. and Hilu, K.W. (2005) Microsatellites and RAPD Markers to Study Genetic Relationships among Cowpea Breeding Lines and Local Varieties in Senegal. *Genetic Resources and Crop Evolution*, **52**, 1057-1067. http://dx.doi.org/10.1007/s10722-004-6107-z

[38] Asare, A.T., Gowda, B.S., Galyuon, K.A., Aboagye, L.L., Takrama, J.F. and Timko, M.P. (2010) Assessment of the Genetic Diversity in Cowpea (*Vigna unguiculata* L. Walp.) Germplasm from Ghana Using Simple Sequence Repeat Markers. *Plant Genetic Resources*, **8**, 142-150. http://dx.doi.org/10.1017/S1479262110000092

[39] Sawadogo, M., Ouedraogo, J.T., Gowda, B.S. and Timko, M.P. (2010) Genetic Diversity of Cowpea (*Vigna unguiculata* L. Walp) Cultivars in Burkina Faso Resistant to *Striga gesnerioides*. *African Journal of Biotechnology*, **9**, 8146-8153.

[40] Badiane, F.A., Gowda, B.S., Cisse, N., Diouf, D., Sadio, O. and Timko, M.P. (2012) Genetic Relationship of Cowpea (*Vigna unguiculata*) Varieties from Senegal Based on SSR Markers. *Genetics and Molecular Research*, **11**, 292-304. http://dx.doi.org/10.4238/2012.February.8.4

[41] Zietkiewicz, E., Rafalski, A. and Labuda, D. (1994) Genome Fingerprinting by Simple Sequence Repeats (SSR)-Anchored Polymerase Chain Reaction Amplification. *Genomics*, **20**, 176-183. http://dx.doi.org/10.1006/geno.1994.1151

[42] Joshi, S.P., Gupta, V.S., Aggarwal, R.K., Ranjekar, P.K. and Brar, D.S. (2000) Genetic Diversity and Phylogenetic Relationship as Revealed by Inter-Simple Sequence Repeat (ISSR) Polymorphism in the Genus *Oryza*. *Theoretical and Applied Genetics*, **100**, 1311-1320. http://dx.doi.org/10.1007/s001220051440

[43] Kojima, T., Nagaoka, T., Noda, N. and Ogihara, Y. (1998) Genetic Linkage Map of ISSR and RAPD Markers in Einkorn Wheat in Relation to That of RFLP Markers. *Theoretical and Applied Genetics*, **96**, 37-45. http://dx.doi.org/10.1007/s001220050706

[44] Singh, B., Pandey, S. and Kumar, J. (2013) A Comparative Study of Inter Simple Sequence Repeat (ISSR), Random Amplified Polymorphic DNA (RAPD) and Simple Sequence Repeat (SSR) Loci in Assessing Genetic Diversity in *Amaranthus*. *Indian Journal of Genetics and Plant Breeding*, **73**, 411-418. http://dx.doi.org/10.5958/j.0975-6906.73.4.062

[45] Štefúnova, V., Bežo, M., Labajová, M. and Senková, S. (2014) Genetic Analysis of Three Amaranth Species Using ISSR Markers. *Journal of Food and Agriculture*, **26**, 35-43.

[46] Ajibade, S.R., Weeden, N.F. and Chite, S.M. (2000) Inter Simple Sequence Repeat Analysis of Genetic Relationships in the Genus *Vigna*. *Euphytica*, **111**, 47-55. http://dx.doi.org/10.1023/A:1003763328768

[47] Vila Nova, M.X., Leite, N.G.A., Houllou, L.M., Medeiros, L.V., Lira Neto, A.C., Hsie, B.S., Borges-Paluch, L.R., Santos, B.S., Araujo, C.S.F., Rocha, A.A. and Costa, A.F. (2014) Genetic Variability and Resistance of Cultivars of Cowpea [*Vigna unguiculata* (L.) Walp] to Cowpea Weevil (*Callosobruchus maculatus* Fabr.). *Genetics and Molecular Research*, **13**, 2323-2332. http://dx.doi.org/10.4238/2014.March.31.12

[48] Poczai, P. and Hyvönen, J. (2011) On the Origin of *Solanum nigrum*: Can Networks Help? *Molecular Biology Reports*, **38**, 1171-1185. http://dx.doi.org/10.1007/s11033-010-0215-y

[49] Sachidanandam, R. and 40 coauthors (2001) A Map of Human Genome Sequence Variation Containing 1.42 Million Single Nucleotide Polymorphisms. *Nature*, **409**, 928-933. http://dx.doi.org/10.1038/35057149

[50] Horton, R., Niblett, D., Milne, S., Palmer, S., Tubby, B., Trowsdale, J. and Beck, S. (1998) Large-Scale Sequence Comparisons Reveal Unusually High Levels of Variation in the HLA-DQB1 Locus in the Class II Region of the Human MHC. *Journal of Molecular Biology*, **282**, 71-97. http://dx.doi.org/10.1006/jmbi.1998.2018

[51] Ching, A.D.A., Caldwell, K.S., Jung, M., Dolan, M., Smith, O.S., Tingey, S., Morgante, M. and Rafalski, A. (2002) SNP Frequency, Haplotype Structure and Linkage Disequilibrium in Elite Maize Inbred Lines. *BMC Genetics*, **3**, 19. http://dx.doi.org/10.1186/1471-2156-3-19

[52] Vignal, A., Milan, D., Sancristobal, M. and Eggen, A. (2002) A Review on SNP and Other Types of Molecular Markers and Their Use in Animal Genetics. *Genetics Selection Evolution*, **34**, 275-305. http://dx.doi.org/10.1186/1297-9686-34-3-275

[53] Govindaraj, M., Vetriventhan, M. and Srinivasan, M. (2015) Importance of Genetic Diversity Assessment in Crop Plants and Its Recent Advances: An Overview of Its Analytical Perspectives. *Genetics Research International*, **2015**, Article ID: 431487. http://dx.doi.org/10.1155/2015/431487

[54] Huynh, B.L., Close, T.J., Roberts, P.A., Hu, Z., Wanamaker, S., Lucas, M.R., Chiulele, R, Cissé, N., Antonio, D., Hearne, S., Fatokun, C., Diop, N.N. and Ehlers, J.D. (2013) Gene Pools and the Genetic Architecture of Domesticated Cowpea. *Plant Genome*, **6**, 1-8. http://dx.doi.org/10.3835/plantgenome2013.03.0005

2.2. Mating biology, nuclear DNA content and genetic diversity in spider plant (*Cleome gynandra*) germplasm from various African countries.

Emmanuel O. Omondi[1], Thomas Debener[2], Marcus Linde[2], Mary Abukutsa-Onyango[3], Fekadu F. Dinssa[4], Traud Winkelmann[1]

[1]Institute of Horticultural Production Systems, Woody Plant and Propagation Physiology, Leibniz Universität Hannover, Hannover, Germany

[2]Institute for Plant Genetics, Molecular Plant Breeding, Leibniz Universität Hannover, Hannover, Germany

[3]Department Horticulture, Jomo Kenyatta University of Agriculture & Technology (JKUAT), Juja, Kenya

[4]AVRDC - The World Vegetable Center, Eastern and Southern Africa, Arusha, Tanzania

Type of authorship	First author
Type of article	Research article
Contribution to the article	Designed and performed all experiments Analysed data Preparation of tables and figures Wrote the manuscript
Contribution of other authors	Traud Winkelmann was involved in the conception of the study, preparation of the tables and figures and writing of the aims Thomas Debener, Marcus Linde assisted in experimental design, data analysis and writing the manuscript. Mary Abukutsa-Onyango and Fekadu F. Dinssa provided the research material and assisted in writing the manuscript.
Journal	Plant Breeding
Date of publication	26 May, 2017
DOI	10.1111/pbr.12485

Received: 28 January 2017 | Accepted: 1 April 2017

DOI: 10.1111/pbr.12485

ORIGINAL ARTICLE

WILEY Plant Breeding

Mating biology, nuclear DNA content and genetic diversity in spider plant (*Cleome gynandra*) germplasm from various African countries

Emmanuel O. Omondi[1] | **Thomas Debener**[2] | **Marcus Linde**[2] | **Mary Abukutsa-Onyango**[3] | **Fekadu F. Dinssa**[4] | **Traud Winkelmann**[1]

[1]Institute of Horticultural Production Systems, Woody Plant and Propagation Physiology, Leibniz Universität Hannover, Hannover, Germany

[2]Institute for Plant Genetics, Molecular Plant Breeding, Leibniz Universität Hannover, Hannover, Germany

[3]Department of Horticulture, Jomo Kenyatta University of Agriculture & Technology (JKUAT), Juja, Kenya

[4]World Vegetable Center, Eastern and Southern Africa, Arusha, Tanzania

Correspondence
Traud Winkelmann, Institute of Horticultural Production Systems, Woody Plant and Propagation Physiology, Leibniz Universität Hannover, Hannover, Germany.
Email: traud.winkelmann@zier.uni-hannover.de

Funding information
German Federal Ministry of Education and Research (BMBF); German Federal Ministry of Economic Cooperation and Development (BMZ)

Communicated by M. Havey

Abstract

Cleome gynandra (L.) Briq. is an African leafy vegetable with a potential to improve food security and micronutrient deficiencies. Cytological traits, breeding biology and genetic diversity of 30 selected entries of *C. gynandra* from six African countries were investigated. The entries consisted of advanced lines, gene bank accessions and farmers' cultivars. Our study revealed chromosome numbers of $2n = 34$ in root tip metaphase cells from one entry. The 30 entries were found to be diploid with genome sizes ranging from 2.31 to 2.45 pg/2C. Hand pollination experiments were carried out to assess self-incompatibility within the entries and revealed that they are self- and cross-compatible. For genetic diversity studies within and among the entries, the pooled data of 499 polymorphic bands from 11 amplified fragment length polymorphism primer combinations and nine simple sequence repeat markers were used. The genetic distance among the entries ranged from 0.13 to 0.77. In a principal coordinate analysis, the farmers' cultivars formed a cluster separate from the advanced lines and the gene bank entries, and the latter were not well resolved.

KEYWORDS
amplified fragment length polymorphism, breeding biology, chromosome numbers, flow cytometry, gene bank accessions, simple sequence repeat

1 | INTRODUCTION

Cleome gynandra (L.) Briq, also known as spider plant, cat's whisker or African cabbage, is reported to have originated from tropical Africa and South-East Asia and has now spread as a weed or cultivated plant in the tropical and subtropical regions of the world (Chweya and Mnzava 1997). It is one of the most important African leafy vegetables after *Amaranthus* spp., cowpea (*Vigna unguiculata*) and African nightshades (*Solanum* spp.). These leafy vegetables have been recognized to play an important role in human nutrition, food security and income generation among poor rural communities and recently in urban communities (Abukutsa-Onyango 2007; Onyango et al. 2013). *Cleome gynandra* leaves are rich in proteins, vitamins and minerals, such as iron and zinc, and have other beneficial health compounds, such as flavonoids and glucosinolates (van Jaarsveld et al. 2014; Odhav et al. 2011; Uusiku et al. 2010). It also has various medicinal applications based on indigenous knowledge that varies among communities. It is important for rural households' food security as this vegetable is available during periods of drought or in seasons when other main food crops do not grow well (Chweya and Mnzava 1997). Despite the considerable benefits of the crop, it is still regarded as a minor crop and has a lack of support from research and development programmes to develop high-yielding cultivars and improved cultural practices (Onyango et al. 2013). Of all studies dealing with the use of molecular markers in genetic diversity studies in African leafy vegetables (reviewed by Omondi et al. 2016),

only one study with limited scope could be listed for *C. gynandra*. Several molecular marker techniques, such as random amplified polymorphic DNAs (RAPD) (Williams et al. 1990), amplified fragment length polymorphic (AFLP) (Vos et al. 1995), inter-simple sequence repeats (ISSR) (Zietkiewicz et al. 1994) and simple sequence repeats (SSR) (Tautz and Renz, 1984), have been applied in molecular diversity studies of African leafy vegetables (Omondi et al. 2016). K'Opondo et al. (2009) assessed the genetic variation of selected *C. gynandra* morphotypes from western Kenya using RAPD markers and were able to cluster the morphotypes into three clusters according to the colour of the stem and petioles, where purple stems, purple petioles (PP), green stems, green petioles (GG) and GP clustered apart from PG morphotypes.

In this study, more robust and highly informative SSR and AFLP markers have been used. The development and testing of SSR primers can be time-consuming, especially in unstudied species such as *C. gynandra*, but they present advantages such as the codominance of alleles, high reproducibility and that low amounts of moderate quality DNA can be used. They are also highly abundant and randomly distributed in the genomes (Agarwal et al. 2008). AFLP markers, on the other hand, do not require prior sequence information and represent a high number of loci distributed randomly throughout the genome concurrently. However, AFLP markers also require high-quality DNA.

The seeds of *C. gynandra* used by farmers are of poor quality as farmers traditionally keep seeds from previous harvests for subsequent seasons. Seeds from agricultural stores in many cases are not superior either because they are not derived from breeding programmes aimed at genetic improvement. The major goals in the breeding and genetic improvement of *C. gynandra* are to combine important traits such as yield, time to maturity, water use efficiency and plant type for different traditional preferences in terms of colour of stems/leaf petioles and bitterness. To enable directed breeding, as well as germplasm conservation strategies, a cytological assessment, for example ploidy level and genome size determinations, knowledge of breeding biology and genetic diversity studies are fundamental prerequisites. Therefore, the objectives of this study were to determine the genome size and ploidy levels and to assess the level of self-compatibility in this outbreeding species and the genetic diversity within and among 30 selected entries (farmers' cultivars, gene bank accessions and advanced breeding lines) of *C. gynandra* from six African countries.

2 | MATERIALS AND METHODS

2.1 | Plant material

Thirty entries of *C. gynandra* were used in this study. The materials comprised farmers' cultivars directly collected from farmers, gene bank accessions and advanced breeding lines developed by mass selection from gene bank accessions, which were kindly provided by the World Vegetable Center (WorldVeg), Arusha, Tanzania; and Jomo Kenyatta University of Agriculture and Technology, Nairobi, Kenya (Table 1).

2.2 | Chromosome analysis

Young root tips, measuring 0.5–1 cm in length, were harvested from actively growing roots approximately 1–2 hr after sunrise from three replicates of five *C. gynandra* breeding line entries and were pretreated with a 2 mM 8-hydroxyquinoline solution at room temperature for 4 hr. The root tips were then fixed with 3:1 (ethanol: acetic acid), after which they were washed in water for 10–30 min before the meristems were dissected and transferred into 0.1 M citric buffer (0.1 M citric acid; 0.1 M sodium citrate) at pH 4.8. An enzymatic treatment as described in the steam drop method by Kirov et al. (2014) followed. Staining was carried out using 4′,6-diamidino-2-phenylindole (DAPI) stain, and the preparations were observed under a fluorescence microscope (Zeiss Axio Scope.A1, Carl Zeiss Microimaging Co. Ltd., Göttingen, Germany) at 1,000× magnification (filter set 02: excitation, 365 nm; beam splitter, 395 nm; emission, 420 nm).

2.3 | Estimation of the nuclear DNA content by flow cytometry

For the isolation of nuclei, 0.5 cm^2 of fresh young leaf tissue from adult plants was cochopped in a plastic Petri dish together with the same amount of leaves of maize (*Zea mays* L., 5.43 pg/2C) (Gatersleben gene bank accession ZEA 3443) as an internal reference standard. A standard protocol as described in the manual for CyStain PI AbsoluteP (Partec Co. Ltd., Germany) was used for nuclei isolation. The fluorescence of the nuclei was then measured using the CyFlow Ploidy analyser (Partec Co. Ltd., Germany). Fluorescence was excited at 400 nm. Three replicate measurements from at least two genotypes of each accession (Table 2) and at least 3,000 nuclei were measured. The 2C DNA content of the sample was calculated as the mean sample peak divided by the mean standard peak and multiplied by the amount of DNA of the standard.

2.4 | Pollination study

The estimations of the level of self-incompatibility were carried out based on hand pollinations in the greenhouse (Table 3). Both self- and cross-pollinations were carried out. At least five flowers per genotype for two to nine genotypes per entry were pollinated. Stigmas from each flower were pollinated by fresh pollen grains from the same flower or from a different flower of the same plant for self-pollination. For cross-pollination, stigmas of emasculated flowers were pollinated with pollen grains from other genotypes or a mix of pollen from other genotypes of the same entry. The flowers of both types of pollination were then bagged for at least 4 days, after which the fruits began to enlarge and the stigmas were no longer receptive for pollen. Fruit and seed set was determined for the controlled pollinations, and the index of self-incompatibility (ISI) was determined by dividing the relative success of seed set from self-pollination by the relative success of seed set from cross-pollination (Arroyo et al. 1993). In addition, pollen tube growth was observed

TABLE 1 Thirty *C. gynandra* farmers' cultivars, gene bank accessions and breeding lines included in the study

Entry	Category	Source	Coordinate	Original entry source
Abuku1	Advanced line	JKUAT	–	Kenya
Abuku2	Advanced line	JKUAT	–	Kenya
Acc20	Farmers' cultivar	Bondo	0°06′07.5″S 34°16′26.9″E	Kenya
Acc21	Farmers' cultivar	Bondo	0°06′07.5″S 34°16′26.9″E	Kenya
Acc26	Farmers' cultivar	Kisii	0°44′48.5″S 34°47′68.5″E	Kenya
Acc28	Farmers' cultivar	Kisii	0°40′51.5″S 34°47′41.1″E	Kenya
Acc30	Farmers' cultivar	Nakuru	0°02′41.9″S 36°01′49.2″E	Kenya
Acc3	Farmers' cultivar	Mbale	0°06′08.0″S 34°42′28.0″E	Kenya
Acc5	Farmers' cultivar	Yala	0°09′24.8″N 34°33′02.2″E	Kenya
Acc6	Farmers' cultivar	Butere	0°10′07.0″N 34°31′36.8″E	Kenya
HTT-Sel[a]	Advanced line	WorldVeg	–	Kenya
GS-Sel	Advanced line	WorldVeg	–	Tanzania
IP7	Gene bank accession	WorldVeg	–	Tanzania
PS	Gene bank accession	WorldVeg	–	Tanzania
Site94	Gene bank accession	WorldVeg	–	Tanzania
ST73-3	Gene bank accession	WorldVeg	–	Tanzania
ST93-1GS	Gene bank accession	WorldVeg	–	Tanzania
IP12	Gene bank accession	WorldVeg	–	S. Africa
IP8-Sel	Advanced line	WorldVeg	–	S. Africa
MLSF12	Gene bank accession	WorldVeg	–	Malawi
MLSF17	Gene bank accession	WorldVeg	–	Malawi
MLSF27	Gene bank accession	WorldVeg	–	Malawi
RWSF2	Gene bank accession	WorldVeg	–	Rwanda
RWSF3	Gene bank accession	WorldVeg	–	Rwanda
UGSF13	Gene bank accession	WorldVeg	–	Uganda
UGSF17-Sel	Advanced line	WorldVeg	–	Uganda
UGSF19	Gene bank accession	WorldVeg	–	Uganda
UGSF2-Sel	Advanced line	WorldVeg	–	Uganda
UGSF26	Gene bank accession	WorldVeg	–	Uganda
UGSF29	Gene bank accession	WorldVeg	–	Uganda

WorldVeg, World Vegetable Center, Arusha, Tanzania; JKUAT, Jomo Kenyatta University of Agriculture and Technology, Nairobi, Kenya.
[a]Entry in which DNA was sequenced for SSR primer design.
'-Sel' indicates that the line was developed by mass selection in the original accession. Colours are according to country of origin and are used throughout the article.

in situ by fluorescence microscopy of aniline blue-stained ovaries (Winkelmann et al. 2010).

2.5 | DNA isolation

Fifty to sixty milligrams of fresh leaf material for DNA isolation was dried overnight over silica at 37°C. The DNA extraction was carried out in duplicate per genotype for five genotypes of each *C. gynandra* entry using the Macherey-Nagel extraction kit (NucleoSpin Plant II, Macherey-Nagel, Düren, Germany) according to the manufacturer's instructions. Buffer PL1 in the extraction kit, which is based on the CTAB method of DNA isolation, was used. The DNA was then eluted in 80 μl TE buffer (10 mM Tris–HCl [pH 7.5], 1 mM EDTA [pH 8]). The integrity of the genomic DNA was assessed by agarose gels, and the purity and quantity were further checked using the NanoDrop spectrophotometer 2000c (PeqLab Biotechnology Co. Ltd., Erlangen, Germany).

2.6 | AFLP and SSR analyses

A set of 11 AFLP primer combinations and nine SSR markers were used to assess the genetic diversity in the 30 *C. gynandra* entries. From a total of 53 *Hind*III and *Mse*I primer pairs tested, 11 were finally used after considering the polymorphisms and reproducibility observed (Table 4). Amplified fragment length polymorphism (AFLP) templates were prepared as described by Vos et al. (1995) with minor modifications. For each sample, 150 ng of DNA was digested with *Hind*III and *Mse*I. The pre-amplification reactions were

performed with primers having no selective nucleotides. Primers with three selective bases were used for the final amplification reactions. The *Hind*IIICAT and *Hind*IIIATA primers were end-labelled with IRD 700 and IRD 800 dyes (Eurofins MWG, Ebersberg, Germany), respectively. The fragments were separated on 6% polyacrylamide gels (Sequagel XR, Hessle, UK) using a DNA analyser (LI-COR, Lincoln, NE, USA) and were automatically documented using e-Seq Software (V3.0, LI-COR).

For SSR amplification reactions, primers were designed from next-generation sequencing (GATC Biotech Company, Constance, Germany) of the genomic DNA from *C. gynandra* entry HTT-Sel. A genomic library was constructed, followed by Illumina HiSeq2500 sequencing of 5 million read pairs, which gave approximately 1.25 gigabases of raw data. Contig assembly from the sequence data was executed using CLC Genomics Workbench software (version 7.5.1), and the contigs were searched for SSRs using the 'SSR locator' software (da Maia et al., 2008). Forward primers designed by the SSR locator were tagged with an M13 primer binding sequence (GTAAAACGACGGCCAGT) at the 5′ end according to Schuelke (2000) to enable a two-step PCR with fluorescently labelled M13 primers (Eurofins Genomics Co. Ltd., Ebersberg, Germany). Thirty-seven primers were designed and tested, but only nine primers were polymorphic and gave products of the expected size (Table 5).

A standard protocol was used for the SSR-PCRs with 10 ng of genomic DNA in a total volume of 10 μl containing 0.25 pmol/μl of forward primer, 2.5 pmol/μl of reverse primer, 1.25 pmol/μl of forward primer labelled with IRD700 (Eurofins MWG, Ebersberg, Germany), 0.5 unit of *Taq* (DNA Cloning Service Co. Ltd., Hamburg, Germany), 0.15 mM deoxy-nucleotide triphosphates and 1 μl of 10× Williams buffer (200 mM Tris-HCl (pH 8.4), 15 mM $MgCl_2$, 500 mM KCl). The conditions for the PCR were as follows: 94°C (5 min), 25 cycles at 94°C (45 s), an annealing step with temperature ranging from 57 to 63°C depending on the primers (Table 5) (1 min), 72°C (1 min), followed by eight cycles at 94°C (30 s), 52°C (45 s), 72°C (1 min), and a final extension at 72°C for 10 min.

All reactions were carried out in an S1000 thermal cycler (Bio-Rad Laboratories GmbH) with two independent DNA extractions per genotype to ensure reproducibility. The PCR products were diluted with equal volumes of formamide loading dye (98% formamide, 10 mmol/L ethylenediaminetetraacetic acid (EDTA) at pH 8.0, 0.05% pararosaniline) for SSR and AFLP, heated for 5 min at 94°C (heating

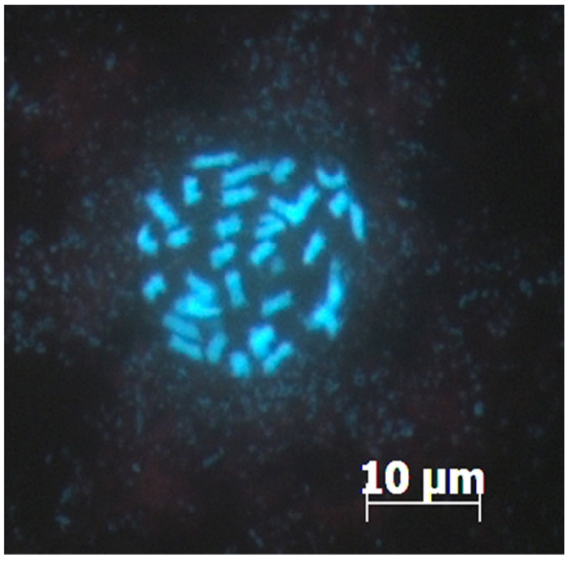


FIGURE 1 Chromosomes in a metaphase cell of a root tip of the entry GS

TABLE 2 Genome sizes of *C. gynandra* entries estimated by flow cytometry against the internal standard *Z. mays*

C. gynandra Entry	No. of genotypes analysed	Genome size (pg/2C)	CV (%) *C. gynandra* $\bar{x} \pm SD$	CV (%) *Z. mays* $\bar{x} \pm SD$
Abuku1	2	2.37 ± 0.01	5.1 ± 0.1	3.9 ± 0.3
Abuku2	3	2.37 ± 0.01	4.5 ± 1.2	3.5 ± 1.1
Acc30	3	2.35 ± 0.01	5.1 ± 0.2	3.5 ± 0.4
Acc20	3	2.37 ± 0.02	4.6 ± 0.4	3.0 ± 0.3
Acc21	4	2.31 ± 0.02	6.9 ± 0.2	4.2 ± 0.3
Acc26	3	2.39 ± 0.01	3.8 ± 0.6	3.3 ± 0.2
Acc28	4	2.35 ± 0.02	7.6 ± 0.4	3.7 ± 0.3
Acc3	3	2.36 ± 0.01	4.2 ± 0.3	2.9 ± 0.3
Acc5	3	2.37 ± 0.01	4.5 ± 0.5	3.3 ± 0.4
Acc6	3	2.35 ± 0.02	4.6 ± 0.3	3.6 ± 0.2
GS-Sel[a]	3	2.35 ± 0.01	4.3 ± 0.4	2.9 ± 0.5
HTT-Sel	3	2.33 ± 0.04	3.7 ± 0.3	3.4 ± 0.1
IP12	2	2.42 ± 0.02	3.7 ± 0.3	3.9 ± 0.3
IP7	2	2.39 ± 0.02	4.1 ± 0.6	3.4 ± 0.6
IP8-Sel	3	2.34 ± 0.01	4.2 ± 0.2	3.2 ± 0.1
MLSF12	3	2.44 ± 0.00	4.1 ± 0.2	3.4 ± 0.2
MLSF17	2	2.34 ± 0.02	3.5 ± 0.4	3.5 ± 0.2
MLSF27	3	2.38 ± 0.03	3.8 ± 0.2	3.3 ± 0.3
PS	2	2.42 ± 0.00	3.9 ± 0.3	3.6 ± 0.2
RWSF2	2	2.39 ± 0.01	4.0 ± 0.2	3.7 ± 0.2
RWSF3	3	2.41 ± 0.00	4.1 ± 0.3	3.6 ± 0.2
Site94	2	2.38 ± 0.01	4.3 ± 0.5	3.9 ± 0.4
ST73-3	2	2.41 ± 0.01	3.3 ± 0.4	3.3 ± 0.2
ST93-1	2	2.45 ± 0.02	4.1 ± 0.3	4.0 ± 0.3
UGSF13	2	2.34 ± 0.01	3.3 ± 0.1	3.8 ± 0.2
UGSF17-Sel	3	2.39 ± 0.02	4.9 ± 0.5	3.1 ± 0.5
UGSF19	2	2.38 ± 0.01	4.6 ± 0.2	4.0 ± 0.2
UGSF2-Sel	3	2.35 ± 0.03	4.1 ± 0.4	3.5 ± 0.4
UGSF26	3	2.41 ± 0.00	4.5 ± 0.4	3.7 ± 0.2
UGSF29	2	2.45 ± 0.02	3.9 ± 0.3	3.7 ± 0.1
Mean		2.38		
SD		0.04		

SD, standard deviation; CV, coefficient of variation.
[a]Standard entry used for chromosome count.

TABLE 3 Controlled pollination study on *C. gynandra* entries grown in the glasshouse

C. gynandra entry	Pollination type	No. of flowers (No. of plants)	No. of fruit sets	% fruit set	Total no. of seed	Average no. of seeds/fruit	ISI
Abuku1	Self	8 (2)	1	12.5	159	159	(SC) 0.9
	Cross	10 (2)	3	30.0	515	172	
Acc20	Self	12 (2)	3	25.0	342	114	(SC) 2.0
	Cross	24 (3)	11	45.8	623	57	
Acc3	Self	6 (3)	0	0.0	0	0	ND
	Cross	16 (2)	6	37.5	508	85	
GS-Sel	Self	7 (2)	5	71.4	1,152	230	(SC) 1.2
	Cross	6 (1)	4	66.7	764	191	
HTT-Sel	Self	6 (2)	4	66.7	790	198	(SC) 1.7
	Cross	12 (2)	9	75.0	1,034	115	
IP12	Self	18 (3)	15	83.3	1,932	129	(SC) 1.0
	Cross	5 (1)	2	40.0	263	132	
IP7	Self	12 (2)	10	83.3	1,044	104	(SC) 0.5
	Cross	8 (2)	6	75.0	1,250	208	
IP8-Sel	Self	17 (3)	14	82.4	1,400	100	(SC) 0.5
	Cross	12 (2)	11	91.7	2,004	182	
MLSF12	Self	12 (2)	10	83.3	1,777	178	(SC) 1.4
	Cross	9 (2)	7	77.8	867	124	
MLSF17	Self	10 (2)	7	70.0	655	94	(SC) 0.6
	Cross	10 (2)	8	80.0	1,186	148	
MLSF27	Self	7 (1)	5	71.4	480	96	(SC) 1.2
	Cross	19 (3)	17	89.5	1,370	81	
PS	Self	12 (2)	7	58.3	760	109	ND
	Cross	6 (2)	0	0.0	0	0	
RWSF2	Self	22 (3)	17	77.3	1,875	110	(SC) 1.0
	Cross	12 (3)	10	83.3	1,109	111	
RWSF3	Self	22 (3)	16	72.7	1,304	82	(SC) 1.0
	Cross	22 (4)	18	81.8	1,481	82	
Site94	Self	11 (2)	6	54.5	564	94	(SC) 2.6
	Cross	5 (1)	4	80.0	145	36	
ST73-3	Self	14 (2)	10	71.4	1,188	119	(SC) 1.8
	Cross	6 (1)	5	83.3	327	65	
ST93-1	Self	26 (4)	22	84.6	1,270	58	(SC) 0.5
	Cross	16 (3)	12	75.0	1,362	113	
UGSF13	Self	5 (1)	2	40.0	108	54	(SC) 0.5
	Cross	5 (1)	4	80.0	447	112	
UGSF17-Sel	Self	17 (3)	14	82.4	950	68	(SC) 0.4
	Cross	10 (2)	9	90.0	1,473	164	
UGSF19	Self	10 (2)	8	80.0	1,460	183	(SC) 1.3
	Cross	8 (2)	7	87.5	969	138	
UGSF2-Sel	Self	15 (3)	14	93.3	1,546	110	(SC) 1.1
	Cross	18 (3)	10	55.6	1,038	104	
UGSF26	Self	37 (5)	30	81.1	2,273	76	(SC) 0.5
	Cross	28 (4)	22	78.6	3,119	142	
UGSF29	Self	5 (1)	2	40.0	344	172	ND
	Cross	4 (1)	4	100.0	0	0	

SC, self-compatible; ND, not determined.

only for the AFLP) and separated on 6% polyacrylamide gels using a Li-Cor 4300 DNA analyser (Li-Cor Biotechnology Co. Ltd., Bad Homburg, Germany).

2.7 | Data analysis

The band patterns for both AFLP and SSR markers were evaluated manually and scored dominantly as either present (1), absent (0) or ambiguous (3) across all genotypes, and the data were transformed into a 0/1 binary matrix for analysis. A band was considered polymorphic if present in at least one of the genotypes and absent in the others and vice versa. The discriminatory power of the AFLP markers was evaluated using the number of polymorphic bands and the percentage of polymorphism, the polymorphic information content (PIC) and the resolving power. The PICs for AFLPs were calculated by applying the formula $PIC_i = 2f_i\ (1-f_i)$, where f_i is the frequency of the amplified allele (band present) and $(1-f_i)$ is the frequency of the null allele (band absent) of marker *i* (Powell et al. 1996). The resolving power was calculated as $R = \Sigma I_b$, where $I_b = 1-(2 \times [0.5-p])$, and *p* is the proportion of the genotypes containing the band (Prevost and Wilkinson 1999).

The marker quality of the SSR markers was described using PIC values and expected heterozygosity (H_e). The PIC values were calculated as $PIC_i = 1-\Sigma P_{ij}^2$ according to Anderson et al. (1993), where P_{ij} is the frequency of the *j*th pattern for marker *i* for *n* patterns, while the heterozygosity was calculated as $H_e = 1-\Sigma P_i^2$, where P_i is the frequency of the *i*th allele (Nagy et al. 2012). The calculations were performed using an online program, PICcalc (https://www.liverpool.ac.uk/~kempsj/pic.html), after calculating the frequencies of the markers using XLSTAT software ver. 18.6 (https://www.xlstat.com/en/).

The cluster analysis for the pooled data from AFLP and SSR presence/absence fragments was carried out based on genetic distances (GD) calculated as GD = 1−Jaccard similarity, and a tree was built using the neighbour-joining method in DARwin ver. 6 software (Perrier & Jacquemoud-Collet, 2006). The robustness of the dendrogram was estimated using bootstrap analysis (Felsenstein, 1985) with 1,000 replicates. Bootstrap values larger than 50% were included in the resulting dendrogram. A principle coordinate analysis (PCoA) was applied to graphically represent the relationship structure in the spider plant germplasm using the same software.

3 | RESULTS

3.1 | Ploidy levels and genome size analysis

A chromosome number of $2n = 34$ was determined in the root tips of the entry GS-Sel (Figure 1). Three consistent counts of this number were obtained from different slides. The nuclear DNA content estimates for the entries ranged from 2.31 ± 0.00 to 2.45 ± 0.01 pg/2C (Table 2, Figure 2). Overall, there were no significant differences for the genome sizes of the *C. gynandra* entries studied. All entries are therefore assumed to be diploid with a chromosome number of $2n = 34$ as the entries had DNA contents in the range of the standard entry GS (2.35 pg/2C) used for the chromosome count.

TABLE 4 Characteristics of 11 AFLP markers, including the numbers of scored and polymorphic bands, polymorphic information content and the resolving power in 30 entries of *C. gynandra* assayed on five genotypes per entry

Primer combination	NSB	NPB	PIC	RP
*Hind*ATA/*Mse*CTT	13	13 (100%)	0.31	19.5
*Hind*ATA/*Mse*GTA	33	27 (81.8)	0.28	42.6
*Hind*ATA/*Mse*GAC	20	20 (100%)	0.29	20.9
*Hind*ATA/*Mse*TCT	20	20 (100%)	0.35	18.1
*Hind*ATA/*Mse*CGT	38	38 (100%)	0.33	44.2
*Hind*ATA/*Mse*CGA	24	24 (100%)	0.31	27.6
*Hind*CAT/*Mse*GCG	35	32 (91.4%)	0.27	31.0
*Hind*CAT/*Mse*GTT	49	37 (75.5%)	0.23	69.5
*Hind*CAT/*Mse*GGT	50	40 (80%)	0.28	67.6
*Hind*CAT/*Mse*GTG	74	74 (100%)	0.33	58.5
*Hind*CAT/*Mse*GTA	97	97 (100%)	0.33	98.4
Average	41	38	0.3	45.3
Total	453	422		

NSB, number of scored bands; NPB, number of polymorphic bands (percentage of polymorphic band); PIC, polymorphic information content; RP, resolving power.

3.2 | Pollination study

No pronounced differences in pollen tube growth in situ between self- and cross-pollinations were observed (Figs S1 and S2). The ISIs ranged from 0.4 for entry UGSF17-Sel to 2.6 for entry Site94 (Table 3). For the entries Acc3, PS and UGSF29, no ISI could be determined as either the self- or the cross-pollinations for these entries failed. For entries whose ISIs were determined, both self- and cross-pollinations produced fruits and/or seeds, and the average number of seeds per fruit for both self- and cross-pollinations was comparable.

3.3 | Genetic marker analyses

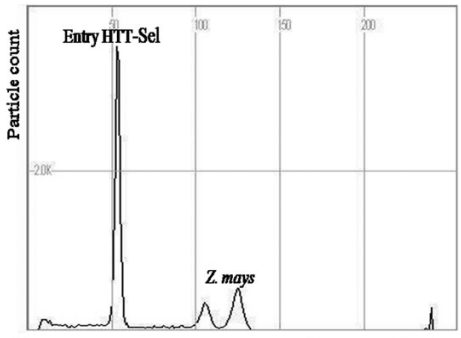
Eleven AFLP primer combinations revealed a total of 453 bands, and 422 (93.2%) were polymorphic between all the 150 genotypes analysed. The average number of scorable bands per primer combination was 41, ranging from 13 (*Hind*ATA/*Mse*CTT) to 97 (*Hind*CAT/*Mse*GTA). The PIC values ranged from 0.23 to 0.35 with an average of 0.30, whereas the resolving power ranged from 19.5 to 98.4 with an average of 45.3 (Table 4).

From the 37 SSR primer pairs, only nine amplified polymorphic products were in the expected band size range. Three of the remaining 28 primer pairs were monomorphic, while the rest did not amplify products in the expected size ranges (13 primers) or any product at all (12 primers) (Table S1). The SSR primer pairs were also

TABLE 5 Characteristics of nine SSR markers, including the number of alleles, PIC and heterozygosity among the 30 *C. gynandra* entries

SSR	Primer sequence		Core motif	Ta (°C)	Expected size (bp)	No. of alleles	H	PPB	PIC
CG001	F	CGTCAGTAGCATTTGGTTCG	(AG) 20	57	215	5	0.78	100	0.74
	R	TTCCAATACAAAGGGTGACAAC							
CG017	F	TTTGAAGTGGCAACAGCGTA	(AACCCTA) 5	60	205	6	0.75	100	0.71
	R	AATGGATTTGGTTCATGTGG							
CG018	F	CGAAATGCTTCACTTGCTCA	(AACCCT) 6	60	276	6	0.66	100	0.59
	R	CCTTCTTCATTCCCAAACGA							
CG022	F	ATGGGCTTTCCGTTTTTCAT	(CAACAC) 7	60	227	4	0.54	100	0.49
	R	CGCTTCCATGGACTGGTAAT							
CG024	F	GGATGCAATTGTACAGCTCG	(TTGTGACCT) 4	60	254	5	0.52	100	0.48
	R	ATGGCGTATGGGTTGAAGAT							
CG027	F	ATATTTGTGTGGGGTGGCTG	(GAATGCTT) 3	60	200	3	0.18	100	0.17
	R	ATTGGAGGCAAACGAATGAG							
CG028	F	ACCTTCGTTTTTGTTGTCGG	(TAGAATTT) 3	63	270	4	0.68	100	0.62
	R	ATCAATTCTCCTGCGCAAAC							
CG032	F	GGGCCTGCAAAAACAAATAA	(AGACC) 7	60	221	6	0.71	100	0.66
	R	TGGACAGATTTTCTGGTGGA							
CG033	F	CCTTAACGATCACGCATTCA	(ATATA) 8	60	184	7	0.73	100	0.69
	R	CTCAACGTTCCACCTCCAAC							
Mean				–	–	5	0.60	100	0.68
Total				–	–	46	–	–	–

H, heterozygosity; PPB, percentage polymorphic bands; PIC, polymorphism information content; F, forward primer; R, reverse primer; Ta, annealing temperature.



	Particles	Mean	CV	Median
Entry HTT	13402	55.61	3.3	53
Z. mays	3839	127.3	2.4	125

FIGURE 2 Flow cytometric histogram for *C. gynandra* entry HTT-Sel cochopped with the standard *Z. mays*

considered single locus markers as they produced one or two fragments for each individual genotype analysed. The primer sequences, allele numbers, heterozygosity and PIC values are shown in Table 5. A total of 46 alleles at nine loci were scored with an average of five alleles, which ranged from three to six alleles. The heterozygosity (H) revealed by the SSR markers ranged from 0.18 to 0.78, whereas the polymorphic information content (PIC) calculated from the allele frequencies ranged from 0.17 to 0.74 with an average of 0.68 (Table 5).

Pairwise genetic distances based on the pooled data from AFLPs and SSRs ranged from 0.13 to 0.77 among the genotypes over all entries with an average of 0.47 (data not shown). The average genetic distance within and among entries varied from 0.24 (within entry RWSF3) to 0.53 (between entries GS-Sel and Abuku1, Acc30, Acc5 and Acc6) with an average of 0.42 (Table 6). The genotypes within entries also showed varying values for average genetic distances, with entry GS-Sel showing the highest value at 0.36 but varying between 0.31 and 0.64, and the lowest average value for RWSF3 at 0.24, with a range of 0.23–0.39 (Table 7). The genetic distance coefficients from the pooled data of AFLP and SSR markers were used for the PCoA (Figure 3). PCoAs 1 and 2 accounted for 9.42 and 7.66% of the total variation, respectively (data not shown). The 30 *C. gynandra* entries were grouped into two major groups (Figure 3), which was also observed in the dendrogram (Fig. S3). Cluster I consisted of the farmers' cultivars (Acc3, Acc5, Acc6, Acc20, Acc21, Acc26, Acc28 and Acc30). However, advanced breeding lines Abuku1 and Abuku2 were also clustered within the Kenyan farmers' cultivars. Cluster II consisted of the gene bank accessions and the advanced lines developed at WorldVeg, Arusha, Tanzania, through original accessions from different African countries. The grouping of entries in both clusters was supported by low bootstrap

TABLE 7 Mean, minimum and maximum genetic distance within each of the 30 entries

Entry	Mean	Minimum	Maximum
Abuku1	0.30	0.35	0.42
Abuku2	0.30	0.31	0.45
Acc20	0.33	0.47	0.31
Acc21	0.31	0.43	0.31
Acc26	0.30	0.31	0.40
Acc28	0.25	0.29	0.35
Acc30	0.26	0.25	0.36
Acc3	0.28	0.31	0.40
Acc5	0.29	0.31	0.42
Acc6	0.31	0.32	0.46
HTT-Sel	0.30	0.32	0.45
GS-Sel	0.36	0.31	0.64
IP7	0.28	0.23	0.48
PS	0.33	0.34	0.50
Site94	0.29	0.29	0.38
ST73-3	0.27	0.23	0.45
ST93-1GS	0.33	0.30	0.49
IP12	0.29	0.32	0.42
IP8-Sel	0.26	0.12	0.39
MLSF12	0.27	0.30	0.38
MLSF17	0.33	0.33	0.51
MLSF27	0.27	0.32	0.38
RWSF2	0.25	0.25	0.37
RWSF3	0.24	0.23	0.39
UGSF13	0.26	0.24	0.37
UGSF17-Sel	0.34	0.12	0.59
UGSF19	0.26	0.28	0.40
UGSF2-Sel	0.33	0.30	0.54
UGSF26	0.27	0.23	0.42
UGSF29	0.26	0.29	0.40

The genetic distances (GD) were calculated as GD = 1−Jaccard similarity. Colour codes represent countries of origin as shown in Table 1; these colours are also used in the dendrogram: White – Kenya; Green – Tanzania; Blue – South Africa; Purple – Malawi; Brown – Rwanda; and Red – Uganda.

values (<50%) in the dendrogram (Fig. S3), suggesting that the entries could not be clearly separated according to their origin.

4 | DISCUSSION

4.1 | Chromosome numbers, genome size and ploidy level in the analysed *C. gynandra* entries

Chromosome numbers, genome size and ploidy levels are important in characterizing germplasm for both classification and breeding purposes, as well as marker development. A chromosome number of $2n = 34$ was observed in this study for entry GS-Sel. Chweya and Mnzava (1997) reported chromosome numbers of $2n = 18, 20, 22, 32$ and 34 for samples of the genus *Cleome* from Asia and South America. A study by Inda et al. (2008) suggested that the ancestral basic chromosome number of the genus *Cleome* was $x = 10$ with occasional reduction to $x = 8$ and $x = 9$ in most of the old world *Cleome*. Higher chromosome base numbers from North American section *Peritoma*, South American section *Melidiscus* and *C. gynandra* might have been due to parallel duplication of this base number to $x = 20$ and later karyotype reduction to $x = 16–17$, where *C. gynandra* is classified. Thus, the actual chromosome number of $2n = 34$ points to a polyploidization event in *C. gynandra* suggesting the species to be allotetraploid. However, a disomic inheritance remains to be proven, although indirect evidence may be derived from our SSR marker data showing only two alleles per genotype at maximum. Our study is in agreement with earlier findings by Áskell (1976), Renard et al. (1983) and Koshy and Mathew (1985) who reported chromosome numbers of $2n = 34$ in *C. gynandra* collections from Burundi and India.

Flow cytometry (FCM), a method that involves staining particles with fluorescent dyes, has been used as the method of choice for genome size and ploidy level estimation in a number of studies (Doležel et al. 2007) because it is easy, convenient and much faster compared to the classical chromosome counting techniques. However, FCM only gives relative information, and a standard with a known chromosome number is necessary for the estimation of the ploidy of any unknown genotype of the same species (Pellicer and Leitch 2014). The ploidy levels of the entries in this study were estimated by comparing peak positions to the standard entry (GS-Sel) with a determined chromosome number. Based on this, the entries are all diploid with a chromosome number of $2n = 34$ and with a mean genome size of 2.38 pg/2C. This is the first estimation of the value (DNA content) for *C. gynandra*; only a rough estimation of the genome size at approximately 1 Gb was carried out (=approximately 0.978 pg/1C following Doležel et al. 2003), and a slightly smaller genome size compared to this study was published previously (van den Bergh et al. 2014).

Chromosome numbers and genome sizes are often stable within species, but some exceptions have been noted; for instance, spring beauty (*Claytonia virginica*) has chromosome numbers that range between $2n = 12$ and 37, and some different cytotypes have even been found in the same population (Lewis and Semple 1977). From our results, the studies of *C. gynandra* entries are considered to have a small but constant genome size when compared with the mean of 5.9 pg/1C in angiosperms (Suda et al. 2015). Vidic et al. (2009) and others have shown that species with small genomes are able to withstand extreme environments compared those with large genome sizes. This could also be a reason for the wide distribution of the *C. gynandra* in many ecological zones in Kenya and why it is considered a weedy plant. Several applications of genome size data are as follows: giving a higher accuracy in taxonomy as a supplement for karyological data, in hybrid identification, where parental species are expected to differ sufficiently in genome size, and as a tool for biological security to identify and manage the spread or introduction of invasive species (Suda et al., 2015). In addition, crosses between

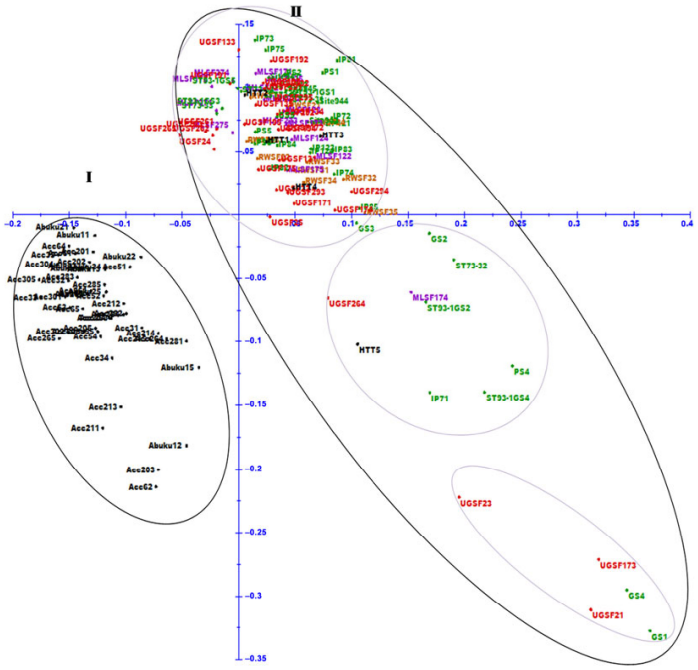


FIGURE 3 Distribution of the 150 genotypes of 30 entries of *C. gynandra* into two major groups based on the first (9.42%) and second (7.66%) coordinates from principal coordinate analysis calculated from the pooled data from eleven AFLP primer combinations and nine microsatellite (SSR) markers

genotypes of different ploidy levels are often associated with low levels of fertility; therefore, information about the ploidy levels is a prerequisite for the analysis of reproductive barriers in breeding programmes.

4.2 | *Cleome gynandra* is self-compatible

The breeding biology in *C. gynandra* has not been studied before in depth, yet this is an important prerequisite to determine the possibility of using different breeding strategies. Through pollination experiments, the extent or presence of self-incompatibility (SI) can be identified. SI is the inability of a plant to produce seed following self-pollination and can be broadly classified as either gametic SI (determined by the genotype of the haploid pollen grain) or sporophytic SI (determined by the genotype of the diploid anther from which the pollen originates) (Jacquemart 2007). Our results show a possibility of seed set after selfing, and average seed numbers per fruit were balanced in self- and cross-pollinations for the entries used (Table 3). Based on the ISI values, the accessions are self-compatible or at least partially self-compatible following the classification used by Lloyd and Schoen (1992) and Arroyo et al. (1993). The floral structure in *C. gynandra*, with its open flowers exposing the anthers and stigmas, colourful petals for pollinator attraction, sticky pollen and elongated stigma may favour cross-pollination more than self-pollination in the open fields. Morphological variations observed in most of the entries also point to the preferential outbreeding in this species. Therefore, isolation in the field and greenhouse is necessary for maintaining advanced breeding lines and gene bank accessions.

4.3 | Genetic diversity

The diversity within *C. gynandra* has been studied previously using morphological and physiological traits (Kiebre et al. 2015; Masuka et al. 2012; Muthoni et al. 2010; Onyango et al. 2016; Stoilova et al. 2015; Wasonga et al. 2015). These studies involved various genotypes from different sources. These kinds of morphological and physiological traits are still useful to distinguish adult plants in the field or to identify unique individuals for selection. However, they are greatly affected by environmental factors and the developmental stage of the plant; hence, classification schemes relying on these visible traits may not be as accurate. Further more, association with genetic markers is therefore necessary to provide more accurate and meaningful information.

In this study, we have been able to establish suitable AFLP primer combinations with high resolving power and PIC values. A higher number of bands were obtained with *HindIII*CAT primers compared to those with *HindIII*ATA primers. Most of the SSR markers have also shown good values for heterozygosity and PIC values therefore representing a good basis for studying the populations of *C. gynandra*. These SSR markers have been developed based on a genomic DNA sequencing approach and are the first to be published in *C. gynandra*. We have pooled data from AFLP and SSR markers to analyse the genetic relations among selected *C. gynandra* entries. The pooling of the data was necessary as only a few SSR markers were available and could not resolve the entries with reliable bootstrap values. The farmers' cultivars in this study were grouped apart from the other entries, which were either advanced lines or gene bank accessions (Figure 3). Advanced lines Abuku1 and Abuku2 clustered together with the

Kenyan farmers' cultivars. This could be due to their development from some of the Kenyan farmers' cultivars. Diversity was also observed among genotypes within and among the farmers' cultivars from different regions in Kenya (Fig. S3). K'Opondo et al. (2009) explained that the diversity among some landraces of *C. gynandra* from the Kakamega and Uasin Gishu districts in western Kenya could be as a result of semi cultivation by farmers without selection. The separate clustering of the farmers' cultivars may also be an indication of rare alleles lost during the selection of advanced lines, which might be useful for the breeding of *C. gynandra*. Advancing on-farm conservation as a way of retaining unique germplasm to guard against the challenges of climatic changes has been proposed by Bellon and Van Etten (2013) and might help to maintain rare alleles found in the farmers' cultivars.

No clear separations according to the entries nor to the countries of origin were noticed, which was shown by the low bootstrap support for the clusters (Fig. S3). In contrast, relatively high variation was observed among the individual genotypes. This may be due to high heterozygosity and variability between the individual genotypes of each of the entries combined with a low degree of differentiation of the entries as no advanced improvement programmes have been carried out on them. The low differentiation among the entries might have been as a result of a low level of selection that was insufficient to reduce the diversity of each of the entries and/or an admixture due to lack of isolation during multiplication in the open fields at WorldVeg. Moreover, bees are common at WorldVeg in some seasons, facilitating cross-pollination. *C. gynandra* is bee-pollinated in the open fields, and Levin (1981) observed that the genotypes of species that are bee-pollinated in a population may share more alleles because pollen are easily transferred among them thereby making them genetically similar. The variations among these entries could presumably be better resolved with a higher number of markers, preferably those that have a good distribution throughout the genomes, such as AFLP and SNPs, combined with a larger number of genotypes.

5 | CONCLUSION

The results of this study for genome sizes and ploidy levels combined with the genetic marker analysis presented useful baseline information for the genetic improvement and germplasm conservation of *C. gynandra*. No variation in ploidy level could be detected, and thus, all accessions are crossable. The C-value of 1.19 pg/1C points to a rather small genome size. Breeding programmes involving selfing steps can be utilized in *C. gynandra* as no or only low levels of self-incompatibility were found. The highly informative SSR and AFLP markers established in this study may be of use in further research and breeding of the crop. The genetic diversity analysis shows relatively high variation within the entries, which resulted in poor resolution among the entries in the dendrogram. This indicates that improvements are needed for germplasm maintenance in the gene banks, whereas the farmers' cultivars were considered a source of genetic variation, and the inclusion of more collections in ex situ conservation would be desirable. Furthermore, it demonstrates that there is genetic variability even in advanced lines that would be useful in future breeding programmes. The maintenance of this genetic diversity is also necessary for a stronger form of selection to avoid genetic erosion.

ACKNOWLEDGEMENTS

The authors are grateful for the funding within the project HORTINLEA by the German Federal Ministry of Education and Research (BMBF) and the German Federal Ministry of Economic Cooperation and Development (BMZ) within the framework of the GlobE-Global Food Security. The authors thank Diro Terefe for assisting in SSR marker designing, David Odongo for farmer's cultivar collection and WorldVeg for supplying plant material.

REFERENCES

Abukutsa-Onyango, M. (2007). The diversity of cultivated African leafy vegetables in three communities in Western Kenya. *African Journal of Food Agriculture Nutrition and Development*, *7*(3), 1–15.

Agarwal, M., Shrivastava, N., & Padh, H. (2008). Advances in molecular marker techniques and their applications in plant sciences. *Plant Cell Reports*, *27*(4), 617–631.

Anderson, J. A., Churchill, G. A., Autrique, J. E., Tanksley, S. D., & Sorrells, M. E. (1993). Optimizing parental selection for genetic linkage maps. *Genome*, *36*(18), 1–186.

Arroyo, M. T. K., & Uslar, P. (1993). Breeding systems in a temperate mediterranean-type climate montane sclerophyllous forest in central Chile. *Botanical Journal of the Linnean Society*, *111*, 83–102.

Áskell, L. (1976). IOPB chromosome number reports LIII. *International Association for Plant Taxonomy (IAPT)*, *25*(4), 483–500.

Bellon, M. R., & Van Etten, J. (2013). Climate change and on-farm conservation of crop landraces in centres of diversity. *Plant Genetic Resources and Climate Change*, 137–150.

Van den Bergh, E., Külahoglu, C., Bräutigam, A., Hibberd, J. M., Weber, A. P. M., Zhu, X.-G., & Schranz, M. E. (2014). Current plant biology gene and genome duplications and the origin of C4 photosynthesis : Birth of a trait in the Cleomaceae. *Current Plant Biology*, *1*, 2–9.

Chweya, J. A., & Mnzava, N. A. (1997). *Cat's whiskers. Cleome gynandra L. Promoting the conservation and use of underutilized and neglected crops*. vol. *11*. Rome, Italy: Institute of Plant Genetics and Crop Plant Research, Gatersleben/International Plant Genetic Resources Institute.

Doležel, J., Bartoš, J., Voglmayr, H., & Greilhuber, J. (2003). Letter to the editor. *Cytometry*, *51A*(2), 127–128.

Doležel, J., Greilhuber, J., & Suda, J. (2007). Estimation of nuclear DNA content in plants using flow cytometry. *Nature Protocols*, *2*(9), 2233–2244.

Felsenstein, J. (1985). Confidence limits on phylogenies: An approach using the bootstrap. *Evolution*, *39*(4), 783–791.

Inda, L. A., Torrecilla, P., Catalan, P., & Ruiz-Zapata, T. (2008). Phylogeny of *Cleome* L. and its close relatives *Podandrogyne* Ducke and *Polanisia* Raf. (Cleomoideae, Cleomaceae) based on analysis of nuclear ITS sequences and morphology. *Plant Systematics and Evolution*, *274*, 111–126.

Van Jaarsveld, P., Faber, M., van Heerden, I., Wenhold, F., van Rensburg, W. J., & van Averbeke, W. (2014). Nutrient content of eight African leafy vegetables and their potential contribution to dietary reference intakes. *Journal of Food Composition and Analysis*, *33*(1), 77–84.

Jacquemart, A. (2007). Methods for determining compatibility and pollinator efficiency in temperate fruit species. *Fruit, Vegetable and Cereal Science and Biotechnology*, *1*(1), 26–38.

Kiebre, Z., Bationo/Kando, P., Sawadogo, N., Sawadogo, S., & Zongo, J.-D. (2015). Selection of phenotypic interests for the cultivation of the plant *Cleome gynandra* L. in the vegetable gardens in Burkina Faso. *Journal of Experimental Biology and Agricultural Sciences*, *3*(3), 288–297.

Kirov, I., Divashuk, M., van Laere, K., Soloviev, A., & Khrustaleva, L. (2014). An easy "SteamDrop" method for high quality plant chromosome preparation. *Annals of Missouri Botanical Gardens*, *68*(2), 233–253.

K'Opondo, F. B. O., Van Rheenen, H. A., & Muasya, R. M. (2009). Assessment of genetic variation of selected spiderplant (*Cleome gynandra* L.) morphotypes from western Kenya. *African Journal of Biotechnology*, *8* (18), 4325–4332.

Koshy, J. K., & Mathew, P. M. (1985). Cytology of the genus Cleome Linn. *Cytologia*, *50*, 283–287.

Levin, D. A. (1981). Dispersal versus gene flow in plants. *Annals of Missouri Botanical Gardens*, *68*(2), 233–253.

Lewis, W. H., & Semple, J. C. (1977). Geography of *Claytonia virginica* cytotypes. *American Journal of Botany*, *64*(9), 1078–1082.

Lloyd, D. G., & Schoen, D. J. (1992). Self- and cross-fertilization in plants. I. functional dimensions. *Journal of Plant Science*, *153*(3), 358–369.

da Maia, L. C., Palmieri, D. A., de Souza, V. Q., Kopp, M. M., de Carvalho, F. I. F., & de Oliveira, A. C. (2008). SSR locator: Tool for simple sequence repeat discovery integrated with primer design and PCR simulation. *International Journal of Plant Genomics*, *2008*, 412696.

Masuka, A., Goss, M., & Mazarura, P. (2012). Morphological characterization of four selected spider plant (*Cleome gynandra* L.) morphs from Zimbabwe and Kenya. *Asian Journal of Agriculture and Rural Development*, *2*(4), 646–657.

Muthoni, J., Nyamongo, D. O., & Silingi, M. (2010). Participatory characterization and evaluation of some African leafy vegetables in Lari, Kiambu West District, Central Kenya. *Journal of Horticulture and Forestry*, *2*(2), 12–16.

Nagy, S., Poczai, P., Cernák, I., Gorji, A. M., Hegedus, G., & Taller, J. (2012). PICcalc: An online program to calculate polymorphic information content for molecular genetic studies. *Biochemical Genetics*, *50*, 670–672.

Odhav, B., Beekrumb, S., Akulaa, U., & Baijnath, H. (2007). Preliminary assessment of nutritional value of traditional leafy vegetables in KwaZulu-Natal, South Africa. *Journal of Food Composition and Analysis*, *20* (5), 430–435.

Omondi, E. O., Debener, T., Linde, M., Abukutsa-Onyango, M., Dinssa, F. F., & Winkelmann, T. (2016). Molecular markers for genetic diversity studies in African leafy vegetables. *Advances in Bioscience and Biotechnology*, *7*, 188–197.

Onyango, C. M., Kunyanga, C. N., Ontita, E. G., Narla, R. D., & Kimenju, J. W. (2013). Current status on production and utilization of spider plant (*Cleome gynandra* L.) an underutilized leafy vegetable in Kenya. *Genetic Resources and Crop Evolution*, *60*, 2183–2189.

Onyango, C. M., Onwonga, R. N., & Kimenju, J. W. (2016). Assessment of spider plant (*Cleome gynandra* L.) germplasm for agronomic traits in vegetable and seed production: A green house study. *American Journal of Experimental Agriculture*, *10*(1), 1–10.

Pellicer, J., & Leitch, I. J. (2014). The application of flow cytometry for estimating genome size and ploidy level in plants. *Molecular Plant Taxonomy: Methods and Protocols*, *1115*, 279–307.

Perrier, X., & Jacquemoud-Collet, J. P. (2006). *DARwin software*. Retrieved from http://darwin.cirad.fr/darwin

Powell, W., Morgante, M., Andre, C., Hanafey, M., Vogel, J., Tingey, S., & Rafalski, A. (1996). The utility of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. *Molecular Breeding*, *2*, 225–238.

Prevost, A., & Wilkinson, M. J. (1999). A new system of comparing PCR primers applied to ISSR fingerprinting of potato cultivars. *Theoretical and Applied Genetics*, *98*, 107–112.

Renard, R., Lambinon, J., Reekmans, M., der Verken, P. V., & Govaert, M. (1983). Nombres chromosomiques de quelques Angiospermes du Rwanda, du Burundi et du Kenya. *Bulletin du Jardin botanique National de Belgique/Bulletin van de Nationale Plantentium va België*, *53*(3/4), 343–371.

Schoefeldt, H. C., & Pretorius, B. (2011). The nutrient content of five traditional South African dark green leafy vegetables – A preliminary study. *Journal of Food Composition and Analysis*, *24*, 1141–1146.

Schuelke, M. (2000). An economic method for the fluorescent labeling of PCR fragments. *Nature Biotechnology*, *18*, 233–234.

Stoilova, T., Dinssa, F. F., Ebert, A. W., & Tenkouano, A. (2015). The diversity of African leafy vegetables: Agromorphological characterization of subsets of AVRDC's germplasm collection. *Acta Horticulturae*, *1102*, 67–74.

Suda, J., Meyerson, L. A., Leitch, I. J., & Pyšek, P. (2015). The hidden side of plant invasions: The role of genome size. *New Phytologist*, *205*(3), 994–1007.

Tautz, D., & Renz, M. (1984). Simple sequences are ubiquitous repetitive components of eukaryotic genomes. *Nucleic Acid Research*, *12*(10), 4127–4138.

Uusiku, N. P., Oelofse, A., Duodu, K. G., Bester, M. J., & Faber, M. (2010). Nutritional value of leafy vegetables of sub-Saharan Africa and their potential contribution to human health: A review. *Journal of Food Composition and Analysis*, *23*(6), 499–509.

Vidic, T., Greilhuber, J., Vilhar, B., & Dermastia, M. (2009). Selective significance of genome size in a plant community with heavy metal pollution. *Ecological Applications*, *19*(6), 1515–1521.

Vos, P., Hogers, R., Bleeker, M., Reijans, M., van de Lee, T., Homes, M., Frijters, A., Pot, J., Peleman, J., Kuiper, M., & Zabeau, M. (1995). AFLP: A new technique for DNA fingerprinting. *Nucleic Acids Research*, *44*(21), 388–396.

Wasonga, D. O., Ambuko, J. L., Chemining'wa, G. N., Odeny, D. A., & Crampton, B. G. (2015). Morphological characterization and selection of spider plant (*Cleome gynandra*) accessions from Kenya and South Africa. *Asian Journal of Agricultural Sciences*, *7*(4), 36–44.

Williams, J. G. K., Kubelik, A. R., Livak, K. J., Rafalski, J. A., & Tingey, S. V. (1990). DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Research*, *18*(22), 6531–6535.

Winkelmann, T., Doil, A., Reinhardt, S., & Ewald, A. (2010). Embryo rescue. In M. Davey, & P. Anthony (Eds.), *Plant cell culture – essential methods* (pp. 79–95). Weinheim: Wiley VCH.

Zietkiewicz, E., Rafalski, A., & Labuda, D. (1994). Genome fingerprinting by simple sequence repeat (SSR)- Anchored polymerase chain reaction amplification. *Genomics*, *20*, 176–183.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Omondi EO, Debener T, Linde M, et al. Mating biology, nuclear DNA content and genetic diversity in spider plant (*Cleome gynandra*) germplasm from various African countries. *Plant Breed*. 2017;00:1–12. https://doi.org/10.1111/pbr.12485

Supplementary material

Table 3: List of designed SSR primers that were either monomorphic or amplified products out of the expected size range or no products.

primer	Forward sequence	Reverse sequence	Repeat motif	Product Size
GC002†	TATCGGTCGATCCCA ATTTC	ACACACACACACA CACACACAC	(CA)30	235
CG003φ	TACCCGACTAAGCC AACCAC	CAATCATGGGAC AAAAACCC	(TC)26	224
CG004φ	GAATAGGTATGGAGA TTGCTTGG	CCTCTTCCGTGAC CTTTTTG	(GA)21	287
CG005φ	GCCAAAGGGGAAAGA AGAAG	AATGTCGCCATTT CTGAAGC	(AG)25	217
CG006†	AGCAAGCTCGATCCC AAGTA	GGCTCATTTTTGAC ACCCCT	(AAT)20	241
CG007*	ATTGATTCTTGGGCA ACGAC	CGAAAAATTCTTG AGACCCAAC	(AAT)28	252
CG008φ	AGTAGATGCGAACCC ACACC	CAGTCACCCATCA CCCTTTT	(TTA)22	276
CG009φ	CTATGACCAATGCGA CCCTT	GGATGGTTCATAG AATGGCG	(CTT)24	243
CG010φ	AACATGTAATGCGGG GTACAA	GCCCACACGTCTT ACACAAA	(AAT)26	212
CG012*	TGTGGGAAAGTGGGA CATTT	ATGTGTTAAGTGT GCTGCCG	(AAG)26	189
CG013φ	GTGAAATTGGGTGAA ATGGG	CCCCTCCGTTAGA CAAACAA	(TGAAATTGGG)3	260
CG014φ	TTTAACCCGATTTCA CCCAA	TTTAACCCGATTTC ACCCAA	(ACCCAATTTC)3	254
CG015φ	TTCGTTATTTGTCCA GCCCT	AAATTTGCACGAA AGATGGC	(GAGCATG)6	268
CG016†	ACCACGTCCCAAGT CCCA	GGGACATGGTACA CGGACA	(GTCCCAT)4-(GTCCCAT)7	240
CG019*	GCAGACCCACTCAGT CCTGT	TGGATACCATCCA CTGACGA	(TGGCAA)7	255
CG020†	GGGCTTGCATTATCAC GTTT	CTCCATAAGGCTA GGCATCG	(TACTCA)6	133
CG021φ	TCCTCCTCCTCCTCCT TCTC	TTGATATGCTTCG TGCGTTC	(CTTCTC)11	223
CG023φ	TGGCTTTTGCAATCT CCTCT	AAGGAGGCCGAG ACCAAG	(CTTGGC)8	247
CG025φ	ACTTGGTGGTGATCT TTGGC	TGAGTCAAAGAGT CGCATGG	(CGACCCATG)9	262
CG026†	GGACATGTTGTCA ATCACATCG	TGGAATGCTAATG GGACAAA	(AAAAAATA)3	202
CG029†	CACAAGCACACAC ACACACG	GGCATTTGAAAA CCCATCAC	(ACACACGC)3	275
CG030†	CATGACATGTCCAG TGAGGG	GCATTCAGTTTCA ATCCCGT	(AAAAG)7	227
CG031†	GAGCCTTTATCTGG GGGAAC	CCTATGGCTGTGC CAACTTT	(TTGAA)6	275
CG034φ	AGGATGCCCGTTAC AAGTTC	TGCTGAAGATGAAA ACGACG	(TATAT)9	223
CG035†	CTTAGATGGCGA GAAATCGG	ATCATTCCCGTT TCCATTTG	(TAAA)9	238

CG036†	GCATACTTCACC ACGGCTTT	AGGCTTCATTCG ATGACTGG	(ATCT)8	242
CG037†	ATCGCTCGTTAC AAATTGCC	TTGTTCATTGGG ATTGGGAT	(TACA)9	252
CG038†	GCCGAGGTGATC ATTTTTGT	TCATAGGGGGAT CACAACCT	(ATAC)7	200

†- no products, *- monomorphic bands, ᵠ- products not in expected ranges

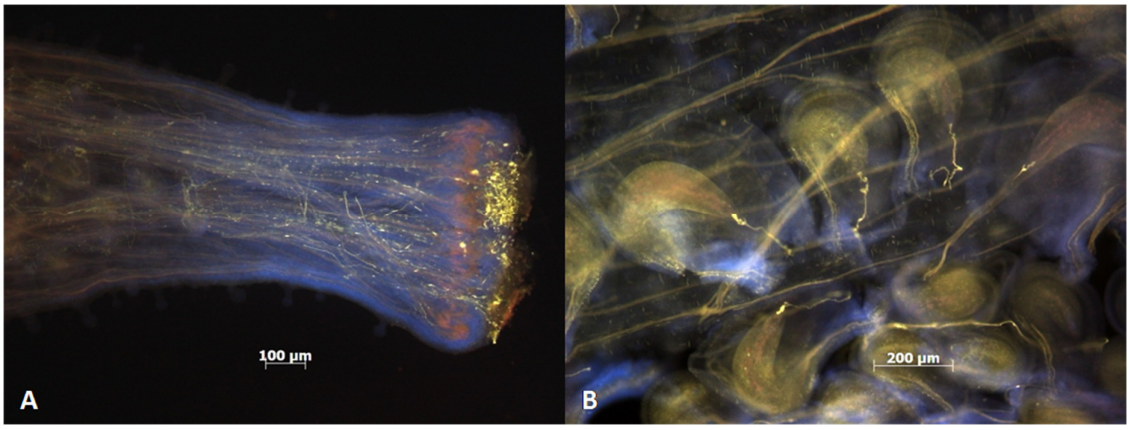


Figure 4: Pollen tube germination in the self-pollinated flower of spider plant entry UGSF2. A: Pollen tube growth on the stigma, B: Pollen tube entering ovule.

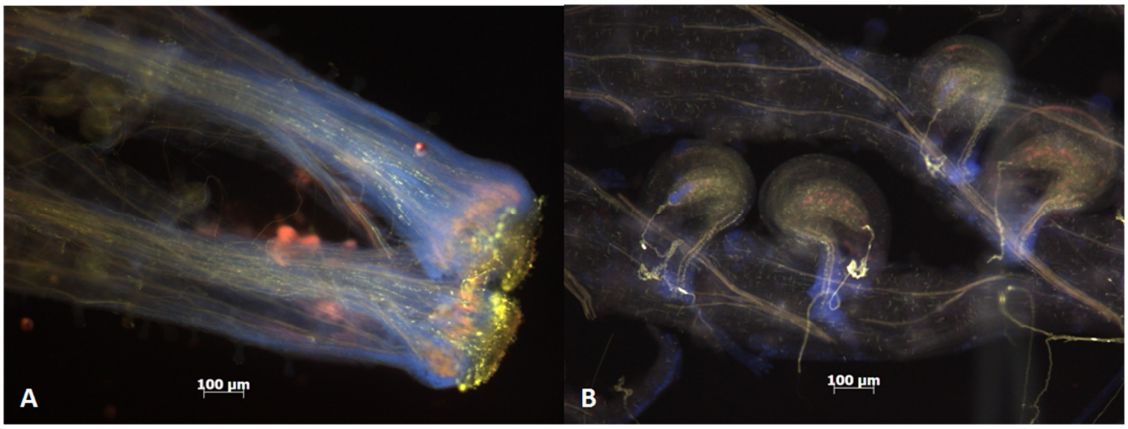


Figure 5: Pollen tube germination in the cross-pollinated flower of spider plant entry UGSF2. A: Pollen tube growth on the stigma, B: Pollen tube entering ovule.

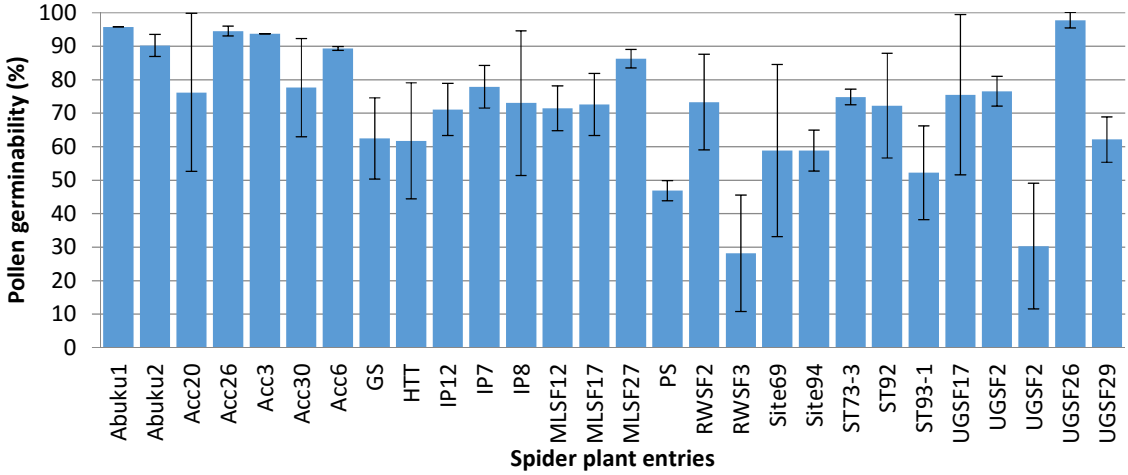


Figure 3: *In vitro* pollen germination percentages for the spider plant entries.

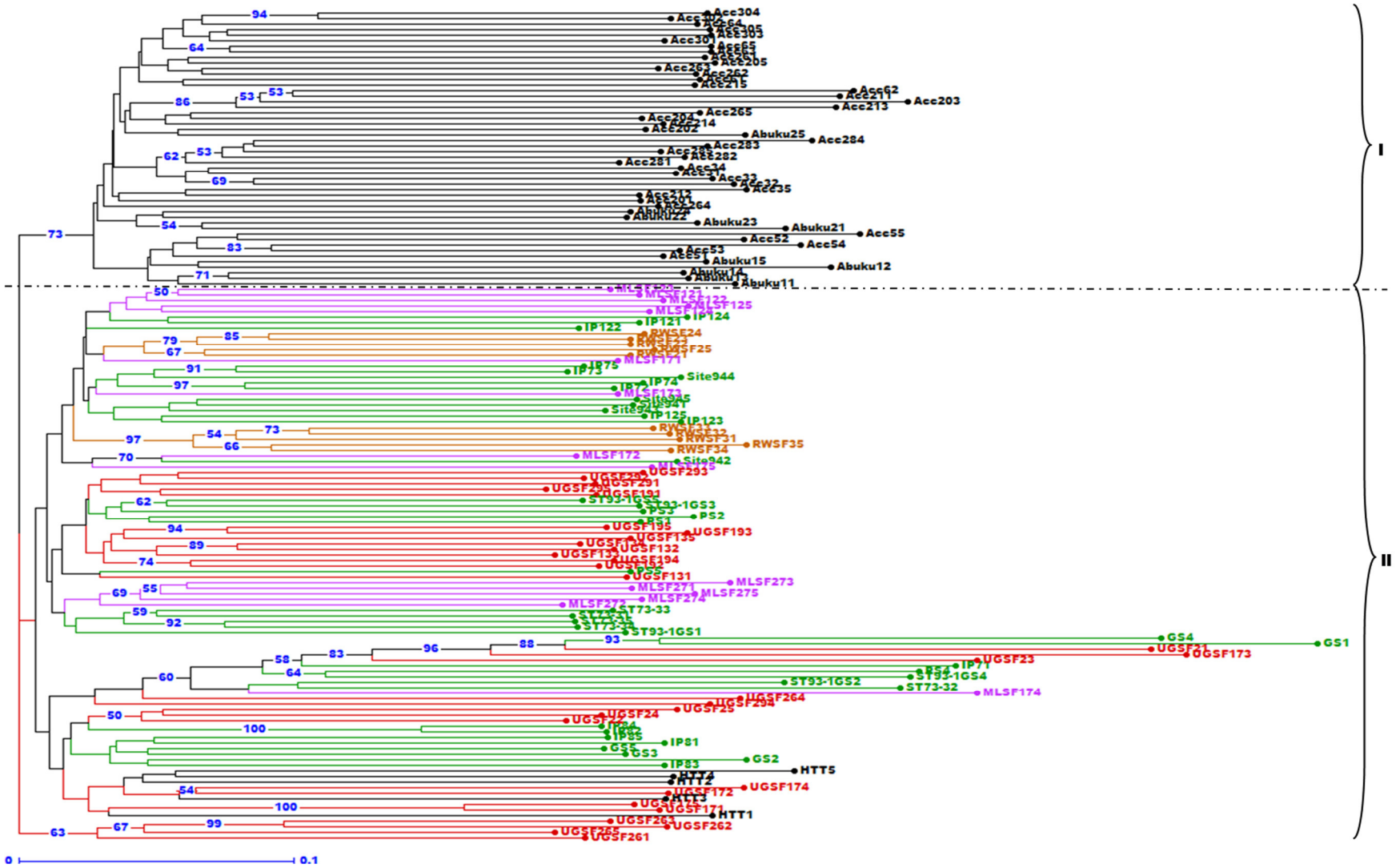


Figure 6: Dendrogram clustering of 150 genotypes of 30 entries of *C. gynandra* as revealed by cluster analysis based on Jaccard's genetic similarity coefficients calculated from the pooled data from eleven AFLP primer combinations and nine microsatellite (SSR) markers. Numbers at the nodes indicate the bootstrap values of 1000 iterations (branches lacking the value received <50% bootstrap support).

2.3. Nutritional compound analysis and morphological characterization of spider plant (*Cleome gynandra*) germplasm from six African countries.

Emmanuel O. Omondi[1], Christof Engels[2], Godfrey Nambafu[2], Monika Schreiner[3], Susanne Neugart[3], Mary Abukutsa[4], Traud Winkelmann[1]

[1]Institute of Horticultural Production Systems, Woody Plant and Propagation Physiology, Leibniz Universität Hannover, Hannover, Germany

[2]Department of Plant Nutrition and Fertilization, Humboldt University of Berlin, Berlin, Germany

[3]Leibniz-Institute of Vegetable and Ornamental Crops, Grossbeeren, Germany

[4]Department Horticulture, Jomo Kenyatta University of Agriculture & Technology (JKUAT), Juja, Kenya

Type of authorship	First author
Type of article	Research article
Contribution to the article	Designed and performed the experiments Analysed data, preparation of tables and figures, Wrote the manuscript
Contribution of other authors	Traud Winkelmann was involved in the conception of the study, preparation of the tables and figures and writing of the aims Monika Schreiner and Susanna Neugart assisted in glucosinolate and flavonoid analysis, data analysis, preparation of tables and figures, and writing the manuscript. Christof Engels and Godfrey Nambafu performed mineral analysis and writing of manuscript. Mary Abukutsa-Onyango provided the seeds, assisted in the field experiment and in writing the manuscript.
Journal	Food International Research
Date of publication	20th June 2017
DOI	10.1016/j.foodres.2017.06.050

Food Research International 100 (2017) 284–295

Contents lists available at ScienceDirect

Food Research International

journal homepage: www.elsevier.com/locate/foodres

Nutritional compound analysis and morphological characterization of spider plant (*Cleome gynandra*) - an African indigenous leafy vegetable

Emmanuel O. Omondi[a], Christof Engels[b], Godfrey Nambafu[b], Monika Schreiner[c], Susanne Neugart[c], Mary Abukutsa-Onyango[d], Traud Winkelmann[a,*]

[a] Institute of Horticultural Production Systems, Woody Plant and Propagation Physiology, Leibniz Universität Hannover, Hannover, Germany
[b] Department of Plant Nutrition and Fertilization, Humboldt University of Berlin, Berlin, Germany
[c] Leibniz Institute of Vegetable and Ornamental Crops, Grossbeeren, Germany
[d] Department Horticulture, Jomo Kenyatta University of Agriculture & Technology (JKUAT), Juja, Kenya

ARTICLE INFO

Keywords:
Glucosinolates
Flavonoids
Health promoting, Mineral concentration
Yield

ABSTRACT

Spider plant is among the important indigenous African leafy vegetables having the potential to contribute to food and nutritional security in sub-Saharan Africa. The main objective of this study was to quantify the mineral concentration, to identify and quantify glucosinolates and flavonoids in spider plant and further to characterize spider plant entries using important morphological traits. Thirty spider plant entries from different African countries, comprising of farmers' cultivars, gene bank accessions and advanced lines were grown in a field experiment and harvested for leaves, stems, flowers and siliques at different developmental stages. Five plant types based on the stem and petiole colorations were identified. Significant genotypic differences were shown for all the morphological traits except for 100 seed weight and silique weight. High mineral concentrations in the leaf tissue were observed especially for potassium, calcium, magnesium, phosphorus, iron, manganese and zinc. The aliphatic 3-hydroxypropyl glucosinolate was the main glucosinolate detected in all tissues with the highest concentrations in the reproductive organs. Glycosides of quercetin, kaempferol and isorhamnetin were the main flavonoids. Isorhamnetin glycosides were detected in trace amounts in both, leaves and inflorescences, while quercetin and kaempferol glycosides were the dominant flavonoids in the leaves and inflorescences, respectively. This knowledge of beneficial nutrient contents is an incentive for promoting spider plant consumption for improved human health while the morphological diversity analysis will be important for the further development of the spider plant germplasm.

1. Introduction

Spider plant (*Cleome gynandra* L. Briq), also known as cat's whisker or African cabbage belongs to the family *Cleomaceae* (Hall, Iltis, & Sytsma, 2004; Hall, Sytsma, & Iltis, 2002). It is among the main indigenous African leafy vegetables (HCDA, 2014) that form an important part of African traditional diets in the rural households where high levels of under-nutrition prevalence have been reported (FAO, 2010). These vegetables are also now popular in urban areas. Leaves, young shoots and sometimes inflorescences are consumed (Chweya & Nzava, 1997). Spider plant, like other African leafy vegetables, has been considered a seasonal food consumed mostly in periods of major food crop shortage or as supplements to major food crops in some parts of Africa (Tabuti, Dhillion, & Lye, 2004; van Rensburg et al., 2004). It has been reported to contain high levels of mineral elements such as calcium, iron, magnesium and zinc (Jinazali, Mtimuni, & Chilembwe, 2017; Nesamvuni, Steyn, & Potgieter, 2001; Odhav, Beekrum, Akula, & Baijnath, 2007). Furthermore, spider plant is adapted to tropical conditions, responds well to organic fertilizers and is able to tolerate biotic and abiotic stresses, is readily available as wild plant and does not require formal cultivation (Abukutsa-Onyango, 2007; Chweya & Nzava, 1997). Spider plant has also been used as traditional medicine to heal a range of ailments in different communities which include food poisoning, rheumatism, inflammation, toothache, headache, bacterial infections, snake bites among others (Chweya & Nzava, 1997; Meda, Bangou, Bakasso, Millogo-Rasolodimby, & Nacoulma, 2013; Mishra, Moharana, & Dash, 2011). These uses as traditional medicine hint at secondary plant metabolites that are health beneficial. Thus recently, interest has increased in the detailed analyses of secondary plant metabolites in African leafy vegetables (Neugart, Baldermann, Ngwene, Wesonga, & Schreiner, 2017). Cancer preventive potential of certain indigenous African leafy

* Corresponding author.
E-mail address: traud.winkelmann@zier.uni-hannover.de (T. Winkelmann).

http://dx.doi.org/10.1016/j.foodres.2017.06.050
Received 20 April 2017; Received in revised form 20 June 2017; Accepted 20 June 2017
Available online 22 June 2017

vegetable, e.g. Ethiopian Kale (*Brassica carinata*), has been reported (Odongo et al., 2017).

The most interesting secondary plant metabolites in spider plant are glucosinolates and flavonoids as recently demonstrated by Neugart et al. (2017). Glucosinolates are sulphur and nitrogen-containing secondary metabolites mainly in members of the order Brassicales (Hanschen, Lamy, Schreiner, & Rohn, 2014). They play a role in plant defense but are also known for their human health promoting properties depending on the hydrolysis products released following their hydrolysis by the plant's myrosinase enzyme (Latté, Appel, & Lampen, 2011).

Flavonoids are also a class of secondary metabolites in plants which have been described to have functions as antioxidants in higher plants experiencing environmental stress. They are the most common phenolic compounds in all plant parts, especially photosynthesizing cells, and flavonols are the most abundant in vegetables (Kumar & Pandey, 2013). Flavonoids act as growth regulators in plants and also combat oxidative stress in plants and humans (Agati, Azzarello, Pollastri, & Tattini, 2012). Flavonoids are health beneficial in several ways to humans due to their antioxidant, anti-inflammatory, anti-diabetic, antibacterial and sometimes even antiviral effects (Amić et al., 2017; Dueñas, Surco-Laos, González-Manzano, González-Paramás, & Santos-Buelga, 2011; Kumar & Pandey, 2013). Generally, spider plant contains high levels of total phenolic compound which are useful antioxidants (Chweya & Nzava, 1997; Kutsukutsa, Gasura, Mabasa, & Ngadze, 2014; Meda et al., 2013). Further knowledge of the qualitative composition and the contents of these health beneficial compounds in spider plant and other African leafy vegetables is necessary to serve as an incentive for greater consumption of these vegetables to promote healthy diets in the sub-Saharan population where the minimum daily intake recommended by WHO/FAO has not been reached (Ruel, Mino, & Smith, 2005).

Besides being cultivated for domestic consumption in home gardens and being collected from the wild, intercropping systems for large-scale production for urban markets is now taking ground since there is a growing demand for pro-health and micro-nutrient dense properties in food by consumers especially in the locally available vegetables (HCDA, 2014). However, poor seed quality and poorly performing farmers' cultivars are a major production constraint (Abukutsa-Onyango, 2007). Besides genetic diversity analyses (Omondi et al., 2017), morphological characterization of germplasm with diverse origin is, therefore, necessary to facilitate the development of superior and pure cultivars. Characterization will also help providing a basis for identifying cultivar-specific traits on nutrient and health protection.

Overall, the analysis of morphological diversity and nutrient content in local spider plant germplasm will be important to enable their exploitation to combat malnutrition and hunger and to deepen the limited knowledge on the secondary metabolites and their contribution to human health. Thus, the present study focuses on various spider plant entries from different African countries (i) to characterize them for important morphological and agronomic traits and (ii) to obtain detailed knowledge about the composition and concentration of potentially protective glucosinolates, flavonoids and mineral elements. Furthermore, (iii) the glucosinolate and flavonoid concentrations of different plant organs were compared, since previous studies on *Brassicaceae* have shown variations in concentration for different plant organs especially for glucosinolates (Brown, Tokuhisa, Reichelt, & Gershenzon, 2003; Sang, Minchinton, Johnstone, & Truscott, 1984; Velasco, Cartea, Gonzalez, Vilar, & Ordas, 2007; Yim et al., 2016).

2. Materials and methods

2.1. Plant materials

Thirty entries (Supplementary Table A1) consisting of 20 advanced lines and gene bank accessions from the World Vegetable Centre, Arusha, Tanzania, eight farmers' cultivars from Kenya and two advanced lines from Jomo Kenyatta University of Agriculture and Technology (JKUAT), Nairobi, Kenya, were used in this study.

A field experiment was carried out at the Jomo Kenyatta University of Agriculture, Juja, Kenya (S01°05.9′ E037°00.8, 1476 m a.s.l) between October 2015 and January 2016. A randomized block design with four blocks was used with each entry represented once in each block. Each entry was sown in four rows per plot with a spacing of 15 cm between rows in the same plot and 30 cm spacing between plots. Seeds were sown in double density and thinning was done approximately three weeks after seedling emergence. Well decomposed cow manure was applied to the rows and mixed with soil during sowing. Weeding was done manually throughout the experimental period. The experiment was conducted under rain fed conditions with supplemental overhead irrigation at least two times per week after planting until flower initiation.

2.2. Morphological characterization

For each entry, plants in the middle rows were randomly used for data collection. Qualitative traits assessed included stem and petiole coloration and were evaluated using 10 plants per accession in each block by recording the frequencies of plants with different stem and petiole colors.

Quantitative traits assessed included: germination percentage, days to 50% flowering, plant height, number of primary branches, 100 seeds weight, number of seeds per silique, total shoot fresh and dry mass as well as leaf fresh and dry mass. Field measurements and observations of traits were done on the same day to avoid any differences due to developmental stages of plants. Percentage germination was determined before thinning as the percentage of the seeds that were germinated based on the number of seeds that were sown for each entry. Days to flowering was recorded as the number of days from sowing to the date when approximately 50% of the plants in each plot flowered. Data for plant height and number of primary branches was taken just after the onset of flowering since the flower develops constantly and after flowering branching for more vegetative biomass production stops. Plant height was measured in centimeters from the base of the plant to the tip of the main stem using a meter rule by selecting ten plants at random from the inner rows of a plot. The number of primary branches was determined as the counts of branches at the main stem. The silique weight was calculated as the mean value of 50 pods harvested from ten plants of each entry for each block. A hundred seeds weight was calculated as the mean value of seeds from fifteen plants of each entry. Dry and fresh mass of the leaves were determined by pooling together twenty plants per plot after 7 weeks. The plants were cut at the base and the leaves separated from the stems.

Sampling for the glucosinolate, flavonoid and mineral concentration analyses was carried out between 09.00 a.m. and 12.00 a.m. in three stages (S1–S3), for the different plant organs: S1; the leaves and stems were collected before the start of the reproductive phase (23 days after sowing); S2, the flowers (36 days after sowing) and; S3, the siliques (63 days after sowing). Sampling was done for fifteen plants from each of the four blocks for each spider plant entry and finally pooled into two with each of the two pooled samples containing the plants from two blocks. In order to avoid degradation or losses in secondary metabolite contents due to microbial activity the samples were transferred immediately to the laboratory and stored at − 80 °C before freeze drying and later milling into powder. All following analyses were carried out on portions of the same homogenized samples for each of the spider plant entries.

2.3. Mineral analysis

Mineral element concentrations in the freeze dried material were

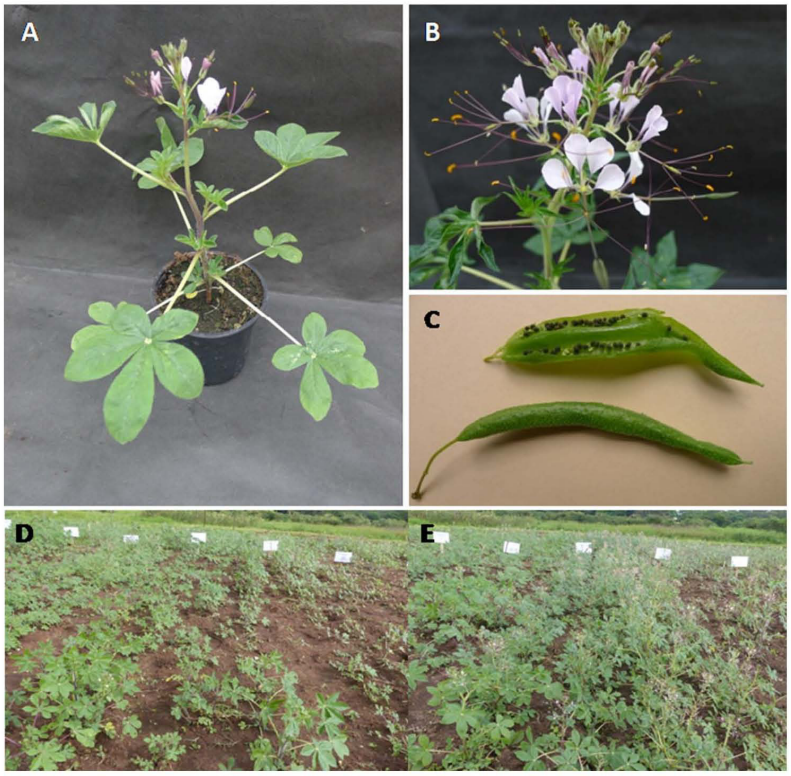


Fig. 1. *C. gynandra*, A - UGSF2 under greenhouse conditions at 5 weeks after sowing, B - close up of the characteristic flowers, C - siliques. D - entry Abuku2 under field conditions at 5 weeks and E - 7 weeks after sowing.

determined with an inductively coupled spectrophotometer (ICP-OES) after ashing at 450 °C for 8 h, and dissolving the ash in 1:3 diluted HNO_3 (Jones, 2001).

2.4. Glucosinolate analyses

The glucosinolate composition of the samples was determined as desulfo-glucosinolates, using a slightly modified method according to Wiesner, Hanschen, Schreiner, Glatt, and Zrenner (2013). The modifications were as follows: The various desulfo-glucosinolates were separated by a UHPLC-DAD device (UHPLC Agilent 1290 Infinity System, Agilent Technologies, Böblingen, Germany) equipped with a Poroshell 120 EC-C18 column of dimension 100 mm × 2.1 mm containing particles of size 2.7 µm (Agilent Technologies). The solvent gradient was formed by water (A) and 40% acetonitrile (B), starting at 0.5% B for 2 min, rising to 49.5% B over the next 10 min, then held for a further 2 min, increased to 99.5% B over the course of 1 min and finally held for a final 2 min. The flow rate was 0.4 mL min^{-1} and the injection volume 5 µL. Desulfo-glucosinolates were identified by comparing retention times and UV absorption spectra with those of known standards. Quantification was done at 229 nm using the response factor of the glucosinolates relative to 2-propenyl glucosinolate (external standard). Determination of glucosinolates was performed in duplicate.

2.5. Flavonoid analyses

Flavonoids were analyzed according to Schmidt et al. (2010) with slight modification. Lyophilized, ground plant material (0.02 g) was extracted with 600 µL of 60% aqueous methanol on a magnetic stirrer plate for 40 min at 20 °C. The extract was centrifuged at 3396 × *g* for 10 min at the same temperature, and the supernatant was collected in a reaction tube. This process was repeated twice with 300 µL of 60% aqueous methanol for 20 min and 10 min, respectively; the three corresponding supernatants were combined. The extract was subsequently evaporated until it was dry and was then suspended in 200 µL of 10% aqueous methanol. The extract was centrifuged at 1509 × *g* for 5 min at 20 °C through a Corning® Costar® Spin-X® plastic centrifuge tube filter (Sigma Aldrich Chemical Co., St. Louis, MO, USA) for the HPLC analysis. Each extraction was carried out in duplicate.

Flavonoid composition (including hydroxycinnamic acid derivatives and glycosides of flavonoids) and concentrations were determined from the filtrate using a series 1100 HPLC (Agilent Technologies, Waldbronn, Germany) equipped with a degaser, binary pump, autosampler, column oven, and photodiode array detector. An Ascentis® Express F5 column (150 mm × 4.6 mm, 5 µm, Supelco) was used to separate the compounds at a temperature of 25 °C. Eluent A was 0.5% acetic acid, and eluent B was 100% acetonitrile. The gradient used for eluent B was 5–12% (0–3 min), 12–25% (3–46 min), 25–90% (46–49.5 min), 90% isocratic (49.5–52 min), 90–5% (52–52.7 min), and 5% isocratic (52.7–59 min). The determination was conducted at a flow rate of 0.85 mL min^{-1} and a wavelength of 280 nm, 320 nm, 330 nm, 370 nm and 520 nm. The hydroxycinnamic acid derivatives and glycosides of flavonoids were tentatively identified as deprotonated molecular ions and characteristic mass fragment ions according to Schmidt et al. (2010) and Neugart, Rohn, and Schreiner (2015) by HPLC-DAD-ESI-MS^n using an Bruker amazon SL ion trap mass spectrometer in negative ionization mode. For the identification of the peaks the data were compared to the literature of the investigated species and their relatives. In the mass spectrometer, nitrogen was used as the dry gas (10 L min^{-1}, 325 °C) and the nebulizer gas (40 psi) with a capillary voltage of −3500 V. Helium was used as the collision gas in the ion trap. The mass optimization for the ion optics of the mass spectrometer

for quercetin was performed at *m/z* 301 or arbitrarily at *m/z* 1000. The MS^n experiments were performed in auto up to MS^3 in a scan from *m/z* 200–2000. Standards (chlorogenic acid, quercertin 3-glucoside, kaempferol 3-glucoside and isorhamnetin-3-glucoside Roth, Karlsruhe, Germany) were used for external calibration curves in a semi-quantitative approach. Results are presented as mg g^{-1} dry weight.

2.6. Statistical analysis

The quantitative trait data was expressed as the mean ± standard deviation (SD) of at least four determinations. Statistical analysis (ANOVA with a statistical significance level set at $p < 0.05$ and means separated using Tukey test) was carried out using XLSTAT ver7.1. The nutritional compound data was also expressed as the mean ± standard deviation (SD) of two determinations. Due to the limited number of replicates ($n = 2$) statistical comparisons were not possible.

3. Results

3.1. Morphological characterization

Characteristics of spider plant, including the inflorescences, seeds and siliques and the field plots are shown in Fig. 1. Spider plant is characterized by polymorphism in hairiness and different colors of the stems and leaf petioles among genotypes. The three main color variations observed among the plants in each entry were purple, green and a mix of purple and green parts (Fig. 2). The stem color for all the entries was dominantly purple (above 70% of plants for each entry) except for Abuku1, Acc26, Acc28, Acc3, Acc30 and GS (Fig. 3). Entry PS had purely purple stems for all the plants assessed which suggests homogeneity with respect to this trait. A majority of the accessions showed leaves with green or purple petioles with few having petioles with a mix of green and purple. Acc28 contained plants with purely purple petioles while Acc6, IP7, HTT, MLSF12, MLSF17, MLSF27, RWSF2, Site 94, ST93-1(GS), ST73-3, UGSF19 and UGSF29 all showed over 70% of the plants having green petioles. The accessions with green petioles were mainly from Tanzania, Rwanda, Malawi and South Africa. Generally, the farmers' cultivars were composed of almost equal proportions of plants with different stem and petiole color variants, whereas the advanced lines and gene bank accessions were mainly dominated by purple stems and green petioles.

Mean values for the quantitative traits measured are shown in Table 1. Percentage germination for the accessions showed significant variation among the entries with an average of 53.4% and a range from 24% (GS) to 77.8% (Acc28). The average number of days to 50% flowering was 36.1 days and was significantly different ($p < 0.05$) among the entries with a range from 31 days for Abuku1 to 40 days for IP8. Number of seeds per silique also showed significant differences ($p < 0.05$) among the entries with a maximum of 199 seeds for Accession 21 and a minimal number of 133 seeds for Acc26. There were no significant differences for 100 seed weight for all the entries. A hundred seed weight ranged from 0.144 g to 0.171 g for Abuku1 and Acc20, respectively. Significant ($p < 0.05$) variation was observed for plant height among the accessions of this study ranging from 53.3 cm to 77.4 cm for entries UGSF2 and IP7, respectively.

Other traits included silique weight, primary branch number, leaf and total fresh and dry mass (see Supplementary Table S2). No significant differences were observed for silique weights ranging between 1.08 g for GS to 0.63 g for ST73-3 with an average of 0.8 g. Primary branch numbers ranged from 13.2 (entry IP7 and ST93-1(GS)) to 8.5 for Acc5.

Leaves are the edible parts of the spider plant. Stems are sometimes consumed together with the leaves, especially for tender young shoots. The average leaf fresh and dry mass per plant were 14.4 g and 2 g, respectively (Fig. 4). The leaf fresh mass per plant ranged from 9 g (Acc28) to 24.4 g (UGSF13). The dry mass per plant, on the other hand, ranged between 1.4 g (Acc28 and Acc5) and 3.7 g (RWSF2) (Fig. 4). Total plants fresh mass per plant ranged between 25.7 g and 70.4 g for Acc5 and Acc20, respectively, while total dry mass per plant ranged from 3.7 g to 12.7 g for Acc28 and Acc20, respectively (see Supplementary Table A2).

3.2. Mineral concentration

Mineral concentrations in leaves of the spider plant entries are shown in Table 2 for macro-elements, micro-elements, and potentially toxic elements. In average of all entries, the macro-element concentration in g kg^{-1} DW was 31 for K, 23 for Ca, 4.3 for Mg, 6.6 for P and 4.0 for S. The variability of macro-element concentrations among all entries as indicated by the standard deviation was lowest for Mg and highest for S. For all macro-elements except Ca, the concentrations of the two advanced lines from JKUAT (Abuku1 and Abuku2) were markedly higher than those of the farmer's cultivars and the entries from WorldVeg, whereas farmer's cultivars in average had similar concentrations as the entries from WorldVeg. There was considerable variance in macro-element concentrations within the entries from farmer's fields and WorldVeg. For example, Mg concentrations varied from 3.5 (Acc20) to 5.0 mg kg^{-1} DW (Acc3) for farmer's cultivars, and from 3.2 (RWSF3) to 5.0 mg kg^{-1} DW (IP12) for the entries from WorldVeg.

In average of all entries, the micro-element concentration in mg kg^{-1} DW was 3603 for Fe, 393 for Mn, 60 for Zn,13 for Cu, 2.5 for

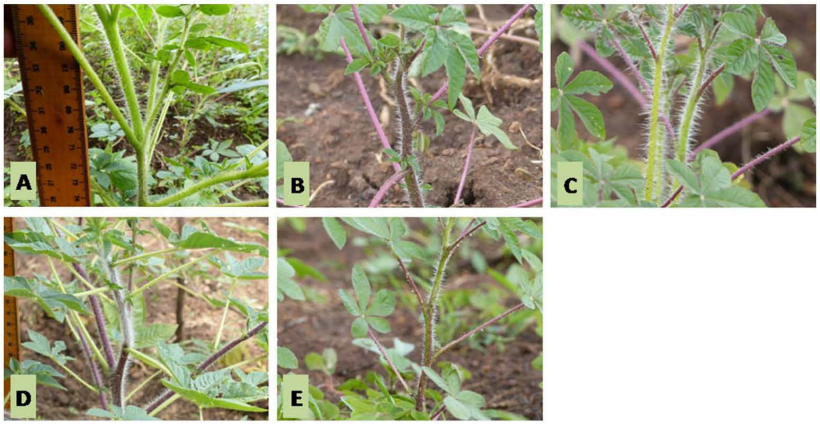


Fig. 2. Stem and petiole color variation in *C. gynandra* entries grown under field conditions. A - UGSF13, green stem/green petiole; B - Acc21, purple stem/purple petiole; C - green stem/purple petiole; D - UGSF29, purple stem/green petiole; E - Site94, purple/green.

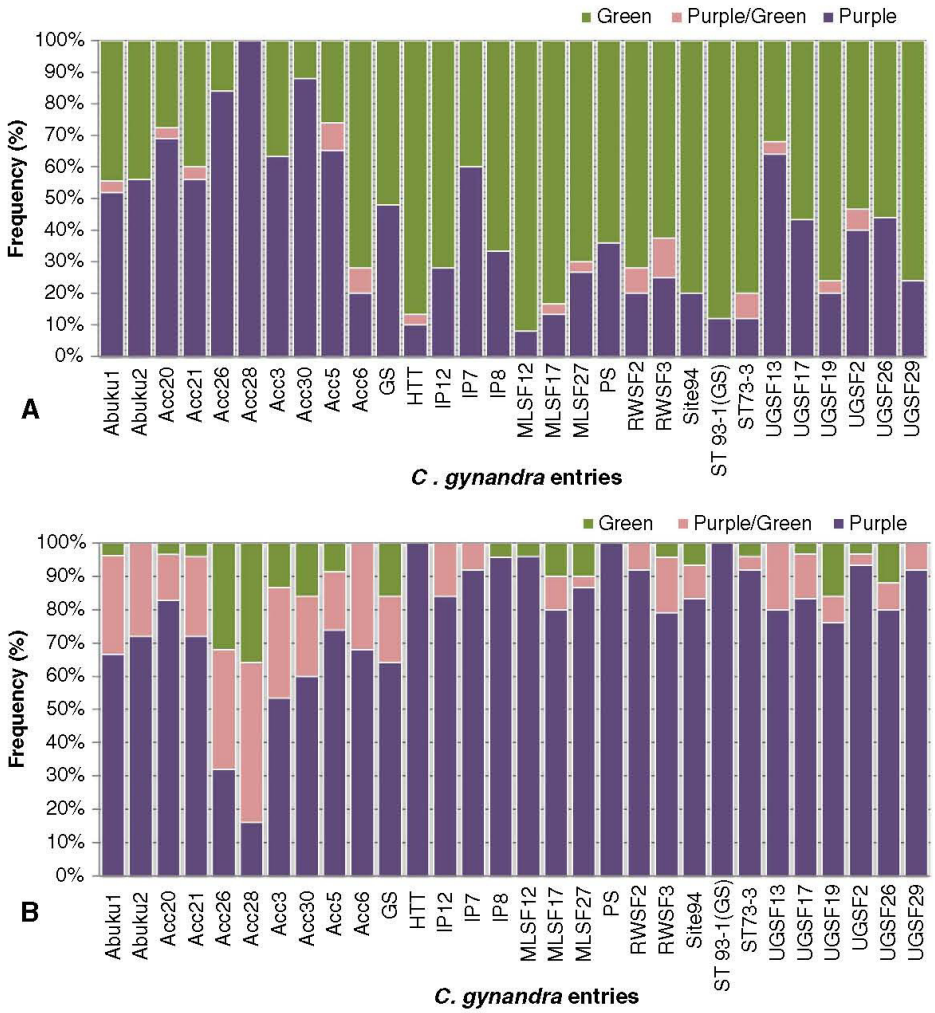


Fig. 3. Stem and petiole color variations within spider plant entries as shown in Fig. 2. A - percentage frequencies of petiole colors and B - percentage frequencies of stem color. $n = 28$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Ni, and 2.1 for Mo. The standard deviation of the mean of all entries, for all micro-elements except Zn was substantially larger than that for the macro-elements. The large standard deviations for the mean of the two analyses for each entry do not allow statements about differences in micro-element concentrations among the different origins (JKUAT, farmer's fields, WorldVeg). For Zn, there was considerable variance among entries, the concentrations varying from 44 (IP7) to 80 mg kg^{-1} DW for MLSF17 and ST93-I(GS).

In average of all entries, the concentrations of potentially toxic elements in mg kg^{-1} DW were 5.1 for Cr, 0.30 for Cd, 3.1 for Pb and 2952 for Al. The standard deviation of the mean of all entries for the toxic elements except Cd was substantially larger than that for the macro-elements. The large standard deviations for the mean of the two analyses for each entry do not allow statements about differences in toxic element concentrations among the different origins (JKUAT, farmer's fields, WorldVeg). There was considerable variance for Cd among entries with concentrations varying from 0.21 (Acc20) to 0.42 mg kg^{-1} DW (UG SF 13).

3.3. Glucosinolates

The aliphatic 3-hydroxypropyl glucosinolate – as predominate glucosinolate in spider plant - was quantitatively determined at all plant organs, especially in the siliques and inflorescences in nearly all investigated entries (Table 3). Also the aliphatic 2-propenyl glucosinolate was assessed, but only in the leaves of one entry (PS) and as a minor glucosinolate (Table 3). In addition, indole glucosinolates were quantitatively measured. The group of indole glucosinolates comprised the indole 3-indolylmethyl glucosinolate and its derivative 4-hydroxy-3-indolylmethyl, but only in trace concentrations (Supplementary Table A3).

The mean values for glucosinolates are presented in Table 3. The 3-hydroxypropyl glucosinolate concentrations varied in different plant organs from 7.8 μmol g^{-1} DW in the leaves to 37.1 μmol g^{-1} DW in the inflorescences. The stems and the siliques had mean values of 19 μmol g^{-1} DW and 19.3 μmol g^{-1} DW, respectively. Acc20 showed the highest 3-hydroxypropyl glucosinolate concentration in leaves at 11 μmol g^{-1} DW of all entries, while ST93-1(GS) had the lowest at

Table 1
Five morphological traits of 30 entries of *C. gynandra*.

Entry	Germination (%)	Days to 50% flowering	No. of seeds/ silique	100 seed weight (g)	Plant height (cm)
Abuku1	71.3^{ab}	31.0^{j}	151.6^{bc}	0.144^{a}	66.1^{abcde}
Abuku2	69.2^{abc}	33.5^{hij}	149.0^{bc}	0.146^{a}	60.3^{bcde}
Acc20	72.5^{ab}	35.5^{efghi}	163.7^{abc}	0.171^{a}	66.0^{abcde}
Acc21	36.9^{gh}	39.5^{ab}	199.0^{a}	0.157^{a}	58.8^{cde}
Acc26	73.4^{ab}	35.5^{efghi}	133.9^{c}	0.153^{a}	76.4^{a}
Acc28	77.8^{a}	35.5^{efghi}	157.4^{abc}	0.155^{a}	76.6^{a}
Acc3	72.2^{abcd}	35.5^{defgh}	134.4^{c}	0.151^{a}	60.0^{bcde}
Acc30	60.4^{abcde}	35.5^{efghi}	157.4^{abc}	0.153^{a}	61.2^{bcde}
Acc5	71.5^{ab}	34.5^{fghi}	147.0^{bc}	0.154^{a}	56.2^{de}
Acc6	69.7^{ab}	35.5^{efghi}	152.8^{abc}	0.159^{a}	58.9^{cde}
GS	24.0^{i}	34.5^{fghi}	182.9^{ab}	0.170^{a}	64.7^{abcde}
HTT	41.7^{efg}	34.5^{fghi}	157.8^{abc}	0.145^{a}	66.8^{abcde}
IP12	57.3^{abcdef}	32.8^{ij}	167.5^{abc}	0.155^{a}	74.7^{ab}
IP7	51.6^{bcdefg}	36.5^{bcdefg}	167.0^{abc}	0.159^{a}	77.4^{a}
IP8	56.8^{abcdef}	40.0^{a}	154.6^{abc}	0.146^{a}	76.2^{a}
MLSF12	54.2^{abcdefg}	39.0^{abc}	164.3^{abc}	0.164^{a}	76.4^{a}
MLSF17	47.2^{cdefg}	38.5^{abcd}	137.4^{c}	0.153^{a}	64.1^{abcde}
MLSF27	68.3^{abc}	34.5^{ghi}	170.1^{abc}	0.160^{a}	71.1^{abc}
PS	66.8^{abcd}	35.5^{efghi}	153.9^{bc}	0.158^{a}	70.0^{abcd}
RWSF2	36.5^{gh}	33.3^{hij}	144.0^{bc}	0.155^{a}	70.3^{abcd}
RWSF3	39.7^{fg}	39.5^{ab}	164.7^{abc}	0.158^{a}	64.3^{abcde}
Site 94	63.3^{abcd}	34.5^{fghi}	164.0^{abc}	0.146^{a}	58.7^{cde}
ST73-3	51.7^{bcdefg}	38.5^{abcd}	134.2^{c}	0.155^{a}	71.4^{abc}
ST93-1 (GS)	54.2^{abcdefg}	36.0^{cdefgh}	156.8^{bc}	0.157^{a}	77.0^{a}
UGSF13	25.0^{hi}	36.8^{bcdefg}	179.9^{ab}	0.151^{a}	68.8^{abcd}
UGSF17	46.3^{efg}	35.0^{efghi}	183.1^{ab}	0.153^{a}	57.8^{cde}
UGSF19	54.1^{abcdefg}	38.3^{abcde}	144.6^{bc}	0.153^{a}	74.5^{ab}
UGSF2	40.7^{fg}	37.5^{abcdef}	144.0^{c}	0.167^{a}	53.3^{e}
UGSF26	46.3^{defg}	38.5^{abcd}	144.8^{c}	0.167^{a}	60.1^{bcde}
UGSF29	51.4^{bcdefg}	36.8^{bcdefg}	181.1^{a}	0.151^{a}	61.0^{bcde}

Given are means of $n = 4$ replicates. Mean values within columns indicated by the same letter are not significantly different according to the Tukey test employing all pair-wise comparisons of the 30 entries ($p \leq 0.05$). Cells with the minimum and maximum values within a column are highlighted by green and yellow color, respectively.

5.1 μmol g^{-1} DW. The 3-hydroxypropyl glucosinolate concentration in stems ranged between 11.5 μmol g^{-1} DW (IP12) and 27.4 μmol g^{-1} DW (UGSF17); between 11.7 μmol g^{-1} DW (Acc26) and 85.3 μmolg^{-1} DW (MLSF17) in inflorescences and between 12.6 μmol g^{-1} DW (Acc26) and 27.5 μmol g^{-1} DW (RWSF2) in siliques. Generally, 3-hydroxypropyl glucosinolate was most abundant in the inflorescences and lowest in the leaves, while the stems and the inflorescences were comparable in amounts for most of the entries.

3.4. Flavonoids

Flavonoid contents in leaves and flowers of the spider plant entries are shown in Fig. 5. The flavonols detected in both, leaves and flowers, were glycosides of quercetin, kaempferol and isorhamnetin. In the leaves, quercetin glycosides were the most abundant compounds and followed by kaempferol and isorhamnetin glycosides in traces. The total leaf flavonoid content among the spider plant entries ranged between

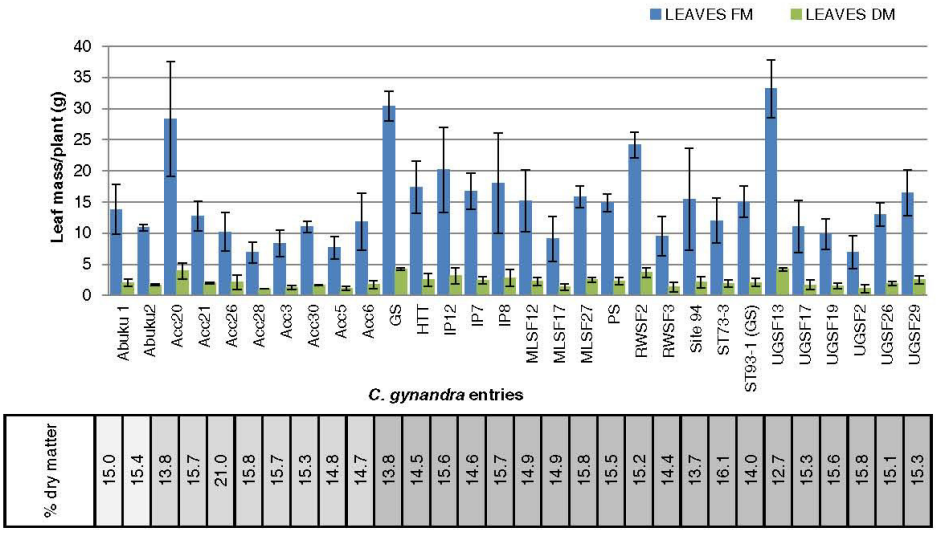


Fig. 4. Leaf fresh and dry mass with percentage of dry matter of *C. gynandra* entries grown in the field. $n = 4$. Results of the statistical analysis are presented in Supplementary Table S2.

Table 2
Mineral concentrations in dry mass of leaves from selected *C. gynandra* entries.

Entry	Macro-elements (g kg^{-1} DW)					Micro-elements (mg kg^{-1} DW)						Potentially toxic elements (mg kg^{-1} DW)			
	K	Ca	Mg	P	S	Fe	Mn	Zn	Cu	Ni	Mo	Cr	Cd	Pb	Al
Abuku1	37±0.4	22±1.8	5±0.1	9±0.6	5±0.9	3457±177	372±18	53±4	10±2	2.3±0.2	0.8±0.0	6.0±0.9	0.30±0.01	1.9±0.3	3081±23
Abuku2	39±0.8	23±0.4	5±0.1	7±0.0	6±0.2	5420±316	542±9	61±6	10±0.2	3.0±0.3	0.7±0.1	6.4±2.3	0.41±0.04	2.9±0.2	4989±337
Acc20	28±1.8	22±0.8	4±0.1	7±0.5	4±0.3	2384±1112	246±73	54±2	10±0.4	1.4±0.2	1.1±0.1	2.7±1.2	0.21±0.00	2.4±0.5	2114±971
Acc21	32±5.4	24±0.5	5±0.5	7±0.1	4±0.4	4089±172	476±35	52±0.2	17±4	2.9±1.4	2.1±1.2	4.9±0.6	0.31±0.08	3.4±1.0	3492±188
Acc26	28±1.1	24±2.7	4±0.6	7±1.3	3±0.1	5601±4690	481±338	58±16	9±0.9	3.3±3.1	3.5±0.6	8.2±7.7	0.28±0.16	3.8±2.5	3536±2224
Acc28	28±1.5	25±1.4	4±0.7	6±0.4	4±0.3	2330±974	302±76	52±4	11±2	1.4±0.2	2.4±1.5	2.9±0.3	0.26±0.01	2.1±0.8	1851±834
Acc3	31±1.4	24±2.4	5±0.4	6±1.5	4±0.6	4366±523	465±71	53±3	23±13	2.9±0.3	0.9±0.4	5.8±1.5	0.42±0.04	3.6±1.1	3880±250
Acc30	30±0.7	23±3.8	4±1.1	6±0.2	3±0.2	4980±869	540±70	49±10	10±3	2.7±0.1	3.8±3.2	6.3±0.7	0.36±0.03	3.4±0.5	4038±249
Acc5	32±0.9	26±1.9	5±0.3	7±0.4	3±0.7	2532±50	309±3	67±24	10±1	1.3±0.5	3.1±2.5	3.4±0.1	0.22±0.00	2.1±0.4	2138±71
Acc6	29±3.3	23±5.9	4±1.5	6±1.9	4±0.8	4409±1491	439±137	54±9	14±6	2.1±0.1	2.0±0.2	6.7±0.9	0.26±0.03	3.6±1.3	3609±910
IP12	33±1.9	22±0.2	5±0.1	8±0.2	4±0.1	4728±373	525±57	62 ±2	n.a.	3.4±0.9	1.1±0.2	8.5±5.0	0.31±0.01	10.8±4.1	3901±100
IP7	27±0.7	22±0.3	4±0.1	5±0.2	4±0.8	1336±18	185±10	44±4	9±0.3	1.9±0.6	1.7±0.7	1.9±0.1	0.34±0.17	1.3±0.04	1096±35
IP8	32±3.2	24±1.1	5±0.0	6±0.2	4±0.3	2132±241	305±27	63±3	26±8	1.8±0.2	0.8±0.2	3.9±0.8	0.27±0.01	3.1±0.3	1722±254
MLSF17	33±5.2	23±0.4	5±0.0	8±0.1	4±0.4	3334±625	416±95	80±13	20±12	2.7±1.3	2.7±2.6	4.4±0.1	0.26±0.06	2.8±1.5	2825±465
MLSF27	27±3.6	23±0.5	4±0.6	6±0.8	4±0.0	2008±776	210±29	57±11	10±3	1.5±0.4	1.0±0.2	2.6±0.2	0.27±0.08	1.7±0.3	1771±726
PS	30±3.9	26±2.1	4±0.8	6±1.8	4±0.2	3399±62	391±31	74±33	n.a.	3.3±2.3	2.8±0.6	5.3±2.3	0.29±0.16	n.a.	2719±271
RWSF3	28±4.3	19±0.6	3±0.0	6±0.7	4±0.3	5892±3476	400±183	56±8	11±3	2.3±0.9	1.2±0.2	5.6±2.2	0.27±0.10	3.3±0.7	3608±854
Site94	34±4.0	23±1.8	5±0.1	7±0.7	3±0.4	5857±1572	662±205	72±23	19±3	3.1±0.1	3.1±2.5	6.9±0.3	0.30±0.02	3.9±1.0	4917±1023
ST73-3	33±2.5	22±1.1	4±0.4	7±0.6	4±0.7	2734±364	361±32	67±3	14±2	1.8±1.0	3.3±2.9	7.0±6.1	0.25±0.11	2.0±0.01	2178±366
ST93-1(GS)	38±1.9	21±1.1	4±0.1	7±0.2	4±0.2	3788±393	431±36	80±30	n.a.	3.9±1.5	0.8±0.0	7.8±4.8	0.37±0.08	n.a.	3289±352
UG SF 13	29±0.4	23±2.2	5±1.1	6±1.2	5±0.2	5455±3014	552±196	64±0.2	12±0.6	4.2±1.3	0.9±0.4	6.0±3.9	0.42±0.06	3.5±1.8	5224±3086
UG SF 19	33±1.3	22±1.2	4±0.7	7±1.4	4±0.1	2521±108	313±30	47±8	9±2	3.6±3.2	1.9±0.8	4.0±2.0	0.24±0.04	1.8±0.1	2006±286
UG SF 2	31±6.5	23±0.6	4±0.2	6±1.4	4±0.2	2693±15	342±14	52±2	24±18	1.9±0.6	2.4±1.9	3.5±0.6	0.25±0.09	3.7±1.1	2128±8
UG SF 29	29±2.1	28±2.1	4±0.1	7±0.0	5±0.1	1036±98	161±5	76±6	8±0.3	0.6±0.1	6.1±0.1	1.4±0.2	0.24±0.02	1.5±0.6	735±98
Mean	31	23	4	7	4	3603	393	60	14	2.5	2.1	5.1	0.30	3.1	2952
SD	3.3	1.8	0.6	0.9	0.7	1464	127	10	5.5	0.9	1.3	2.0	0.1	1.9	1213

Given are means ± SD. n = 2 biological replicates. n.a. = not analyzed. Cells with the minimum and maximum values within a column are highlighted by green and yellow color, respectively. The three gray scales represent the source of the seeds used as JKUAT, farmers' cultivars and WorldVeg respectively.

2.2 mg g^{-1} DW (MLSF27) and 12.5 mg g^{-1} DW (MLSF17). With respect to the type of flavonols detected, quercetin glycoside content ranged from 2.12 mg g^{-1} DW (MLSF27) to 12.19 mg g^{-1} DW (MLSF17); isorhamnetin glycosides ranged from 0.04 mg g^{-1} DW (MLSF27) to 0.41 mg g^{-1} DW (Abuku1); and kaempferol glycosides ranged between 0.00 mg g^{-1} DW (Acc26 and Acc30) and 0.47 mg g^{-1} DW (Abuku1).

In more details, the main glycosides of quercetin, kaempferol and isorhamnetin were diglucosides and rutinosides (Fig. 6). In leaves, quercetin-3-diglucoside was the main compounds except for Abuku1 in which quercetin-3-rutinoside (rutin) was the main quercetin glycoside. The corresponding kaempferol glycosides were present in 17 of the 19 entries analyzed with kaempferol-3-diglucoside as main compound whereas for isorhamnetin glycosides, mainly isorhamnetin-3-rutinoside was found the in spider plant entries (Fig. 6).

Unlike the leaves, kaempferol glycosides were the most abundant flavonoids while isorhamnetin glycosides being the least in the flower tissue (Fig. 5B). Total flavonoids in flowers of spider plant varied between 4.3 mg g^{-1} DW (Acc20) and 12.22 mg g^{-1} DW (MLSF17). The single flavonoids were represented among the entries analyzed as follows: kaempferol glycosides ranged between 3.46 mg g^{-1} DW (Acc20) and 8.75 mg g^{-1} DW (UGSF26); quercetin ranged between 0.13 mg g^{-1} DW (Abuku1) and 4.6 mg g^{-1} DW (Acc5); and isorhamnetin ranged between 0.03 mg g^{-1} DW (Acc21) and 0.52 mg g^{-1} DW (Acc 5).

4. Discussion

4.1. Phenotypic characterization

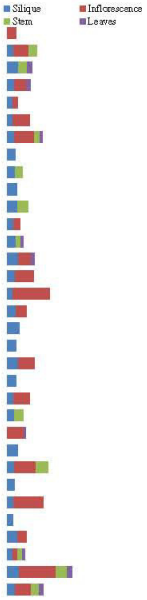
Spider plant germplasm is continuously collected from farm populations and from the wild and transferred to the gene banks and institutions for propagation and development. Morphological characterization in spider plants is important for proper systematic identification of traits to presumably categorize this germplasm into similar groups and to describe morphological diversity to enable exploitation in genetic development programs. Our findings here show a mix of plant types with regard to stem and petiole colors (Fig. 2), an observation also made in previous studies including spider plant germplasm from different sites and populations (K'Opondo, 2011; Stoilova, Dinssa, Ebert, & Tenkouano, 2015; Wasonga, Ambuko, Chemining'wa, Odeny, & Crampton, 2015). The coloration is due to anthocyanins accumulation (Chalker-Scott, 1999). These anthocyanins are sometimes stress indicators in response to environmental stressors such as strong light, temperature extremes, drought, nitrogen and phosphorus deficiencies (Chalker-Scott, 1999; Hatier & Gould, 2009), but they are also attractants for insect pollinators (Dasgupta & De, 2007). The farmer's cultivars show more variations in the color pigmentation than the gene bank accessions and the advanced lines indicating more diversity among the farm collections (Fig. 3).

The results for most of the quantitative traits in this study are comparable to results from other studies (Onyango, Onwonga, & Kimenju, 2016; Stoilova et al., 2015; Wasonga et al., 2015). Germination percentage number of seeds per silique were higher than those reported by Onyango et al. (2016). This might be explained by favorable environmental conditions, since temperature ranges of 20–30 °C and darkness are optimal for spider plant germination (K'Opondo, Muasya, & Kiplagat, 2005). Poorly germinating entries resulted in weaker plants and poor yield, strongly pointing to the importance of high quality seed.

The period between sowing and flowering significantly limits leaf productivity and yield (Ojiewo, Mwai, Abukutsa-Onyango,

Table 3
Contents of 3-hydroxypropyl glucosinolate (μmol g^{-1} dry weight) in various organs of selected *C. gynandra* entries.

Accession	Silique	Inflorescence	Stem	Leaves
Abuku1	n.a	20.7±10.8	n.a.	n.a.
Abuku2	15.7±1.6	35.6±18.3	17.2±5.2	n.a.
Acc20	25.7±2.4	n.a.	20.5±1.2	11.0±1.6
Acc21	17.1±11.5	28.7±3.8	n.a.	8.9±2.1
Acc26	12.6±7.3	11.7±0.4	n.a.	n.a.
Acc28	13.4±6.1	38.2±25.3	n.a.	n.a.
Acc 3	17.8±7.7	46.1±26.0	12.7±1.4	6.3±1.5
Acc30	18.5±1.0	n.a.	n.a.	n.a.
Acc5	18.1±2.0	n.a.	17.0±1.7	n.a.
Acc6	23.2±0.7	n.a.	n.a.	n.a.
GS	24.9±2.0	n.a.	24.5±2.9	n.a.
HTT	15.0±3.0	14.4±9.8	n.a.	n.a.
IP12	20.0±3.7	n.a.	11.5±3.7	6.9±2.9
IP7	27.2±13.9	36.0±1.4	n.a.	n.a.
IP8	26.4±7.9	29.3±3.2	n.a.	8.6±2.1
MLSF12	18.2±8.0	44.0±19.1	n.a.	n.a.
MLSF17	14.1±5.4	85.3±18.4	n.a.	n.a.
MLSF27	20.0±5.8	25.5±14.6	n.a.	n.a.
PS	21.1±0.6	n.a.	n.a.	n.a.
RWSF2	27.5±9.9	n.a.	n.a.	n.a.
RWSF3	21.0±4.8	n.a.	n.a.	n.a.
Site94	14.4±3.4	37.6±0.4	n.a.	n.a.
ST73-3	17.1±6.0	n.a.	20.9±1.0	n.a.
ST93-1(GS)	n.a.	38.5±0.4	n.a.	5.1±1.5
UGSF13	23.8±7.8	n.a.	n.a.	n.a.
UGSF17	16.9±10.6	51.1±24.7	27.4±1.9	n.a.
UGSF19	17.4±1.1	n.a.	n.a.	n.a.
UGSF2	15.1±2.6	69.1±19.0	n.a.	n.a.
UGSF26	14.0±3.3	n.a.	n.a.	n.a.
UGSF29	25.1±4.1	19.1±5.4	n.a.	n.a.
Min	**12.6**	**11.7**	**11.5**	**5.1**
Max	**27.5**	**85.3**	**27.4**	**11**
Mean	**19.3**	**37.1**	**19**	**7.8**



Given are mean ± SD. n = 2. n.a. = samples could not be analyzed in at least one replicate.
The three gray scales represent the source of the seeds used as JKUAT, farmers' cultivars and WorldVeg respectively.

Agong, & Nono-Womdim, 2013). Late flowering is, therefore, an advantage since late flowering allows a prolonged period of harvesting. Flower removal as soon they appear has been suggested to increase leaf yield and longer periods of harvesting (Ojiewo et al., 2013; Wangolo, Onyango, Gachene, & Mong'are, 2015). Leaf yield per plant in the present study (14.11 g) was much lower compared to 135.65 g reported by Stoilova et al. (2015) who carried out three successive harvests.

The number of seeds per silique is important for seed production while 100 seed weight is a criterion for seed quality determination since seed mass affects seed germination and emergence (Li, Lu, Yang, Kong, & Deng, 2015). A higher number of seed per silique and higher seed mass are preferred for commercial seed production because the seeds are sold by weight.

4.2. Mineral concentration

Mineral elements are essential for normal and healthy growth and development in both children and adults (FAO/WHO, 2004). This study is the first dealing mineral analyses in a wide range of spider plant genotypes with different genetic backgrounds. The spider plant entries were rich in potassium, calcium, magnesium, phosphorus, iron, manganese and zinc which were also comparable to the levels previously reported for *Cleome* species for some of the elements (Jinazali et al., 2017; Odhav et al., 2007) (see Supplementary Table A5 for a comparison of our data to previously published data). However, higher Fe and Cr concentrations were realized in this study and this could be attributed to surface contamination by dust (soil), as also indicated by the high Al concentrations since Al transfer from the roots to shoots in plants is usually negligible in herbaceous species except for Al accumulators (Klug, Kirchner, & Horst, 2015; Metali, Salim, & Burslem, 2012). A strong correlation for the three elements (Fe, Al and Cr) (Supplementary Table A4) may further point to contamination as a source of their elevation.

Minerals concentrations in the edible leaf portions of spider plant are mainly dependent on the soil mineral concentrations. Various mineral concentrations could be altered by different fertilizer applications, for instance, increased calcium ammonia nitrate fertilizer enhanced calcium in spider plant (Hutchinson, 2011). Our results showed higher amounts for lead and cadmium exceeding the allowable limits of 0.30 mg kg^{-1} and 0.20 mg kg^{-1} respectively in leafy vegetables (Commission Regulation No 629/2008, 2008). The bioavailability of mineral nutrients in these vegetables remains a challenge considering that some traditional methods of cooking involve boiling to eliminate bitter taste and most of these minerals leach to the cooking water which is normally discarded (Hutchinson, 2011; Mziray, 1999).

In order to make use of the variation detected in mineral concentrations for the different entries in future breeding programs, in-depth analyses of single genotypes, preferably under different nutrient supplies, would be the first step, followed by investigations in progenies of genotypes with high and low concentrations of the most relevant nutrients.

4.3. Glucosinolates

Cleomeaceae family, to which spider plant belongs, is a family of the overall order Brassicales (Hanschen et al., 2014). The present study shows 3-hydroxypropyl glucosinolate as the main glucosinolate present in the selected spider plant entries, but the aliphatic methyl glucosinolate (glucocapparin), the 2-hydroxy-2-methylbutyl glucosinolate (glucocleomin) and the methyl glucosinolate have also been previously identified in the genus *Cleome* (Abdullah, Elsayed, Abdelshafeek, Nazif, & Singab, 2016; Neugart et al., 2017; Songsak & Lockwood, 2002). The health-promoting effects of *Cleome* has recently been studied suggesting to be partly effective against malaria, tuberculosis,

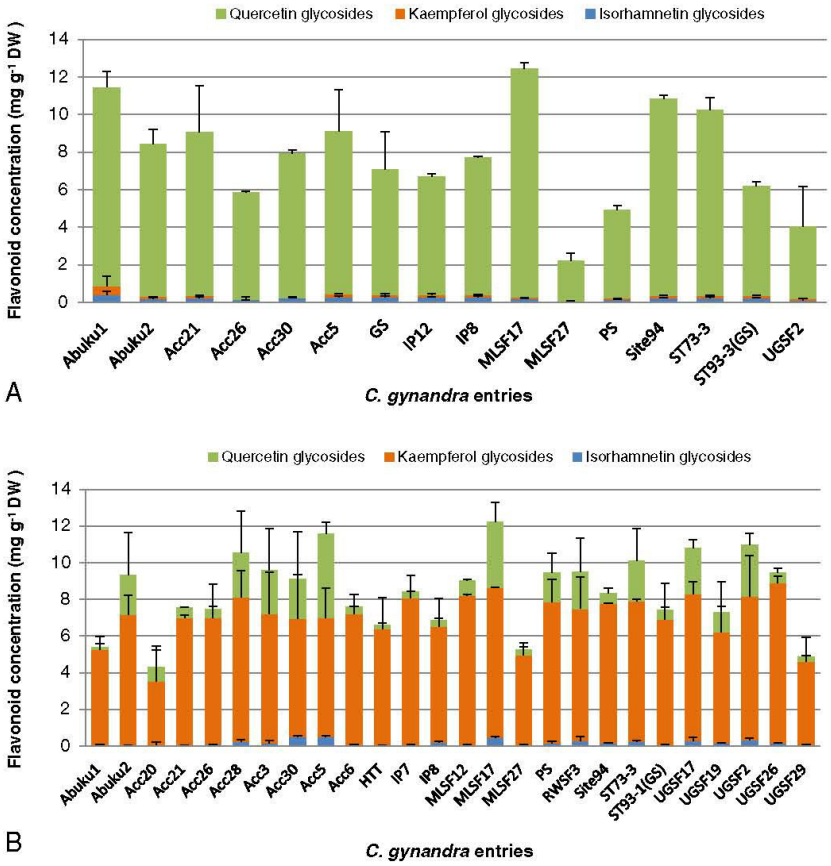


Fig. 5. Flavonoids in A - leaves and B - flowers from selected *C. gynandra* entries. Whiskers indicate standard deviations. $n = 2$.

pneumonia and fungal skin infection (Chinsembu, 2016) that might be also induced by its glucosinolates and their corresponding hydrolyses products.

Our preliminary glucosinolate quantification for the same spider plant entries grown in the greenhouse or in the field (Supplementary Table A3) generally indicated higher amounts of total glucosinolates and indole glucosinolates of the field-grown plants which might be due to the different climate conditions, e.g. higher irradiation intensities especially of UVB radiation (Mewis et al., 2012) (Table 3). The indole glucosinolates identified in greenhouse-grown plants were 3-indolylmethyl glucosinolate and its hydroxylated derivative 4-hydroxy-3-indolylmethyl glucosinolate (Supplementary Table A3) indicating that in spider plant indole glucosinolate biosynthesis is strongly environmentally regulated as demonstrated in other Brassicales (e.g. Verkerk et al., 2009).

The glucosinolate concentrations among the selected entries were different and most likely due to genotypic effects. However, the standard deviations point to variation among the biological replicates clearly indicating the need for further analyses employing more biological replicates to allow detailed statistical analyses. The total glucosinolate concentration was lowest in leaves and highest in the inflorescences in the present study (Table 3). This is in agreement with previous studies that have demonstrated glucosinolate transport through the phloem from the mature leaves to floral tissues and fruits (Chen, Petersen, Olsen, Schulz, & Halkier, 2001). The reproductive organs can also synthesize additional glucosinolates (Reintanz, 2001). High glucosinolate concentrations in reproductive organs (seeds, inflorescences and fruits) is predicted to be for maximum defense since these organs contribute most to plant fitness (Brown et al., 2002).

4.4. Flavonoids

The major flavonoids identified in spider plant were glycosides of the flavonols kaempferol, quercetin and isorhamnetin. These are generally the main flavonoids in food plants. Comparable concentrations to our results were reported for kale (*Brassica oleracea* var. *sabellica*) that ranged between 6.0 and 14.8 mg g^{-1} DW (Schmidt et al., 2010). Differences in the concentration of these flavonoids in leaf and floral tissues were observed. Quercetin glycosides were dominant in the leaves with rutin as main compound in most genotypes while kaempferol glycosides were dominant in the flowers. These variations in quantities of the flavonols in the different plant parts could be explained by the fact that kaempferol is a precursor to quercetin while isorhamnetin is a methylated product of quercetin. A higher antioxidant activity for quercetin glycosides compared to kaempferol glycosides in curly kale reported by Zietz et al. (2010) could explain its high content in the photosynthetic tissues. No correlation was noted, however, for the flavonoid content in leaves and flowers for the specific entries. Important biological activities of flavonoids contributing to human health have been pointed out in several studies. An antioxidant effect of quercetin and kaempferol on the nasal mucosa, synergistic effects in reducing cell proliferation hence reducing the risk of cancer, and anti-inflammatory effects in microphages have been reported (Ackland, van de Waarsenburg, & Jones, 2005; Hämäläinen, Nieminen, Vuorela,

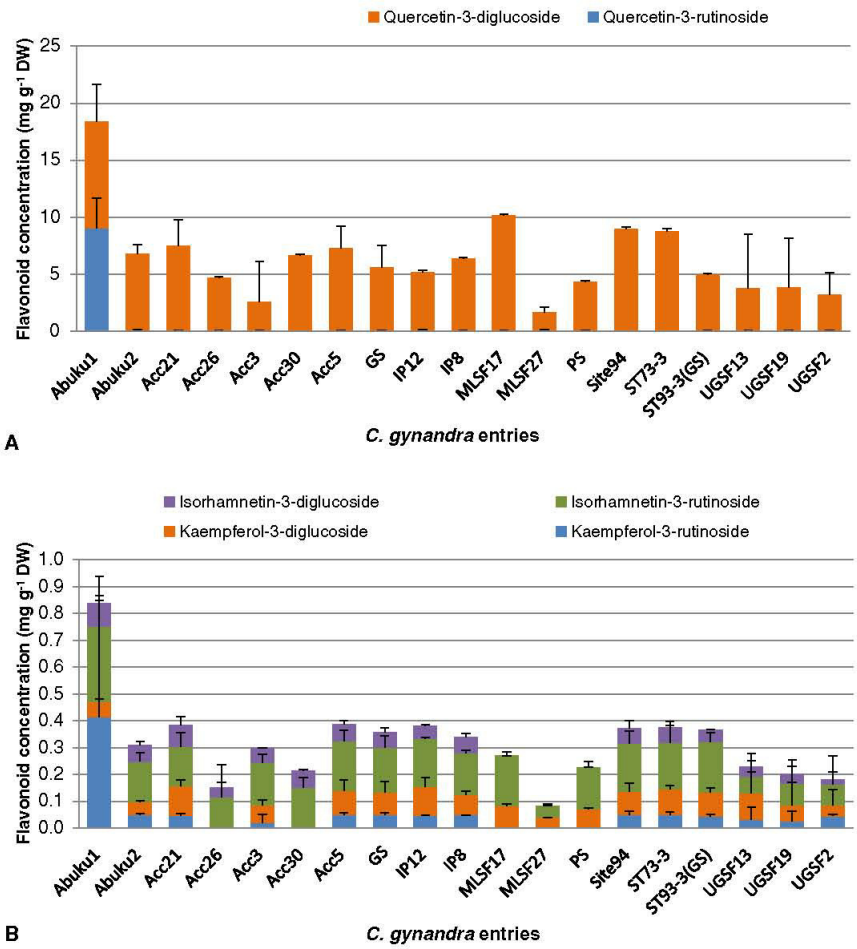


Fig. 6. Main glycosides of: A - quercetin; B - kaempferol and isorhamnetin in the leaves of *C. gynandra* entries. $n = 2$; given are means + SD.

Heinonen, & Moilanen, 2007; Lin et al., 2007; Reiter, Rupp, Baumeister, Zieger, & Harreus, 2009). Higher quercetin intake was associated with reduced mortality from ischemic heart diseases and protection for several types of cancers (Knekt et al., 2002), blood pressure reduction, vascular protective effect and improved vasodilator functions (Ibarra et al., 2003). Quercetin-3-rutinoside (rutin) is nowadays studied for its antioxidant, anti-inflammatory, anti-allergy, and anti-tumor activity (Habtemariam, 2016; Hosseinzadeh & Nassiri-Asl, 2014; Koval'skii et al., 2011). The other identified quercetin glycosides are rare and not yet studied but might also have an effect on the antioxidant activity of spider plant.

Flavonoid concentrations are affected by genetics, temperature and radiation as reported for kale (Schmidt et al., 2010). Species and cultivars usually show similar flavonoid profiles while differing in quantities (Krumbein, Saeger-Fink, & Schonhof, 2007). The entries in this study are of the same species and generally showed the same flavonoid profiles while differing in the concentrations. The farmers' accessions collected directly from the farmers and markets in Kenya did not show noticeable flavonoid profile variation compared to the rest of the entries sourced from the WorldVeg except for accession Abuku1 that showed unique profiles for the flavonoid glycosides (Fig. 6). This could be due to a genetic influence since the plants were grown in the same environmental condition. Higher concentrations of flavonoids have been associated with low temperatures leading to higher quantities of reactive oxygen species (Klimov et al., 2008) thereby necessitating an increase in antioxidants in the plant. Increased radiation also increases the flavonoid concentration since flavonoids are shielding components against radiation in plants (Neugart et al., 2012). No obvious correlation between flavonoid contents and coloration of stems and petioles was observed in the present study.

4.5. Conclusion

As main indigenous African leafy vegetables spider plant is rich in minerals, glucosinolates and flavonoids and hence can greately contibute to nutrition and health promotion in the diet particularly of African consumers in rural and urban areas. The concentrations of the glucosinolates and flavonoids vary in different plant organs and also among the selected entries analyzed. This study serves as a valuable knowledge base for improvement of the nutrient diets and food selection. Further in-depth analyses with a higher number of replicates and also single plant analyses to identify variarions between genotypes and accessions coupled with health/medical studies are needed before recommendations of specific entries for breeding focusing on enhancing these nutrients can be concluded. Our previous genetic diversity study of these entries (Omondi et al., 2017) and also the high variation within

entries regarding the morphological data presented in this study show a pronounce variability between the genotypes within the entries, thus demonstrating the great potential of spider plant as nutritious and protective food plant accompanied with a distinct environmental plasticity for growing also in marginal areas.

Acknowledgements

The authors are grateful for the funding within the project HORTINLEA by the German Federal Ministry of Education and Research (BMBF), grant number 031A248E, within the framework of the GlobE-Global Food Security. We also thank from David Odongo, Francis Wetende, Rose Khayanga, Merina Ingwe, Jane Were and Alfred Masaba for technical assistance in the field, Andrea Jankowsky and Sarah Farrherr for support in lab analysis, JKUAT and WorldVeg Arusha, Tanzania, especially Dr. Fekadu Dinssa, for providing the seed material.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.foodres.2017.06.050.

References

Abdullah, W., Elsayed, W. M., Abdelshafeek, K. A., Nazif, N. M., & Singab, N. B. (2016). Chemical constituents and biological activities of *Cleome* genus: A brief review. *International Journal of Pharmacognosy and Phytochemical Research*, *8*(5), 777–787.

Abukutsa-Onyango, M. (2007). The diversity of cultivated African leafy vegetables in three communities in western Kenya. *African Journal of Food Agriculture Nutrition and Development*, *7*(3), 1–15. http://www.bioline.org.br/request?nd07021.

Ackland, M. L., van de Waarsenburg, S., & Jones, R. (2005). Synergistic antiproliferative action of the flavonols quercetin and kaempferol in cultured human cancer cell lines. *In Vivo*, *19*(1), 69–76.

Agati, G., Azzarello, E., Pollastri, S., & Tattini, M. (2012). Flavonoids as antioxidants in plants : Location and functional significance. *Plant Science*, *196*, 67–76. http://dx.doi.org/10.1016/j.plantsci.2012.07.014.

Amić, A., Lučić, B., Stepanić, V., Marković, Z., Marković, S., Dimitrić Marković, J. M., & Amić, D. (2017). Free radical scavenging potency of quercetin catecholic colonic metabolites: Thermodynamics of 2H +/2e − processes. *Food Chemistry*, *218*, 144–151. http://dx.doi.org/10.1016/j.foodchem.2016.09.018.

Brown, P. D., Tokuhisa, J. G., Reichelt, M., & Gershenzon, J. (2003). Variation of glucosinolate accumulation among different organs and developmental stages of *Arabidopsis thaliana*. *Phytochemistry*, *62*, 471–481. http://dx.doi.org/10.1016/S0031-9422(02)00549-6.

Brown, A. F., Yousef, G. G., Jeffery, E. H., Klein, B. P., Wallig, M. A., Kushad, M. M., & Juvik, J. A. (2002). Glucosinolate profiles in broccoli: Variation in levels and implications in breeding for cancer chemoprotection. *Journal of the American Society for Horticultural Sciences*, *127*(5), 807–813. http://journal.ashspublications.org/content/127/5/807.short.

Chalker-Scott, L. (1999). Environmental significance of anthocyanins in plant stress sesponses. *Photochemistry and Photobiology*, *70*(1), 1–9. http://dx.doi.org/10.1562/00318655(1999)070<0001:ESOAIP>2.3.CO;2.

Chen, S., Petersen, B. L., Olsen, C. E., Schulz, A., & Halkier, B. A. (2001). Long-distance phloem transport of glucosinolates in Arabidopsis. *Plant Physiology*, *127*(1), 194–201. http://dx.doi.org/10.1104/pp.127.1.194.

Chinsembu, K. C. (2016). *Ethnobotanical study of plants used in the management of HIV/AIDS-related diseases in Livingstone, Southern Province, Zambia. Evidence-based complementary and alternative medicine, 2016*. http://dx.doi.org/10.1155/2016/4238625.

Chweya, J. A., & Nzava, A. M. (1997). *Cat's whiskers. Cleome gynandra L. promoting the conservation and use of underutilized and neglected crops. 11*. Rome, Italy: Institute of Plant Genetics and Crop Plant Research, Gatersleben/International Plant Genetic Resources Institute.

Commission Regulation (EC) No 629/2008 (2008). Amending regulation (EC) no 1881/2006 setting maximum levels for certain contaminants in foodstuffs (text with EEA relevance). *Official Journal of European Union L*, *173*, 6–9. http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex:32008R0629.

Dasgupta, N., & De, B. (2007). Antioxidant activity of some leafy vegetables of India: A comparative study. *Food Chemistry*, *101*(2), 471–474. http://dx.doi.org/10.1016/j.foodchem.2006.02.003.

Dueñas, M., Surco-Laos, F., González-Manzano, S., González-Paramás, A. M., & Santos-Buelga, C. (2011). Antioxidant properties of major metabolites of quercetin. *European Food Research and Technology*, *232*(1), 103–111. http://dx.doi.org/10.1007/s00217-010-1363-y.

FAO (2010). The state of food insecurity in the world addressing food insecurity in protracted crises 2010 key messages. Notes http://www.fao.org/docrep/013/i1683e/i1683e.pdf.

FAO/WHO (2004). *Vitamin and mineral requirements in human nutrition*. World Health Organization and Food and Agriculture Organization of the United Nations.

Habtemariam, S. (2016). Rutin as a natural therapy for Alzheimer's disease: Insight into its mechanisms of action. *Current Medical Chemistry*, *23*(9), 860–873.

Hall, J. C., Iltis, H. H., & Sytsma, K. J. (2004). Molecular phylogenetics of Core Brassicales, placement of orphan genera *Emblingia*, *Forchhammeria*, *Tirania*, and character evolution. *Systematic Botany*, *29*(3), 654–669. http://dx.doi.org/10.1600/0363644041744491.

Hall, J. C., Sytsma, K. J., & Iltis, H. H. (2002). Phylogeny of Capparaceae and Brassicaceae and based on chloroplast sequence data 1. *American Journal of Botany*, *89*(11), 1826–1842.

Hämäläinen, M., Nieminen, R., Vuorela, P., Heinonen, M., & Moilanen, E. (2007). Anti-inflammatory effects of flavonoids: Genistein, kaempferol, quercetin, and daidzein inhibit STAT-1 and NF-κB activations, whereas flavone, isorhamnetin, naringenin, and pelargonidin inhibit only NF-κB activation along with their inhibitory effect on iNOS expression and NO production in active macrophages. *Mediators of Inflammation*, *2007*. http://dx.doi.org/10.1155/2007/45673.

Hanschen, F. S., Lamy, E., Schreiner, M., & Rohn, S. (2014). Reactivity and stability of glucosinolates and their breakdown products in foods. *Angewandte Chemie - International Edition*, *53*(43), 11430–11450. http://dx.doi.org/10.1002/anie.201402639.

Hatier, J. B., & Gould, K. S. (2009). *Anthocyanin function in vegetative organs. Anthocyanins*. http://dx.doi.org/10.1007/978-0-387-77335-3.

Hosseinzadeh, H., & Nassiri-Asl, M. (2014). Review of the protective effects of rutin on the metabolic function as an important dietary flavonoid. *Journal of Endocrinological Investigation*, *37*(9), 783–788. http://dx.doi.org/10.1007/s40618-014-0096-3.

HCDA (Horticultural Crops Development Authority) (2014). Horticulture Data 2011–2013 Validation Report. *Horticultural Crops Development Authority*, *2014*, 26–27.

Hutchinson, M. J. (2011). The effect of farmyard manure and calcium ammonium nitrate fertilizers on micronutrient density (iron, zinc, manganese, calcium and potassium) and seed yields of *Solanum villosum* (Black nightshade) and *Cleome gynandra* (Cat's whiskers) on eutric nitisol. *Journal of Agriculture Science and Technology*, *13*(1), 35–52.

Ibarra, M., Moreno, L., Vera, R., Cogolludo, A., Duarte, J., Tamargo, J., & Perez-Vizcaino, F. (2003). Effects of the flavonoid quercetin and its methylated metabolite isorhamnetin in isolated arteries from spontaneously hypertensive rats. *Planta Medica*, *69*(11), 995–1000. http://dx.doi.org/10.1055/s-2003-45144.

Jinazali, H., Mtimuni, B., & Chilembwe, E. (2017). Nutrient composition of Cat's whiskers (*Cleome gynandra* L.) from different agro ecological zones in Malawi. *African Journal of Food Science*, *11*(1), 24–29. http://dx.doi.org/10.5897/AJFS2016.1478.

Jones, J. B. (2001). *Laboratory guide for conducting soil tests and plant analysis*. Boca Raton, Fl: CRC Press.

Klimov, S. V., Burakhanova, E. A., Dubinina, I. M., Alieva, G. P., Sal'nikova, E. B., Olenichenko, N. A., ... Trunova, T. I. (2008). Suppression of the source activity affects carbon distribution and frost hardiness of vegetating winter wheat plants. *Russian Journal of Plant Physiology*, *55*(3), 308–314. http://dx.doi.org/10.1134/S1021443708030035.

Klug, B., Kirchner, T. W., & Horst, W. J. (2015). Differences in aluminium accumulation and resistance between genotypes of the genus *Fagopyrum*. *Agronomy*, *5*, 418–434. http://dx.doi.org/10.3390/agronomy5030418.

Koval'skii, I. V., Krasnyuk, I. I., Krasnyuk, I. I., Jr., Nikulina, O. I., Belyatskaya, A. V., & Kharitonov, Yu. Ya., Feldman, N. B. and Lutsenko, S. V. (2011). Molecular-biological problems of drug design and mechanism of drug action. *Pharmaceutical Chemistry Journal*, *44*(12), 651–653. http://dx.doi.org/10.1007/s11094-011-0539-5.

Knekt, P., Kumpulainen, J., Järvinen, R., Rissanen, H., Heliövaara, M., Reunanen, A., ... Aromaa, A. (2002). Flavonoid intake and risk of chronic diseases. *The American Journal of Clinical Nutrition*, *76*(3), 560–568.

Kumar, S., & Pandey, A. K. (2013). Chemistry and biological activities of flavonoids: An overview. *The Scientific World Journal*, *2013*, 16. http://dx.doi.org/10.1155/2013/162750.

Krumbein, A., Saeger-Fink, H., & Schonhof, I. (2007). Changes in quercetin and kaempferol concentrations during broccoli head ontogeny in three broccoli cultivars. *Journal of Applied Botany and Food Quality*, *81*(2), 136–139.

Kutsukutsa, R. T., Gasura, E., Mabasa, S., & Ngadze, E. (2014). Variability in condensed tannins and bitterness in spider plant genotypes. *African Crop Science Journal*, *22*(4), 275–280.

K'Opondo, F. B. O., Muasya, R. M., & Kiplagat, O. K. (2005). A review on the seed production and handling of indigenous vegetables (spider plant, jute mallow and African nightshade complex). In M. O. Abukutsa-Onyango, A. N. Murithi, V. E. Anjichi, K. Ngamau, S. G. Agong, A. Fricke, B. Hau, & H. Stutzel (Eds.), *Proceedings of the third horticulture workshop on sustainable horticultural production in the Tropics* (pp. 42–48). Maseno: Maseno University.

K'Opondo, F. B. O. (2011). Morphological characterization of selected spiderplant (*Cleome gynandra* L.) types from western Kenya. *Annals of Biological Research*, *2*(2), 54–64.

Latté, K. P., Appel, K. E., & Lampen, A. (2011). Health benefits and possible risks of broccoli — An overview. *Food and Chemical Toxicology*, *49*(12), 3287–3309. http://dx.doi.org/10.1016/j.fct.2011.08.019.

Li, Z., Lu, W., Yang, L., Kong, X., & Deng, X. (2015). Seed weight and germination behavior of the submerged plant *Potamogeton pectinatus* in the arid zone of northwest China. *Ecology and Evolution*, *5*(7), 1504–1512. http://dx.doi.org/10.1002/ece3.1451.

Lin, J., Rexrode, K. M., Hu, F., Albert, C. M., Chae, C. U., Rimm, E. B., ... Manson, J. E. (2007). Dietary intakes of flavonols and flavones and coronary heart disease in US women. *American Journal of Epidemiology*, *165*(11), 1305–1313. http://dx.doi.org/10.1093/aje/kwm016.

Meda, N. T. R., Bangou, M. J., Bakasso, S., Millogo-Rasolodimby, J., & Nacoulma, O. G.

(2013). Antioxidant activity of phenolic and flavonoid fractions of *Cleome gynandra* and *Maerua angolensis* of Burkina Faso. *Journal of Applied Pharmaceutical Science*, *3*(2), 36–42. http://dx.doi.org/10.7324/JAPS.2013.30207.

Metali, F., Salim, K. A., & Burslem, D. F. R. P. (2012). Evidence of foliar aluminium accumulation in local, regional and global datasets of wild plants. *New Phytologist*, *193*(3), 637–649. http://dx.doi.org/10.1111/j.1469-8137.2011.03965.x.

Mewis, I., Schreiner, M., Nguyen, C. N., Krumbein, A., Ulrichs, C., Lohse, M., & Zrenner, R. (2012). UV-B irradiation changes specifically the secondary metabolite profile in broccoli sprouts: Induced signaling overlaps with defense response to biotic stressors. *Plant and Cell Physiology*, *53*(9), 1546–1560. http://dx.doi.org/10.1093/pcp/pcs096.

Mishra, S. S., Moharana, S. K., & Dash, M. R. (2011). Review on *Cleome gynandra*. *International Journal of Research in Pharmacy and Chemistry*, *1*(3), 681–689.

Mziray, S. R. (1999). *Nutrient and antinutrient of raw, cooked, sundried and stored vegetable Amaranth grown in Dar es salaam, Tanzania. PhD thesis.* University of Nairobi.

Nesamvuni, C., Steyn, N. P., & Potgieter, M. J. (2001). Nutritional value of wild, leafy plants consumed by the Vhavenda. *South African Journal of Science*, *97*, 51–54.

Neugart, S., Kläring, H. P., Zietz, M., Schreiner, M., Rohn, S., Kroh, L. W., & Krumbein, A. (2012). The effect of temperature and radiation on flavonols aglycones and flanonol glycosides of kale (*Brassica oleracea* var. *sabellica*). *Food Chemistry*, *133*(4), 1456–1465.

Neugart, S., Baldermann, S., Ngwene, B., Wesonga, J., & Schreiner, M. (2017). *Indigenous leafy vegetables of Eastern Africa — A source of extraordinary secondary plant metabolites*. Food Research Internationalhttp://dx.doi.org/10.1016/j.foodres.2017.02.014.

Neugart, S., Rohn, S., & Schreiner, M. (2015). Identification of complex, naturally occurring flavonoid glycosides in *Vicia faba* and *Pisum sativum* leaves by HPLC-DAD-ESI-MSn and the genotypic effect on their flavonoid profile. *Food Research International*, *76*(P1), 114–121. http://dx.doi.org/10.1016/j.foodres.2015.02.021.

Odhav, B., Beekrum, S., Akula, U., & Baijnath, H. (2007). Preliminary assessment of nutritional value of traditional leafy vegetables in KwaZulu-Natal, South Africa. *Journal of Food Composition and Analysis*, *20*(5), 430–435. http://dx.doi.org/10.1016/j.jfca.2006.04.015.

Odongo, G. A., Schlotz, N., Herz, C., Hanschen, F. S., Baldermann, S., Neugart, S., ... Lamy, E. (2017). The role of plant processing for the cancer preventive potential of Ethiopian kale (*Brassica carinata*). *Food & Nutrition Research*, *61*(1), 1–11. http://dx.doi.org/10.1080/16546628.2017.1271527.

Ojiewo, C. O., Mwai, G. N., Abukutsa-Onyango, M. O., Agong, S. O., & Nono-Womdim, R. (2013). Exploiting the genetic diversity of vegetable African nightshades. *Bioremediation, Biodiversity and Bioavailability*, *7*(1), 6–13. http://www.researchgate.net/publication/235322354_Exploiting_the_Genetic_Diversity_of_Vegetable_African_Nightshades/file/32bfe510bcae797a30.pdf.

Omondi, E. O., Debener, T., Linde, M., Dinssa, F. F., Abukutsa-Onyango, M., & Winkelmann, T. (2017). Mating biology, nuclear DNA content and genetic diversity in spider plant (*Cleome gynandra*) germplasm from various African countries. *Plant Breeding*. http://dx.doi.org/10.1111/pbr.12485.

Onyango, C., Onwonga, R., & Kimenju, J. (2016). Assessment of spider plant (*Cleome gynandra* L.) germplasm for agronomic traits in vegetable and seed production: A green house study. *American Journal of Experimental Agriculture*, *10*(1), 1–10. http://dx.doi.org/10.9734/AJEA/2016/20209.

Reintanz, B. (2001). Bus, a bushy *Arabidopsis* CYP79F1 knockout mutant with abolished synthesis of short-chain aliphatic glucosinolates. *The Plant Cell Online*, *13*(2), 351–367. http://dx.doi.org/10.1105/tpc.13.2.351.

Reiter, M., Rupp, K., Baumeister, P., Zieger, S., & Harreus, U. (2009). Antioxidant effects of quercetin and coenzyme Q10 in mini organ cultures of human nasal mucosa cells. *Anticancer Research*, *29*, 33–39.

Ruel, M. T., Mino, N., & Smith, L. (2005). Patterns and determinants of fruits and vegetable consumption in sub-Saharan Africa: A multicountry comparison. *FAO/WHO Workshop on Fruit and Vegetables for Health, 1–3 September 2004, Kobe, Japan*.

Sang, J. P., Minchinton, I. R., Johnstone, P. K., & Trus-Cott, R. J. W. (1984). Glucosinolate profiles in the seed, root and leaf tissue of cabbage, mustard, rapeseed, radish and swede. *Canadian Journal of Plant Science*, *64*, 77–93.

Schmidt, S., Zietz, M., Schreiner, M., Rohn, S., Kroh, L. W., & Krumbein, A. (2010). Genotypic and climatic influences on the concentration and composition of flavonoids in kale (*Brassica oleracea* var. *sabellica*). *Food Chemistry*, *119*(4), 1293–1299. http://dx.doi.org/10.1016/j.foodchem.2009.09.004.

Songsak, T., & Lockwood, G. B. (2002). Glucosinolates of seven medicinal plants from Thailand. *Fitoterapia*, *73*(3), 209–216. http://dx.doi.org/10.1016/S0367-326X(02)00061-8.

Stoilova, T., Dinssa, F. F., Ebert, A. W., & Tenkouano, A. (2015). The diversity of African leafy vegetables: Agromorphological characterization of subsets of AVRDC's germplasm collection. *Acta Horticulturae*, *1102*, 67–74. http://dx.doi.org/10.17660/ActaHortic.2015.1102.7.

Tabuti, J. R. S., Dhillion, S. S., & Lye, K. A. (2004). The status of wild food plants in Bulamogi County, Uganda. *International Journal of Food Sciences and Nutrition*, *55*(6), 485–498. http://dx.doi.org/10.1080/09637480400015745.

van Rensburg, J. W. S., Venter, S. L., Netshiluvhi, T. R., van den Heever, E., Vorster, H. J., & de Ronde, J. A. (2004). Role of indigenous leafy vegetables in combating hunger and malnutrition. *South African Journal of Botany*, *70*(1), 52–59. http://dx.doi.org/10.1016/S0254-6299(15)30268-4.

Velasco, P., Cartea, M. E., Gonzalez, C., Vilar, M., & Ordas, A. (2007). Factors affecting the glucosinolate content of kale (*Brassica oleracea* acephala group). *Journal of Agricultural and Food Chemistry*, *55*(3), 955–962. http://dx.doi.org/10.1021/jf0624897.

Verkerk, R., Schreiner, M., Krumbein, A., Ciska, E., Holst, B., Rowland, I., ... Dekker, M. (2009). Glucosinolates in Brassica vegetables: The influence of the food supply chain on intake, bioavailability and human health. *Molecular Nutrition and Food Research*, *53*(2), 219–265. http://dx.doi.org/10.1002/mnfr.200800065.

Wangolo, E., Onyango, C., Gachene, C., & Mong'are, P. (2015). Effects of shoot tip and flower removal on growth and yield of spider plant (*Cleome gynandra* L.) in Kenya. *American Journal of Experimental Agriculture*, *8*(6), 367–376. http://dx.doi.org/10.9734/AJEA/2015/17271.

Wasonga, D. O., Ambuko, J. L., Chemining'wa, G. N., Odeny, D. A., & Crampton, B. G. (2015). Morphological characterization and selection of spider plant (*Cleome gynandra*) cccessions from Kenya and South Africa. *Asian Journal of Agricultural Science*, *7*(4), 36–44.

Wiesner, M., Hanschen, F. S., Schreiner, M., Glatt, H., & Zrenner, R. (2013). Induced production of 1-methoxy-indol-3-ylmethyl glucosinolate by jasmonic acid and methyl jasmonate in sprouts and leaves of pak choi (*Brassica rapa ssp. chinensis*). *International Journal of Molecular Sciences*, *14*(7), 14996–15016. http://dx.doi.org/10.3390/ijms140714996.

Yim, B., Hanschen, F. S., Wrede, A., Nitt, H., Schreiner, M., Smalla, K., & Winkelmann, T. (2016). Effects of biofumigation using *Brassica juncea* and *Raphanus sativus* in comparison to disinfection using Basamid on apple plant growth and soil microbial communities at three field sites with replant disease. *Plant and Soil*, *406*, 389–408. http://dx.doi.org/10.1007/s11104-016-2876-3.

Zietz, M., Weckmüller, A., Schmidt, S., Rohn, S., Schreiner, M., Krumbein, A., & Kroh, L. W. (2010). Genotypic and climatic influence on the antioxidant activity of flavonoids in kale (*Brassica oleracea* var. *sabellica*). *Journal of Agricultural and Food Chemistry*, *58*(4), 2123–2130. http://dx.doi.org/10.1021/jf9033909.

Supplementary materials

Table A1: Thirty *C. gynandra* farmers' cultivars, advanced lines and gene bank accessions used in the study.

Entry	Category	Source	Coordinate	Original entry source
Abuku1	Advanced line	JKUAT	-	Kenya
Abuku2	Advanced line	JKUAT	-	Kenya
Acc20	Farmers' cultivar	Bondo	0°06'07.5"S 34°16'26.9"E	Kenya
Acc21	Farmers' cultivar	Bondo	0°06'07.5"S 34°16'26.9"E	Kenya
Acc26	Farmers' cultivar	Kisii	0°44'48.5"S 34°47'68.5"E	Kenya
Acc28	Farmers' cultivar	Kisii	0°40'51.5"S 34°47'41.1"E	Kenya
Acc30	Farmers' cultivar	Nakuru	0°02'41.9"S 36°01'49.2"E	Kenya
Acc3	Farmers' cultivar	Mbale	0°06'08.0"S 34°42'28.0"E	Kenya
Acc5	Farmers' cultivar	Yala	0°09'24.8"N 34°33'02.2"E	Kenya
Acc6	Farmers' cultivar	Butere	0°10'07.0"N 34°31'36.8"E	Kenya
HTT	Advanced line	WorldVeg	-	Kenya
GS	Advanced line	WorldVeg	-	Tanzania
IP7	Gene bank accession	WorldVeg	-	Tanzania
PS	Gene bank accession	WorldVeg	-	Tanzania
Site94	Gene bank accession	WorldVeg	-	Tanzania
ST73-3	Gene bank accession	WorldVeg	-	Tanzania
ST93-1GS	Gene bank accession	WorldVeg	-	Tanzania
IP12	Gene bank accession	WorldVeg	-	S. Africa
IP8-Sel	Advanced line	WorldVeg	-	S. Africa
MLSF12	Gene bank accession	WorldVeg	-	Malawi
MLSF17	Gene bank accession	WorldVeg	-	Malawi
MLSF27	Gene bank accession	WorldVeg	-	Malawi
RWSF2	Gene bank accession	WorldVeg	-	Rwanda
RWSF3	Gene bank accession	WorldVeg	-	Rwanda
UGSF13	Gene bank accession	WorldVeg	-	Uganda
UGSF17	Advanced line	WorldVeg	-	Uganda
UGSF19	Gene bank accession	WorldVeg	-	Uganda
UGSF2	Advanced line	WorldVeg	-	Uganda
UGSF26	Gene bank accession	WorldVeg	-	Uganda
UGSF29	Gene bank accession	WorldVeg	-	Uganda

WorldVeg- World Vegetable Center, Arusha, Tanzania. JKUAT-Jomo Kenyatta University of Agriculture and Technology, Nairobi, Kenya. The three gray scales represent the source of the seeds used as JKUAT, farmers' cultivars and WorldVeg respectively.

Table A2: Mean±SD for six morphological traits of 30 entries of *C. gynandra*.

Entry	Sil. W (g)	PBN	Leaf FM(g)	Leaf DM (g)	Total FM (g)	Total DM (g)
Abuku1	0.79^{a}±0.3	9.9^{efgh}±2.1	13.8^{bcde}±4.0	2.1^{cdefg}±0.5	39.1^{def}±12.0	6.4^{bcdef}±2.0
Abuku2	0.83^{a}±0.2	9.2^{gh}±1.5	10.9^{de}±0.5	1.7^{efg}±0.1	32.0^{ef}±3.9	5.3^{cdef}±1.4
Acc20	0.91^{a}±0.0	8.7^{h}±2.0	23.7^{a}±9.3	3.3^{ab}±1.3	70.4^{a}±7.9	12.7^{a}±3.1
Acc21	0.91^{a}±0.3	9.5^{fgh}±1.5	12.8^{bcde}±2.4	2.0^{cdefg}±0.2	33.6ef±6.2	4.7^{cdef}±1.0
Acc26	0.72^{a}±0.1	$10.5^{bcdefgh}$±2.2	13.9^{bcde}±3.1	2.1^{bcdefg}±1.1	41.1^{cdef}±18.5	7.7^{bcdef}±5.4
Acc28	0.94^{a}±0.1	12.3^{abcd}±2.1	9.0^{e}±1.7	1.4^{g}±0.0	28.1^{f}±6.1	3.7^{f}±0.7
Acc3	0.78^{a}±0.1	9.9^{efgh}±1.8	11.2^{de}±2.1	1.7^{efg}±0.31	32.7^{ef}±3.7	4.5^{def}±0.4
Acc30	0.74^{a}±0.2	10.4^{cdefgh}±2.2	9.8^{de}±0.9	1.5^{fg}±0.1	30.3^{ef}±2.9	4.5^{def}±0.9

Acc5	0.85^{a}±0.2	8.5^{h}±1.7	9.2^{e}±1.8	1.4^{g}±0.4	25.7^{f}±3.0	3.9^{ef}±0.8
Acc6	0.84^{a}±0.2	8.8^{h}±1.8	11.8^{cde}±4.6	1.7^{efg}±0.7	32.7^{ef}±13.6	4.8^{cdef}±2.2
GS	1.08^{a}±0.4	9.6^{fgh}±1.8	20.4^{ab}±2.4	3.0^{abcd}±0.2	58.8^{abcd}±21.5	8.3^{bcd}±0.6
HTT	0.87^{a}±0.2	10.4^{cdefgh}±2.3	13.9^{bcde}±4.2	2.1^{cdefg}±1.1	38.8^{def}±12.7	6.6^{bcdef}±3.3
IP12	0.78^{a}±0.1	11.4^{abcdefg}±2.8	20.2^{abc}±6.8	3.1^{abc}±1.2	59.1^{abcd}±22.0	8.1^{bcde}±3.0
IP7	0.86^{a}±0.2	13.2^{a}±2.2	14.9^{bcde}±2.9	2.3^{bcdefg}±0.6	41.0^{cdef}±2.5	5.2^{cdef}±0.2
IP8	0.70^{a}±0.1	11.7^{abcdef}±2.9	18.0^{abcd}±8.0	2.8^{abcde}±1.3	53.2^{abcde}±24.4	8.9^{abc}±5.4
MLSF12	0.77^{a}±0.2	12.7^{abc}±2.1	17.6^{abcd}±5.0	2.6^{abcdef}±0.6	50.2^{abcdef}±14.6	6.8^{bcdef}±2.0
MLSF17	0.65^{a}±0.1	11.6^{abcdef}±2.2	11.7^{de}±3.6	1.8^{defg}±0.5	30.4^{ef}±3.8	4.2^{def}±0.3
MLSF27	0.80^{a}±0.2	11.4^{abcdefg}±2.3	15.9^{abcde}±1.8	2.5^{abcdefg}±0.4	49.8^{abcdef}±6.6	7.7^{bcdef}±1.7
PS	0.90^{a}±0.2	11.4^{abcdefg}±2.4	12.7^{bcde}±1.4	2.0^{cdefg}±0.6	35.8^{def}±4.8	5.3^{cdef}±2.1
RWSF2	0.84^{a}±0.2	11.7^{abcdef}±3.0	24.2^{a}±2.0	3.7^{a}±0.7	64.7^{abc}±6.6	10.1^{ab}±1.7
RWSF3	0.87^{a}±0.2	10.3^{cdefgh}±2.2	12.0^{bcde}±3.2	1.8^{defg}±0.8	34.3^{def}±6.2	4.9^{cdef}±0.8
Site 94	0.76^{a}±0.1	11.7^{abcdef}±2.5	13.2^{bcde}±8.2	1.9^{defg}±0.9	37.2^{def}±11.9	5.1^{cdef}±1.3
ST73-3	0.63^{a}±0.1	12.8^{ab}±2.8	10.3^{de}±3.7	1.6^{efg}±0.5	32.7^{ef}±14.7	5.0^{cdef}±1.9
ST93-1(GS)	0.79^{a}±0.3	13.2^{a}±3.0	15.1^{bcde}±2.5	2.1^{cdefg}±0.6	44.1^{bcdef}±11.8	6.2^{bcdef}±2.0
UGSF13	0.79^{a}±0.3	10.0^{defgh}±1.9	24.4^{a}±4.6	3.2^{abc}±0.3	68.3^{ab}±21.3	10.6^{ab}±4.6
UGSF17	0.81^{a}±0.2	10.3^{defgh}±2.1	11.0^{de}±4.2	1.7^{efg}±0.8	29.2^{ef}±9.5	4.7^{cdef}±2.2
UGSF19	0.73^{a}±0.2	12.0^{abcde}±1.8	11.9^{cde}±2.4	1.8^{defg}±0.4	32.7^{ef}±2.5	4.8^{cdef}±00.4
UGSF2	0.93^{a}±0.0	8.6^{h}±2.2	9.7^{de}±9.7	1.5^{fg}±0.6	26.2^{f}±7.7	3.8^{ef}±1.7
UGSF26	0.83^{a}±0.2	9.6^{fgh}±2.1	13.1^{bcde}±1.9	2.0^{cdefg}±0.3	37.8^{def}±2.4	6.0^{bcdef}±0.8
UGSF29	0.72^{a}±0.2	10.7^{bcdefgh}±2.7	17.8^{abcd}±3.7	2.5^{abcdefg}±0.6	43.4^{bcdef}±9.4	5.6^{cdef}±1.1

Mean value followed by the same letter in columns belong to the same category in accordance with Tukey test (p≤0.05). Sil. W= Silique weight, PBN= Primary brunch number, FM= Fresh mass, and DM= Dry mass. n=4.Cells with the minimum and maximum values within a column are highlighted by green and yellow color, respectively. The three gray scales represent the source of the seeds used as JKUAT, farmers' cultivars and WorldVeg respectively.

Table A3: Glucosinolate(GS) concentrations (µmol/g dw) of selected *C. gynandra* entries grown under greenhouse conditions.

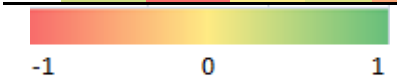
Entry	Plant part	Methyl GS	2-propenyl GS	4-hydroxy-3-indolylmethyl GS	Indolyl-3-methyl GS	Total GS
ML-SF-12	leaves	36.46	nd	nd	0.03	36.50
ML-SF-17	leaves	61.52	nd	0.04	0.11	61.67
ML-SF-27	leaves	45.04	nd	nd	0.06	45.10
UG-SF-13	leaves	41.25	nd	nd	0.03	41.29
UG-SF-17	leaves	13.36	nd	nd	0.04	13.40
UG-SF-19	leaves	65.71	nd	nd	0.28	65.99
UG-SF-2	leaves	19.56	nd	nd	0.01	19.57
UG-SF-26	leaves	39.41	nd	nd	0.12	39.53
UG-SF-29	leaves	26.57	nd	nd	0.14	26.71
GS	leaves	28.04	nd	nd	0.02	28.06
PS	leaves	33.47	0.52	nd	0.07	34.07
HTT1	leaves	30.72	nd	nd	0.01	30.73
HTT2	leaves	38.70	nd	0.03	0.03	38.76
HTT3	leaves	31.75	nd	0.02	0.02	31.78

HTT4	leaves	39.98	nd	0.03	0.02	40.03
HTT5	leaves	43.83	nd	0.02	0.01	43.86
HTT6	leaves	32.08	nd	nd	0.01	32.09
HTT7	leaves	44.08	nd	nd	0.05	44.12
HTT8	leaves	42.95	nd	0.05	0.03	43.03
HTT9	leaves	62.59	nd	0.07	0.05	62.71
HTT10	leaves	33.83	nd	nd	0.01	33.84
HTT11	leaves	32.08	nd	nd	0.03	32.11
***HTT_{mean}**	leaves	**39.33±9.3**	**-**	**0.04±0.02**	**0.03±0.02**	**41.32±11.78**
†HTT_A	leaves	49.32	nd	0.03	0.05	49.41
HTT	Inflores-cences	57.75	nd	0.44	0.05	58.24
HTT	stem	67.60	nd	0.87	0.24	68.72
ST 73-3	leaves	31.93	nd	nd	0.01	31.94
ST 92	leaves	74.91	nd	nd	0.02	74.93
ST 93-1 (GS)	leaves	49.77	nd	0.02	0.11	49.90
RW SF 2	leaves	40.93	nd	nd	0.02	40.95
RW SF 3	leaves	27.57	nd	nd	0.04	27.61
Site 69	leaves	26.44	nd	nd	0.01	26.45
Site 94	leaves	41.25	nd	nd	0.04	41.29
IP7	leaves	41.94	nd	0.01	0.21	42.16
IP8	leaves	22.53	nd	nd	0.03	22.56
IP12	leaves	38.32	nd	nd	0.01	38.34

*Mean value ± SD for leaves of 11 individual plant samples shown as HTT1-HTT11†HTT pooled leaves sample from 10 plants. GS= Glucosinolate. nd=not detected.

Table A4: Heat map for the pair-wise Pearson correlations between different minerals form leaves of selected *C. gynandra* entries

	K	Ca	Mg	P	S	Al	Fe	Mn	Zn	Cu	Ni	Mo	Cd	Pb
Ca	-0.14													
Mg	**0.61****	0.32												
P	**0.72****	-0.12	**0.48***											
S	0.32	-0.03	0.09	0.24										
Al	0.37	-0.23	0.27	0.20	0.10									
Fe	0.24	-0.29	0.07	0.14	-0.04	**0.94****								
Mn	**0.42***	-0.16	0.36	0.26	-0.09	**0.95****	**0.91****							
Zn	0.33	0.24	0.23	0.31	0.01	0.13	0.10	0.20						
Cu	0.18	0.11	0.05	0.04	-0.12	0.05	0.03	0.11	**0.47***					
Ni	0.39	-0.28	0.28	0.22	-0.10	**0.73****	**0.68****	**0.74****	0.19	0.39				
Mo	-0.26	**0.60****	-0.11	-0.13	-0.34	-0.26	-0.16	-0.10	0.30	-0.05	-0.30			
Cd	0.29	-0.12	0.37	-0.11	0.24	**0.67****	**0.53****	**0.57****	0.04	0.14	**0.64***	-0.35		
Pb	0.24	0.00	0.06	0.09	-0.12	0.18	0.17	0.23	**0.49**	**0.97****	**0.50***	-0.10	0.23	
Cr	**0.46***	-0.32	0.26	0.39	-0.07	**0.78****	**0.81****	**0.83****	0.23	0.25	**0.72****	-0.17	**0.46***	**0.40***



The significance levels represented as follows: * represents p<0.05, ** p<0.01, *** p<0.001, yellow-green: positive correlation, red: negative correlation**.**

Table A5: Average mineral composition (mg/100g) in *C. gynandra* germplasm in the present study in comparison with previous studies based on dry weight.

Mineral element	Present study	Jinazali et al., (2017)	(Odhav et al., 2007)	(Agbo et al., 2014)	
				Dry season	Rainy season
Potassium	3123	-	-	109.38	58.96
Calcium	2316	2210	3203	22.65	45.25
Magnesium	440	-	371	7.28	15.08
Phosphorus	659	-	784	34.0	4.19
Sulphur	404	-	-		
Aluminium	295	-	-		
Iron	358	36	21	23.98	10.17
Manganese	39	-	10		
Zinc	5.9	2.6	5		
Copper	3.3	-	2		
Nickel	0.3	-	-		
Molybdenum	0.2	-	-		
Cadmium	0.03	-	-		
Lead	0.5	-	-		
Chromium	0.5	-	-		

2.4. Genetic and morphological diversity of cowpea (*Vigna unguiculata* (L.) Walp.) entries from East Africa

Max Menssen[1,5], Marcus Linde[2], Emmanuel Omondi[1,] Mary Abukutsa-Onyango[3], Fekadu Fufa Dinssa[4], Traud Winkelmann[1]

[1]Institute of Horticultural Production Systems, Woody Plant and Propagation Physiology, Leibniz Universität Hannover, Hannover, Germany

[2]Institute for Plant Genetics, Molecular Plant Breeding, Leibniz Universität Hannover, Hannover, Germany

[3]Department Horticulture, Jomo Kenyatta University of Agriculture & Technology (JKUAT), Juja, Kenya

[4]AVRDC - The World Vegetable Center, Eastern and Southern Africa, Arusha, Tanzania

[5]Leibniz Universität Hannover, Institute of Biostatistics (present address)

Type of authorship	Third author
Type of article	Research article
Contribution to the article	Assisted in the experiments and writing the manuscript
Contribution of other authors	Max Menssen performed all the experiments for his masters thesis research Marcus Linde and Traud Winkelmann assisted in experimental design, data analysis and writing the manuscript. Mary Abukutsa-Onyango and Fekadu F. Dinssa provided the research material and assisted in writing the manuscript.
Journal	Scientia Horticulturae
Date of publication	1st August 2017
DOI	10.1016/j.scienta.2017.08.003

Scientia Horticulturae 226 (2017) 268–276

Contents lists available at ScienceDirect

Scientia Horticulturae

journal homepage: www.elsevier.com/locate/scihorti

Research Paper

Genetic and morphological diversity of cowpea (*Vigna unguiculata* (L.) Walp.) entries from East Africa

Max Menssen[a,e,*], Marcus Linde[b], Emmanuel Otunga Omondi[a], Mary Abukutsa-Onyango[c], Fekadu Fufa Dinssa[d], Traud Winkelmann[a]

[a] Leibniz Universität Hannover, Institute of Horticultural Production Systems, Germany
[b] Leibniz Universität Hannover, Institute of Plant Genetics, Germany
[c] Jomo Kenyatta University of Agriculture and Technology, Department of Horticulture, Kenya
[d] World Vegetable Center, Eastern and Southern Africa, Arusha, Tanzania
[e] Leibniz Universität Hannover, Institute of Biostatistics, Germany

ARTICLE INFO

Keywords:
African indigenous vegetables
DNA content
Genetic diversity
Microsatellites
Molecular marker
Morphological diversity
Yield

ABSTRACT

Cowpea (*Vigna unguiculata* (L.) Walp.) is widely grown by subsistence farmers in West and East Africa where its grain and leaves are sources of highly valuable food, due to their high contents of proteins, minerals and vitamins. Therefore, cowpea could play a significant role in mitigating malnutrition such as micronutrient deficiencies. The objectives of this study were to evaluate the performance in agronomic traits of cowpea entries (released and farmers' cultivars, and gene bank accessions) from different sources, and to assess the extent of genetic diversity in this material to provide basic information for its use in breeding programs.

A total of 15 entries were evaluated. All entries were morphologically uniform, except the farmers' local cultivars. Fresh leaf yield varied from 34.6 to 52.8 g per plant and days to 50% flowering from 64 to 82 days. Hundred seed weight ranged from 7.67 to 15.12 g. On average, the number of pods ranged from 4.8 to 15.6 pods per plant. No correlation between fresh leaf yield and other traits was detected, whereas the number of pods per plant and the hundred seed weight were negatively correlated.

Genetic diversity was assessed on five genotypes per entry using 544 Amplified Fragment Length Polymorphism (AFLP) and 18 microsatellite markers. Genetic distances calculated using the Jaccard algorithm ranged from 0.002 to 0.193 among genotypes of the same entry and from 0.098 to 0.301 for genotypes from different entries. A principle coordinate analysis separated four entries from the rest. Although the consensus tree based on Neighbor Joining trees was unable to resolve the whole cluster, an assignment of most of the entries into entry-specific clades was possible.

1. Introduction

Cowpea is a dicotyledonous grain legume that belongs to the genus *Vigna* within the family of *Fabaceae* and is native to Africa (Padulosi and Ng, 1997). Pasquet (1993a, 1993b, 1997, 1998) divided *V. unguiculata* into eleven perennial and one annual subspecies (*ssp. unguiculata*) that includes both, the cultivated forms (*var. unguiculata*) and a wild form (*var. spontanea)* which is treated as the wild progenitor of cultivated cowpea. The cultivated cowpea forms are grouped into five cultigroups (cv.-gr.) (Pasquet, 1998): The cultigroups *Textilis*, *Unguiculata* and *Melanophthalmus* grow in Africa, whereas *Sesquipedalis* occurs in Asia. *Biflora* is cultivated in northern Africa, on the Arabian Peninsula and in Asia (Pasquet, 2000). Members of the cv.-gr. *Unguiculata,* to which the entries used in the present study belong, are late flowering even under inductive conditions.

Until now the center of domestication is quite unclear. Ng and Marechal (1985) postulated two independent centers for primary domestication of cowpea: one in the Zambezian region in eastern Africa and another in western Africa. Because the morphological diversity of *var. spontanea* is highest between Ethiopia and South Africa, the whole eastern part of the African continent was supposed to be the center of domestication (Baudoin and Marechal, 1985). Due to chloroplast DNA polymorphisms observed in wild *(var. spontanea)* and cultivated cowpea (*var. unguiculata*), Vaillancourt and Weeden (1992) postulated Nigeria as another center of domestication.

Vigna unguiculata is a diploid species with 2n = 2x = 22 chromosomes,a total nuclear DNA content of 1.27 pg/2C and a genome size of 613 Mbp (Arumuganathan and Earle, 1991; Faris, 1964). Due to the

* Corresponding author. present address: Leibniz Universität Hannover, Institute of Biostatistics, Germany.
E-mail address: menssen@biostat.uni-hannover.de (M. Menssen).

http://dx.doi.org/10.1016/j.scienta.2017.08.003
Received 30 March 2017; Received in revised form 27 July 2017; Accepted 1 August 2017

cleistogamous flower structure and pollen shed occurring simultaneously with stigma receptivity, cowpea is treated as a highly self-pollinated plant (Ehlers and Hall, 1997). Nevertheless, outcrossing occurs at different rates depending on the subspecies (Pasquet, 1996). In wild cowpea populations from West Africa, Kouam et al. (2012) observed outcrossing rates varying between 1 and 9.5% as well as gene flow from domesticated to wild cowpea (*var. spontanea*).

Due to the high seed protein (~30%) and carbohydrate (50–60%) content, cowpea plays a fundamental role for the human diet in many developing countries, therefore it is being referred to as "poor man's meat" (Diouf and Hilu, 2005; Nielsen et al., 1993). Ragab et al. (2004) fractionated cowpea protein into albumin (71.4%), globulin (11.1%), prolamin (2.2%) and glutelin (11.0%). Compared to cereal grains, the storage proteins of cowpea are rich in the amino acids lysine and tryptophan, but low in methionine and cysteine compared to animal protein. Cowpea grain is also rich in minerals and vitamins and has a very high level of folic acid (Timko and Singh, 2008). The leaves are also rich in vitamins and minerals. According to Ahenkora et al. (1998), the dried leaves contain 3.0–6.7 mg/g phosphorus, 0.3–1.5 mg/g ascorbic acid and 27–35% protein. The phosphorus content of cowpea leaves is twice the amount of cabbage (*Brassica oleracae,* 1.79–2.51 mg/g (Warman and Havard 1996)). With 0.5–0.6 mg/g, the amount of ascorbic acid in cabbage leaves is comparable to that in cowpea leaves (Ogbede et al., 2015). A detailed overview of the nutritional composition of both leaves and seeds is given by Stadlmayr et al. (2012).

The aims of cowpea breeding are high yields, early maturity for grain production, long vegetative period for vegetable production, high leaf and grain nutrient contents, cooking quality and high germination rate. In order to provide farmers with quality seed of improved cultivars, breeding programs and seed systems should be based on information on the genetic diversity available in the germplasm. Therefore, the aim of this study was to assess the genetic and morphological diversity of 15 East African cowpea entries representing farmers' local cultivars, gene bank accessions and released cultivars.

2. Materials and methods

2.1. Plant material

For this study, 15 cowpea entries – accessions, advanced lines and released cultivars – were used (Table 1): 10 entries from different East African countries were provided by the World Vegetable Center (WorldVeg) and five Kenyan entries (three entries originally collected from farmers' field, and two commercial cultivars from seed companies) were obtained from Jomo Kenyatta University of Agriculture and Technology (JKUAT). The five entries Dakawa, Tumaini, Ex-Iseke, Ngoji and UG-CP-4 were characterized for their DNA content and ploidy according to Braun and Winkelmann (2016).

2.2. Morphological characterization

Field assessment was carried out at JKUAT in Juja, Kenya (01°05.9′S, 037°00.8′E and 1476 m a.s.l.) from October to December 2015. A randomized complete block design with four blocks was used with each block representing one randomized replicate of the complete set of entries. The plot size was 6 m by 1.2 m with 50 cm spacing between two consecutive plots. On October 7th 2015, sowing was carried out in double density in four rows with 15 cm between the plants within a row and 30 cm between the rows. Every second plant was removed 28 days after sowing, leaving 40 plants per row. At planting time, well rotten goat manure was applied at the rate of 20 tons per hectare. The field was irrigated at least twice per week until the rainy season started in the mid of November. No chemical fertilizers or pesticides were applied. Data were collected on plants obtained in the central two rows. Twenty plants from one of the two central rows per plot were harvested 35 days after sowing and were separated into leaves and stems for biomass measurement. To determine the dry mass, the probes were dried at 90 °C for seven days. Days to 50% flowering – the period between sowing and the time when 50% of the plants per plot flowered – was recorded. Ten randomly picked pods per plot were used to assess pod length, pod width and number of locules per pod. Seed length and width were measured on 10 randomly picked seeds per plot, and hundred seed weight was measured on 10 random samples of 100 seeds each.

2.3. Statistical analyses

Statistical analyses were performed on the quantitative traits – fresh leaf yield, days to 50% flowering, number of pods per plant and 100 seed weight – using the statistical software "R" (version 3.3.0). These traits were considered as the most important traits analyzed regarding the yield. Therefore, these parameters were fitted in linear models and mean comparisons (Tukey-Tests) were performed using the package "lsmeans" (Russell, 2016). Since the fresh mass of the leaves was determined on the basis of 20 plants which were pooled to one sample, the mean leaf fresh mass per plant was computed prior to the statistical analysis.

Pearson correlation coefficients were calculated among the various quantitative traits using the function "rcorr" of the package "Hmisc" (Harrell, 2016). The construction of plots and graphics was done using packages "ggplot2" (Wickham, 2009) and "PerformanceAnalytics" (Peterson and Carl, 2014).

2.4. Molecular characterization

Approximately 70–80 mg of young leaves from five randomly chosen plants per entry were used for DNA extraction. The leaf materials were dried overnight on silica gel at 30 °C followed by extraction using the NucleoSpin Plant II Kit® (Macherey and Nagel, Düren, Germany) according to the standard protocol with the following modifications: the leaf samples were homogenized for one minute at a frequency of 30 revolutions per second and the incubation for cell lysis was carried out for 15 min. The DNA was eluted in 100 µl PE buffer.

For the AFLP analysis 100 ng of DNA were digested overnight at 37 °C using *HindIII* and *Mse*I (Thermo Fisher Scientific, Waltham, USA) followed by an adaptor ligation at 37 °C for 3.5 h. The pre-amplification reaction was performed with non-selective *HindIII* and *Mse*I primers. The thermocycler (C1000 Thermal Cycler, Biorad, Hercules, USA) was set to the following conditions: 94 °C for 5 min, 20 cycles of 94 °C for 30 s, 60 °C for 30 s, 72 °C for 1 min, and 72 °C for 10 min.

The final reaction was performed according to Klie et al. (2017) as a double run with IRD marked *HindIII* primers using two different wavelengths (700 nm and 800 nm). For most reactions, *HindIII* primers using CAT (700 nm) and ACG (800 nm) as selective bases were applied in combination with different *Mse*I primers. In Table A.1 (Supplementary material) the sequences of the oligonucleotides are listed. The PCR conditions for the final reaction were set to 94 °C for 5 min, one cycle of 94 °C for 30 s, 65 °C for 30 s, 72 °C for 1 min; 11 cycles with 0.7 °C decrement down to the annealing temperature, 24 cycles of 94 °C for 30 s, 56 °C for 30 s, 72 °C for 1 min and a final elongation at 72 °C for 10 min.

The SSR analysis was performed following the approach of Schuelke (2000) using 10 ng/µl of genomic DNA. A M13 tail (5‘ GTAAAACGACGGCCAGT 3‘) was appended to the forward primers. The reaction mix of the M13-PCR was composed of 20 ng DNA, 0.25 µM M13 tailed forward primer, 1.25 µM labelled forward primer, 5 µM reverse primer, 1.5 µl 2 mM dNTPs and 1 U Taq polymerase (DCS Hamburg, Germany). The PCR conditions were set to 94 °C for 5 min, 20 cycles of 94 °C for 45 s, T_A °C (depending of the respective primer) for 1 min, 72 °C for 1 min, eight cycles of 94 °C for 30 s, 52 °C for 45 s, 72 °C for 1 min, and 72 °C for 10 min.

Six different SSR primers were used, five (VM19, VM31, VM36,

Table 1
List of cowpea entries used for evaluation under field conditions and for molecular characterization.

Entry name	Entry type[1]	Source[2]	Origin	Color code[3]
Ex-Iseke	Released cultivar	WorldVeg	Tanzania	
Dakawa	Released cultivar	WorldVeg	Tanzania	
Ngoji	Released cultivar	WorldVeg	Tanzania	
Tumaini	Released cultivar	WorldVeg	Tanzania	
9334	Accession	WorldVeg	Tanzania	
UG-CP-4	Accession	WorldVeg	Uganda	
UG-CP-10	Accession	WorldVeg	Uganda	
GKK-CP-2	Accession	WorldVeg	Malawi	
GKK-CP-3	Accession	WorldVeg	Malawi	
RW-CP-7	Accession	WorldVeg	Rwanda	
Simlaw[4]	Released cultivar	JKUAT	Kenya	
EASEED[5]	Released cultivar	JKUAT	Kenya	
Accession 6	Local cultivar	JKUAT	Kenya	
Accession 20	Local cultivar	JKUAT	Kenya	
Accession 25	Local cultivar	JKUAT	Kenya	

[1]Local cultivar: cultivar developed by farmers from landraces through many years of selection, released cultivar: cultivar officially released from research system, accession: germplasm collection hold in gene bank. [2]WorldVeg: World Vegetable Center, JKUAT: Jomo Kenyatta University of Agriculture and Technology, [3]Different colors are used to label the different entries according to their origin, blue: Tanzania, orange: Uganda, magenta: Malawi, brown: Rwanda, green: Kenya. [4]Cultivar 'K/80' from the Simlaw Seeds Company Ltd, Nairobi, [5]Cultivar: 'Ken-Kunde 1' from the East African Seed Company Ltd, Nairobi

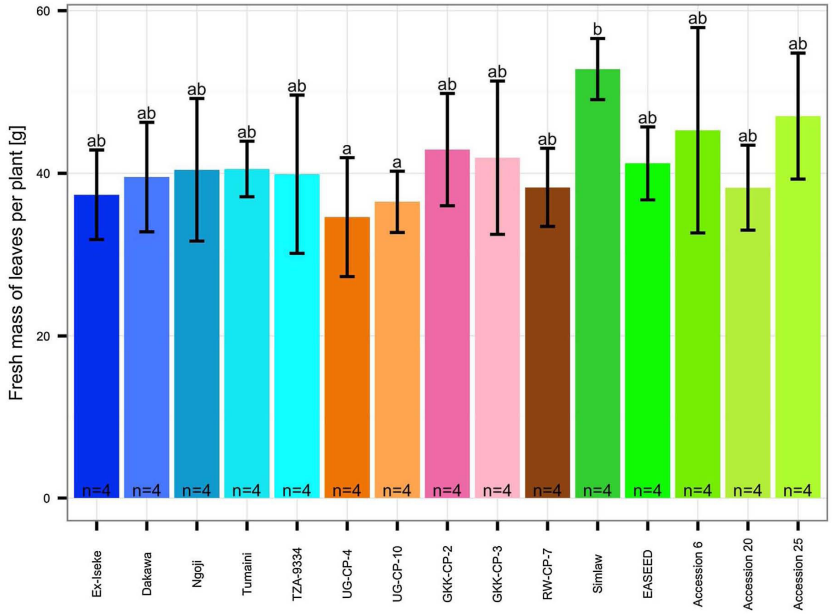


Fig. 1. Fresh leaf yield (g/plant) of 15 cowpea entries evaluated under field conditions in Kenya, 2015.
Means ± standard deviation; each sample contained leaves of 20 plants per block; significant differences between the means are indicated by different letters, Tukey-Test, $\alpha = 5\%$. Different colors are used to label the different entries and their origin, blue: Tanzania, orange: Uganda, magenta: Malawi, brown: Rwanda, green: Kenya.

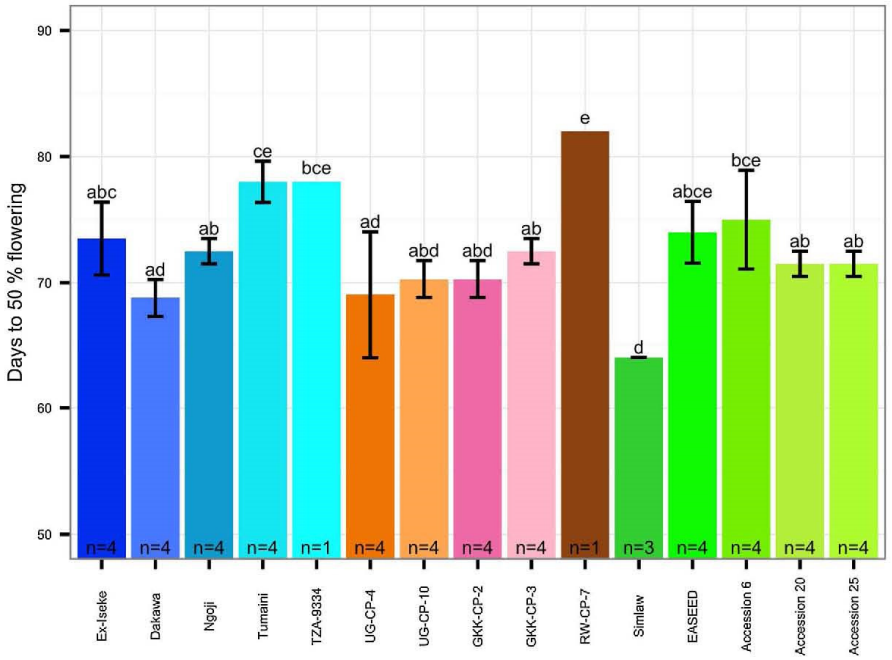


Fig. 2. Days to 50% flowering of 15 cowpea entries evaluated under field conditions in Kenya, 2015.
Means ± standard deviation, if n ≠ 4 the information was not available for all blocks; significant differences between the means are indicated by different letters, Tukey-test, α = 5 %. Different colors are used to label the different entries and their origin, blue: Tanzania, orange: Uganda, magenta: Malawi, brown: Rwanda, green: Kenya.

VM70, VM71) published by Li et al. (2001) and one (SSR-6436) by Asare et al. (2010). The primer sequences and their annealing temperatures are given in Table A.2 (Supplementary Material).

The PCR fragments for both marker types were separated on a 6% polyacrylamide gel (SequaGel Solution, National Diagnostics, Atlanta, USA) in 1 x TBE buffer (89 mM Tris, 89 mM Boric acid, 2 mM Na_2EDTA, pH 8.0) using a LiCor4200 and a LiCor4300 Sequencer (LiCor Biosciences, Bad Homburg, Germany).

The resulting banding patterns were transferred to a presence-absence matrix in which "1" stands for the presence of a band and "0" for its absence. Ambiguous fragments were labelled as missing values. The analysis of the matrix was done by using the R-packages "vegan" (Oksanen et al., 2015) and "ape" (Paradis et al., 2004). Based on the Jaccard distances, a Principle Coordinate Analysis (PCoA) was carried out. 1000 bootstraps were performed and dendrograms were computed using the neighbor joining algorithm (Studier and Keppler, 1988). Afterwards, the single neighbor-joining trees were added to one majority rule consensus tree, and bootstrap values were computed for each node.

3. Results

3.1. Morphological characterization

The commercial cultivar from Simlaw Seeds Company gave the highest fresh leaf yield (52.8 g/plant; Fig. 1). However, it was only significantly higher in yield when compared to the entries UG-CP-4 and UG-CP-10, originally collected from Uganda.

The number of days to 50% flowering is an important trait to assess the length of the vegetative growth phase (Fig. 2). The Rwandan entry RW-CP-7 was the latest in flowering (82 days). It formed a significance group with the accession 9334 and the cultivar Tumaini (both from Tanzania) as well as with Accession 6 and the commercial cultivar from EASEED (both from Kenya). The commercial cultivar from Simlaw Seeds Company was the earliest to flower after 64 days.

The entry RW-CP-7 had the highest seed weight (15.12 g/100 seeds), but this was not significantly different from the 100 seed weights of the entries Ex-Iseke, Ngoji and GKK-CP-2 (Fig. A.1, Supplementary Material). The entry UG-CP-4 had the lowest 100 seed weight (7.67 g) followed by accession 9334 (8.26 g).

The significantly lowest number of pods per plant was observed in entry RW-CP-7, whereas the highest number was found in accession UG-CP-4 (15.62) (Fig. A.2, Supplementary Material). Accession GKK-CP-3, originating from Malawi, gave the second highest number of pods per plant (14.28).

The correlations between fresh leaf yield and all other morphological traits measured were not significant (Fig. 3). The number of pods per plant was negatively correlated with all traits measured, but just the correlations to the 100 seed weight and the length and width of both, pods and seeds, were significant. Days to 50% flowering and pod length, as well as seed length and seed width were significantly and positively correlated.

3.2. Molecular characterization

For five of the entries the DNA content and ploidy was analyzed. They were all diploid with 2x = 22 chromosomes (counted in root tip metaphase cells) and the DNA content was estimated at 1.29–1.32 pg/2C (Table A.3, Supplementary Material). In total 42,525 single bands of 544 AFLP loci were analyzed, and on average 51.1% of them were polymorphic (Table 2). The total percentage of missing values accounted to 6.87%.

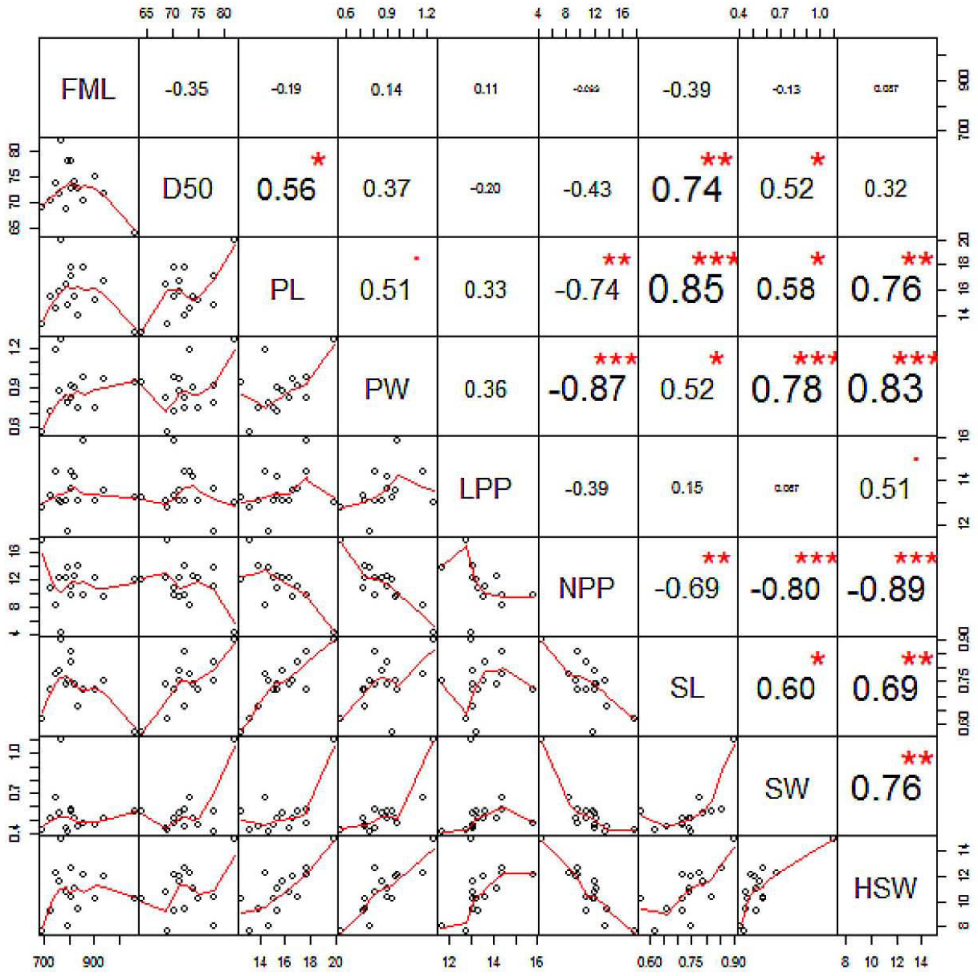


Fig. 3. Correlation plot for pairwise Pearson correlations among various quantitative morphological traits.
FML: Fresh mass of leaves, D50: Days to 50% flowering, PL: Pod length, PW: Pod width, LPP: Locules per pod, NPP: Number of pods per plant, SL: Seed length, SW: Seed width, HSW: 100 seed weight. Above the diagonal the correlation values are given. Significance levels are indicated by the stars: * represents α = 5%, ** represents α = 1%, *** represents α = 0.1%. Below the diagonal, the data underlying the correlation is plotted.

Table 2
Scored AFLP fragments and their level of polymorphism.

No.	Primer pair	No. of markers scored	No. of polymorphic markers	Percent polymorphic markers
1	CAA/GTG	77	43	55.8
2	CAT/GTG	55	31	56.4
3	CAT/AGC	49	24	49.0
4	CAT/AGG	58	30	51.7
5	CAT/AGA	86	47	54.7
6	ATA/GTG	28	14	50.0
7	ACG/AGT	44	24	54.5
8	ACG/AGC	28	11	39.3
9	ACG/AGG	51	24	47.1
10	ACG/AGA	68	36	52.9
		Σ 544	**Σ 284**	**Average 51.1**

Primer coding: Selective bases of the *HindIII* primer/selective bases of the *MseI* primer.

With the six SSR-primers used, a total number of 18 different alleles were scored, with the number of alleles per SSR-primer varying between two and five. The allele frequency ranged from 0.01 to 0.84, with a mean of 0.28. The highest PIC-value was 0.71 for primer VM 70 and the lowest was 0.28 for primer VM 19 (Table 3). The microsatellite primer SSR-6436 reproducibly showed a maximum of five bands per genotype and was thus scored like the AFLP markers, but excluded from the calculation of the allelic frequency and the PIC-values. After three repetitions, two genotypes from the entries Dakawa and 9334 and one genotype of the entries GKK-CP-2 and Ex Iseke each, did not show any bands after the amplification using the SSR primer VM 19.

The smallest Jaccard distance (0.002) was observed between the genotypes 4 and 5 of the entry Ngoji which is a released cultivar from Tanzania (Table 4, Table A.4 Supplementary Material). Between the genotypes 1 and 4 of Ex-Iseke, an advanced line from Tanzania, the highest within-accession genetic distance was observed (0.193). The lowest mean distance of 0.029 between its genotypes was recorded for RW-CP-7 whereas Accession 6, which is one of the landraces from Kenya, expressed the highest mean distance of 0.109. The mean

Table 3
Number of alleles and their frequencies for five SSR-primers originating from Li et al. (2001).

Primer	Heterozygosity	PIC	No of allele	Allele frequency
VM19	0.040	0.28	1	0.84
			2	0.13
VM31	0.000	0.63	1	0.21
			2	0.45
			3	0.35
VM36	0.027	0.30	1	0.83
			2	0.10
			3	0.09
VM70	0.053	0.71	1	0.22
			2	0.35
			3	0.31
			4	0.15
			5	0.01
VM71	0.067	0.60	1	0.12
			2	0.02
			3	0.05
			4	0.45
			5	0.42
Average	**0.037**	**0.50**	**3.6**	**0.28**

Table 4
Minimum, maximum and mean Jaccard distances inside the 15 analyzed accessions.

Entry name	Entry type[1]	Origin	Distance (min.)	Distance (max.)	Distance (mean)
Ex-Iseke	Released cultivar	Tanzania	0.041	0.193	0.084
Dakawa	Released cultivar	Tanzania	0.055	0.114	0.070
Ngoji	Released cultivar	Tanzania	0.002	0.136	0.073
Tumaini	Released cultivar	Tanzania	0.044	0.129	0.065
9334	Accession	Tanzania	0.056	0.126	0.080
UG-CP-4	Accession	Uganda	0.032	0.110	0.058
UG-CP-10	Accession	Uganda	0.005	0.066	0.035
Gkk-CP-2	Accession	Malawi	0.100	0.164	0.105
Gkk-CP-3	Accession	Malawi	0.051	0.114	0.067
RW-CP-7	Accession	Rwanda	0.007	0.074	0.029
Simlaw	Released cultivar	Kenya	0.023	0.085	0.042
EASEED	Released cultivar	Kenya	0.038	0.150	0.072
Accession 6	Local cultivar	Kenya	0.115	0.163	0.109
Accession 20	Local cultivar	Kenya	0.007	0.145	0.063
Accession 25	Local cultivar	Kenya	0.058	0.142	0.087

[1] Local cultivar: cultivar developed by farmers from landraces through many years of selection, released cultivar: cultivar officially released from research system or companies, accession: germplasm collection hold in gene bank.

distances within the other two landraces were considerably lower and comparable with the mean distances observed in most of the other entries (0.063 for Accession 20 and 0.087 for Accession 25).

The smallest distance between genotypes of different entries (0.098) was observed between 9334 1 (Tanzania) and GKK-CP-2 2 (Malawi). On average of all genotypic comparisons, UG-CP-10 from Uganda and GKK-CP-3 from Malawi were the genetically closest entries (0.132), whereas 9334 and Simlaw (Kenya) had the highest mean genetic distance (0.273). With a distance of 0.301, the two most distant genotypes (9334 4 and Simlaw 5) belonged to these accessions too (Table A.4, Supplementary Material).

Using the Jaccard distances, a PCoA was computed (Fig. 4) in which the first axis could explain 23.3% of the variance. In the first dimension it was possible to separate the genotypes of accession Simlaw from the remaining accessions (cluster I). Along with the first and second genotype from the entry Ngoji, the genotypes of the entries Ex-Iseke and Dakawa grouped together within the same cluster (III) in which Ex Iseke formed its own sub-cluster (III-b). Together with the genotypes of Ex Iseke, the two genotypes of Ngoji were found in a second sub-cluster (III-a). Furthermore, all the genotypes of 9334 were grouped together into cluster II.

Based on the distance-matrices resulting from 1000 bootstraps, neighbour joining trees were put together in a majority rule consensus tree (Fig. 5) in which most entries are represented by one single clade. One neighbor joining tree is depicted in Fig. A3 (Supplementary Material). With a bootstrap value of 92%, the entries Ex-Iseke and Dakawa were grouped to one clade wherein the five genotypes of Dakawa and one of Ex-Iseke formed a sub-clade.

One clade showing a bootstrap value of 90% was formed by the genotypes of Accession 20, GKK-CP-2 and 9334. This clade was split into two sub-clades with the five genotypes of 9334 and the genotype GKK-CP-2 2 clustering into the same sub-clade. The other sub-clade was formed by the remaining genotypes, Accession 20 1 and GKK-CP-2 3. Supported with a low bootstrap value of 53%, also the genotypes of UG-CP-4 and Tumaini were grouped into their own clade. All genotypes of Accession 6, two genotypes of Accession 25, one genotype of GKK-CP-2 and one of Ngoji could not be grouped to any clade of the consensus tree.

4. Discussion

The Kenyan cultivar Simlaw showed the highest fresh leaf yield (Fig. 1) and the shortest period to 50% flowering (Fig. 2). This cultivar could be used in breeding programs focusing on these traits. The entry RW-CP-7 had the highest 100 seed weight (Fig. A.1, Supplementary Material). The observed seed weights of the majority of the present entries were in agreement with results of 8.0–32.0 g recorded by Timko and Singh (2008). Also, Demooy and Demooy (1998) reported a 100 seed weight of 11.48 g for a breeding line from Botswana.

In the present study, the 100 seed weight was negatively correlated with the number of pods per plant (Fig. 3). This result, however, was not in agreement with Aryeetey and Laing (1973) who characterized various yield components in an F_2 population descending from a cross between two cultivars, one from Ghana and the other from USA, and reported no correlation between these traits. Likewise, in another set of 10 different accessions from Nigeria, Manggoel et al. (2012) could not detect a correlation between these traits. In the present study, entry RW-CP-7 was the most extreme entry driving this correlation, because it gave the highest 100 seed weight and the lowest number of pods per plant. Excluding RW-CP-7 from the analysis resulted in a slightly lower, but still significant correlation coefficient (data not shown). According to the present result, breeding for cultivars combining a high number of pods per plant and a high 100 seed weight would be difficult. Due to the fact that the 100 seed weight was positively correlated to the width and length of pods and seeds, it can be rated as a valid descriptor for seed and pod size.

Cowpea leaves are rich sources of minerals and vitamins and therefore can play a role in combating malnutrition in different African countries (Grubben et al., 2014). Furthermore, fresh above ground vegetative biomass serves as hay and is an important source of animal feed (Badiane et al., 2014). This makes high vegetative yield and late flowering important breeding goals of vegetable cowpea. The data set could not reveal any correlation between fresh leaf yield and all the other investigated traits. Consequently, it could be possible to breed for cultivars incorporating both high leaf yield and high seed weight. For such breeding programs, molecular markers would be helpful. Recently, QTLs for domestication related traits in cowpea, such as seed weight, number of ovules, germination percentage and days to flowering were mapped employing 202 SSR markers (Andargie et al., 2014). Single nucleotide polymorphism (SNP) markers are also available in cowpea

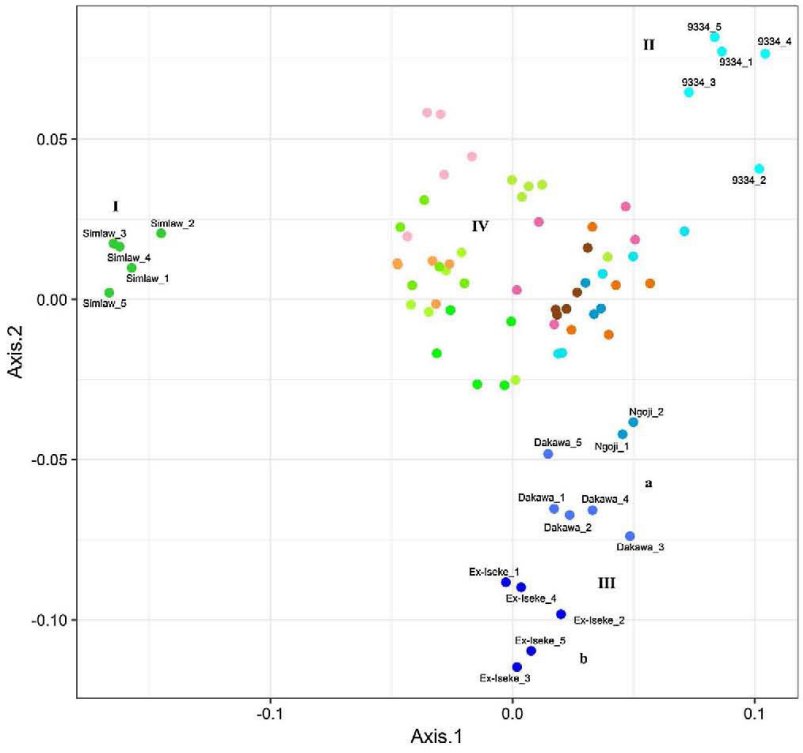


Fig. 4. Principle Coordinate Analysis for the analyzed entries.
Four main clusters are present: cluster I is composed of the genotypes from entry Simlaw, cluster II is composed of the genotypes of the entry 9334, cluster IIIa is composed of the genotypes of the entry Dakawa and the genotypes Ngoji 1 and Ngoji 2, cluster IV comprises all the remaining genotypes. Different colors are used to label the different entries and their origin, blue: Tanzania, orange: Uganda, magenta: Malawi, brown: Rwanda, green: Kenya.

and SNP markers associated with seed size have been identified (Lucas et al., 2013; Egbadzor et al., 2013). The use of cowpea as leafy vegetable has been far less considered in literature so far and would require additional efforts in future in order to provide breeders with molecular markers associated with fresh leaf yield as well as characters defining the nutritional value.

The Jaccard genetic distances obtained in the present study, ranging between 0.002 and 0.301, are in line with the results obtained by other researchers. Based on AFLP markers, Fang et al. (2007) observed a maximum Jaccard distance of 0.14 between 87 accessions including both, landraces from Africa, Asia and South America, and breeding lines from West Africa and the USA. However, between 90 breeding lines originating from the International Institute of Agriculture in Nigeria, Li et al. (2001) reported higher Jaccard distances of up to 0.56, but in most lines they were lower than 0.55. These results obtained on a broad range of accessions from several countries illustrate that genetic diversity is generally low in cultivated cowpea. The wild relatives of the cultivars show higher degrees of genetic variability (mostly above 0.4; Pasquet, 1999, 2000). The loss of variability in the cultivated forms could be due to domestication which may have occurred on a maximum of two limited populations of *var. spontanea* in East and West Africa (Huynh et al., 2013; Pasquet, 1999).

The lack of variability in the cultivated forms is a limiting factor for plant breeding. As described by Badiane et al. (2014), Barone et al. (1992) and Gomathinayagam et al. (1998), uncultivated *V. unguiculata* subspecies can be used for an enrichment of the gene pool to raise the level of heterozygosity. Additionally, they could be assumed to serve as a source of several resistance genes.

In the neighbor-joining tree (Fig. A3, Supplementary Material), two main clusters were obtained but the clustering did correspond neither to the country of origin of the entries nor to any morphological trait or the level of breeding input. Madode et al. (2012) performed AFLP analyses on twenty morphologically different landraces from West Africa resulting in Jaccard distances ranging between 0.01 and 0.22. They constructed a dendrogram which clustered landraces without any pigmentation of the seed testa into one clade. Cultivated and wild cowpea accessions formed two main clusters in a dendrogram computed by Coulibaly et al. (2002) who observed Nei's genetic distances ranging between 0.02 and 0.24. Moreover, the regional origin of the different wild cowpea accessions was also resolved in their study.

In our analyses, a distinction between most of the entries was revealed in both dendrograms (Fig. 5 and Fig. A.3, Supplementary Material), although the consensus tree was unable to reveal the relationships between the assessed genotypes completely. However, the clustering was reliable since most clades were formed by single entries showing bootstrap values in the consensus tree that were considered as reasonable.

The Tanzanian entries Ex-Iseke and Dakawa clustered together in both trees. This cluster was also found in the PCoA (Fig. 4). It can be assumed that they come from the same narrow gene pool, eventually also from the same breeding program because of their common origin and the fact that both entries have already been processed by breeders.

A second cluster comprised of entry 9334 from Tanzania, together with two genotypes of GKK-CP-2 from Malawi and one genotype of Accession 20. Due to the high bootstrap values, and the fact that 9334 formed its own cluster in the PCoA it could be assumed that this entry is genetically different from the remaining set of entries.

An explanation for the clustering of genotypes originating from different countries to the same clade could be found in the fact that cowpea played a role in the human history (Flight, 1976). Due to the evidence of prehistoric trade of crops between West and East Africa (Baudoin and Marechal, 1985), it can be assumed that cowpea seeds

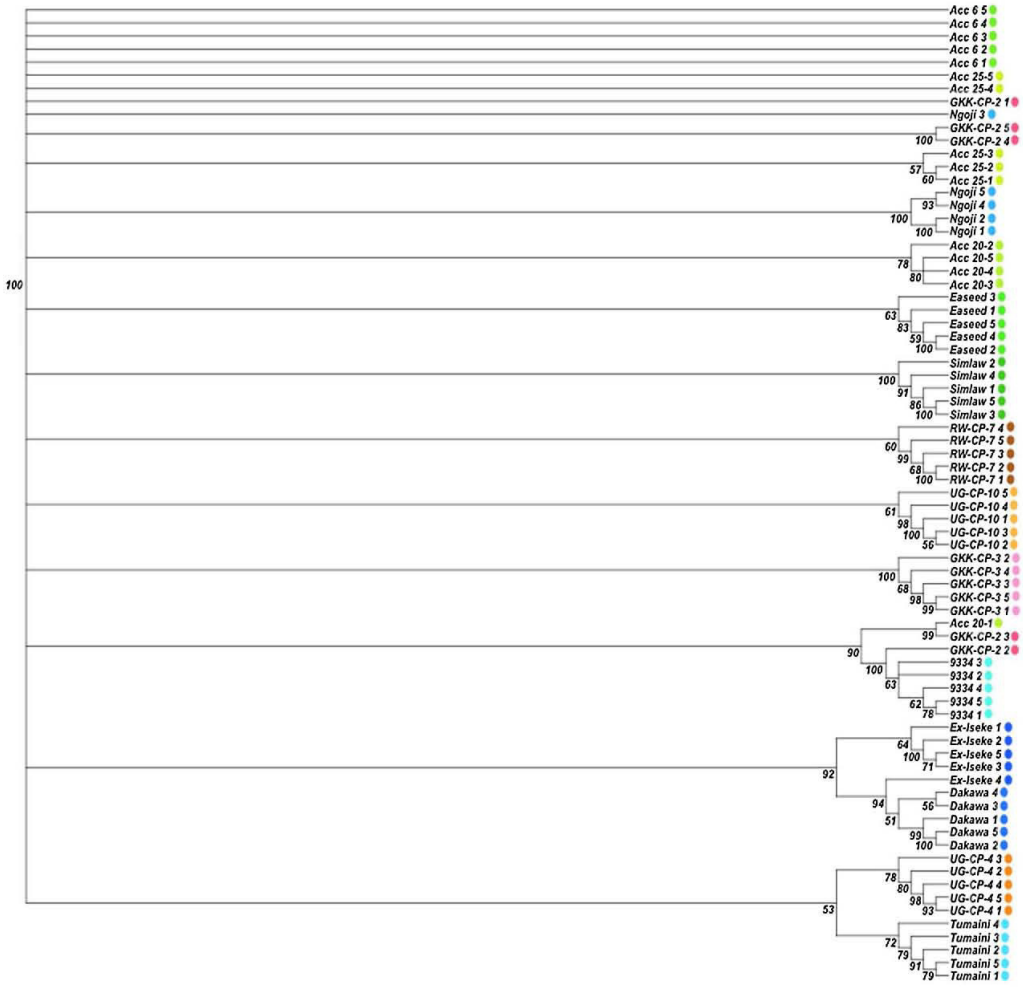


Fig. 5. Majority Rule consensus tree including all 75 genotypes from 15 different entries.
The tree is based on 1000 bootstrap replications. Bootstrap values above 50% are plotted to each node. Different colors are used to label the different entries and their origin: blue: Tanzania, orange: Uganda, magenta: Malawi, brown: Rwanda, green: Kenya.

were and still are traded between different regions in East Africa.

In the consensus tree, genotypes of the Kenyan local cultivars did not cluster together, showing that they were less uniform than the other entries. This was not unexpected, since they are developed from landraces which are compositions of different genotypes lacking crop improvement (Villa et al., 2006).

In conclusion, despite rather low genetic diversity within and between the analyzed cowpea entries, significant differences in several morphological traits were detected. It may be recommended to preserve the germplasm for further breeding programs by collecting landraces in different geographical regions and by preserving high numbers of entries represented by a relatively small number of seeds or plants.

Acknowledgements

The authors gratefully acknowledge funding within the project HORTINLEA by the German Federal Ministry of Education and Research (BMBF) and the German Federal Ministry of Economic Cooperation and Development (BMZ) within the framework of the program GlobE – Global Food Security. We acknowledge World Vegetable Center (WorldVeg) and Jomo Kenyatta University of Agriculture and Technology (JKUAT) for providing germplasm. We appreciate technical assistance from David Odongo, Sylviah Buleti, Betty Orangi, Francis Wetende, Rose Khayanga and Jane Were. Finally, we thank Vitalina Karfik for assistance in the cytological and flow cytometric analyses.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.scienta.2017.08.003.

References

Ahenkora, K., AduDapaah, H.K., Agyemang, A., 1998. Selected nutritional components and sensory attributes of cowpea (*V. unguiculata* (L.) Walp.) leaves. Plant Foods Hum. Nutr. 52, 221–229.

Andargie, M., Pasquet, R.S., Gowda, B.S., Muluvi, G.M., Timko, M.P., 2014. Molecular mapping of QTLs for domestication-related traits in cowpea (*V. unguiculata* (L.)

Walp.). Euphytica 200, 401–412.
Arumuganathan, K., Earle, E.D., 1991. Nuclear DNA content of some important plant species. Plant Mol. Biol. Report. 9, 208–218.
Aryeetey, A.N., Laing, E., 1973. Inheritance of yield components and their correlation with yield in cowpea (*Vigna unguiculata* (L.) Walp.). Euphytica 22, 386–392.
Asare, A.T., Bhavani, S.G., Isaac, K.A.G., Lawrence, L.A., Jemmy, F.T., Timko, M.P., 2010. Assessment of the genetic diversity in cowpea (*Vigna unguiculata* (L.) Walp.) germplasm from Ghana using simple sequence repeat markers. Plant Genet. Resour.: Character. Utili. 8, 142–150.
Badiane, F.A., Diouf, M., Diouf, D., 2014. Cowpea. In: Singh, M., Bisht, I.S., Dutta, M. (Eds.), Broadening the Genetic Base of Grain Legumes. Springer India (pp 95–115).
Barone, A., Del Giudice, A., Ng, N.Q., 1992. Barriers to interspecific hybridization between *Vigna unguiculata* and *Vigna vexillata*. Sex. Plant Reprod. 5, 195–200.
Baudoin, J.P., Marechal, R., 1985. Genetic diversity in Vigna. In: Singh, S.R., Rachie, K.O. (Eds.), Cowpea research, production and utilization. John Wiley and Sons, New York, NY pp 3–9.
Braun, P., Winkelmann, T., 2016. Flow cytometric analyses of somatic and pollen nuclei in midday flowers (Aizoaceae). Caryologia 69, 303–314.
Coulibaly, S., Pasquet, R.S., Papa, R., Gepts, P., 2002. AFLP analysis of the phonetic organization and genetic diversity of *Vigna unguiculata* (L.) Walp. reveals extensive gene flow between wild and domesticated types. Theor. Appl. Genet. 104, 358–366.
Demooy, B.E., Demooy, C.J., 1998. Effects of leaf-harvesting practices on yield and yield components of ER-7 cowpea (*Vigna unguiculata*) in semi-arid Botswana. Field Crops Res. 22, 27–31.
Diouf, D., Hilu, K.W., 2005. Microsatellites and RAPD markers to study genetic relationships among cowpea breeding lines and local varieties in Senegal. Genet. Res. Crop Evol. 52, 1057–1067.
Egbadzor, K.F., Yeboah, M., Danquah, E.Y., Ofori, K., Offei, S.K., 2013. Identification of SNP markers associated with seed size in cowpea (*Vigna unguiculata* (L) Walp.). Intern. J. Plant Breed. Genet. 7, 115–123.
Ehlers, J.D., Hall, A.E., 1997. Cowpea (*Vigna unguiculata* (L.) Walp.). Field Crops Res. 53, 187–204.
Fang, J., Chao, T.C.C., Roberts, P.A., Ehlers, J.D., 2007. Genetic diversity of cowpea (*Vigna unguiculata* (L.) Walp.) in four West African and USA breeding programs as determined by AFLP analysis. Genet. Res. Crop Evol. 54, 1197–1209.
Faris, D.G., 1964. The chromosome number of *Vigna sinensis* (L.) Savi. Can. J. Genet. Cytol. 6, 255–258.
Flight, C., 1976. The Kintampo culture and its place in the economic prehistory of West Africa. In: Harlan, R., de Wet, J.M.J., Stemler, A.B.L. (Eds.), Origins of African Plant Domestication. The Hague, Netherlands pp. 212–221.
Gomathinayagam, P., Ganeshram, S., Rathnaswamy, R., Ramaswamy, N.M., 1998. Interspecific hybridization between *Vigna unguiculata* (L.) Walp. and *V. vexillata* (L.) A. rich. through in vitro embryo culture. Euphytica 102, 203–209.
Grubben, G., Klaver, W., Nono-Womdim, R., Everaarts, A., Fondio, L., Nugteren, J.A., Corrado, M., 2014. Vegetables to combat the hidden hunger in Africa. Chronica Horti. 54, 24–32.
Harrell, F.E., 2016. Hmisc: Harrell Miscellaneous. R Package Version 3. pp. 4–17 https://CRAN.R-project.org/package=Hmisc (visited 4 July 2016).
Huynh, B.L., Close, T.J., Roberts, P.A., Hu, Z., Wanamaker, S., Lucas, M.R., Chiulele, R., Cissé, N., David, A., Hearne, S., Fatokun, C., Diop, N.N., Ehlers, J.D., 2013. Gene pools and the genetic architecture of domesticated cowpea. Plant Genome 6, 1–8.
Klie, M., Menz, I., Linde, M., 2017. Lack of structure in the gene pool of the highly polyploid ornamental chrysanthemum. Mol Breeding. http://dx.doi.org/10.1007/s11032-013-9874-4.
Kouam, E.B., Pasquet, R.S., Campagne, P., Tignegre, J.B., Thoen, K., Gaudin, R., Ouedraogo, J.T., Salifu, A.B., Muluvi, G.M., Gepts, P., 2012. Genetic structure and mating system of wild cowpea populations in West Africa. BMC Plant Biol. 12, 113.
Li, C., Fatokun, C.A., Ubi, B., Singh, B.B., Scoles, G.J., 2001. Determining genetic similarities and relationships among cowpea breeding lines by microsatellite markers. Crop Sci. 41, 189–197.
Lucas, M.R., Huynh, B.L., da Silva Vinholes, P., Cisse, N., Drabo, I., Ehlers, J.D., Roberts, P.A., Close, T.J., 2013. Association studies and legume synteny reveal haplotypes determining seed size in Vigna unguiculata. Front. Plant Sci. 4, 95.
Madode, Y.E., Linnemann, A.R., Nout, M.J.R., Vosman, B., Hounhouigan, D.J., van Boekel, M.A.J.S., 2012. Nutrients, technological properties and genetic relationships among twenty cowpea landraces cultivated in West Africa. Internat. J. Food Sci. Technol. 47, 2636–2647.
Manggoel, W., Uguru, M.I., Ndam, O.N., Dasbak, M.A., 2012. Genetic variability: correlation and path coefficient analysis of some yield components of ten cowpea (*Vigna unguiculata* (L) Walp.) accessions. J. Plant Breed. Crop Sci. 4, 80–86.
Ng, N.Q., Marechal, R., 1985. Cowpea taxonomy, origin and germplasm. In: Singh, S.R., Rachie, K.O. (Eds.), Cowpea Research, Production and Utilization. John Wiley and Sons, Chichester, UK pp. 11–21.
Nielsen, S.S., Brandt, W.E., Singh, B.B., 1993. Genetic variability for nutritional composition and cooking time of improved cowpea lines. Crop Sci. 33, 469–472.
Ogbede, S.C., Saidu, A.N., Kabiru, A.Y., Busari, M., 2015. Nutrient and anti-nutrient compositions of *Brassica oleracae* var *capitata* L. IOSR J. Pharm. 5, 19–25.
Oksanen, J., Guillaume, B.F., Kindt, R., Legendre, P., Minchin, P.R., O'Hara, R.B., Simpson, G.L., Solymos, P., Henry, M., Stevens, H., Wagner, H., 2015. Vegan: Community Ecology Package. R Package Version 2. pp. 1–2 http://CRAN.Rproject.org/package=vegan (visited 24 July 2016).
Padulosi, S., Ng, N.Q., 1997. Origin, taxonomy and morphology of *Vigna unguiculata* (L.) Walp. In: Singh, B.B., Mohan, D.R., Dashiell, K.E., Jackai, L.E.N. (Eds.), Advances in Cowpea Research, IITA. International Institute of Tropical Agriculture, Ibadan, Nigeria pp. 1–12.
Paradis, E., Claude, J., Strimmer, K., 2004. APE: analyses of phylogenetics and evolution in R language. Bioinformatics 20, 289–290.
Pasquet, R.S., 1993a. Classification infraspecifique des forms spontanees de *Vigna unguiculata* (L.) Walp. (*Fabaceae*) a partir de donnees morphologiques. Buletin du Jardin Botanique National de Belgique 62, 127–173.
Pasquet, R.S., 1993b. Two new subspecies of *Vigna unguiculata* (L.) Walp. (*Leguminosae: papilionideae*). Kew Bull. 48, 805–806.
Pasquet, R.S., 1996. Wild cowpea (*Vigna unguiculata*) evolution. In: Pickersgill, B., Lock, J.M. (Eds.), Advances in Legume Systematics 8: Legumes of Economic Importance. Royal Botanic Gardens, pp. 95–100.
Pasquet, R.S., 1997. A new subspecies of *Vigna unguiculata* (*Leguminosae-Papilionoideae*). Kew Bull. 52, 840.
Pasquet, R.S., 1998. Morphological study of cultivated cowpea (*Vigna unguiculata* (L.) Walp.) Importance of ovule number and definition of cv.-gr. Melanophthalmus. Agronomie 18, 61–70.
Pasquet, R.S., 1999. Genetic relationships among subspecies of *Vigna unguiculata* (L.) Walp. based on allozyme variation. Theor. Appl. Genet. 98, 1104–1119.
Pasquet, R.S., 2000. Allozyme diversity of cultivated cowpea *Vigna unguiculata* (L.) Walp. Theor. Appl. Genet. 101, 211–219.
Peterson, B.G., Carl, P., 2014. Performance Analytics: Econometric tools for performance and risk analysis. R package version 1.4.3541. (visited 15 March 2017). https://CRAN.R-project.org/package=PerformanceAnalytics.
Ragab, D.D.M., Babiker, E.E., Eltinay, A.H., 2004. Fractionation, solubility and functional properties of cowpea (*Vigna unguiculata*) proteins as affected by pH and/or salt concentration. Food Chem. 84, 207–212.
Russell, V.L., 2016. Least-squares means: The R package lsmeans. J. Statist. Softw. 69, 1–33.
Schuelke, M., 2000. An economic method for the fluorescent labelling of PCR fragments. Nature Biotechnol. 18, 233–234.
Stadlmayr, B., Charrondiere, U.R., Enujiugha, V.N., Bayili, R.G., Fagbohoun, E.G., Samb, B., Addy, P., Barikmo, I., Ouattara, F., Oshaug, A., Akinyele, I., Annor, G.A., Bomfeh, K., Ene-Obong, H., Smith, I.F., Thiam, I., Burlingame, B., 2012. West African Food Composition Table. FAO, Rome (visited 15 September 2016). http://www.fao.org/docrep/015/i2698b/i2698b00. pdf.
Studier, J.A., Keppler, K.J., 1988. A note on the neighbor-joining algorithm of Saitou and Nei. Mol. Biol. Evol. 5, 729–731.
Timko, M.P., Singh, B.B., 2008. Cowpea, a multifunctional legume. In: Moore, P.H., Ming, R. (Eds.), Genomics of Tropical Crop Plants. Springer pp. 227–258.
Vaillancourt, R.E., Weeden, N.F., 1992. Chloroplast DNA polymorphism suggests Nigerian center of domestication for the cowpea, *Vigna unguiculata* (*Leguminosae*). Am. J. Bot. 79, 1194–1199.
Villa, T.C.C., Maxted, N., Scholten, M., Ford-Lloyd, B., 2006. Defining and identifying crop landraces. Plant Genet. Res. 3, 373–384.
Warman, P.R., Havard, K.A., 1996. Yield, vitamin and mineral contents of organically and conventionally grown carrots and cabbage. Agriculture. Ecosyst. Environ. 61, 155–162.
Wickham, H., 2009. ggplot2: Elegant graphics for data analysis. Springer-Verlag, New York.

7.1 Supplementary material

Table A.1: DNA contents of specimen of five cowpea entries determined by flow cytometry employing propidium iodide staining.

Entry	DNA content [pg/2C] $\bar{x} \pm s$	Coefficient of variance *Cowpea* [%] $\bar{x} \pm s$	Coefficient of variance *Solanum* [%] $\bar{x} \pm s$
Dakawa (n=9)	1.29 ± 0.02	9.2 ± 0.7	7.3 ± 0.8
Tumaini (n=10)	1.30 ± 0.02	9.3 ± 0.5	7.8 ± 0.6
Ex-Iseke (n=10)	1.31 ± 0.02	9.0 ± 0.5	7.2 ± 1.1
Ngoji (n=9)	1.30 ± 0.02	9.4 ± 0.6	7.8 ± 1.2
UG-CP-4 (n=10)	1.32 ± 0.01	9.4 ± 0.3	7.6 ± 0.4

The reference standard used was a *Solanum lycopersicum* cultivar with known genome size (`Stupické Polní Rane´, 1.96 pg/2C, Genbank Gatersleben acc. no. LYC 418).

Table A.2: Oligonucleotide sequences used for AFLP analysis.

		Sequences
DNA adapter	*MseI* O	5'-GACGA TGAGTCCTGAG-3'
	MseI U	5'-TACTCAGGACTCAT-3'
	HindIII O	5'-CTCGTAGACTGCGTACC-3'
	HindIII U	5'-AGCTGGTACGCAGTCTAC-3'
Pre Amplification Primer	*MseI*	5'-GACGATGAGTCCTGAGTAA-3'
	HindIII	5'-AGACTGCGTACCAGCTT-3'
Final Reaction Primer	*MseI*+ XXX[a]	5'-GATGAGTCCTGAGTAA **XXX**-3'
	HindIII + XXX	5'-GACTGCGTACCAGCTT **XXX**-3'

[a]XXX: The selective bases of the used *MseI* primers were GTG, AGT, AGG, AGA and AGC. One reaction was performed with *HindIII* Primers CAA (700 nm) and ATA (800 nm) and the GTG *MseI* primer.

Table A.3: Annealing temperature (T_A), product size and sequences without the M13 tail for the six SSR-primers.

Primer	T_A	Product size	Sequence	Reference
SSR-6436	60 °C	80 - 500 bp	Forward 5`-CAGAATCCTTGTGAACCTG Reverse 5`-TTTCGCAATATGCCCTTTTC	ASARE ET AL. (2010)
VM 19	60 °C	241 bp	Forward 5`- ATTC TGCGCCGTGACACTA Reverse 5`-TCGTGGCACCCCCTATC	LI ET AL. (2001)
VM 31	66 °C	200 bp	Forward 5`- CGCTCTTCGTTGATGGTTATG Reverse 5`-GTGTTCTAGAGGGTGTGATGGTA	LI ET AL. (2001)
VM36	66 °C	160 bp	Forward 5`-ACTTTCTGTTTTACTCGACAACT Reverse 5`-GTCGCTGGGGGTGGCTTATT	LI ET AL. (2001)
VM70	66 °C	186 bp	Forward 5`-AAAATCGGGGAAGGAAACC Reverse 5`GAAGGCAAAATACATGGAGTCAC	LI ET AL. (2001)
VM 71	66 °C	225 bp	Forward 5`-CGTGGCAGAGAATCAAAGACAC Reverse 5`-TGGGTGGAGGCAAAAACAAAAC	LI ET AL. (2001)

Table A.4: Minimum, maximum and mean Jaccard distances among the 15 analyzed entries

Entry	Entry	Distance (min)	Distance (max)	Distance (mean)
Ex-Iseke	Dakawa	0.125	0.202	0.144

Ex-Iseke	Ngoji	0.158	0.213	0.17
Ex-Iseke	Tumaini	0.151	0.21	0.185
Ex-Iseke	9334	0.197	0.243	0.22
Ex-Iseke	UG-CP-4	0.152	0.209	0.185
Ex-Iseke	UG-CP-10	0.155	0.205	0.178
Ex-Iseke	GKK-CP-2	0.179	0.205	0.186
Ex-Iseke	GKK-CP-3	0.174	0.216	0.193
Ex-Iseke	RW-CP-7	0.159	0.215	0.175
Ex-Iseke	Simlaw	0.207	0.244	0.23
Ex-Iseke	EASEED	0.158	0.211	0.18
Ex-Iseke	Accession 6	0.153	0.214	0.188
Ex-Iseke	Accession 20	0.179	0.207	0.191
Ex-Iseke	Accession 25	0.165	0.203	0.175
Dakawa	Ngoji	0.138	0.184	0.157
Dakawa	Tumaini	0.135	0.202	0.173
Dakawa	9334	0.176	0.199	0.187
Dakawa	UG-CP-4	0.147	0.189	0.17
Dakawa	UG-CP-10	0.151	0.188	0.17
Dakawa	GKK-CP-2	0.148	0.177	0.162
Dakawa	GKK-CP-3	0.161	0.195	0.175
Dakawa	RW-CP-7	0.149	0.19	0.169
Dakawa	Simlaw	0.195	0.252	0.232
Dakawa	EASEED	0.141	0.168	0.161
Dakawa	Accession 6	0.15	0.194	0.176
Dakawa	Accession 20	0.153	0.183	0.164
Dakawa	Accession 25	0.159	0.178	0.168
Ngoji	Tumaini	0.12	0.178	0.152
Ngoji	9334	0.155	0.2	0.178
Ngoji	UG-CP-4	0.155	0.186	0.17
Ngoji	UG-CP-10	0.135	0.173	0.166
Ngoji	GKK-CP-2	0.128	0.149	0.147
Ngoji	GKK-CP-3	0.144	0.189	0.168
Ngoji	RW-CP-7	0.13	0.175	0.158
Ngoji	Simlaw	0.232	0.254	0.239
Ngoji	EASEED	0.144	0.178	0.151
Ngoji	Accession 6	0.158	0.183	0.169
Ngoji	Accession 20	0.128	0.17	0.153
Ngoji	Accession 25	0.123	0.184	0.162
Tumaini	9334	0.169	0.194	0.18
Tumaini	UG-CP-4	0.118	0.191	0.145
Tumaini	UG-CP-10	0.137	0.216	0.176
Tumaini	GKK-CP-2	0.15	0.206	0.177
Tumaini	GKK-CP-3	0.16	0.221	0.184
Tumaini	RW-CP-7	0.148	0.176	0.166
Tumaini	Simlaw	0.216	0.275	0.232
Tumaini	EASEED	0.16	0.226	0.186
Tumaini	Accession 6	0.162	0.215	0.184
Tumaini	Accession 20	0.155	0.197	0.174
Tumaini	Accession 25	0.141	0.2	0.17
9334	UG-CP-4	0.161	0.208	0.182
9334	UG-CP-10	0.167	0.223	0.202
9334	GKK-CP-2	0.098	0.201	0.163
9334	GKK-CP-3	0.172	0.206	0.18
9334	RW-CP-7	0.162	0.188	0.174
9334	Simlaw	0.252	0.301	0.273

9334	EASEED	0.195	0.221	0.203
9334	Accession 6	0.182	0.211	0.199
9334	Accession 20	0.159	0.194	0.167
9334	Accession 25	0.184	0.213	0.195
UG-CP-4	UG-CP-10	0.111	0.165	0.142
UG-CP-4	GKK-CP-2	0.148	0.186	0.175
UG-CP-4	GKK-CP-3	0.16	0.192	0.18
UG-CP-4	RW-CP-7	0.156	0.186	0.175
UG-CP-4	Simlaw	0.223	0.256	0.242
UG-CP-4	EASEED	0.156	0.203	0.175
UG-CP-4	Accession 6	0.164	0.188	0.175
UG-CP-4	Accession 20	0.134	0.165	0.158
UG-CP-4	Accession 25	0.156	0.187	0.184
UG-CP-10	GKK-CP-2	0.146	0.168	0.158
UG-CP-10	GKK-CP-3	0.111	0.15	0.132
UG-CP-10	RW-CP-7	0.143	0.18	0.157
UG-CP-10	Simlaw	0.178	0.215	0.193
UG-CP-10	EASEED	0.128	0.158	0.144
UG-CP-10	Accession 6	0.131	0.154	0.135
UG-CP-10	Accession 20	0.13	0.157	0.151
UG-CP-10	Accession 25	0.116	0.168	0.143
GKK-CP-2	Ex-Iseke	0.162	0.197	0.186
GKK-CP-2	GKK-CP-3	0.145	0.185	0.169
GKK-CP-2	RW-CP-7	0.145	0.196	0.166
GKK-CP-2	Simlaw	0.198	0.25	0.228
GKK-CP-2	EASEED	0.133	0.18	0.157
GKK-CP-2	Accession 6	0.152	0.202	0.174
GKK-CP-2	Accession 20	0.1	0.152	0.136
GKK-CP-2	Accession 25	0.127	0.191	0.163
GKK-CP-3	RW-CP-7	0.153	0.186	0.167
GKK-CP-3	Simlaw	0.169	0.22	0.193
GKK-CP-3	EASEED	0.143	0.178	0.162
GKK-CP-3	Accession 6	0.125	0.175	0.144
GKK-CP-3	Accession 20	0.139	0.17	0.159
GKK-CP-3	Accession 25	0.135	0.165	0.149
RW-CP-7	Simlaw	0.22	0.238	0.228
RW-CP-7	EASEED	0.163	0.199	0.181
RW-CP-7	Accession 6	0.151	0.185	0.165
RW-CP-7	Accession 20	0.151	0.164	0.165
RW-CP-7	Accession 25	0.132	0.17	0.147
Simlaw	EASEED	0.18	0.229	0.211
Simlaw	Accession 6	0.167	0.207	0.194
Simlaw	Accession 20	0.198	0.227	0.214
Simlaw	Accession 25	0.158	0.226	0.196
EASEED	Accession 6	0.139	0.201	0.172
EASEED	Accession 20	0.132	0.15	0.15
EASEED	Accession 25	0.141	0.177	0.162
Accession 6	Accession 20	0.153	0.173	0.166
Accession 6	Accession 25	0.114	0.164	0.152
Accession 20	Accession 25	0.137	0.18	0.161

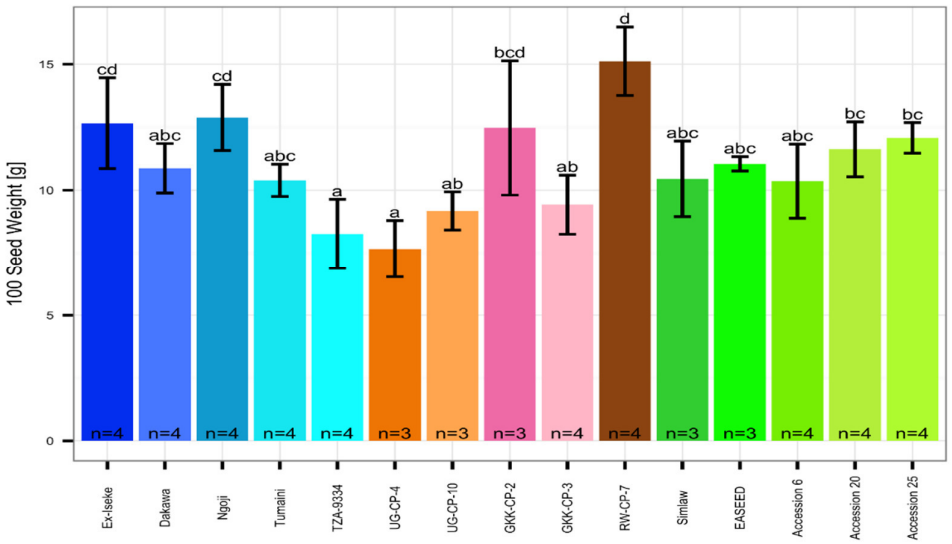


Fig.A.1: The averaged 100 seed weight of 15 cowpea entries evaluated under field conditions, Kenya, 2015.

Averages ± standard deviation, if n ≠ 4 the information was not available for all blocks; significant differences between the means are indicated by different letters, Tukey-test, α = 5 %. Different colors are used to label the different entries and their origin, blue: Tanzania, orange: Uganda, magenta: Malawi, brown: Rwanda, green: Kenya.

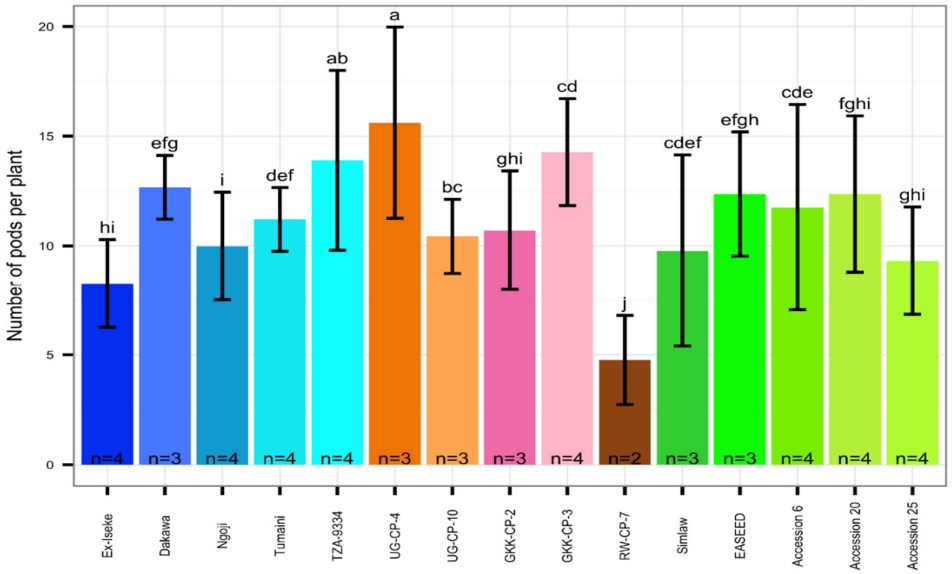


Fig.A2: Averaged number of pods per plant of 15 cowpea entries evaluated under field conditions, Kenya, 2015.

Averages ± standard deviation, if n ≠ 4 the information was not available for all blocks; significant differences between the means are indicated by different letters, Tukey-test, α = 5 %. Different colors are used to label the different entries and their origin, blue: Tanzania, orange: Uganda, magenta: Malawi, brown: Rwanda, green: Kenya.

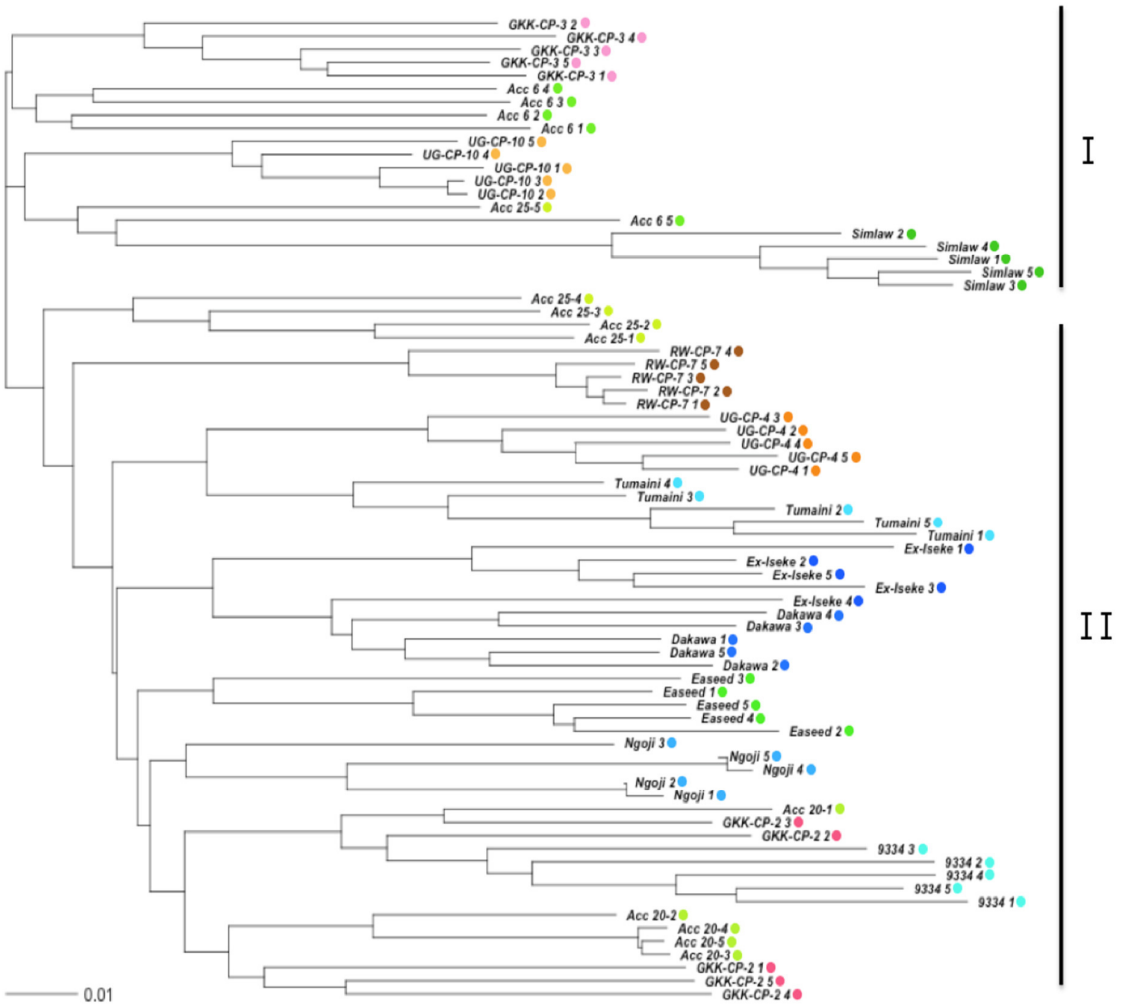


Fig.A.3: Neighbor joining tree including all 75 genotypes from 15 different entries.
Different colors are used to label the different entries and their origin, blue: Tanzania, orange: Uganda, magenta: Malawi, brown: Rwanda, green: Kenya, the genetic distance is represented by the bar

.

3. General discussion, conclusion and outlook.

This study was carried out to contribute important basic information for genetic improvement of spider plant and cowpea to help maximize their potential as source of food and income in addition to possible health benefits to the consumers. This study generally focused on investigating the diversity within the germplasm originally collected from different countries and continually reproduced and maintained at WorldVegetable Center in Arusha, Tanzania and cultivars' collected directly from farmers and markets in Kenya. The detailed results and discussions of the study are already discussed in the three manuscripts in chapter 2 and therefore this chapter shall focus on a general discussion and conclusions then finally give recommendations derived from the main findings for future research.

Genetic diversity is important for survival and also provides the basis for selection of material in breeding for traits of interest. It is also essential for core collection management in genetic resource preservation programs. Several studies based on the traditional field characters have covered diversity of spider plant using different spider plant germplasm (Chapter 2 section 2.3) but to the best of our knowledge, Omondi et al. (2017) (chapter 2.3) is the first study to utilize the robust and informative AFLP and SSR markers for genetic diversity analysis of spider plant. Unlike spider plant, genetic diversity in cowpea germplasm from various parts of the world has been well studied using molecular markers. Nevertheless, regular genetic evaluations of the genetic diversity of germplasms will greatly contribute to improvement of seed systems by identifying parental variants and ensuring that enough diversity is available among the germplasm for use in breeding programmes. The advantages of the AFLP and SSR markers are already discussed in the previous sections (Chapter 2, section 2.1) but it is worth noting that AFLP and SSR markers are currently the most commonly used in many studies assessing genetic diversity. Like the common schedule for any genetic diversity study, this study involved designing (for the SSRs) and testing of the markers, estimation of the relatedness of the germplasm and the final calculation of distance coefficients using a single joint matrix for the two markers and finally partitioning the genotypes in a dendrogram and in an ordination method (principal coordinate analysis - PCoA).

3.1. Establishment of new SSR markers for spider plant

SSRs are usually transferable within species or between close species (Rallo et al., 2003). In our study, initial testing of inter-specific SSR markers designed from genome sequence of the close by related species *Cleome hassleriana* showed non-amplification and sometimes weak unclear bands for our germplasm. During the testing stage for the SSRs designed from our material, various challenges including non-amplification of genomic DNA templates, multiple fragments (for SSRs) and absence of polymorphism were experienced. Non-amplification could be due to mutation within the primer binding sites in certain alleles (null alleles) (Pemberton et al., 1995). Redesigning of the primers in other stretches of the flanking DNA, avoiding the mutated binding sites has been proposed as a

solution for non-amplification (Callen et al., 1993; Esselink et al., 2004; Ishibashi et al., 1996; Paetkau and Strobeck, 1995). Multiple fragments may occur when a locus or parts of a large repeat are duplicated or when primers are insufficiently specific (Weising et al., 2005). A positive correlation between levels of polymorphism and the total size of perfect repeats has been reported (Webber, 1990). Higher mutation rates have also been observed in longer repeat arrays (Brohede et al., 2002; Jin et al., 1996; Schlötterer, 2004). However, according to Eichler et al.(1994), mutability reduces considerably when the arrays are interrupted. Over 50% of SSR primers on average is usually lost at the testing stage (Squirrell et al., 2003). Reduction rates of 24.7% in barley microsatellites and 67.8% for microsatellites in *Brassica* crop species for instance were reported by Ramsay et al. (2005) and Lowe et al. (2004) respectively. We experienced a reduction of 76.9% from 39 primers tested to 9 primers that were finally used. Very long SSRs may also make it difficult for Taq DNA polymerase to run through long stretches of repeats (Areshchenkova and Ganal, 2002). Based on these information, effort required for genetic marker design for crops such as the spider plant, whose genome studies are scarce, should not be underestimated especially with regard to time and resources.

3.2. Estimation of the genetic distances of spider plant and cowpea entries

Challenges in scoring the markers AFLPs were also experienced. The smallest and largest band for AFLP markers were sometimes difficult to score and unreliable hence excluded from the final analysis. Sufficient minimum number of polymorphic bands and a good genome coverage of a primer or a combination of primers is necessary for effective genetic distance estimations (Rallo et al., 2003; Fossati et al., 2001; Goulao et al., 2001; Garcia-Mas et al, 2000). Here, a combined matrix from the AFLP and SSR markers was used for estimating the genetic distances for both spider plant and cowpea. Double runs for each of the genotypes were performed and only strong bands scored to increase the reliability of the markers. Low resolution of the spider plant entries according to the countries of origin could be as a result of the self- and cross-compatible nature of spider plant allowing gene flow through pollen among the entries in the propagation fields at Worldveg center where the germplasm was sourced. This could further be explained by the lack of advanced genetic improvement in spider plant entries. The farmers' accessions too could not be resolved according to the regions of collection in Kenya. Diversity among landraces was also observed among spider plant landraces from Western Kenya by K'Opondo et al. (2009) and was attributed to cultivation without selection by the farmers. The spider plant populations may closely be related since farmers usually exchange among themselves or trade these seeds in the markets. Some genotypes from GS-sel, UGSF17-sel and UGSF2-sel entries showed higher genetic distances to the other entries and farmers' accessions shown by the average genetic distances **(Table 6chapter 2.2)**. the lowest genetic distances of 0.12 within the entries IP8-sel and UGSF17-sel **(Table 7 chapter 2.2)** could be due to selection process since these entries were developed by mass selection at WorldVeg center from the original collections.

For cowpea, the clades of the consensus tree were formed by single accessions except for Ex-Iseke and Dakawa **(figure 5 chapter 2.4).** The uniformity of genotypes in clades and very low genetic distances among genotypes within the accessions could be attributed to the autogamous nature of cowpea. Clustering of the accession Ex-Iseke and Dakawa may be due to a common origin of gene pool. Compared to the WorldVeg accessions, the genotypes of farmers' cultivars did not cluster in same clades indicating lack of uniformity due to the composition of different genotypes lacking any form of genetic improvement (Villa et al., 2005). In a study suggesting East Africa to be the origin of *V. unguiculata* var. spontanea, landraces were also found to be more diverse in East Africa using RAPD markers (Ba et al., 2004).Generally, the unresolved clustering of the genotypes of the cowpea accessions according to countries of origin can be assumed to be due to exchange of seeds between different regions in East Africa.

While the low genetic variability the within the cowpea accessions would be a limiting factor for breeding (Gepts, 2006), the variability among within the spider plant would be a good source of material for genetic improvement. Low genetic diversity observed in cowpea also puts the populations at risk of destruction by biotic and abiotic stresses (Gepts, 2006). Farmers cultivars for both crops could still be a good source of diversity in breeding programs.

3.3. Morphological diversity in spider plant and cowpea.

Morphological characterization has implications for future germplasm collection and in cultivar/ genotype selection for recommendations in breeding research. Morphological traits were significantly different among the spider plant entries and the cowpea accessions assessed in this study. Some traits varied even within the entries as shown by the standard deviations. For spider plant the significant variations are also in agreement with other studies with respect to traits such as plant height, days to 50% flowering (Onyango et al., 2016; Wasonga et al., 2015; Stoilova et al., 2015; K'Opondo, 2011). Leaf yield is an important aspect of these vegetables and the mean leaf fresh mass per plant of spider in the present study (14.11 g) was much lower compared to 135.65 g reported by Stoilova et al. (2015) who carried out three successive harvests. The highest leaf fresh mass per plant was recorded for entry UGSF13 (24.4 g) while the lowest was for Acc28 at 9.0 g. Flower removal in spider plant produced significantly increased plant height, leaf yield and fresh and dry shoot weights (Wangolo et al., 2015). A general increase in plant height, leaf number and area with application of organic, semi-organic and inorganic fertilizer in spider plant has been observed in spider plant (Jusoh et al., 2015; Masinde and Agong, 2011)but semi-organic fertilizer gave the best results compared to the other fertilizer treatments (Jusoh et al., 2015). For cowpea, the highest fresh mass per plant in this study was 52.83 g for Simlaw accession.

In spider plant, a mix of morphotypes within the entries based on stem or leaf petiole color was noted. However, the farmers' accessions mostly showed plants with purple or purple green stems compared to the advance lines and the gene bank entries. The farmers' accessions were also highly mixed in

terms of the leaf petiole colors compared to the gene bank entries and the advanced lines and this was particularly noted for accession 26 and accession 28 collected from the Kisii region in Kenya. The variations in stem and petiole colours are in agreement with findings in other studies (Wasonga et al., 2015; Wenyika et al., 2015; Kiebre et al., 2015 ; K'Opondo, 2011). A majority of the farmers (83.3%) in Burkina Faso preferred the green spider plant variety with a long cycle and that the green spider plant meets the consumers' need for organoleptic quality and great biomass (Kiebre et al., 2015). Great pubescence and green morphotypes have not been correlated in other studies of spider plant, in fact Masuka et al. (2012) observed late flowering for the purple stem compared to the green stem. This differences between the farmers' accessions and the gene bank entries and the advanced lines from the WorldVeg correlates with the clustering of the genotypes from the genetic markers analysis in the two main groups of the farmers' accessions and the WorldVeg sourced entries.

Selection of cultivars based on the phenotypic differences from these findings will permit varietal improvement taking into account the needs of both the farmers and consumers. The phenotypic traits are easy to score and remember (Wenyika et al., 2015). Cultivars bearing certain traits such as organoleptic quality for instance, can therefore be discriminated easily using traits such as stem colour. This would help to reduce the loss of thermo labile nutrients such as vitamins (Kutsukutsa et al, 2014; Wenyika et al., 2015) when consumers have to boil the vegetable to reduce the bitterness.

3.4. Cytological basis and reproductive biology of spider plant

Flow cytometry (FCM) is now widely used for nuclear DNA amounts estimation and ploidy level determination because it is convenient, fast and reliable (Dolezel et al, 2007). Low plant tissue is required for FCM making the it a non-destructive method and suitable for the analysis of limited material. A further important advantage of FCM for ploidy estimation is that tissues containing dividing cells are not required unlike in the laborious and destructive method of counting chromosomes in the root and shoot tips. However, to establish the number of chromosomes remains the unambiguous method of choice for ploidy determination.

The FCM results showed that the spider plant entries of this study were diploid with a chromosome number of 2n=34 established from root tips for one of the entries. This is also indirectly suggested by the SSR maker data that showed at most two alleles per genotype. The chromosome number points to a polyploidization event in *C. gynandra* suggesting the species to be allotetraploid.DNA content is an important character with many applications in biology and biodiversity (Bennett et al., 2000). For instance, genome size has been reported to be important in deciding the AFLP protocol to use i.e. for small genomes, the number of interpretable bands can be increased by decreasing the number of selective bases(Fay et al., 2005). Moreover, genome size was an important factor in decision making with regards to which taxa were chosen as first candidates for genome sequencing and even which chromosome(s) were first sequenced for instance, *Arabidopsis thaliana* because of its small genome size(Bennett et al., 2000). A relatively bigger genome size of 2.38 pg/2C was recorded for the spider

plant entries in this study compared to a previously published size of approximately 1 Gb (= approximately 0.978 pg/1C according to Doležel et al., 2003)by van den Bergh et al. (2014).This genome size falls in the ranges (1C= 0.16 – 1.31) reported for some *Brassicaceae* species (Johnston et al., 2005) and is generally on the lower side of the range of DNA amounts (1C = 0.1 – 125 pg) in angiosperms (Bennett et al., 2000). The DNA amount recorded for spider plant in this study is also slightly lower than 1.4 pg/ 1C that is reported for the ancestral genome size of angiosperms and other flowering plants (Soltis et al., 2003)**.**

The spider plant entries are self and cross-compatible as has been shown by the self-incompatibility indices calculated from the seed sets in our study. Self-incompatibility is an intra-specific pollination barrier to promote out-breeding (De Nettancourt, 2001). The self-incompatibility response is common *Brassicaceae*, the sister family to *Cleomaceae* where spider plant belongs and also in the genus Arabidopsis, including the model plant *A. thaliana* (Kitashiba and Nasrallah, 2014). Viable pollen is necessary to test self-compatibility in plants. Several methods could be used to test for pollen viability and these include use of vital stains, *in vitro* pollen germination, in vivo pollen germination, and seed set analysis (Dafni and Firmage, 2000).

In this study, *in vitro, in vivo* pollen germination and seed set analysis were used as already discussed in chapter 2 (2.2). However, it is worth mentioning that the *in vivo* situation simulates natural pollination and is more reliable than the *in vitro* pollen germination method, and that seed set method remains the most natural and reliable method to examine pollen viability. *In vivo* and seed set methods are notably laborious. Incompatibility is shown either *in vivo* when pollen fail to germinate and produce pollen tubes that elongate into the stigma to enable egg cell and central cell fertilization thereby giving a viable seed set or when abnormalities of pollen tube growth is observation in the stigma under an inflorescence microscope (Bots and Mariani, 2005). Incompatibility at the level of seed set is determined by theindex of self-incompatibility (ISI). This is calculated by dividing the relative success of seed set from self-pollination by the relative success of seed set from cross-pollination (Arroyo and Uslar, 1993).*In vitro* pollen germination was also carried out (Brewbaker and Kwack, 1963) and viable pollen counted under a light microscope and all the entries showed above 60% pollen viability **(see chapter 2.2 supplementary data Fig. 3)**. Several factors such as heat, UV light and water as reported by Bots and Mariani, (2005) may influence pollen viability in the field. Self-compatibility in these spider plant entries would be important for inheritance studies and would also enable pure-line development by selfing where uniformity of traits such as stem and petiole color is desirable. On the other hand, cross-compatibility will also allow development of hybrid cultivars from planned mating of selected superior genotypes in the general populations.

3.5. Nutritional content of selected spider plant entries

Nutritional content analysis is important for the promotion ALVs as a solution for hunger, malnutrition and non-communicable diseases in sub-Sahara regions (FAO, 2015). Detailed analysis of

the secondary metabolites is needed to estimate the health protecting properties in this vegetable which will enable proper food choices among consumers. The presence of important glucosinolates and flavonoid glycosides reported in this study will promote spider plant for more research interests especially in health. Glucosinolates are among bioactive compounds of interest in health as evidence has been found of their properties as anti-inflammatory, carcino preventive (Lippmann et al., 2014) and antidiabetogenic (Waterman et al., 2014).Aliphatic methyl glucosinolates (glucocapparin) seem to be the predominate glucosinolates in all organs of the spider plant. This seems to be unique for Cleome species since previous studies have also reported the absence of indole glucosinolates in *C. gynandra*(Neugart et al., 2017; Songsak and Lockwood, 2002).However, the indole glucosinolates were identified in low amounts in some greenhouse grown plants in this study **(see chapter 2.4 supplementary Table A3)**and this could imply that they could be environmentally regulated as have also been shown in other Brassicales (Verkerk et al., 2009). 2-hydroxy-2-methylbutyl glucosinolate (glucocleomin) has also been previously identified in the genus *Cleome* (Neugart et al., 2017; Songsak and Lockwood, 2002). Higher concentrations reported here is also in agreement with previous studies where the floral tissues, seeds and fruits had higher glucosinolate concentrations compared to the leave (Brown et al., 2003; Chen et al., 2001; Reintanz et al., 2001)a phenomenon that is thought to be due to protection against attack by pests since these organs contributed most to the plant survival.

Glycosides of kaempferol, quercetin and isorhamnetin varied in different organs of spider plant and are the main flavonoids in most food plants. The leaves of spider plant mostly contain glycosides of quercetin with Quercetin-3-rutinoside (rutin) as main compound. On the contrary, kaempferol glycosides were the dominant flavonoid in the flowers. These varying concentrations could be attributed to a higher antioxidant activity of quercetin glycosides (Zietz et al., 2010) and higher oxidation would be expected more in the photosynthetic tissues of plants. Several benefits of flavonoids to human health such as anti-oxidants, anti-carcinogens, vascular protective effect and improved vasodilator functions have been reported in previous studies and have already been discussed here (see chapter 2.3).

The richness of spider plant in macro- and micro-element established in this study will also be important for the consumers in combating ailments associated with lack of sufficient minerals. These macro- and micro-elements have also been shown to have antioxidant activity in spider plant (Sowunmi and Afolayan, 2015b). Surface contamination by dust and other source could also elevate other toxic elements for instance in this study higher concentrations of chromium and aluminum were established. Lead and cadmium content in this study also exceeded the allowable limits of 0.30 mg kg^{-1} and 0.20 mg kg^{-1}respectively in leafy vegetables (Commission Regulation No 629/2008, 2008). Soil mineral content and different fertilizer applications affect mineral component in spider plant, for instance, Hutchinson (2011) reported an increase of calcium ammonia nitrate fertilizer enhanced calcium in spider plant. According to Agbo et al. (2014), mineral concentrations in spider plant varied

between rainy and dry seasons. Iron, potassium and phosphorus content was elevated in the dry season while magnesium and calcium content was elevated in the rainy season.

Generally, high antioxidant activity (Mibei et al., 2012; Rao and Kumar, 2015) and antimicrobial activity (Khan et al., 2015; Rajaselvam and Rose, 2016; Ranjitha et al., 2014)has been reported for spider plant and these has been attributed to the secondary phytochemicals and the mineral content.

3.6. Conformance with the thesis objectives

This chapter refers back to the set out objectives of this thesis described in chapter 1.6 to explain the extent at which the objectives were reached.

1. Analysis of the genetic diversity in selected spider plant entries and farmers' cultivars using morphological and molecular markers.

The highly informative SSR and AFLP markers for spider plant and cowpea were established in this study. These markers will form an important basis in further research and breeding of the crop, for instance in marker association studies.The genetic diversity of the spider plant is found within the entries rather than the geographical origin. The farmers' cultivars for spider plant were generally considered an independent genetic pool. These farmers' cultivars/ landraces could be important to breeders as a source material to increase genetic variability in the crops. Inclusion of more collections of the landraces in *ex situ* conservation programs would therefore be prudent to avoid genetic erosion. Significant variations in the morphological traits were realized among the spider plant entries. The traits also vary within the entries as shown by sometimes high standard deviations.

2. Analysis of the cytological basis and reproductive biology of spider plant.

This study established same ploidy level among the spider plant entries suggesting the possibility of crossing all the entries in breeding. The spider plant entries showed a relatively small genome size of 1.19 pg/1C. The spider plant entries were also established to have no or only low levels of self-incompatibility and therefore selfing steps in breeding programmes involving these spider plant entries is possible since.

3. To analyze the nutritional content of selected spider plant entries.

The leaf tissues of the selected entries of spider plant in this study were rich in mineral content especially potassium, calcium, magnesium, phosphorus, iron, manganese and zinc and therefore have a potential to alleviate ailments due to malnutrition. The entries were also potent with aliphatic methyl glucosinolates in all the tissues and glycosides of quercetin, kaempferol and isorhamnetin as the main flavonoids. The health benefits of these secondary metabolites make these spiders plant a potential choice for healthy diets. However, recommendations of specific entries for breeding focusing on enhancing these nutrients can not be concluded due to mixed genotypes within the entries. Due to limited replicates in this study, statistical analyses were not possible so far.

3.7. Outlook

A high genetic diversity resolution of the spider plant entries in a cluster could presumably obtained by a higher number of markers with a good distribution throughout the genomes, such as SSRs, AFLPs and SNPs. This could be combined with a higher number of representative genotypes of each entry. Association of the AFLP and SSR markers to the morphological traits is desirable to enable marker assisted selection in breeding of these vegetables.

Morphological and genetic variability established within entries in this study permit breeding for pure plant types based on market demands or traits of interest. Genotypic selection is also needed to develop pure uniform genotype of the spider plant from the mixed morphotypes observed in this sudy. This would open up further research on spider plant on aspects such as nutritional content, disease and pest resistance, water stress tolerance among others based on pure lines.

Variations in flavonoid quantities and profile patterns in kales have been reported to be affected by factors including genetics, temperature and radiation (Schmidt et al., 2010) while for glucosinolate patterns, factors include species and ecotypes, individual plants variation, development stage, plant tissue, season and photoperiod (Yan and Chen, 2007; Petersen et al., 2002; Chen et al., 2001; Rosa, 1997). Further in-depth analyses with a higher number of replicates and also single plant preferably cloned plant material to identify effectsdue to these factors between genotypes and accessions of spider plant are necessary.These kind of analysis would also inform recommendations of specific accessions for breeding focusing on enhancing these nutrients.

A recent study has reported acancer preventive potential of *Brasica carinata*, an African indiginous leafy vegetable, using human liver cancer cells (HepG2) (Odongo et al., 2017). This kind of analysis would also be fitting for spider plant to increase the scientific knowledge of its health potential.

4. References

Abukutsa-Onyango M. (2007). The Diversity of cultivated African leafy vegetables in three communities in western Kenya. *African Journal of Food Agriculture Nutrition and Development*, *7*(3), 1–15.

Adeniji OT, Aloyce A. (2012). Farmer's knowledge of horticultural traits and participatory selection of African eggplant varieties (*Solanum aethiopicum*) in Tanzania. *Tropicultura*, *30*(3), 185–191.

Afari-Sefa V, Tenkouano A, Ojiewo CO, Keatinge JDH, Hughes Jd'A. (2012). Vegetable breeding in Africa: constraints, complexity and contributions toward achieving food and nutritional security. *Food Security*, *4*(1), 1–13.

Agbo AE, Kouame C, Anin AOL, Soro LC, N'Zi JC, Fondio L, Gnakri D. (2014). Seasonal variation in nutritional compositions of spider plant (*Cleome gynandra* L.) in south Côte d'Ivoire. *International Journal of Agricultural Policy and Research*, *2*(11), 406–413.

Areshchenkova T, Ganal MW. (2002). Comparative analysis of polymorphism and chromosomal location of tomato microsatellite markers isolated from different sources. *Theoretical and Applied Genetics*, *104*(2–3), 229–235.

Arroyo MT, FLS, Uslar P. (1993). Breeding systems in a temperate mediterranean-type climate montane sclerophyllous forest in central Chile. *Botanical Journal of the Linnean Society*, *111*, 83–102.

Ba FS, Pasquet RS, Gepts P. (2004). Genetic diversity in cowpea [*Vigna unguiculata* (L.) Walp] as revealed by RAPD markers. *Genetic Resources and Crop Evolution*, *51*, 539–550.

Baudoin JP, Marechal R. (1985): Genetic diversity in Vigna. In: Singh S. R., Rachie K. O. *Cowpea research, production and utilization*. John Wiley and Sons, New York, NY, Pages: 3-9.

Bennett MD, Bhandol P, Leitch IJ. (2000). Nuclear DNA Amounts in Angiosperms and their modern Uses-807 new estimates. *Annals of Botany*, *86*, 859–909.

Bots M, Mariani C. (2005). *Pollen viability in the field*. Radboud Universiteit Nijmegen.

Boukar O, Fatokun CA, Huynh BL, Roberts PA, Close TJ. (2016). Genomic tools in cowpea breeding programs: status and perspectives. *Frontiers in Plant Science*, *7:* 757.

Brewbaker JL, Kwack BH. (1963). The essential role of calcium ion in pollen germination and pollen tube growth. *Source: American Journal of Botany*, *50*(9): 859–865.

Brohede J, Primmer CR, Møller A, Ellegren H. (2002). Heterogeneity in the rate and pattern of germline mutation at individual microsatellite loci. *Nucleic Acids Research*, *30*(9): 1997–2003.

Brown PD, Tokuhisa JG, Reichelt M, Gershenzon J. (2003). Variation of glucosinolate accumulation among different organs and developmental stages of *Arabidopsis thaliana*. *Phytochemistry*, *62*, 471–481.

Callen DF, Thompson AD, Shen Y, Phillips HA, Richards RI, Mulley JC, Sutherland GR. (1993). Incidence and origin of "null" alleles in the (AC)n microsatellite markers. *American Journal of Human Genetics*, *52*(5), 922–927.

Centre for Agricultural and Rural Cooperation (CTA). (2014). *Seed Systems, Science and policy in East and Central Africa.* Wageningen, Netherlands.

Cernansky R. (2015). Super vegetables: Long overlooked in parts of Africa, indigenous greens are now capturing attention for their nutritional and environmental benefits. *Nature*, *522*, 146.

Chen S, Petersen BL, Olsen CE, Schulz A, Halkier BA. (2001). Long-distance phloem transport of glucosinolates in Arabidopsis. *Plant Physiology*, *127*(1), 194–201.

Chen X, Laudeman TW, Rushton PJ, Spraggins TA, Timko MP. (2007). CGKB: an annotation knowledge base for cowpea (*Vigna unguiculata* L.) methylation filtered genomic genespace sequences. *BMC Bioinformatics*, *8*, 129.

Chinsembu KC. (2016). Ethnobotanical study of medicinal flora utilised by traditional healers in the management of sexually transmitted infections in Sesheke District, Western province, Zambia. *Brazilian Journal of Pharmacognosy*, *26*(2), 268–274.

Chongori GO, Nyalala S, Mwangi M. (2016). African spider plant (*Cleome gynandra* L.) as biofumigant against weeds during turfgrass (*Paspalum notatum*) establishment. *Journal of* Horticulture and Forestry, 8(6), 44–50.

Christiansen M, Andersen S, Ortiz R. (2002). Diversity changes in an intensively bread wheat germplasm during the 20th century. *Molecular Breeding*, *9*, 1–11.

Chweya JA, Mnzava NA. (1997). Cat's whiskers. *Cleome gynandra* L. Promoting the conservation and use of underutilized and neglected crops. (Vol. 11). Rome, Italy.

Commission Regulation (EC) No 629/2008 (2008). Amending Regulation (EC) No 1881/2006 setting maximum levels for certain contaminants in foodstuffs (Text with EEA relevance). *Official Journal of European Union L* 173, p. 6–9. *http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex:32008R0629*

Dafni A, Firmage D. (2000). Pollen viability and longevity: practical, ecological and evolutionary implications. *Plant Systematics and Evolution*, *222*(1–4), 113–132.

De Nettancourt D. (2001). *Incompatibility and Incongruity in Wild and Cultivated Plants* (2nd Editio). New York: Springer-Verlag Berlin Heidelberg.

De Riek J, Dendauw J, Mertens M, De Loose M, Heursel J, Van Blockstaele E. (1999). Validation of criteria for the selection of AFLP markers to assess the genetic variation of a breeders' collection of evergreen azaleas. *Theoretical and Applied Genetics*, *99*(7–8), 1155–1165.

Dinssa FF, Hanson P, Dubois T, Tenkouano A, Stoilova T, Hughes JA, Keatinge JDH. (2016). AVRDC - The world vegetable center's women-oriented improvement and development strategy for traditional African vegetables in sub-Saharan Africa. *European Journal of Horticultural Science*, *81*(2), 91–105.

Doležel J, Greilhuber J, Suda J. (2007). Estimation of nuclear DNA content in plants using flow cytometry. *Nature Protocols*, *2*(9), 2233–2244. http://doi.org/10.1038/nprot.2007.310

Doležel J, Bartoš J, Voglmayr H, Greilhuber J. (2003). Letter to the editor. *Cytometry*, *51A*(2), 127–128.

Ehlers JD, Hall AE. (1997). Cowpea (*Vigna unguiculata* L. Walp.). *Field Crop Research*, *53*(1997), 187–204.

Eichler EE., Holden JJA, Popovich BW, Reiss AL, Snow K, Thibodeau SN, Thibodeau SN, Richards CS, Ward PA, Nelson DL. (1994). Length of uniterrupted CGG repeats determines instability in the *FMR1* gene. *Nature*, *8*(4), 340–344.

Ekpong B. (2009). Effects of seed maturity, seed storage and pre-germination treatments on seed germination of cleome (*Cleome gynandra* L.). *Scientia Horticulturae*, *119*(3), 236–240.

El Bakkali A, Haouane H, Moukhli A, Costes E, van Damme P, Khadari B. (2013). Construction of Core Collections Suitable for Association Mapping to Optimize Use of Mediterranean Olive (*Olea europaea* L.) Genetic Resources. *PLoS ONE*, *8*(5), 23–24.

Ayua E, Mugalavai V, Simon J, Weller S, Obura P, Nyabinda N. (2016). Ascorbic acid content in leaves of Nightshade (*Solanum* spp.) and spider plant (*Cleome gynandra*) varieties grown under different fertilizer regimes in Western Kenya. *African Journal of Biotechnology*, *15*(7), 199–206.

Esselink GD, Nybom H, Vosman B. (2004). Assignment of allelic configuration in polyploids using the MAC-PR (microsatellite DNA allele counting-peak ratios) method. *Theoretical and Applied Genetics. Theoretische Und Angewandte Genetik*, *109*(2), 402–408.

Estoup A, Presa P, Krieg F, Vaiman D, Guyomard R. (1993). (CT)n and (GT)n microsatellites: a new class of genetic markers for Salmo trutta L. (brown trout). *Heredity*, *71*(February), 488–496.

FAO, IFAD, & WFP. (2015). *State of food insecurity in the world In brief*.

Farris JS, Albert VA, Kallersjo M, Lipscomb D, Kluge AG. (1996). Parsimony Jackknifing Outperforms Neighbor-Joining. *Cladistics*, *12*(2), 99–124.

Fay MF, Cowan RS, Leicht IJ. (2005). The Effects of Nuclear DNA Content (C-value) on the Quality and Utility of AFLP Fingerprints. *Annals of Botany*, *95*, 237–246.

Felsenstein, J. (1985). Confidence Limits on Phylogenies : An Approach Using the Bootstrap. *Society for the Study of Evolution*, *39*(4), 783–791.

Field D, Wills C. (1996). Long, polymorphic microsatellites in simple organisms. *Proceedings of the Royal Society B: Biological Sciences*, *263*, 209–215.

Food and Agriculture Organization of the United Nations. (2010). *The State of Food Insecurity in the World Addressing food insecurity in protracted crises 2010 Key messages*. *Notes*. http://www.fao.org/docrep/013/i1683e/i1683e.pdf

Fossati T, Labra M, Castiglione S, Failla O, Scienza A, Sala F. (2001). The use of AFLP and SSR molecular markers to decipher homonyms and synonyms in grapevine cultivars: The case of the varietal group known as “Schiave.” *Theoretical and Applied Genetics*, *102*(2–3), 200–205.

Fu YB. (2015). Understanding crop genetic diversity under modern plant breeding. *Theoretical and Applied Genetics*, *128*(11), 2131–2142.

Fufa F, Baum M, Grando S, Kafawin O, Ceccarelli S. (2007). Consequences of a decentralized participatory barley breeding programme on changes in SSR allele frequency and diversity in one cycle of selection. *Plant Breeding*, *126*(5), 527–532.

Garcia-Mas J, Oliver M, Gomez-Paniagua H, de Vicente MC. (2000). Comparing AFLP, RAPD and RFLP markers for measuring genetic diversity in melon. *Theoretical and Applied Genetics*, *101*, 860–864.

Gepts P. (2006). Plant genetic resources conservation and utilization: The accomplishments and future of a societal insurance policy. *Crop Science*, *46*(5), 2278–2292.

Gevorgyan E, Losenge T, Gefäller L, Elsen M, Cronjaeger P. (2016). *HORTINLEA/SLE-Survey conducted in 2014 within the frame of the SLE-Study-S260 analysing the Innovation System of African Leafy Vegetables in Kenya-(unpublished data)*. Nairobi/ Berlin.

Goulao L, Cabrita L, Oliveira CM, Leitao JM. (2001). Comparing RAPD and AFLP (TM) analysis in discrimination and estimation of genetic similarities among apple (Malus domestica Borkh.) cultivars - RAPD and AFLP analysis of apples. *Euphytica*, *119*, 259–270.

Govindaraj M, Vetriventhan M, Srinivasan M. (2015). Importance of genetic diversity assessment in crop plants and its recent advances: an overview of its analytical perspectives. *Genetics Research International*, *2015*(Figure 1), 1–14.

Greenbaum G, Templeton AR, Zarmi Y, Bar-David S. (2014). Allelic richness following population founding events - A stochastic modeling framework incorporating gene flow and genetic drift. *PLoS ONE*, *9*(12), 1–23.

Gur-Arie R, Cohen CJ., Eitan Y, Shelef L, Hallerman EM, Kashi Y. (2000). Simple Sequence Repeats in Escherichia coli: Abundance, Distribution, Composition, and Polymorphism. *Genome Research*, *10*(1), 62–71.

Hajjar R, Jarvis DI,Gemmill-Herren B. (2008). The utility of crop genetic diversity in maintaining ecosystem services. *Agriculture, Ecosystems and Environment*, *123*(4), 261–270.

Hall JC, Sytsma KJ, Iltis HH. (2002). Phylogeny of and Based on Chloroplast Sequence Data 1. *American*, *89*(11), 1826–1842.

Hall JC, Iltis HH, Sytsma KJ. (2004). Molecular phylogenetics of Core Brassicales, placement of orphan genera Emblingia, Forchhammeria, Tirania, and character evolution. *Systematic Botany*, *29*(3), 654–669.

Hamada H, Petrino MG, Kakunaga T. (1982). A novel repeated element with Z-DNA-forming potential is widely found in evolutionarily diverse eukaryotic genomes. *Proceedings of the National Academy of Sciences of the United States of America*, *79*(21), 6465–9.

Horticultural Crops Development Athority (HCDA). (2014). *Horticulture validated report 2014*.

Hughes AR, Inouye BD, Johnson MTJ, Underwood N, Vellend M. (2008). Ecological consequences of genetic diversity. *Ecology Letters*, *11*(6), 609–623.

Hughes AR,Stachowicz JJ. (2004). Genetic diversity enhances the resistance of a seagrass ecosystem to disturbance. *Proceedings of the National Academy of Sciences*, *101*(24), 8998–9002.

Hutchinson MJ. (2011). The effect of farmyard manure and calcium ammonium nitrate fertilizers on micronutrient density (iron, zinc, manganese, calcium and potassium) and seed yields of *Solanum villosum* (Black nightshade) and *Cleome gynandra* (Cat's whiskers) on eutric nitisol. *Journal of Agriculture Science and Technology*, *13*(1), 35–52.

Imanirampa L, Alele PE. (2016). Antifungal activity of *Cleome gynandra* L. aerial parts for topical treatment of *Tinea capitis*: An in vitro evaluation. *BMC Complementary and Alternative Medicine*, *16*(1), 1–8.

Inda LA, Torrecilla P, Catalán P, Ruiz-Zapata T. (2008). Phylogeny of *Cleome* L. and its close relatives Podandrogyne Ducke and Polanisia Raf. (Cleomoideae, Cleomaceae) based on analysis of nuclear ITS sequences and morphology. *Plant Systematics and Evolution*, *274*(1–2), 111–126.

Ishibashi Y, Saitoh T, Abe S, Yoshida MC. (1996). Null microsatellite alleles due to nucleotide sequence variation in the grey-sided vole Clethrionomys rufocanus. *Molecular Ecology*, *5*(4), 589–90.

Jaccard, P. (1901). Etude de la distribution florale dans une portion des Alpes et du Jura. *Bulletin de La Societe Vaudoise Des Sciences Naturelles*, *37*(142), 547–579.

Jin L, Macaubas C, Hallmayer J, Kimura A, Mignot E. (1996). Mutation rate varies among alleles at a microsatellite locus: Phylogenetic evidence. *Proceedings of the National Academy of Sciences of the United States of America*, *93*(26), 15285–15288.

Jinazali H, Mtimuni B, Chilembwe E. (2017). Nutrient composition of cat ' s whiskers (*Cleome gynandra* L .) from different agro ecological zones in Malawi. *African Journal of Food Science*, *11*(1), 24–29.

Johnston JS, Pepper AE, Hall AE, Chen ZJ, Hodnett G, Drabek J, Lopez R, Price, H. J. (2005). Evolution of Genome Size in Brassicaceae. *Annals of Botany*, *95*, 229–235.

Jusoh FM, Shamsuri S, Ayob MA, Johar MN. (2015). Effects of different types of foliar fertilizer on growth performance of spider Plant (*Cleome gynandra*). *Australian Journal of Basic & Applied Sciences*, *9*(24), 44–48.

Kalia RK, Rai MK, Kalia S, Singh R, Dhawan AK. (2011). Microsatellite markers: An overview of the recent progress in plants. *Euphytica*, *177*(3), 309–334.

Keller, G. B. (2004). African nightshade, eggplant , spiderflower et al.- production and consumption of traditional vegetables in Tanzania from the farmers point of view. Georg-August Universität Göttingen.

Khan ME, Yakubu UD, Kubmarawa D. (2015). Determination of mineral composition , mnti-aicrobial activity and anti-nutrient screening of *Cleome gynandra* (Cat's Whiskers). *British Journal of Applied Science and Technology*, *5*(4), 2231.

Kiebre Z, Bationo/Kando P, Sawadogo N, Sawadogo M, Zongo JD. (2015). Selection of phenotypic interests for the cultivation of the plant *Cleome gynandra* L . in the vegetabble gardens in Burkina Faso. *Journal of Experimental Biology and Agricultural Sciences*, *3*(3), 288–297.

Kimiywe J, Waudo J, Mbithe D, Maundu P. (2007). Utilization and medicinal value of indigenous leafy vegetables consumed in urban and peri-urban Nairobi. *African Journal of Food Agriculture Nutrition and Development*, *7*(4), 1–16.

King KC, Lively CM. (2012). Does genetic diversity limit disease spread in natural host populations? *Heredity*, *109*(4), 199–203.

Kitashiba H, Nasrallah JB. (2014). Self-incompatibility in Brassicaceae crops: lessons for interspecific incompatibility. *Breeding Science*, *64*(1), 23–37.

Kouam EB, Pasquet RS, Campagne P, Tignegre JB, Thoen K, Gaudin R., Ouedraogo JT, Salifu AB, Muluvi GM, Gepts P. (2012). Genetic structure and mating system of wild cowpea populations in West Africa. *BMC Plant Biology*, *12*(1), 113.

Krause H, Faße A, Grote U. 2017. African indiginous vegetable production - a pathway out of food insecurity? (In press)

Kutsukutsa RT, Gasura E, Mabasa S, Ngadze E. (2014). Variability in condensed tannins and bitterness in spider plant. *African Crop Science Journal*, *22*(4), 275–280.

K'Opondo FBO. (2011). Morphological characterization of selected spiderplant (*Cleome gynandra* L .) types from western Kenya. *Annals of Biological Research*, *2*(2), 54–64.

K'Opondo FB, van Rheenen HA, Muasya RM. (2009). Assessment of genetic variation of selected spiderplant (*Cleome gynandra* L.) morphotypes from western Kenya. *African Journal of Biotechnology*, *8*(18), 4325–4332.

Lippmann D, Lehmann C, Florian S, Barknowitz G, Haack M, Mewis I, Wiesner M, Schreiner M, Glatt H,Brigelius-Flohe R, Kipp, A. P. (2014). Glucosinolates from pak choi and broccoli induce enzymes and inhibit inflammation and colon cancer differently. *Food & Function*, *5*(6), 1073–81.

Liu J, Wang ZR, Li C, Bian YB, Xiao Y. (2015). Evaluating genetic diversity and constructing core collections of Chinese Lentinula edodes cultivars using ISSR and SRAP markers. *Journal of Basic Microbiology*, *55*(6), 749–760.

Lowe AJ, Moule C, Trick M, Edwards KJ. (2004). Efficient large-scale development of microsatellites for marker and mapping applications in Brassica crop species. *Theoretical and Applied Genetics*, *108*(6), 1103–1112.

Maguire TL, Sedgley M. (1997). Genetic diversity in Banksia and Dryandra (Proteaceae) with emphasis on Banksia cuneata, a rare and endangered species. Heredity, 79, 394–401.

Marshall DM, Muhaidat R, Brown NJ, Liu Z, Stanley S, Griffiths H, Sage RF, Hibberd, J. M. (2007). *Cleome*, a genus closely related to Arabidopsis, contains species spanning a developmental progression from C_3 to C_4 photosynthesis. *Plant Journal*, *51*(5), 886–896.

Masinde PW, Agong SG. (2011). Plant growth and leaf N of spider plant (*Cleome gynandra* L.) genotypes under varying nitrogen supply. *African Journal of Horticultural Science*, *5*, 36–49.

Masuka A, Goss M, Mazarura U. (2012). Morphological Characterization of four selected spider plant (*Cleome gynandra* L.) morphs from Zimbabwe and Kenya. *Asian Journal of Agriculture and Rural Development*, *2*(4), 646–657.

Mauyo LW, Anjichi VE, Wambugu GW, Omunyini ME. (2008). Effect of nitrogen fertilizer levels on fresh leaf yield of spider plant (*Cleome gynandra*) in Western Kenya. *Scientific Research and Essay*, *3*, 240–244.

McCouch SR, Chen X, Panaud O, Temnykh S, Xu Y, Cho YG, Huang N, Ishii T, Blair M. (1997). Microsatellite marker development, mapping and applications in rice genetics and breeding. *Plant Molecular Biology*, *35*(1–2), 89–99.

Miah G, Rafii MY, Ismail MR, Puteh AB, Rahim HA, Islam NK, Latif MA. (2013). A review of microsatellite markers and their applications in rice breeding programs to improve blast disease resistance. *International Journal of Molecular Sciences*, *14*(11), 22499–22528.

Mibei EK, Ojijo NKO, Karanja SM, Kinyua JK. (2012). Phytochemical and antioxidant analysis of methanolic extracts of four african indegenous leafy vegetables. *Annals. Food Science and Technology*, *13*(1).

Modi M, Modi A, Hendriks S. (2006). Potential role for wild vegetables in household food. *African Journal of Food Agriculture and Nutrition and Development*, *6*(1), 1–13.

Morgante M, Olivieri AM. (1993). PCR-amplified microsatellites as markers in plant genetics. *The Plant Journal : For Cell and Molecular Biology*, *3*(1), 175–182.

Mort ME, Soltis PS, Soltis DE, Mabry ML. (2000). Comparison of three methods for estimating internal support on phylogenetic trees. *Systematic Biology*, *49*(1), 160–171.

Motsa MM, Slabbert MM, van Averbeke W, Morey L. (2015). Effect of light and temperature on seed germination of selected African leafy vegetables. *South African Journal of Botany*, *99*(July), 29–35.

Mueller UG, Wolfenbarger LL. (1999). AFLP genotyping and fingerprinting. *Trends in Ecology and Evolution*, *14*(10), 389–394.

Muhanji G, Roothaert RL, Webo C, Stanley M. (2011). African indigenous vegetable enterprises and market access for small-scale farmers in East Africa. *International Journal of Agricultural Sustainability*, *9*, 194–202.

Muñoz-Amatriaín M, Mirebrahim H, Xu P, Wanamaker SI, Luo M, Alhakami H, Alpert M, Atokple I, Batieno BJ, Boukar O,Bozdag S, Cisse N, Drabo I, Ehlers JD, Farmer A, Fatokun C, Gu YQ, Guo YN, Huynh BL, Jackson SA, Kusi F, Lawley CT, Lucas MR, Ma Y, Timko MP, Wu J, You F, Barkley NA, Roberts PA, Lonardi S, Close, T. J. (2016). Genome resources for climate-resilient cowpea, an essential crop for food security. *The Plant Journal*, 1042–1054.

Ndengwa EA, Achigan-Dako EG, Mbugua G, Maye D, Ojanji W. (2013). Agricultural diversification with indigenous vegetables for cash cropping and nutrition: Examples from rift valley and central provinces in Kenya. *Acta Horticulturae*, *979*, 549–558.

Nei M, Li WH. (1979). Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proceedings of the National Academy of Sciences of the United States of America*, *76*(10), 5269–5273.

Nesamvuni C, Steyn NP, Potgieter MJ. (2001). Nutritional value of wild, leafy plants consumed by the Vhavenda. *South African Journal of Science*, *97*(1–2), 51–54.

Neugart S, Baldermann S, Ngwene B, Wesonga J, Schreiner M. (2017). Indigenous leafy vegetables of Eastern Africa — A source of extraordinary secondary plant metabolites. *Food Research International*. (In press). https://doi.org/10.1016/j.foodres.2017.02.014.

Ng NQ, Marechal R. (1985): Cowpea taxonomy, origin and germplasm, in: Singh S. R., Rachie K. O. (1985): *Cowpea research, production and utilization,* John Wiley and Sons, Chichester, UK, Pages: 11-21

Ngwediagi P, Maeda E, Kimwone H, Kamara R, Massawe S, Akonaay HB, Mapunda LND. (2009). *Tanzania report on the state of plant genetic resources for food and agriculture*.

Nyalala S, Grout B. (2007). African spider flower (*Cleome gynandra* L./*Gynandropsis gynandra* (L.) Briq.) as a red spider mite (*Tetranychus urticae* Koch) repellent in cut-flower rose (*Rosa hybrida* L.) cultivation. *Scientia Horticulturae*, *114*(3), 194–198.

Nyalala SO, Petersen MA, Grout BWW. (2011). Acetonitrile (methyl cyanide) emitted by the African spider plant (*Gynandropsis gynandra* L . (Briq)): Bioactivity against spider mite (*Tetranychus urticae* Koch) on roses. *Scientia Horticulturae*, *128*(3), 352–356.

Ochuodho JO, Modi AT. (2006). Statistical evaluation of the germination of *Cleome gynandra* L. seeds. *South African Journal of Plant and Soil*, *23*(4), 310–315.

Odhav B, Beekrum S, Akula U, Baijnath H. (2007). Preliminary assessment of nutritional value of traditional leafy vegetables in KwaZulu-Natal, South Africa. *Journal of Food Composition and Analysis*, *20*(5), 430–435.

Odongo GA, Schlotz N, Herz C, Hanschen FS, Baldermann S, Neugart S, Trierweiler B, Frommherz L, Franz CMAP, Ngwene B, Luvonga WA, Schreiner M, Rohn S, Lamy E. (2017). The role of plant processing for the cancer preventive potential of Ethiopian kale (*Brassicacarinata*). *Food and Nutrition Research*, *61*(1), 1–11.

Omondi, E. O., Debener, T., Linde, M., Abukutsa, O. M., Dinssa, F. F., and Winkelmann, T. (2017). Mating biology, nuclear DNA content and genetic diversity in spider plant (*Cleome gynandra*) germplasm from various African countries. *Plant Breeding,* in press.

Onyango C, Onwonga R, Kimenju J. (2016). Assessment of spider plant (*Cleome gynandra* L.) germplasm for agronomic traits in vegetable and seed production: A Green House study. *American Journal of Experimental Agriculture*, *10*(1), 1–10.

Padulosi S, Ng NQ. (1997): Origin, taxonomy and morphology of *Vigna unguiculata* (L.) Walp., In: Singh B. B., Mohan D. R., Dashiell K. E. & Jackai L. E. N.: Advances in Cowpea Research, IITA, Ibadan Nigeria, International Institute of Tropical Agriculture, Pages: 1-12

Paetkau D, Strobeck C. (1995). The molecular basis and evolutionary history of a microsatellite null allele in bears. *Molecular Ecology*, *4*(4), 519–520.

Pasquet RS. (1999). Genetic relationships among subspecies of *Vigna unguiculata* (L.) Walp. based on allozyme variation. *Theoretical and Applied Genetics*, *98*, 1104–1119.

Pemberton JM, Slete J, Bancroft DR, Barrett JA. (1995). Non-amplifying alleles at microsatellite loci: a caution for percentage and population studies. *Molecular Ecology*, *235*, 1616–1622.

Petersen BL, Chen S, Hansen CH, Olsen CE, Halkier BA. (2002). Composition and content of glucosinolates in developing *Arabidopsis thaliana*. *Planta*, *214*(4), 562–571.

Rajaselvam J, Rose SMRB. (2016). In vitro Antimicrobial activity and Phytochemical analysis of *Cleome gynandra* Linn Leaf Extracts Against Human Pathogens . *International Journal for Research in Applied Science and Engineering Technology*, *4*(viii), 793–796.

Rallo P, Tenzer I, Gessler C, Baldoni L, Dorado G, Martín A. (2003). Transferability of olive microsatellite loci across the genus Olea. *Theoretical and Applied Genetics*, *107*(5), 940–946.

Ramsay L, Macaulay M, Ivanissevich S, Maclean K, Cardle L, Fuller J, Edwards KJ, Tuvesson S, Morgante M, Massari A, Maetri E, Marmiroli N, Sjakste T, Ganal M, Powell W, Waugh R. (2005). A simple sequence repeat-based linkage map of barley, *2005*, 1997–2005.

Ranjitha J, Shalma M, Donatus M, Vijayalakshmi S. (2014). Isolatio on of novel phytoconstituents from the stem part of *Cleome gynandra* Linn and their antimicrobial activity. *International Journal of Phytomedicine*, *6*, 341–345.

Rao MRK,Kumar SS. (2015). Preliminary phytochemical analysis and antioxidant properties of *Gynandropsis pentaphyla*. *Der Pharmacia Lettre*, *7*(12), 20–24.

Rauf S, Jaime A, da Silva T, Khan AA, Naveed A. (2010). Consequences of plant breeding on genetic diversity. *International Journal of Plant Breed*, *4*(Troyer 1996), 1–21.

Reintanz B, Lehnen M, Reichelt M, Gershenzon J, Kowalczyk M, Sandberg G, Godde M, Uhl R, Palme K. (2001). *bus*, a bushy Arabidopsis *CYP79F1* knockout mutant with abolished synthesis of short-chain aliphatic glucosinolates. *The Plant Cell Online*, *13*(2), 351–367.

Rosa EAS. (1997). Daily variation in glucosinolate concentrations in the leaves and roots of cabbage seedlings in two constant temperature regimes. *Journal of the Science of Food and Agriculture*, *73*(3), 364–368.

Saitou N, Imanishi T. (1989). Relative efficiencies of the Fitch-Margoliash, maximum-parsimony, maximum-likelihood, minimum-evolution, and neighbor-joining methods of phylogenetic tree construction in obtaining the correct tree. *Molecular Biology and Evolution*, *6*(5), 514–525.

Saitou N, Nei M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution*, *4*(4), 406–425.

Saroop S, Kaul V. (2015). Correlation patterns among floral traits in *Cleome viscosa* L., a sexually porlymorphic species. *Current Science*, *109*(4), 208–214.

Schlötterer C. (2004). The evolution of molecular markers--just a matter of fashion? *Nature Reviews. Genetics*, *5*(1), 63–69.

Schmidt S, Zietz M, Schreiner M, Rohn S, Kroh LW, Krumbein A. (2010). Genotypic and climatic influences on the concentration and composition of flavonoids in kale (*Brassica oleracea* var. sabellica). *Food Chemistry*, *119*(4), 1293–1299.

Semagn K, Bjornstad A, Ndjiondjop MN. (2006). An overview of molecular marker methods for plants. *African Journal of Biotechnology*, *5*(25), 2540–2568. http://doi.org/10.5897/AJB2006.000-5110

Shanmugam S, Rajendran K, Suresh K. (2012). Traditional uses of medicinal plants among the rural people in Sivagangai district of Tamil Nadu, Southern India. *Asian Pacific Journal of Tropical Biomedicine*, *2*(1 SUPPL.), S429–S434.

Shashidhara G, Hema MV, Koshy B, Farooqi AA. (2003). Assessment of genetic diversity and identification of core collection in sandalwood germplasm using RAPDs. *Journal of Horticultural Science and Biotechnology*, *78*(4), 528–536.

Singh BB, Mohan Raj DR, Dashiell KE, Jackai LEN. (1997). Advances in Cowpea Research. Copublication of International Institute of Tropical Agriculture (UTA) and Japan International Research Center for Agricultural Sciences (JIRCAS). UTA, Ibadan, Nigeria. 375.

Smith FI, Eyzaguirre P. (2007). African Leafy Vegetables : Teir role in the World Health Organization's global fruit and vegetables initiative. *African Journal of Food Agriculture Nutrition and Development*, *7*(3), 1–17.

Soltis DE, Soltis PS, Bennett MD, Leitch IJ. (2003). Evolution of genome size in the angiosperms. *American Journal of Botany*, *90*(11), 1596–1603.

Songsak T, Lockwood GB. (2002). Glucosinolates of seven medicinal plants from Thailand. *Fitoterapia*, *73*(3), 209–216. http://doi.org/10.1016/S0367-326X(02)00061-8

Sowunmi LI, Afolayan AJ. (2015). Effects of environmental factors and sowing depth on seed germination in *Cleome gynandra* L. (Capparaceae). *Pakistan Journal of Botany*, *47*(6), 2189–2193.

Sowunmi LI,Afolayan AJ. (2015). Phytochemical constituents and antioxidant properties of acetone extract of Cleome gynandra (L.) growing in the Eastern cape, South Africa. African Journal of Traditional, Complementary and Alternative Medicine. 12 (3): 1-8, *12*, 1–8.

Squirrell J, Hollingsworth PM, Woodhead M, Russell J, Lowe AJ, Gibby M, Powell W. (2003). How much effort is required to isolate nuclear microsatellites from plants? *Molecular Ecology*, *12*(6), 1339–1348.

Stoilova T, Dinssa FF, Ebert AW, Tenkouano A. (2015). The diversity of African leafy vegetables: agromorphological characterization of subsets of AVRDC's germplasm collection. *Acta Horticulturae*, *1102*, 67–74.

Tautz D, Renz M. (1984). Simple sequences are ubiquitous repetitive components of eukaryotic genomes. *Nucleic Acid Research*, *12*(10), 4127–4138.

Timko MP, Singh BB. (2008). Cowpea, a Multifunctional Legume. *Genomics of Tropical Crop Plants*, *XXXIII*(2), 227–258.

Vaillancourt RE, Weeden NF. (1992). Chloroplast DNA polymorphism suggests Nigerian center of domestication for the cowpea, *Vigna unguciulata* (Leguminosae). *American Journal of Botany*, *79*(10), 1194–1199.

Van Belkum A, Scherer S, van Alphen L, Verbrugh H. (1998). Short-sequence DNA repeats in prokaryotic genomes. *Microbiology and Molecular Biology Reviews*, *62*(2), 275–293.

van den Bergh E, Külahoglu C, Bräutigam A, Hibberd JM, Weber APM, Zhu X, Schranz ME. (2014). Current Plant Biology Gene and genome duplications and the origin of C_4 photosynthesis : Birth of a trait in the *Cleomaceae*. *Current Plant Biology*, *1*, 2–9.

Van Jaarsveld P, Faber M, Van Heerden I, Wenhold F, Van Rensburg JW, Van Averbeke W. (2014). Nutrient content of eight African leafy vegetables and their potential contribution to dietary reference intakes. *Journal of Food Composition and Analysis*, *33*(1), 77–84.

Van Rensburg JWS, Venter SL, Netshiluvhi TR, Van Den Heever E, Vorster HJ, De Ronde JA. (2004). Role of indigenous leafy vegetables in combating hunger and malnutrition. *South African Journal of Botany*, *70*(1), 52–59.

Vasconcelos EV, de Andrade Fonsêca AF, Pedrosa-Harand A, de Andrade Bortoleti KC, Benko-Iseppon AM, da Costa AF, Brasileiro-Vidal AC. (2015). Intra- and interchromosomal rearrangements between cowpea [*Vigna unguiculata* (L.) Walp.] and common bean (*Phaseolus vulgaris* L.) revealed by BAC-FISH. *Chromosome Research*, *23*(2), 253–266.

Verkerk R, Schreiner M, Krumbein A, Ciska E, Holst B, Rowland I, De Shrijver R, Hansen M, Gerhauser C, Mithen R, Dekker M. (2009). Glucosinolates in Brassica vegetables: The influence of the food supply chain on intake, bioavailability and human health. *Molecular* Nutrition and Food Research, 53(SUPPL. 2), 219–265.

Villa TCC, Maxted N, Scholten M, Ford-Lloyd B. (2005). Defining and Identifying Crop Landraces. *Plant Genetic Resources*, *3*(3), 373–384.

Vorster HJ, Stevens JB, Steyn G. (2008). Production systems of traditional leafy vegetables: Challenges for research and extension. *South African Journal of Agricultural Extension*, *37*(1), 85–96.

Vos P, Hogers RMB, Reijans M,van de Lee T, Hornes M, Fritjers A, Pot J, Peleman J, Kuiper M, Zabeau M. (1995). AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research*, *44*(21), 388–396.

Wang ML, Barkley NA, Jenkins TM. (2009). Microsatellite markers in plants and insects. Part I: applications of biotechnology. *Genes, Genomes And, Genomics*, *3*, 54–67.

Wangolo E, Onyango C, Gachene C, Mong'are P. (2015). Effects of shoot tip and flower removal on growth and yield of spider plant (*Cleome gynandra* L.) in Kenya. *American Journal of Experimental Agriculture*, *8*(6), 367–376.

Wasonga DO, Ambuko JL, Chemining'wa GN, Odeny DA, Crampton BG. (2015). Morphological characterization and celection of spider plant (*Cleome gynandra*) cccessions from Kenya and South Africa. *Asian Journal of Agricultural Science*, *7*(4), 36–44.

Waterman C, Rojas-silva P, Tumer TB, Kuhn P, Richard J, Wicks S, Stephens JM, Wang Z, Mynatt R, Cefalu W, Raskin, I. (2014). Isothiocyanate-rich *Moringa oleifera* extract reduces weight gain, insulin resistance and hepatic gluconeogenesis in mice. *Molecular Nutrition and Food Research*, *56*(6), 1013–1024.

Weber JL, May PE. (1989). Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. *American Journal of Human Genetics*, *44*(3), 388–96.

Weinberger K, Msuya J. (2004). Indigenous vegetables in Tanzania- Significance and Prospects. *The World Vegatable Center. Technical Bulletin No. 31. Taiwan*.

Weising K, Nybom H, Wolf K, Kahl G. (2005). *DNA Fingerprinting in Plants: Principles, Methods, and Applications, Second edition* (2nd ed.). New York: CRC Press, Taylor & Francis Group.

Wenyika P, Gasura E, Kafesu N, Goss M, Matikiti A, Kujeke G. (2015). Discriminating phenotypic markers reveal low genetic diversity in spider plant. *African Crop Science Journal*, *23*(4), 305.

Yan X, Chen S. (2007). Regulation of plant glucosinolate metabolism. *Planta*, *226*(6), 1343–1352.

Zedler B, Srinivasan R, Su F. (2016). Journal of Asia-Paci fi c Entomology Assessing the potential of spider plant (*Cleome gynandra* L.) as a trap crop for the management of specialist feeders on vegetable brassicas, *19*, 477–485.

Zietz M, Weckmüller A, Schmidt S, Rohn S, Schreiner M, Krumbein A, Kroh LW. (2010). Genotypic and climatic influence on the antioxidant activity of flavonoids in kale (*Brassica oleracea* var. sabellica). *Journal of Agricultural and Food Chemistry*, *58*(4), 2123–2130.

5. Appendices

5.1. AFLPadapters preparation

***Mse*I adapter (MseI O + U) [50 pmol/µL]**

Dilute the *Mse*I O and the *Mse*I U oligonucleotides, separatly synthesized by the manufacturer, to 100pmole/µL (= 100 µM).
Mix equal volumes of both oligonucleotides. This results in a final concentration of 50 pmol/µL (one oligo is diluted by the other 1:1). Prepare aliquotes of 50µl.
AdapMseI-oben 5'-GAC GAT GAG TCC TGA G-3'
AdapMseI-unten 5'-TA CTC AGG ACT CAT-3'

***Hin*dIII adapter (HindIII O + U) [5 pMol/µL]**

Dilute the *Hin*dIII O and the *Hin*dIII U oligonucleotides, separatly synthesized by the manufacturer, first to 100 pmole/µL (= 100µM). Prepare aliquotes of 50µl in 1.5ml tubes. Dilute one tube each to to 10 pmole/µL (+ 450µl H_2O).
Mix equal volumes of both oligonucleotides. This results in a final concentration of 50 pMol/µL. Prepare aliquotes of 50µl.
AdapHind-oben 5'-CTC GTA GAC TGC GTA CC-3'
AdapHind-unten 5'-AGC TGG TAC GCA GTC TAC-3'

Oligonucleotides sequences

DNA-Adapter	Sequences
MseI O	5'-GACGATGAGTCCTGAG-3'
MseI U	5'-TACTCAGGACTCAT-3'
HindIII O	5'-CTCGTAGACTGCGTACC-3'
HindIII U	5'-AGCTGGTACGCAGTCTAC-3'

AFLP-Primer	Sequences
MseI+0	5'-GACGATGAGTCCTGAGTAA-3'
HindIII+0	5'-AGACTGCGTACCAGCTT-3'
MseI+XXX	5'-GATGAGTCCTGAGTAA XXX-3'
HindIII+XXX (IRD 700 or 800)	5'-GACTGCGTACCAGCTT XXX-3'

5.2. Buffers, dyes and working solutions used

10x RL Buffer
100 mM Tris HCL
100 mM MgAc
500 mM KAc
50 mM DTT (77 mg/10 mL)
pH 7.5

Loading dye for the Licor Sequencers
49 ml of 98% Formamide
0.5 ml of 10 mM EDTA
25 mg of Pararosanilin (0.05%)

1x TBE buffer(pH 8.0)
89 mM Tris
89 mM Boric acid
2 mM Na_2EDTA

10 x Williams buffer
10 ml of 100 mM Tris/HCL, pH 8,3
50 ml of 500 mM KCL
2 ml of 20 mM $MgCl_2$
10 ml of Gelatine (0.1 %)
28 ml of H_2O

1x TAE buffer
40 mM Tris
1mM EDTA
pH 8.0 (adjust with concentrated acetic acid)

10x DNA loading buffer
0.25 % (w/v) Orange G (Merck Millipore, Darmstadt, Germany)
30.00 % (w/v) Glycerin
1mM EDTA pH 8.0

5.3. Components of the acrylamide gels

Thecomponentsandtheiramountusedforthepreparationof6%acrylamidegelsusingtheSequaGelXRsolution are listed in Table 24.

Componentsusedforthepreparationof6%acrylamidegels.

	SSR	**AFLP**
Monomer Solution	12 ml	16 ml
Complete Buffer	3 ml	4 ml
APS (10 %)	120 µl	160 µl

5.4. The principle of M13-SSR-PCR

To consider:
Order the forward Primer with a M13-tail, i.e. add the sequence **GTAAAACGACGGCCAGT** to the 5'-end of the forward Primer. (Example RMS015: 5`**GTA AAA CGA CGG CCA GTT AAT GTA GGC AGA TAT AAA GGA GT 3**`)
The 700 or 800 IRD labelled M13-Primer (forward = uni) can be used. The concentration of the stock solution is 100 pmol/µL. The dilution is 1:40 (1 µL Primer + 39 µL H_2O) to obtain a working solution with a concentration of 2.5 pmol/µL. **Avoid exposure to light!!!** (Note: there are 10 µL aliquots of stock solution in the freezer. Please dilute these completely and make 40 µL aliquots. Finish these aliquots before new dilutions are made.)

M13-SSR-PCR

	Single SSR Reaction (10 µl) in tubes
H2O **[µl]**	**5,65**
10x Williams Buffer **[µl]**	**1**
dNTPs **(2mM) [µl]**	**0,75**
Taq-Polymerase DCS **(5U/µl) [µl]**	**0,1**
M13 tailed forward primer **(0,5pmol/µl** = 1:200**) [µl]**	**0,5**
M13 forward labelled (2,5pmol/µl = 1:40 **) [µl]**	**0,5**
reverse primer **(5,0 pmol/µl** = 1:20**) [µl]**	**0,5**
Reaction Volume [µl]	**9,00**
DNA (10ng/µl) [µl]	**1,0**

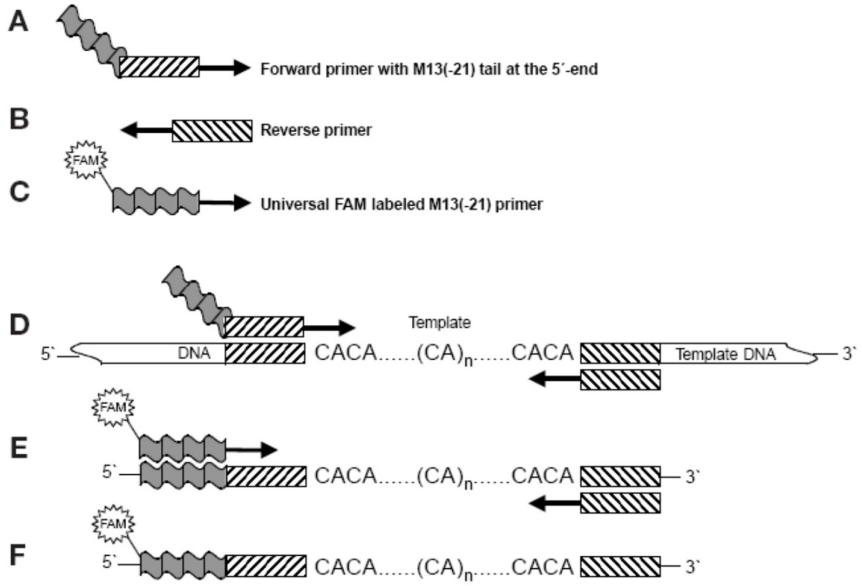


Figure 7: Principle of M13-SSR-PCR(from Schuelke, 2000)

5.5. AFLP and SSR markers gel profiles

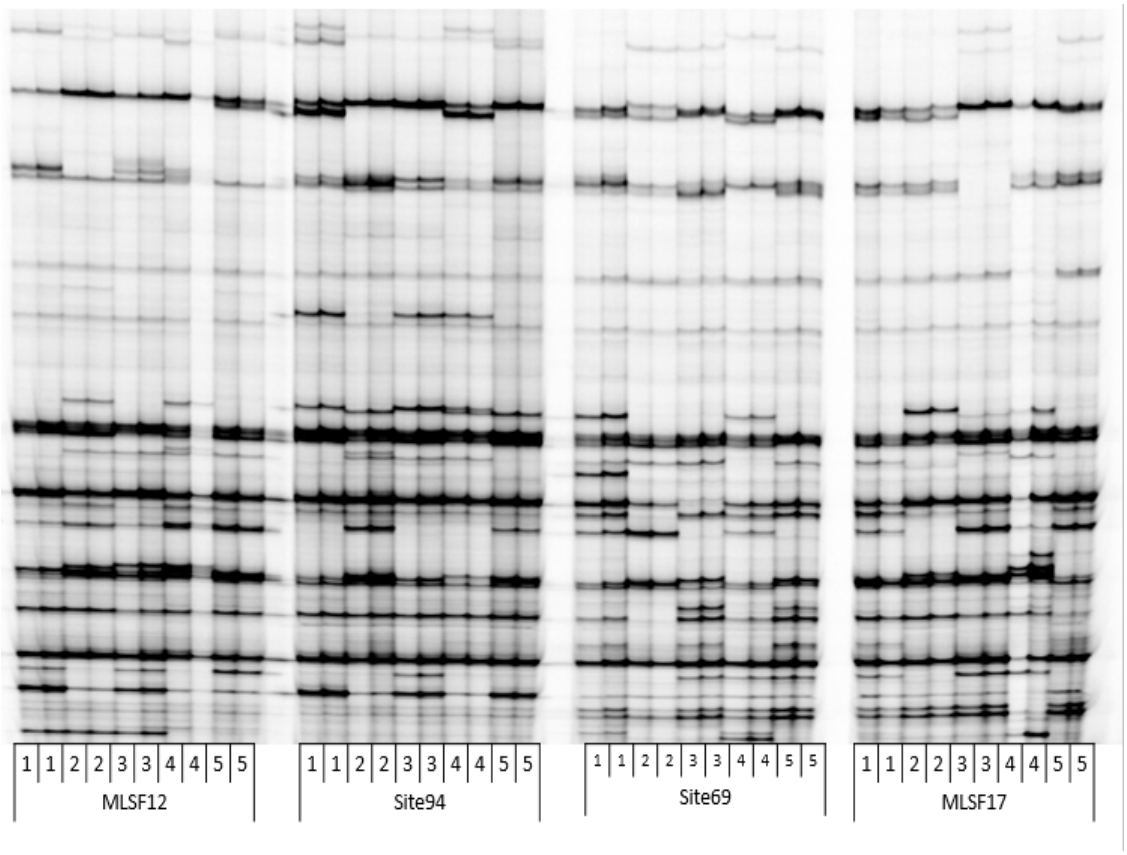


Figure 8: Amplified fragment length polymorphism (AFLP) banding patterns for the spider plant genomic DNA from 4 accessions. Hind III (ATA) and Mse I (GAC)

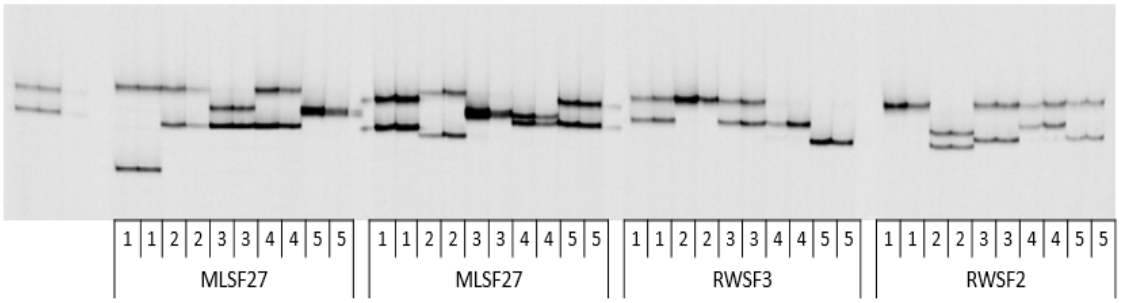


Figure 9: Simple sequence repeat (SSR) marker CG001 banding patterns for the spider plant genomic DNA from 4 accessions.

Curriculum vitae

Emmanuel Gordon Omondi Otunga

Date of birth	21st April, 1984
Place of birth	Kisumu, Kenya

EDUCATION

2014 - 2017	PhD student at the Gottfried Wilhelm Leibniz University Hannover, Germany Institute of Horticulture Production Systems, Department of Woody Plant and Propagation Physiology.
2008 - 2011	Jomo Kenyatta University of Agriculture and Technology, Nairobi, Kenya Masters of Science in Botany (Genetics) Research topic: 'Comparative Analysis of Genetic Variability, Water stresses Tolerance and Nutritional Properties of Selected Sorghum (*Sorghum bicolor* (L) Moench) varieties in Kenya.'
2004 - 2008	Jomo Kenyatta University of Agriculture & Technology, Nairobi Kenya Bachelor of Science in Botany Major in Biotechnology. Research topic: 'In vitro propagation of the money plant (*Epipremnum auream*)'
1999 - 2002	Maranda Boys High School, Bondo, Kenya Kenya Certificate of Secondary Education (K.C.S.E.)
1991 - 1998	Shaurimoyo Primary School, Kisumu, Kenya Kenya Certificate of Primary Education (K.C.P.E.)

WORK EXPERIENCE

2014-2017	Research Assistant Leibniz University Hannover, Germany Institute of Horticultural Production, Department of Woody Plant and Propagation Physiology Systems
2012-2014	Part-time Teaching Assistant Jomo Kenyatta University of Agriculture and Biotechnology, Department of Botany and University of Eldoret , Department of Biological Sciences.
2009-2011	Part-time Lab Technician Jomo Kenyatta University of Agriculture and Technology, Nairobi Department of Botany/ Institute of Biotechnology Research (IBR-JKUAT)

PUBLICATIONS

Peer reviewed

Emmanuel O. Omondi, Christof Engels, Godfrey Nambafu, Monika Schreiner, Susanne Neugart, Mary Abukutsa-Onyango, Traud Winkelmann. 2017. Nutritional compound analysis and morphological characterization of spiderplant (*Cleome gynandra*) - an African indigenous leafy vegetable. Food Research International 100: 284-295.

Max Menssen, Marcus Linde, **Emmanuel Otunga Omondi**, Mary Abukutsa-Onyango, Fekadu Fufa Dinssa, Traud Winkelmann. 2017. Genetic and morphological diversity of cowpea (*Vigna unguiculata* (L.)Walp.) entries from East Africa. Scientia Horticalturae 226: 268-276.

Emmanuel O. Omondi, Thomas Debener, Marcus Linde, Mary Abukutsa-Onyango, Fekadu F. Dinssa, Traud Winkelmann. 2017. Mating biology, nuclear DNA content and genetic diversity in spider plant (*Cleome gynandra*)germplasm from various African countries. Plant Breeding 2017;1-12.

Emmanuel O. Omondi, Thomas Debener, Marcus Linde, Mary Abukutsa-Onyango, Fekadu F. Dinssa, Traud Winkelmann. 2016. Molecular Markers for Genetic Diversity Studies in African Leafy Vegetables. Advances in Bioscience and Biotechnology 7: 188-197

Omondi, E.G.O., Makobe, M.N., Onyango, C.A., Matasyoh, L.G., Imbuga, M.O. and Kahangi, E.N. 2013. Evaluation of Nutritional Properties of Tissue Cultured Sorghum (*Sorghum bicolor*). Journal Of Agriculture Science and Technology 15(1).

Omondi, E.G.O., Makobe, M.N., Matasyoh, L.G. and Onyango, C. A. 2013. Genetic variability of tissue cultured *Sorghum bicolor* (L) Moench as revealed by morphological traits and simple sequence repeats (SSR) markers. African journal of biotechnology 12(1):1-7.

CONFERENCE PRESENTATIONS

Oral presentations

***2nd*International Conference on Biodiversity for Food&Nutrition,**23-24th November, 2015*,* Nairobi*, Kenya.*
"Analysis of genetic and reproductive characteristics of spider plant (*Cleome gynandra*) - an African vegetable"

Poster presentations

Tropentag 2016, 18-21st September, 2017, University of Natural Resources and Life Sciences (BOKU Vienna), Germany.
"Morphological diversity and performance of *Cleome gynandra* (L.) Briq.-an African leafy vegetable germplasm collections"

5th Quedlinburger Pflanzenzüchtungstage /18th Kurt von Rümker Vorträge / GPZ Meeting of AG Genome Analysis (QPzT-GPZ-KvR-2017), 01.-03rd March, 2017, The Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Gatersleben.
"Genetic diversity in Spider plant (*Cleome gynandra*) germplasm from various African countries"

Acknowledgements

I would like to sincerely thank the German Federal Ministry of Education and Research (BMBF) and the German Federal Ministry of Economic Cooperation and Development (BMZ) for the funds to have made this work possible within the project HORTINLEA in the framework of the program GlobE – Global Food Security.

My deepest and sincere gratitude goes to my PhD advisor Prof. Dr. Traud Winkelmann, for allowing me the PhD opportunity in her working group, for her immense contribution that gave my work value, for nurturing my skills in research and inspiring me through her detailed discussion and for support throughout my study. My sincere thanks also goes to my co-advisor Prof. Dr. Thomas Debener and Dr. Marcus Linde for allowing me perform my molecular work in their lab besides offering useful comments, suggestions and critiques during our fruitful discussion, Prof. Mary Abukutsa-Onyango and Dr. Fekadu Dinsa for providing the spider plant and cowpea seed material, and for their very helpful comments, critiques and suggestions.

I would also like to appreciate our project cooperation partners Prof. Dr. Monika Schreiner and Dr. Sussane Neugart. for their help in determining the secondary metabolites and for their very helpful suggestions and comments in discussing the results, Prof. Christof Engels and Mr. Godfrey Nambafu for their contribution in the mineral analysis.

My special appreciation also goes to Bunlong Yim for sharing his expertise in data analysis and kind technical assistance and Dr. Phillip Braun for introducing me into flow cytometry analysis and pollination experiments, Ute Meisel and Sabine Reinecke for helping me with administrative work, Mr. Herbert Geyer and the greenhouse technicians for making my experiments possible in the greenhouse.
I am indebted to all my colleagues at Leibniz Universität Hannover, Department of Horticultural Production Systems, Section of Woody Plant and Propagation Physiology especially my office mates Dr. Clarissa Caprestano, Dr. Mona Quambusch, Svenja Ratjens, Felix Mahnkopp and Annmarie-Deetja Rohrfor their support in discussions, for the friendship that made me feel at home. I also want to thank my Kenyan and other international friends for their warmth and fun moments that made me feel home especially Tracy Cherogony for her encouragements, prayer, kind support and sincere friendship. “Baraka tele watu wangu”.

My utmost gratitude goes to God for my beloved wife, Mrs. Everline Atieno Omondi and our son Andrew Emmanuel Amani. They made a huge sacrifice to allow me leave for my study far away from home and making me feel loved even when I was far off. Last yet not least, I cannot thank enough my entire family especially my dad Mr. Francis Otunga and mum Mrs. Dorcas Otunga for their unconditional support, encouragement, counsel and sacrifice to help me be successful and responsible.

I dedicate this thesis to
my parents, Francis and Dorcas, my wife, Eve, and my beloved son Amani
for their constant support and unconditional love.
I love you all dearly.