GENETIC DIVERSITY STUDY OF *Vanilla **planifolia*** G. Jackson, syn. *V. fragrans* CROP GROWN IN TANZANIA USING MOLECULAR TECHNIQUES.

BY

CONSTANTINE BUSUNGU

A DISSERTATION SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE IN CROP SCIENCE OF SOKOINE UNIVERSITY OF AGRICULTURE. MOROGORO, TANZANIA.

2009

ABSTRACT

Natural vanilla (*Vanilla planifolia G. Jackson, syn. V. fragrans)* is native to the tropic forest of Mexico. It is now cultivated in humid tropical areas of Africa, America, Asia and Australia continents. In Tanzania it had been cultivated since 1940s in Kagera region but decline in coffee prices in the world market in the 1990s resulted into more cultivation of the vanilla crop as an alternate crop in Kagera and Kilimanjaro regions. Vanilla which had a remarkable high price compared to coffee was consequently spread and grown in Tanga, Morogoro regions and Zanzibar Island. A study was undertaken to identify cultivars and examine the extent of genetic diversity of *Vanilla planifolia* using molecular technique.

A total of 126 samples were randomly taken from Bulinda, Bakabuye and Kibona villages in Kagera Region, Mkunazini and Donge village in Zanzibar Island and Mudio and Kidia villages in Kilimanjaro Region. The deoxyribonucleic acid (DNA) extraction method that was used in this research was the Dellaporta protocol.

DNA quantification was done by comparing band intensity of different concentrations of standard genomic DNA markers using agarose gel electrophoresis. For the optimization of PCR, three different components; dNTPs, Taq DNA polymerase and MgCl2 were tested in different combinations.Ten primers were selected for the Random Amplified Polymorphic DNA (RAPD) reactions.

The results revealed genetic distances of 0.667, 0.705 and 0.805 for Kagera, Kilimanjaro and Zanzibar respectively for intra population diversity and 1.28 for interpopulation diversity. This indicates a narrow genetic diversity. Similarly altogether the 126 cultivars included in this study generated 53 bands of which only 27 were polymorphic this indicates a low level polymorphism.

The fact that the vanilla cultivars from Kilimanjaro and Kagera did not markedly diverge genetically from the vanilla cultivars in Zanzibar island suggests a narrow genetic diversity of populations and probably present cultivars have been derived from common source parents and maintained over several decades. Exchange of cultivars between plantations and seedling selection may also have played a role.

DECLARATION

I, **CONSTANTINE BUSUNGU;** do hereby declare to the Senate of Sokoine University of Agriculture that this dissertation is my own original work, and has not been submitted for a degree award in any University.

_____ _____

CONSTANTINE BUSUNGU

Date

(MSc. Candidate)

_____ _____

Dr.C.L.RWEYEMAMU Date

(Supervisor)

_____ _____

Prof. R.N.MISANGU Date

(Supervisor)

COPYRIGHT

ACKNOWLEDGEMENTS

I would like to express my gratitude to many individuals who assisted me throughout the course of my MSc research. First and foremost, I would like to express my gratitude to my advisor, Dr.C. Rweyemamu and Prof. R. Misangu, for the wonderful opportunity of participating in this exciting research endeavour and for constant guidance, unwavering moral support during my years as MSc student.

My sincere gratitude also goes to Dr. Ndunguru Joseph and Dr. Tairo Fred for invaluable assistance, suggestions and for special attention, personal help, using his time and effort to show me different computer programs. I would be remiss not to mention all staff at International Institute of Tropical Agriculture (IITA) Tanzania office who in one way or another have contributed a lot for my life in the institute to be so enjoyable and unforgettable, and making this study a successful experience.

I also owe heartfelt exceptional thanks to my wife, Margareth Mageni, for her patient love. Without her support and love, I would have not been able to complete this work. My sincere thanks are to my family: my mother for educating me and for her unconditional support and encouragement throught out my life; my brothers and my sisters for their ever-loving care towards me

Finally, all this could not be possible without the fund for this research which was provided by SUA -PANTIL Project so I am very thankful to them for helping me.

DEDICATION

This dissertation is dedicated to my wife Margareth, my beloved mother Bernadette Maria and my late Father Athanasius Kwilasa whom I saw once but his memories still lives on.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF PLATES

LIST OF ACRONYMS

AFLP	Amplified Fragment length Polymorphism
AMOVA	Analysis of Molecular Variance
ARI	Agriculture Research Institute
CMV	Cucumber Mosaic Virus
CYMV	Cymbidium mosaic Virus
dNTPs	Deoxyribo Nucleotide Tri -Phosphate
DNA	Deoxy ribo nucleic acid
EDTA	Ethylene DiamineTetraacetic Acid
FAO	Food and Agriculture organization
GD	Genetic Distance
IITA	International Institute of Tropical Agriculture
ITC	International Trade Commission
INV	Invitrogen
ISSR	Inter Short Simple Repeats
Kg	Kilograms
MAYAWA	Maendeleo Ya Wakulima
NBS	National Bureau of Statistics
NTSYS	Numerical Taxonomy and multivariate Analysis System
PCR	Polymerase Chain Reaction
PIC	Polymorphism Information Content per primer
ORSV	Odontoglossum Ring Spot virus

RAPD	Random Amplified polymorphic DNA
RFLP	Restriction Fragment Length Polymorphism
RPM	Rotation Per Minute
SDS	Sodium Dodecyl Sulfate
SM	Simple Matching
SNP	Single Nucleotide Polymorphism
SSR	Simple Sequence Repeates marker
STS	Sequence Tagged Sites
TAE	Tris-Acetate-Edta
Taq	Thermus aquas
TAS	Tanzanian Shillings
TUKI	Taasisi ya Ukuaji wa Kiswahili
UPGMA	Unweighted Pair Group Method with Arithmetic mean
UN	United Nations
URT	United republic of Tanzania
UV	Ultra violet
VMC	Vanilla Mosaic Virus
VNV	Vanilla Necrosis Virus

CHAPTER ONE

1. 0. INTRODUCTION

Vanilla planifolia G. Jackson, syn. V. fragrans, arguably a highly prized tropical spice crop is a creeping perennial plant growing in the humid tropics in America, Asia and Africa (FAO, 2006). According to FAO the worldwide vanilla production in 2007 was 8438 tonnes from 19000ha area harvested. The flavouring material is obtained from the dried, cured, full sized but not fully ripe fruits (beans). Vanilla is generally used as "vanilla extract", also as "vanilla sugar"(finely ground beans mixed with sugar) (FAO, 2006 Soto *et al*., 2003). Vanilla is one of the world's most popular flavours, with an estimated annual worldwide consumption of over 8000 tons (FAO, 2007).

Natural vanilla flavour is obtained from beans of the vanilla plant whereby the major flavour compound is vanillin (4-hydroxy-3-ethoxybenzaldehyde), although over 250 different compounds have been isolated from vanilla beans, including a 4-hydroxy benzaldehyde, 4-hydroxybenzoic acid and 4-hydroxy-3-methoxybenzoic(Guarino and Brown,1985).The vanilla beans are harvested up to 8 months post-pollination. At this stage, the green beans are flavourless but contain large quantities of glucosides of the various flavour compounds. The characteristic flavour develops on “curing” of the beans, a process that can last for as much as 6 months and which is associated with increases in hydrolytic enzymes which include glycosidases,

esterases, proteases and lipases, as well as oxidative enzymes such as polyphenol oxidases and peroxidases (Dignum et al., 2001; Soto *et al*., 2003).

The vanilla crop which is also known as “Lavani” in Swahili language (TUKI, 2002) is grown on a plantation scale in Java, Mauritius, Madagascar, China, India, Uganda, Tahiti, Seychelles, Reunion, Brazil and Jamaica and other islands of the West Indies (FAO, 2006).Vanillin ($C_8H_8O_3$), which is extracted from vanilla, is mainly used in flavouring ice creams, soft drinks, condiments, oleoresins, beverages, pudding, cakes, chocolates, baked goods, syrups, candies, and pharmaceuticals (Purseglove *et al*., 1981).

Most Tanzanians who live in rural areas where they largely depend on subsistence agriculture are poor (NBS 2002; URT 2005). For many years people living in rural areas have depended largely on major cash crops such as sisal, cashew nuts, cotton and coffee as a source of income. Farmers in the coffee growing areas have started growing vanilla as an additional cash crop. Good prices of vanilla in 2002, 2003 and 2004 which was 30,000 -.40, 000 TAS per Kg of green vanilla beans also stimulated more investment in vanilla as an additional cash crop. However in 2005 the price of vanilla dropped significantly in the world market and local market. In the world market the price dramatically slumped from 450-550$ per Kg of cured beans of vanilla in previous years to 20$ per Kg of cured vanilla in 2005 while in the local market in Tanzania the local price of green vanilla slumped significantly from TAS 40,000 to TAS 6000 per Kg for grade one to three and TAS 500 for lower grades (FAO 2006; Ngaga *et al.*, 2006).

Despite the decline of vanilla price in the world market the growing of vanilla in Tanzania has not been affected and in fact the production of vanilla crop is increasing slowly year after year (Rweyemamu *et al*., 2007).

Vanilla still commands the highest mean price (TAS 5,562 per kg) by a larger margin compared with other cash and food crops, followed by groundnuts (TAS 800 per kg),coffee (TAS 635 per kg) cotton (TAS 500 per kg),beans (TAS 375 per kg),banana (TAS 284 per kg) and maize (TAS 214 per kg) and another reason for increasing cultivation of vanilla despite decline of its price in the world market is the fact that vanilla cultivation in the regions where they grow banana and coconut is incorporated with banana and coconut as mixed cropping without needing additional land(Ngaga *et al.,* 2006*).*

In Bukoba district the crop is also locally referred to as green gold or "mkombozi". According to the District Agriculture and Livestock Development Officer in Bukoba district, by June 2006 there were 5500 vanilla farmers who are organized in 85 groups and each group consists of 10-30 farmers. The vanilla production in Bukoba is three tons of green beans (or 0.8 tons of cured beans) annually (Rweyemamu *et al*., 2007). Also, apart from Bukoba district, in the past three years, new vanilla plots have been initiated in some parts of Kilimanjaro, Kigoma, Tanga, Mbeya, Morogoro regions and in Zanzibar Island. In all new vanilla production locations in Tanzania Bukoba district has been the main source of planting materials (Ngaga *et al., 2006;* MAYAWA, 2001).

The history of vanilla cultivation in Bukoba district spans back since colonial period and up to now the crop has been a very important source of income as a cash crop; The crop has been used as means of employment for the resource poor farmers and is utilized in domestic uses like flavouring in ice cream, beverages, cakes, candy, baked goods, yoghurt and liquor, including local gin known as "enkonyagi" (MAYAWA, 2001)

The source of vanilla that is grown in Bukoba district is known to be Uganda where it is believed to have come from Madagascar and Reunion through Christian missionaries in Uganda (MAYAWA, 2001). Sixty five percent of vanilla farmers in Bukoba and Kilimanjaro areas did not know the type of vanilla varieties they were growing and studies indicated that these vanilla varieties were probably from a very limited resource base hence with narrow genetic base (Ngaga *et al., 2006).* Vanilla is clonally propagated and has shown a narrow genetic base (Fouche and Jouve, 1999).

Little is known about the levels of genetic variability within and between *Vanilla planifolia* cultivars/varieties which are cultivated in Tanzania (MAYAWA, 2001, MAFS, 2004) therefore proper germplasm characterizations is essential to meet legal requirements for cultivar registration or for its protection under the plant breeders Protection Act (Smith and Helentjaris, 1996).

RAPD assay in conjunction with botanical descriptions can be useful in the documentation of the distinctness, uniformity, and stability of the available elite cultivars of *vanilla spp*. In addition, these molecular markers can also provide

valuable information about the phylogenetic relationships within and between *Vanilla planifolia.*

Preliminary investigation currently going on at Sokoine University of Agriculture indicates that there are no substantive morphological differences between vanilla cutting materials collected from Bulinda,Kibona and Bakabuye villages in Bukoba district (Mategu, 2007).The study of genetic structure of vanilla has not been done in Tanzania but is of paramount importance for guiding parental choice in breeding programs and the success of any breeding program will depend on the availability of extensive level of genetic variability (Rweyemamu *et al., 2007).*

The results will be beneficial from both academic and economic point of view due to the fact that it will reveal the genetic diversity information among and within the vanilla germplasm used in Tanzania as some species of *Vanilla planifolia* may represent valuable sources of useful traits for the improvement of *Vanilla planifolia* with regards to disease resistance, self-pollination, or aromatic quality (Soto *et al*., 2003; Minoo *et al*., 2006; Bory *et al*., 2007).

1.1. GENERAL OBJECTIVE

1.1.1. To characterize the vanilla cultivars grown in Tanzania using molecular technique.

1.2. SPECIFIC OBJECTIVES

1.2.1. To determine the extent of intra-population genetic diversity of vanilla cultivars grown In Tanzania using RAPD markers.

1.2.2. To determine the extent of inter-population genetic diversity of vanilla cultivars grown In Tanzania using RAPD markers.

1.2.3. To identify potential parental genotypes for mapping and marker assisted selection that maybe used in vanilla national crop improvement.

1.3 HYPOTHESIS

1.3.1 Null Hypothesis-There is a wide genetic variation in Vanilla cultivars grown in Tanzania

CHAPTER TWO

2. 0. LITERATURE REVIEW

2. 1. Origin and distribution of Vanilla planifolia

Vanilla planifolia originates from the tropically-humid regions of Mexico and Central America, yet also grows wild in the virgin forests of South America where the native populations use it as a spice and also as perfume (Purseglove *et al*., 1981). Vanilla arrived in Europe during the Spanish conquest in the 15th century, and was later spread throughout Africa and Asia. It is particularly popular for fine, aromatic taste (Soto, 1999).

There are 110 described species of vanilla, distributed in the tropics of both the world and the New World. They belong to the orchid family, Orchidaceae, which is the largest family of flowering plants, with about 700 genera and 20,000 species (Soto *et al*., 2003).

The Orchidaceae comprise a very natural, distinctive and highly advanced group of monocotyledons. They are perennial herbs which are widely distributed throughout the world with the greatest number in the tropics as they exhibit a wide range of life form and have terrestrial, climbing, epiphytic and saprophytic species. Apart from the large number of ornamental species which are grown for the flowers, vanilla is the only genus which has species of economic importance (Porteres, 1954).

2. 2. Ecological Requirements

2. 2.1. Altitude

Vanilla is an epiphytic orchid and in a wild state V. planifolia usually grows climbing on trees in wet tropical lowland forests. The aerial roots of epiphytes obtain much of their nutrients from the air and from debris and mosses collected in the trees upon which they grow. Vanilla is successfully grown from sea level to 600 metres. The crop thrives in hot, moist, insular climates, with frequent, but not excessive rain (Nielsen, 2000).

2. 2.2. Temperature

The optimum temperature is 21-32 degrees Celsius, with an average around 27 degrees Celsius. However a distinctly cooler period, with temperatures towards 20 °C will also help induce flowering (Minoo 2002).

2. 2. 3. Rain fall

Rainfall should be between 1700 - 2000 mm, and evenly distributed. However, two drier months are required (with precipitation considerably lower than evaporation) to check vegetative growth and induce flowering (Kumar, 2004).

2. 2. 4. Shading

The thumb rule is that vanilla requires about 50 % shade. Therefore in location where vanilla is grown it is necessary to retain the natural shade provided by lofty trees that allow penetration of sunlight to the ground level and to leave the soil or the rich humus layer on top undisturbed (Pearson ***et al.,*** **1991).** Vanilla requires support and shade to be grown successfully. It is planted from cuttings once a suitable

support tree has been established. A considerable amount of inert compost/mulch (coconut husks, cocoa shells, dry leaves, etc.) is required for healthy root development. As the vanilla vine grows, it is looped around the support tree with the ends placed into the mulch (Fouche and Jouve, 1999).

2. 2. 5. Soil

Vanilla is cultivated in a wide range of soil types provided there is plenty of organic matter. Vanilla grows well in well-drained, sandy loam soil, rich in humus with pH ranging between 6 and 7 and on a gently sloping land. Clayey soils and water logged areas are not suitable for the plant. The land should be gentle sloped with light friable sandy loam to laterite soils, adequate but not excessive drainage, and a thick surface layer of humus or mulch in which the roots can spread (Purseglove ***et al.,* 1981 and** Soto *et al.*, 2003). Growing vanilla does not affect production of companion crops like bananas, coconut, clove and coffee. It can therefore be grown as an additional crop to the main crops on the same piece of land (Fouche and Jouve, 1999).

2. 3. Botany and Cultivation

2. 3. 1. Vanilla species

Orchid species are grown for their flowers. *Vanilla* is the only genus of the orchid family whose species produce a commercially important flavouring material. Out of about 110 known species of *Vanilla* only *Vanilla planifolia* is commercially important today (Stern and Judd, 1999). The species produces narrowly cylindrical, ill-defined three sided pods, which are 14 to 22 cm long and 1 to 1.5 cm in diameter.

Two other species of minor commercial importance are *Vanilla tahitensis* and *Vanilla pompona* (Soto *et al*., 2003).

V. tahitensis is indigenous to Tahiti, and may be of manmade hybrid origin. Vanilla beans of *V. tahitensis* have distinct floral aroma and contain anisyl alcohol, anisyl aldehyde, and anisic acid in relative abundance in contrast to their virtual absence in widely used *V. planifolia* beans (Guarino and Brown, 1985). The beans from this species also contain a much higher level of *p*-hydroxybenzoic acid than either *V. planifolia* or *V. pompona*. The delicate and more fragrant character of this vanilla is capitalized on to create interesting and exotic culinary fares. It is sold in the specialized market for much higher price compared to *V. planifolia* beans from Madagascar. It has slender stems, narrower leaves and shorter pods than *V. planifolia*. The indehiscent pods are 12-14 cm long and about 1.5-1.8 cm wide, which turn reddish brown rather than dark chocolate brown upon curing (Stern and Judd, 1999).

V. pompona, also known as vanillon, produces inferior quality vanilla beans, but is the hardiest species. *V. pompona* can grow under more unfavourable moisture conditions and poorer soil than *V. planifolia*. It is also more resistant to *Fusarium batatatis*, the root rot disease. This species might be very useful in the future for cross-breeding disease resistant varieties of *V. planifolia*. The pods of *V. pompona* are almost cylindrical, short (10-12.5 cm long) and fat (2-2.5 cm in diameter) (Soto, 1999).

2. 3. 2. Vanilla plant

Vanilla planifolia is a fleshy perennial vine, which climbs trees or other supports by means of adventitious roots called holders (aerial roots). If not trained, it can climb 10 to 15 meters. However, in commercial vanilleries it is trained to a height which allows hand pollination and harvesting carried out without using a ladder. The roots 2mm in diameter, long, aerial, adventitious are produced singly opposite the leaves and adhere firmly to the support. The roots at the base are shallow rooted and ramify in the humus or mulch, which are very important for that reason. **. (Pearson *et al.*, 1991).** The stems are long, cylindrical, simple or branched, succulent and brittle. They are 1-2 cm in diameter and are dark green. The internodes are at 5-15 cm intervals. The leaves are large, flat, fleshy, alternate and oblong-elliptic. They are 8-20 cm long and 3-6 cm broad. At the joint of each stem and a leaf is a bud (Nielsen, 2000). This bud can develop into a root, a new growing shoot or inflorescence. Inflorescences are stout and simple and about five to eight cm long. Each inflorescence can bear from 15 to 30 flowers, which open from the base upwards. From one to three flowers open at a time and last only for one day. The large, waxy, pale greenish flowers are about seven to 10 cm in diameter (Sengupta, 2004).

2. 3. 3. Planting

The vanilla seedlings are planted 6-12 months after the site has already been planted with trees. The best time to plant is directly before the rainy season, on sites which lay slightly higher, and are in no danger of becoming water-logged. **(Pearson *et al.*, 1991).** Two seedlings are planted beside each tutor tree. The seedlings should be

planted in rills that are 30 cm long and 10 cm deep, and then covered over with top soil. Finally, the planting hole is covered over to an area of at least 1 m^2 with a layer of organic material mulch. The seedling is then attached to the tutor. The density of plants can be up to 400 and 800 plants/ha (Kumar, 2004).

2. 3. 4. Crop nutrition

Basically there are three approaches in crop nutrition. Organic farming, integrated nutrient management and chemical farming are practiced to supply nutrients for crop growth. In organic farming, organic manures, both bulky and concentrated besides bio inoculants are liberally used to sustain soil health. Whereas in integrated nutrient management all the three components, which are organic manures, inorganic fertilizers and bio inoculants are integrated. Chemicals alone are used in chemical farming which results in pollution of the environment and several other ill effects (Kumar, 2004; Sengupta, 2004).

Organic farming approach is ideal for vanilla gardens. Farmyard manure, vermicompost, coir pith compost, neem cake, wood ash, bone meal, dry leaves etc. can be used for manuring. Bio inoculants such as, Fluorescent pseudomonad's, Azospirillum and phosphorus solubilising bacteria are also beneficial for promotion of plant growth. NPK requirement of vanilla varies and the general recommendation is 40 to 60 g of N, 20 to 30 g of P2O5, and 60 to 100 g of K2O per vine per year. Organic manure is incorporated with the onset of south west monsoon in June and fertilizers are applied in two to three splits (Kumar, 2004).

2. 3. 5. Trailing and Coiling

A special system of trailing is practiced to facilitate induction of flowering and promotion of vegetative growth. When the vines attain a height of 150 cm, the after growth has to be trailed downwards very close to the ground leaving a gap of 30 cm and again coiled up like a loose loop (Purseglove *et al*., 1981). Support trees are suitably pruned at a convenient height for easy trailing of vines. Bamboo poles or split areca nut stump fixed at a convenient height on the support tree can also be used for trailing of vines.

This is an essential farm operation for flowering. Besides, it helps for easy hand of aerial roots. Production of a large number of yielding vines is also possible when this type of trailing is adopted. The top most leaf of the hanging vine which is positioned at the point of coiling for instance 150 cm height from the ground is folded backwards and the auxiliary bud is exposed to sunlight. A new shoot starts developing from the auxiliary bud. After a few years, the support tree may not be able to bear the weight of vines and pods and it breaks off. So, it is necessary to remove the yielded senile portion of the vine periodically to maintain the support tree in good condition (Ranadive, 1994, 1996).

2. 3. 6. Weeding, Mulching and Irrigation

Vanilla is a surface feeder and the surface soil should not be disturbed once the crop is established. However, weeds should be removed by slashing and the dry leaves preferably lopping from the support trees is widely practiced in vanilla gardens mainly for enriching soil organic matter (Purseglove *et al*., 1981).

Mulching with coconut husk is ideal for soil moisture conservation. Besides, it provides a favourable soil microclimate for root development. Coconut husk acts as an insulator for the surface feeding roots of vanilla from adverse weather conditions. Plant water potential for instance internal plant water status decides the pattern of growth in vanilla. A lower plant water potential result in induction of flowering and a higher potential encourages vegetative growth. Growth is arrested under certain moisture regimes (Walton *et al*., 2003). Being an orchid, it is able to absorb atmospheric moisture if the atmospheric moisture potential is higher than that of vanilla. So, sprinkler irrigation is ideal for promotion of vine growth since it provides a favourable microclimate for absorption of moisture through leaves besides supplementing soil moisture. Modern micro irrigation systems can also be installed depending upon the availability of water, edaphic and topographic features (Nielsen, 2000).

Drip irrigation can also be practiced. As far as possible, flooding should be avoided as vanilla can not withstand water logging. Weed biomass is spread around vanilla basins as mulch. Pruning or lopping of shade trees or standards is carried out to provide 50 % shade to vanilla vines. The pruning or lopping is also used for mulching the vines. A slight reduction in shade is necessary one or two months prior to flower bud initiation. The practice of spreading organic waste materials such as dry leaves, bushy, chopped dry stems etc, around the base of plants is called mulching (Sengupta, 2004).

2. 3. 7. Flowering and Artificial pollination

Juvenile phase of vanilla lasts for about two to three years. Depending upon the length of the planting material, vanilla starts flowering from the second or third year of planting. Regulation of shade and slight wilting of vines encourage flowering. Partial lopping of support trees and avoiding irrigation and removal of the top 15 cm of vine one to two months prior to flower bud initiation are ideal. Lopping of shade trees should be done without causing any damage to the hanging and fruiting vines. It takes 45 days from the initiation of inflorescence to opening of its first flower. The inflorescence in the leaf axils consists of around 15 to 20 greenish yellow flowers (Soto, 1999).

Self pollination never happens in vanilla because a structure called rostellum prevents stigma coming into direct contact with the pollen grains. Hence, artificial hand pollination is resorted to by pushing back the rostellum with the help of a tooth pick or a pointed bamboo splinter. The pollen sac is further pressed to spread pollen over the stigma. Artificial pollination starts from early morning and completed before noon as the flower closes in the after noon. The retention of flower and enlargement ovary is indications of the success of artificial pollination. A skilled worker can pollinate 1000 flowers a day. Even though every plant produces 18 to 20 inflorescences, only 10 to 12 Inflorescences are allowed to develop and in each inflorescence only 10 to 15 flowers alone are hand pollinated to ensure the formation of high quality beans. Generally only one flower opens a day and flowering continues for about three weeks. Pods take about eight to nine months to

attain maturity. During flowering chemicals should not be sprayed as it may result in scorching of ovaries and subsequent scab formation in beans (Sengupta, 2004; Kumar, 2004).

2. 4. Pests and Diseases

Vanilla plants are, in general, free from any major pests and disease incidence. However, conditions leading to weakened plants by drought, lack of nutrients, too much sun, over pollination and bean production certainly favour the incidence of diseases. Environmental conditions favouring diseases are excessive moisture, prolonged rainy weather, and insufficient drainage, too much shade damage to roots and over planting of vanilla (Wen and Li, 1992).Fungal diseases like shoot tip rot stem and bean rot caused by *Phytophthora sp*. as well as immature bean dropping are sometimes noticed. The cultivation of vanilla is threatened by root rot caused by *Fusarium batatis* and anthracnose caused by *Calospora vanillae* and more recently *Phytophthora* infection has also been reported in India (Joseph and Bhai, 2001).

Incidence of diseases can be reduced by maintaining conditions leading to vigorous growth of vines such as adequate shade, heavy mulch especially during the dry season, moderate to light pollination, irrigation during extended dry periods, and application of fertilizers and providing adequate drainage. Over crowding of vines in a single support tree also should be avoided. The disease affected portions should be removed regularly and fungicides such as one per cent Bordeaux mixture or 0.2 per-cent Indofil-MA5 (200 g in 100 litres of water) may be applied in a need based manner to reduce the spread of diseases (Ruan *et al*., 1998).

Table 1. Overview of Vanilla Fungal Disease

Disease and Pathogens	Country the disease reported
1. Stem rot (Basal rot) (Fusarium)	India, Indonesia
2. Root rots (Fusarium)	India, PuertoRico, Polynesia
3. Stem blight (Phytophthora.)	India, French Polynesia
4. Shoot tip rot (Fusarium, Colletotrichum)	India
5. Beans rot (Fusarium, Phytophthora,)	India
6. Yellowing & shedding of young beans (Fusarium)	India
7. Inflorescence rots (Fusarium)	India
8. Leaf rot, Leaf Spot (Colletotrichum)	India
9. Horse hair blight (Marasmius)	India
10. Storage rots (Aspergillus flavus, A. parasiticus, Penicillium sp.)	India, Indonesia, and Tahiti
11. Anthracnose (Callospora)	Madagaskar, Seychelles, Mauritius
12. Black crust (Mycoleptodiscus)	Madagaskar, Tahiti

Source : Joseph *et al*. (2003)

Table. 2 Overview of Vanilla Viral diseases

Viral disease	Country Reported
1. Vanilla mosaic (VMV)	India, Reunion, Fiji, Cooks Island, Vanuatu
2.Vanilla Necrosis (VNV)	Tonga, Fiji, Vanauta, India
3. Cymbidium mosaic (CYMV)	S. Pacific Islands, Tahiti
4. Odontoglossum Ring Spot virus (ORSV)	Tonga
5. Unidentified (Poty Virus)	Fiji, Vanuatu
6. Cucumber mosaic virus (CMV)	Fiji, Vanuatu

Source: Joseph and Bhai. (2001)

2. 5 Yield

The yield of vanilla varies depending upon the age of vines and the method of cultivation (Kumar, 2004). Normally it starts yielding from the third year and the yield goes on increasing till the seventh or eighth year. Thereafter the yield slowly starts declining till the vines are replanted after another seven to ten years. In one acre you can plant about 3000 vanilla plants (Ranadive, 1996).

Each plant is expected to yield about 500-800 grams of green beans per year. Under reasonable level of management, the yield range of a middle aged plantation will be

about 500 kg-1000 kgs of green beans per acre. For every 5 kgs of green beans you can get around 1 kg of cured beans (Kumar, 2004).

Table. 3 Leading countries in Vanilla production in 2006

Country	Production(tonnes)	Area harvested(ha)
China	1350	1500
Comoros	90	450
Guadeloupe	8	40
Indonesia	3700	9000
Kenya	8	40
Madagascar	2600	6400
Malawi	20	80
Mexico	350	750
Portugal	15	50
Reunion	25	250
Turkey	192	670
Uganda	70	140
Zimbabwe	10	45
Total	8438	19415

Source: FAO. (2007)

2. 6 Harvesting and processing

Vanilla pods are ready for harvest after eight to nine months of flowering. Generally harvesting begins in October and ends in December. As there is no uniformity in the development of pods even in a single inflorescence due to variations in flower bud initiation, there is no synchronization in the maturity of pods and each pod is harvested separately by looking at its colour (Guarino and Brown, 1985). Immature dark green pods are not harvested. Pale yellow discoloration which commences at the distal end of the beans is an indication of the maturity of pods. Mature pods alone are picked and its commercial value is fixed based on the length of the pod. If the pod length is more than 15 cm it is categorized under first quality. 10 to 15 cm

long pods are put under second quality and pods having less than 10 cm length under third quality (Pearson ***et al.*, 1991).**

Harvest should be done at the appropriate stage as over mature pods are likely to split causing a reduction in market value. Green vanilla pods are odourless and flavourless and contain little vanillin. Every hanging vine has the potential to produce an inflorescence in every leaf axils. Once the hanging vine has fully transformed into the fruiting vine and all the leaf nodes have produced a flower bunch each, which may take about two to three years, the vine can be cut and removed to reduce the biomass load on the support tree (Nielsen, 2000).

The pod yield depends on the care and management given to the hanging and fruiting vines. Any practice directed to stimulate aerial root production has a direct bearing on vine productivity.

A five year old vine can produce 1.5 to three kg pods which may increase steadily to six kg after a few years. The harvested green pods can be sold as such or processed for getting a better market price (Walton *et al*., 2003).

Four steps are involved in the curing process of harvested vanilla pods. The process consists of a) killing by dipping the pods in hot water (63 to 65°C) for three minutes to arrest the vegetative growth of pods and initiation of enzymatic reactions for the development of aroma, b) Sweating by wrapping in woollen clothes to raise the temperature of beans and alternate storing in air tight wooden boxes during night and exposing to sun during day for about eight to ten days, c) Slow drying by

spreading beans in a wooden rack at room temperature for three to four weeks which results in considerable reduction in bean weight.(Walton *et al*.,2003). The weight is reduced to one third and the beans become flexible and can be twisted on a finger and d) conditioning of beans by storing in closed boxes for a few months. The processed beans are sorted, graded and bundled and wrapped in paraffin paper and preserved for the development of desired bean qualities, especially flavour and aroma (Ranadive, 1994, 1996).

2. 7. Vanilla market

Natural vanilla is mainly used for high-quality confectionery and for baking; it is also becoming increasingly popular in manufacturing ice-cream. Expensive brands of chocolates made of organically cultivated ingredients are also flavoured with vanilla. (Walton *et al*., 2003).

In addition to the processed capsules (vanilla sticks), the ground down fruits are also marketed as vanilla powder, or, mixed with sugar, as vanilla sugar. Other products include vanilla extract, an alcoholic extract (35% alcohol) made with sugar and binding substances, which is used in a variety of dilutions. (Walton *et al*., 2003). Along with vanillin (0.75-3.7%), vanilla also contains vanillic acid, alcohols, ester of cinammic acid, phydroxybenzaldehyde and other taste and odour substances, in addition to sugar, resins, mucilage, tanning agents and fat. Besides vanillin, Vanilla tahitensis also contains piperonal, which is mainly used in the manufacturing of perfumes (Guarino and Brown, 1985).

It is estimated that more than 500 tonnes of vanilla beans is used in USA every year in the preparation of cola type drinks. The major industrial purchasers of vanilla are pharmaceutical companies and soft drink companies like Coke and Pepsi. However the fact remains that market for natural vanilla essence is today largely only confined to the western countries (Todd, 1999).

Table. 4. World leading Countries in Vanilla Importation

Importing Country	Kg	% of world import
United States	1,111,137	62.1
France	323,400	18.1
Germany	148,300	8.3
United Kingdom	62,088	3.5
Japan	59,768	3.3
Canada	31,767	1.8

Australia	19,234	1.1
Switzerland	9,623	0.5
Netherlands	8,100	0.5
Austria	5,300	0.3
Norway	4,555	0.2
Italy	4,431	0.2
South Africa	1,739	0.1
Others	1,495	0.1
Total	1,790,637	100

Source: Manceau. (2003)

Vanilla imports are dominated by three countries- USA, France and Germany (Table 4). Importers in Germany and France are suppliers to other markets especially in Europe. Europe imports generally high quality beans while USA accepts low quality beans too (Manceau, 2003).

There is an understanding between Bourbon vanilla producing countries which includes Madagascar, Comoro and Reunion, and importers of France and Germany in the marketing of vanilla beans. Ugandan and Tanzanian vanilla seems to be earning a good reputation as a Bourbon-like vanilla. In 2002, 2003, 2004 and 2005 Ugandan exports amounted to 70, 120, 90,185 tons respectively of cured vanilla worth about 35 million US$ and continued to grow. Tanzanian vanilla is not yet visible in the market, since the small available quantity has so far been sold through

Uganda. It seems logical for both countries to capitalize on Uganda's current position, and even try to define a joint marketing strategy, which should lead to increased visibility and recognition of East-African vanilla (Paul, 2003).

The international vanilla market is characterized by its extreme volatility. Three features dictate the international vanilla market, viz., highly speculative cycles, raw material quality and competition from synthetic vanillin. Prices can fluctuate enormously over a period of few months. The world production was of the order of 5,400 tonnes in 2004(Manceau, 2003). The production has been increasing rather heavily in recent years with an average annual growth at 4 % over a 10-year period. Increase in production without corresponding growth in demand has pushed cured vanilla prices in the international market down to US $ 35 a kg from over US$ 500 a kg two years back. Shooting prices during 2003-04 was due to the shortage of supply from Madagascar, owing to their political instability and crop failure as a result of repeated cyclones (Paul, 2003).

High prices in the recent years have enthused farmers in countries such as Uganda, Papua New Guinea, India, Costa Rica, China and Colombia to take up vanilla cultivation. The current fall of prices can be attributed to the recovery of the Madagascar crop, introduction of new entrants in the market and shift in demand to alternate products such as synthetic vanillin and natural identical vanillin as a result of skyrocketing prices of natural vanillin. Although vanilla extract from beans is still used by the food industry, this accounts for less than 1 % of vanillin production. Remaining 99% is obtained synthetically (Manceau, 2003).

The current depression in the world vanilla market is expected to continue to prevail until companies who shifted to alternate products revert to using natural vanillin from beans. It is estimated that in the international market, the price of cured beans may range from 20-500 USD ($) per kilogram. On the commodities market, vanilla beans cost importers about $33 per kilogram in 2000 and by 2003 were at an all-time high of $400-$500 per kilogram (Paul, 2003). According to ITC/UN statistics the total global demand for vanilla is about 2000 to 3000 metric tonnes a year with the world market for vanilla beans highly concentrated in a few developed countries (Partos, 2006).

2. 8. Characterization

Characterization is the description of a character or quality of an individual (Merrium, 1991). The word ‘characterize’ is also a synonym of ‘distinguish’, that is, to mark as separate or different, or to separate into kinds, classes or categories. Thus, characterization of genetic resources refers to the process by which accessions are identified or differentiated. This identification may, in broad terms, refer to any difference in the appearance or make-up of an accession. In the agreed terminology of gene banks and germplasm management, the term ‘characterization’ stands for the description of characters that are usually highly heritable, easily seen by the eye and equally expressed in all environments (Papa and Gepts, 2003).

2.8.1 Morphological Characterization

The term morphological characterization refers to a broad subject that involves taxonomy or botanical descriptors, which distinguish one plant from another or one accession against other accessions. Morphological descriptors which are usually included in morphological characterization are mature plant height, plant spread, panicle height and length, and flag leaf height, width, length, flower colour and root type (Altaf *et al*., 1997).

The importance of morphological features was demonstrated by Mendel's work. Mendel used simple morphological traits and studied their inheritance in hybrids produced from distinct homozygous parents.

The outcome of his work was the demonstration that traits are transferred from parents to offspring by factors called genes. The discipline of genetics was opened by a work based on morphological traits. Nowadays, Mendel's work and morphological traits are still used to understand the mechanism of inheritance and genetics of traits in many different species (Rudall *et al*., 2000).

Morphological characteristics were among the earliest genetic markers used for assessment of variation and are still of great importance. Usually, these characters are inexpensive and simple to score. The sharing of physical features is also often accepted as an indication of relatedness. There are several sets of physical character assessment for different crops at different developmental stages for

instance in vanilla physical characters such as flowering time, flower number, umbel form, distance stigma, pedicel orientation, pedicel colour, pedicel length, leaf orientation, leaf sheath colour, leaf margin, leaf apex, aerial stem, endocarp colour distribution, pod size, pod length and width are often used in studying morphological diversity (Porteres, 1954; Purseglove *et al*., 1981)). However, these sets of characters lack adequate coverage of the genome, are strongly influenced by environmental factors, and are apparently controlled by several genes (Wang and Tanksley, 1989). Besides, assessment of morphological characters in perennial plants such as vanilla, often require a lengthy and expensive evaluation during the whole vegetative growth. (Bekele, 2005).

Morphological and physiological features of plants have been used to understand the genetic variation. Though morphological features are indicative of the phenotype, they are affected by environmental factors and growth practices (Capdevielle, 2001; Bekele, 2005). Many studies have been done on morphological diversity study of *vanilla spp* particularly on *Vanilla aphyla*, *Vanilla planifolia*, *Vanilla tahitiensis*, *Vanilla pompona* and their hybrids, which showed low level similarity (Fouche and Jouve 1999). Preliminary results in the study of *Vanilla planifolia* cultivated in Tanzania using morphological descriptors also show greater similarity (Mategu, 2007).

Tanksley (1983) listed five properties that distinguish molecular markers from morphological markers. These properties are (1) genotypes can be determined at the

whole plant, tissue and/or cellular level, (2) a relatively larger number of naturally occurring alleles exists at many loci, (3) phenotypic neutrality i.e., deleterious effects are not usually associated with different alleles, (4) alleles at many loci are co dominant, thus all possible genotypes can be distinguished and (5) few epistatic or pleiotrophic effects are observed.

2.8.2 Molecular Characterization

In genetic terms, characterization refers to the detection of variation as a result of differences in either DNA sequences or specific genes or modifying factors. Standard characterization and evaluation of accessions may be routinely carried out by using different methods, including traditional practices such as the use of descriptor lists of morphological characters. They may also involve evaluation of agronomic performance under various environmental conditions. In contrast, genetic characterization refers to the description of attributes that follow a Mendelian inheritance or that involve specific DNA sequences (Wang and Tanksley, 1989).

In this context, the application of biochemical assays such as those that detect differences between isozymes or protein profiles, the application of molecular markers and the identification of particular sequences through diverse genomic approaches all qualify as genetic characterization methods. The genetic diversity has an impact on the higher levels of biodiversity (Templeton 1991, 1993).

Understanding the patterns of genetic diversity of species is fundamental not only for addressing questions concerning evolutionary processes and the development of

conservation strategies, but also as a prerequisite for the efficient use of genetic resources in breeding programs (Sun *et al*., 1998}. The interest in the genetic structure of natural populations has increased in the last few years owing to the necessity to broaden the knowledge of genetic variation in cultivated species. New approaches to the study of genetic variation from wild species to cultivated varieties, mediated by information on molecular markers, are promising avenues to exploit wild genetic resources (Excoffier *et al*., 1992).

Genetic biodiversity finds its natural resources in wild species for which is important to find out the amount of genetic variability by the way of morphological, biochemical and molecular markers, besides some interesting physiological turns. Characterization of diversity has long been based on morphological traits mainly. However, morphological variability is often restricted, characters may not be obvious at all stages of the plant development and appearance may be affected by environment. Nowadays, a variety of different genetic markers has been proposed to assess genetic variability as a complementary strategy to more traditional approaches in genetic resources management (Dos Santos *et al*., 1994).

Molecular tools provide valuable data on diversity through their ability to detect variation at the DNA level. Identification is of fundamental importance in diversity studies in a variety of different ways. For evaluation of species diversity, it is essential that individuals can be classified accurately. The identification of taxonomic units and endangered species, whose genetic constitution is distinct from their more abundant relatives, is important in the development of appropriate

conservation strategies (Zhang *et al.*, 1999). In population studies, molecular tools are being used to identify whether two individuals are from the mating of specific parents and estimating the degree of relatedness among individuals, and are helpful for the determination of social behaviour, reproductive success and mating choice. Identification is of importance in studying clonal organisms and the roles of sexual and asexual reproduction, where the definition of population units and different genetic lines may be difficult. From a more practical standpoint, identification of who breeds with whom is important for the management of small populations (Excoffier *et al*., 1992).

Moreover, being able to detect gene flow, introgression and contamination from unwanted out crossing is crucial for the attainment of purity and has become a central issue in debate over the safety of genetically modified organisms. A number of techniques are available for identifying genetic differences between organisms and choice for any one specific use will depend upon the material being studied and the nature of the questions being addressed. Molecular techniques differ in the way they sample within the genome and in the type of data that they generate (Templeton, 1993).

2.8.3 Molecular Markers

Numerous molecular markers have been developed and applied for analyses of genetic diversity and relatedness, and all marker systems have their strengths and weaknesses (Schlötterer, 2004; Weising *et al*., 2005). This Substantial progress which has been made in recent years in mapping, tagging and isolating many

agriculturally important genes using molecular markers due in large part to improvements in the techniques that have been developed to help find markers of interest. Among the techniques that are particularly promising are Restriction Fragment Length Polymorphism (RFLPs), Amplified Fragment Length Polymorphism (AFLPs), Random Amplified Polymorphic DNA (RAPDs), Microsatellites and PCR based DNA markers such as Sequence Characterized Amplified Regions (SCARs) or Sequence Tagged Sites (STS), Single Nucleotide Polymorphism (SNP), Short Simple Repeats (SSR) and Inter Short Simple Repeats (ISSR). The following are the most extensively used techniques and their potential in future plant genomics study (Schlötterer, 2004).

2.8.3.1 Isozymes

The term ‘isozymes’ was introduced by Markert and Moller (1959) and refers to protein forms of an enzyme with the same catalytic activity, and converting the same substrate, but differing in molecular weight or in electric charge. The differences in size and charge can be caused by amino-acid substitutions or posttranslational modifications. Protein variants in isozyme analysis are distinguished by gel electrophoresis and visualized by an enzyme-specific staining mixture, where

substrate, co-factor and an oxidized salt are included. Isozymes are not necessarily products of the same gene, and they can be active at different life stages or in different cell compartments. Isozymes encoded by the same locus but by different alleles are usually referred to as allozymes (Weising *et al.*, 2005).

The main advantage of isozyme analysis, especially when allozymes are used, is their co-dominant inheritance, which allows the discrimination between homozygous and heterozygous genotypes. This is necessary for precise estimations of population genetics parameters, especially in cross-pollinated species (Zhang *et al.*, 1999).

The main disadvantage is that isozyme analysis detects variation only in protein coding loci and therefore provides fewer markers compared to DNA-based methods. Furthermore, an isozyme system often show low variability and in some cases no variability at all due to a low rate of mutational events. Isozymes have not been used extensively in the study of genetic diversity of Vanilla *spp*. (Dean *et al* 1999).

2.8.3.2 Amplified Fragment Length Polymorphism (AFLP)

The AFLP technique is used to visualize numerous amplified DNA restriction fragment simultaneously (Vos *et al*., 1995). An AFLP protocol includes several steps: (i) the digestion of genomic DNA with two restriction endonucleases, (ii) the ligation of digested DNA to double stranded nucleotide adoptors (iii) pre-selective amplification of genomic fragments containing an adopter at each end and (iv)

selective amplification using primers with base extensions. The selective primers under stringent annealing conditions will amplify only those fragments with complementary nucleotides extending beyond the restriction site and this reduces the complexity of the mixture (Zhang *et al.*, 1999; Vos *et al*., 1995, Xu *et al*., 2000).

This technique has a number of advantages (i) minimal amount of genomic DNA are required (ii) no prior genome sequence data is necessary for primer construction (iii) markers are randomly distributed throughout the genome and (iv) many bands are generated that potentially provide a large number of polymorphism (Zhang *et al.*, 1999).

There are also some disadvantages to this technique: (i) it requires extremely pure, high molecular weight genomic DNA (ii) co-migration of non-allelic fragment of the same size can occur (Mueller and Wolfenbarger, 1999) and (iii) AFLP markers are dominant and therefore it is not well possible to distinguish between the homozygous state of the dominant allele and the heterozygote (Zhang *et al.*, 1999; Vos *et al*., 1995, Xu *et al*., 2000).

High – throughout AFLPs have been used as a powerful tool to perform detailed analysis on the genetic diversity within population of *vanilla spp*. (Dean *et al*., 1999; Bory *et al*., 2007b; Besse *et al*., 2004).

2.8.3.3 Microsatellites

Microsatellites, also known as short simple repeats (SSRs) are short segments of DNA that contain a core motif which consist of a few nucleotides that are normally 1-6bp in length. This core motif can be repeated up to a maximum of 60 times in tandem (Williams *et al*., 1990).

The regions are dispersed throughout the entire genome; both in protein-coding and non-coding regions although it has been shown that microsatellite are less abundant in coding than in non-coding regions. The number of tandem repeats at a specific locus can vary between two individuals and it is possible to screen for these differences using primer that anneal to the flanking regions of the specific microsatellite regions. The generated PCR products can be size-fractionated on a polyacrylamide gel, allowing the discrimination of alleles with a size difference of up to one bp. Because of their high mutability, microsatellites are thought to play a significant role in gene evolution by creating and maintaining genetic variation (Williams *et al*., 1990; DeBustos *et al*., 1998).

Microsatellite markers have several advantages that make them excellent molecular markers to use (i) low quantities of template genomic DNA are necessary (ii) it is usually highly polymorphic (iii) it is highly reproducible (iv) the possibility exist to score more than one microsatellite at the same time by perfoming a multiplex PCR and (V) it represent co-dominant marker. Unfortunately there are also few disadvantages:(i) an allele can be absent due to deletion or mutations in the flanking DNA that contains the primer binding site and (ii) the identification and development of these markers are expensive and time consuming (DeBustos *et al*., 1998).

Microsatellite markers have been used in genetic diversity study and DNAfingerprinting in *vanilla spp* especially in *Vanilla planifolia*, *Vanilla pompona* and *Vanilla tahitiensis* (Sun *et al*., 1998: Cibrian, 1999).

2.8.3.4 Restriction Fragment Length Polymorphism (RFLP)

Restriction Fragment Length Polymorphism was the first DNA-based marker to be developed (Botstein *et al*., 1980). The original method includes several steps: (i) digestion of genomic DNA with or more restriction endonucleases (ii) separation of digested fragments on high percentage agarose gels (iii) fixation of the digested DNA into a membrane (v) hybridization with fluorescently or radioactively labelled fragment of which the sequence correlates with a region of interest within a genome, and (v) the detection of the bands of interest. Difference within restriction endonuclease recognition sites between the genomes of two or more individuals can be scored as the presense or absence as well as size differences of the fragments which are homologous to probe used (Williams *et al*., 1990; DeBustos *et al*., 1998).

A variant of the hybridization-based method is polymerase chain reaction/RFLP (PCR/RFLP)-based method also known as cleaved amplified polymorphic sequence (CAPS). This method is much easier and faster and is based on the amplification of a specific locus in a genome, followed by the restriction endonuclease digestion of the amplicon using a specific restriction enzyme. Agarose or polyacrylamide gel electrophoresis can be used to reveal differences in a fragment size after staining the gel with ethidium bromide or silver nitrate (Prenner et al., 1999).

.

Similarly to this technique, the fragment results from mutations of nucleosites within the restriction enzyme recognition sequence. This method is generally less laborious than the original one and provides results faster. A negative aspect of the PCR/RFLP

method is that it requires the presence of mutation, which can be recognized by restriction endonuclease for at least one allele, and the availability of PCR primers for the amplification reaction. Unfortunately, many point mutations do not result in a restriction site or abolish it and the polymorphisms can therefore not be detected by this approach (Zhang *et al.*, 1999; Vos *et al*., 1995, Xu *et al*., 2000).

Both RFLP techniques have been successfully used in fingerprinting, phylogenetic studies and genetic diversity studies of *Vanilla planifolia* (Besse *et al*., 2004; Bory *et al*., 2007).

2.8.3.5 Randomly Amplified Polymorphic DNA (RAPD)

The randomly amplified polymorphic DNA techniques (Williams *et al*., 1990;Welsh and Mc Clelland, 1990) is a PCR – based method that involves the binding of a short single arbitrary primer of approximately 10bp to arbitrary sequences in the genome under relatively low stringency conditions with annealing temperature as low as 36ºC. A fragment is amplified when the primer binds on both ends of the genomic DNA strands, but in opposite orientation and with an amplifiable distance between the two primer binding sites (Smulders *et al*., 1997; Sun *et al*., 1998).

The variations in band patterns between different isolates result from sequence differences in the primer binding sites or sometimes from deletions and insertions within specific regions between the primers binding sites. These differences are visible as presence or absence of particular bands or difference in band sizes. The amplified products can be sized-fractionated on either agarose or polyacrylamide gels (Smulders *et al*., 1997).

This marker system has the advantage that: (i) low quantities of template genomic DNA is required (ii) prior knowledge of the genome sequence is not necessary, (iii) markers are usually highly abundant within genomes and can reveal a number of polymorphisms at the same time, (iv) it is easy and (v) cheap method (Dos Santos *et al.,* 1994).

Unfortunately this techniques also has some disadvantages:(i) it is very sensitive to reaction conditions (ii) it has low reproducibility (iii) the primer are able to amplify DNA fragments from every type of genome and it is therefore essential to avoid contamination with undesired genomic DNA (Smulders *et al*., 1997; Sun *et al*., 1998).

Despite this fact, RAPD methodology has provided informative data consistent with other marker, especially at the intraspecific and it is cost effective for large-scale population analysis level (Dos Santos *et al.,* 1994; Lerceteau *et al*., 1997).

Randomly Amplified Polymorphic DNA (RAPD) has extensively been used to study the genetic diversity of Vanilla *spp* particularly *Vanilla pompona*, *Vanilla* aphyla,

Vanilla tahitiensis and *Vanilla planifolia* in India, Mexico, France and Reunion Republic (Duval *et al*., 2006: Schuluter, 2002).

The criterion that was used to choose RAPD marker for the current research on genetic diversity study of Vanilla planifolia was based on the facilities and chemicals that was readily available at ARI-Mikocheni and IITA Laboratory in Dar es Salaam and the time frame of completion according to the Sokoine University of Agriculture (SUA).

CHAPTER THREE

3. 0. MATERIAL AND METHODS

3. 1. Sample Collection

Vanilla cultivars from Kibona, Bulinda, and Bakabuye which was used in this study was collected in Bukoba on 18th 20th December 2006 and planted at Sokoine University of Agriculture horticulture unit for morphological diversity study (Mategu.M, 2007).

The collection was done following a plant extraction mini preparation protocol elaborated by dellaporta (Dellaporta *et al.,* 1983). A clean, undamaged, young and

tender (2nd - 4th fully open leaf), *Vanilla* leaves were collected early morning from the horticulture unit at Sokoine University of Agriculture in Morogoro (22/06/2008), Kidia and Mudio in Kilimanjaro (03/07/2008), Mkunazini and Donge in Zanzibar (12/07/2008).

The Leaf samples were collected separately and each leaf was placed in a micro centrifuge tube and kept in a special cool box (-20ºC) then transported to ARI-Mikocheni laboratories in Dar-es-Salaam kept at -80ºC to stop cell activity which might cause deterioration of the DNA.

3. 2. Stock solution preparation

The Dellaporta extraction buffer (500 mls) was prepared by mixing 6.05g of 100mM Trisma base, 1.58g of 8.5mM EDTA, 14.61g of 500mM NaCl, 390.0µl of 10mM B-mercaptoethanol, and distilled water was added to fill 500mls volume in flask and the pH was adjusted at 8.0.

Likewise Lauryl Sulphate (SDS) working stock was prepared by dissolving 20.0g of lauryl sulphate in 100mls of distilled water and similarly the working stock solution for potassium acetate was made by dissolving 49.08g of 5M potassium acetate (CH_3CO_2K) in 100mls of distilled water then stirred gently to form a clear solution. Isopropanol (Propan-2-ol, was kept in -20°C in freezer) ready for being used whereas 70% ethyl alcohol was prepared by dissolving 140mls of 99.9% absolute ethanol in 60ml of distilled water to make 150mls of 70% ethyl alcohol.

3. 3. DNA Extraction

DNA extraction was done according to a modified Dellaporta plant DNA extraction protocol (Dellaporta *et al.*, 1983) where by 500µl of Dellaporta extraction buffer was measured and poured into each tube containing a leaf sample and the leaf was ground finely with the Kontes pestles. Thereafter 42µl of SDS was drawn using an automatic pipettor and added to each tube. The tubes were turned upside down to mix the content properly; and incubated at 65˚C 10 minutes. After 10 minutes the tubes were removed from the incubator then using an automatic pipettor 160µl of Potassium acetate was pipetted into each tube and mixed thoroughly and the tube was placed in freezer for 10minutes to accelerate crystallization.

The tubes were then removed from the freezer and instantly centrifuged at 13000 rotations per minute (rpm) for 10 minutes; then carefully 450µl of the supernatant was drawn and transferred into new tubes (a new pipette tip for each sample was used to avoid contamination). Then 450µl of ice cold Isopropanol was added into each tube with supernatant and mixed well by turning tubes top bottom (ratio of volume supernatant: Isopropanol = 1: 1).

The tubes were then again centrifuged at 13000rpm for 10 minutes followed by pouring out the supernatant carefully leaving behind the DNA pellets. The DNA pellet was washed with 500µl of 70% ethyl alcohol followed by centrifugation at full speed for 5minutes. After that the supernatant was removed and the DNA pellet was air dried then suspended in 200µl of sterile distilled water. Following this, 2µl of RNase A was added to each DNA tubes to digest all RNA. The tube containing DNA and RNase were incubated at 37°C for one hour as optimum temperature for

enzymes to digest the RNA and lastly the tubes containing DNA were stored at minus 80C° freezer.

3.4. DNA Quantification and Quality checking

Quantification of DNA was performed by electrophoresis method whereby extracted DNA along with a dilution series of a standard DNA (phage lambda) at λ 50, λ100 and λ150 with respective concentration of 265ng, 270ng and 275ng and after staining with ethidium bromide were run in an 1.2% agarose gel at 80Volts for 20 minutes and thereafter the gel was exposed to UV-light. The result as indicated (plate 1) showed that DNA concentration was (270ng) and this was obtained through estimation by comparison of the band's intensity.

DNA quality was checked at the same time by checking the band characteristics. High quality DNA should give a sharp, high molecular weight band, whereas sheared, smeared or DNase contaminated bands should not be used for further analysis since the fragmentation is arbitrary.

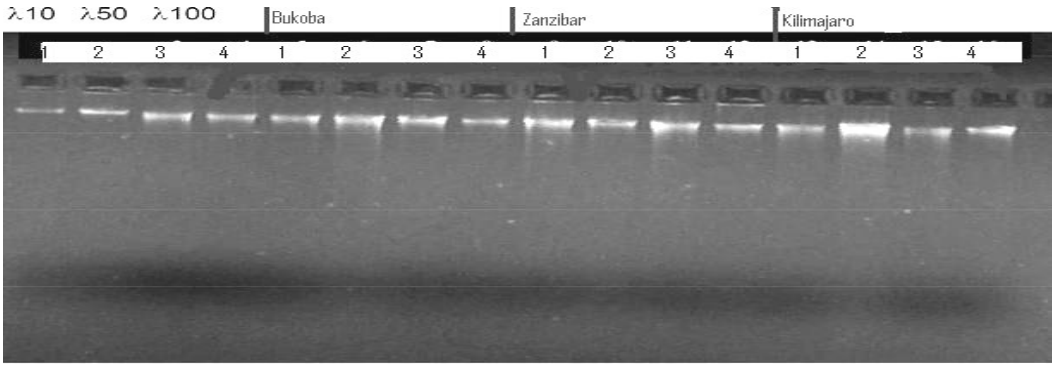


Plate 1.DNA amplification and quality checking.

3.5. Polymerase Chain reaction of Primers

.

RAPD amplifications were run following the procedure of (Welsh *et al*., 1990; Rohlf, 1998) with some modifications, using a Gene Amp PCR System 9700 (Applied Bio systems) thermo cycler. Ten decamer oligonucleotide arbitrary primers from the Invitrogen Kit (Invitrogen Technologies Inc.), selected from a preliminary screening, were used (Invitrogen INV-11, INV-23, INV-14, INV-15, INV-27, INV-33, INV-21, INV-16, INV-19 and INV-20). All samples were run in duplicate (with 5µl of DNA and 25 µl of PCR mix). The optimization of RAPD reaction was performed by varying the concentration of MgCl2 (2.0, 2.5 and 3.0 mM), DNA template (10 and 20 ng) and *Taq* DNA polymerase (0.5, and 1.0 U) while fixing the Concentration of RAPD primer and dNTPs at the concentration of 0.6 uM and 200 uM respectively. The PCR master mix was prepared as shown in Table 5.The DNA was amplified in Gene amp thermo-circlers at the conditions given below; Initial denaturing at 94°C for 4 minutes, annealing at 36°C for 40 seconds and extension at 72°C for 1 minute, for 1 cycle. Then, 93°C for 20 seconds, 36°C for 40 seconds min and 72°C for 1 minute for 45 cycles and finally the last step was 94°C for 4 minutes, 36°C for 40 seconds and 72°for 4 minutes for 1 cycle. Then the samples kept cool at 4°C.

The amplified DNA products were then electrophoresed in a 1.0% prior prepared agarose gel stained with ethidium bromide and run at 100 volts for 40 minutes in x 0.5 Tris-Acetate-Edta (TAE) buffer at pH 8 (Soto *et al.*,2003).The gel was then visualized under UV- light and photographed and documented using Syngene-InGenius Imaging system.

Table 5. Reagents at known concentration and volume that were used as PCR master mix for amplification process

	Conc. Stock so	Volume	Final Conc
PCR-Buffer + $Mgcl_2$	10 ×(15 mM)	3 µl	1×(1.5mM)
Primer	15µM	0.4 µl	0.2µM
dNTPs-Mix	10mM	0.51 µl	0.17mM
Taq-polymerase	5U/µl	0.25 µl	1.25 U
Distilled water		28.84 µl	
Tween 20	5%	2.00 µl	
Final Volume		25 µl	

3.6. Data Collection

Data of the RAPD assay were collected through visualization, marking and counting of the amplified bands. RAPD assay represent a consensus of at least two replicates per plant. Only reproducible well-marked amplified bands were scored.

Faint amplified products and unreproducible fragments of more than 3 kb were not considered for analysis. Bands with the same mobility were considered as identical fragments, irrespective of band intensity. The fragment data were entered in a computer file as binary matrices, where 0 and 1 represented absence and presence of a band, respectively.

3.7 Data Analysis

Analysis of molecular variance (AMOVA) and statistically tested with F-Test

< 0.001 at 95% level of confidence. After finishing scoring all of the profiles for all plants using the 10 primers that were used during this research work then genetic analysis began by doing a pair-wise comparison.

The comparison was made between each plant and all others in the same locality (intrapopulation) and then between each plant and all others in different localities (interpopulation). Only the shared presence of a band was counted, and not the shared absence of a band. Next, comparison was made between to plant one and two, three, and so on, until comparison was made between each plant and all the others. A table was constructed follong this comparison (Table 6).

This shows how all of the plants across the top and the plants down the column, which helped to keep track of which plants had already been compared (Williams *et al*., 1990). The polymorphism information content (PIC) for each marker was determined separately for the three location of cultivar sampling using the following equation:

$$PIC=1-\sum_{i=1}^{n}p_i^2-2\left[\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}p_i^2p_j^2\right]$$

Where p_i is the frequency of the i^{th} polymorphic band per cultivar, and n is the number of samples (Botstein *et al.*, 1980).

Table 6. Pair-wise Comparisons Table

Band	1	2	3	4	5	6	7	8	9	10
1	7	7	7	7	7	7	7	7	6	7
2	7	7	7	7	7	7	7	7	7	7
3	7	7	7	7	7	7	7	7	7	7
4	7	7	7	7	7	7	7	7	7	7
5	7	7	7	7	7	7	7	7	7	7
6	7	7	7	7	7	7	7	7	7	7
7	7	7	7	7	7	7	7	7	7	7
8	7	7	7	7	7	7	7	7	7	7
9	6	7	7	7	7	7	7	7	7	7
10	6	6	6	6	6	6	6	6	7	6

A similarity coefficient was calculated using the simple matching (SM) index, with SMxy = [n11 + n00]/[n11 + n10 + n01 + n00] (Welsh and Mc Cleland ,1990;Soto *et al*.,2003).

Where n_{11} is the number of bands shared by individuals *x* and *y*, n_{10} the number of bands present in *x* and absent in *y*, n_{01} the number of bands present in *y* and absent in *x*, and n_{00} the number of bands absent both in *x* and *y*.

Genetic distance was calculated using the NTSYs pc software (Rohlf, 1998) using the below equation;

$$D = 1 - \frac{N_{xy}}{(N_x + N_y - N_{xy})}$$

where:

D_{xy} = the genetic distance between plant "x" and plant "y"

N_{xy} = the number of bands shared by plant "x" and plant "y"

N_x = the number of bands in plant "x"

N_y = the number of bands in plant "y"

CHAPTER FOUR

4.0 RESULTS

4.1 RAPD bands

The RAPD bands patterns generated for DNA from Bukoba and Kilimanjaro regions and from Zanzibar Island were photographed and recorded as shown (Plate2-19). Generally the banding patterns for all samples did not significantly differ from each other. Banding patterns using primer Inv 27, Inv 33, Inv 15 and Inv 16 while primer Inv 14 and Inv 20 appeared to correspond to other and on the other hand primer Inv 19, Inv 21 and Inv 23 had banding pattern a little bit close to each other while prime Inv 11 was different to all others.

Primer Inv 11(Plate No.15) showed the least number of band patterns with only one band patterns while primer Inv 27 and primer Inv 15(Plate 2 and plate 16) showed the highest band patterns with 7 band patterns.

The average number of band patterns produced by vanilla cultivar DNA was four banding patterns in most of the primers used, also most of banding pattern occurred between 600 – 24 50 base pairs (Plates 2-19).

4.2 Genetic distance among and within population

The genetic distance of samples taken from Kagera region (Bakabuye, Bulinda, and kibona), Zanzibar Island (Donge, Mkunazini) and Kilimanjaro region (Mudio and kidia) that was calculated by the NTSYSpc software using the Unweighted pair

group method with arithmetic mean (UPGMA) and then analysed using simple matching matrix were found to be be 0.667, 0.705 and 0.805 respectively. The overall (interpopulation) genetic distance for all the samples (Bukoba, Kilimanjaro, Zanzibar) was 1.80. The obtained distances were then used to calculate the similarity coefficient/similarity percentage which was found to be 98.20 which represent a great uniformity of the sample.

Analysis of molecular variance (AMOVA) was perfomed and the results were tested using F-test < 0.001(Appendix 1) at 95% level of confidence which showed no statistically significance difference in intra and inter genetic diversity of all the samples used.

A dendrogram was constructed following cluster analysis using NTSYS software from which the Dendrogram graph presented in (Fig.1a, 1b, and 1c). The grouping – association identified by classification was rather weak as the 126 samples analysed failed to cluster in separate distinct groups. The resulting Dendrogram was then compared with the Eigen values (Appendix 5) based on simple matching matrix. The was no significant difference n the general grouping of the main clusters using Dendrogram or relative distances within the clusters comparison using the Eigen values.

4.3 Polymorphic products and polymorphic bands percentage

The random primers used in this study was from Invitrogen Technology Inc. (UK) produced different banding patterns which was obtained from different primers used

(Table 7). All of the primers used were distinguishable by their band patterns produced (plate 2 -19). Both RAPD primers generated reliable and reproducible polymorphic patterns.

The total number of bands amplified for all primers was 53 (Table 7), of all the 126 samples of vanilla cultivar used the lowest number of polymorphic band per primer recorded was 1 and this was produced by primers Inv -11, Inv -14 and Inv -16 while the highest number recorded for polymorphic band per primer was 5 and this was produced by primer Inv-21(Table 21).Also the polymorphic information content per primer(PIC) was highest in primer Inv-21(0.57) and lowest in Inv-27 meanwhile the average PIC values for all the 10 primer used in this study was 0.31(Table 7).

The polymorphic bands percentage is illustrated by a histogram (Fig 3) where by the primer with highest polymorphic bands percentage was Inv-21 (83.34%) and the primer with the lowest polymorphic band percentage was primer Inv-11 with only (14.28%). Band frequency of the entire collection was depicted in a frequency graph by comparing with the standard graph given by the NTSYS software and the graph shows that the frequency did not deviate from the standard therefore the frequency was good (fig2).

The principal coordinate analysis (PCA) was generated using similarity matrix (fig 4a and 4b) shows a two and three dimensional presentation of the relationship and distribution of vanilla cultivar sample in correspondence to the location where sampling took place.

The PCA revealed that several samples collected in the same locality are more closely related genetically than those from different localities but the closeness that was revealed was not significantly at F < 0.001and 95%level of confidence. The numbers 1-126 represent the total number of sample that was sampled in given localities. Euglian distance value (appendix 5) also shows the percentage and cumulative percentage value of genetic distance of vanilla with respect to each vanilla cultivar from which sample was taken.

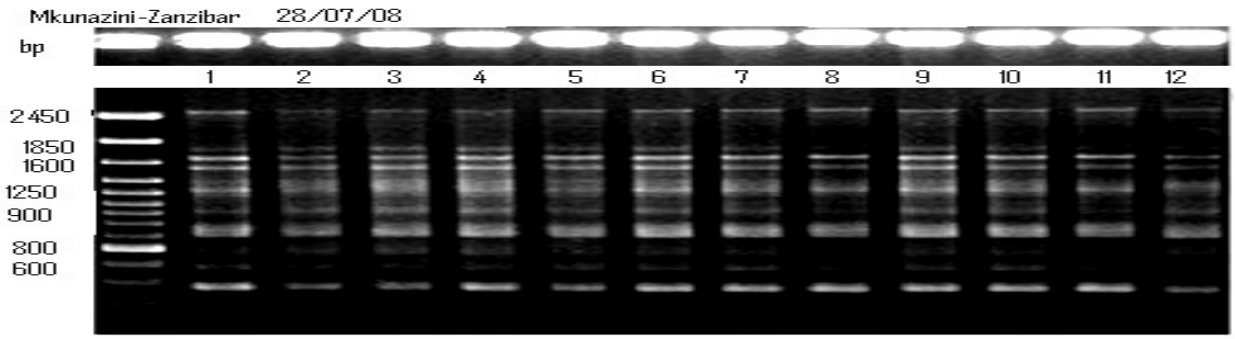
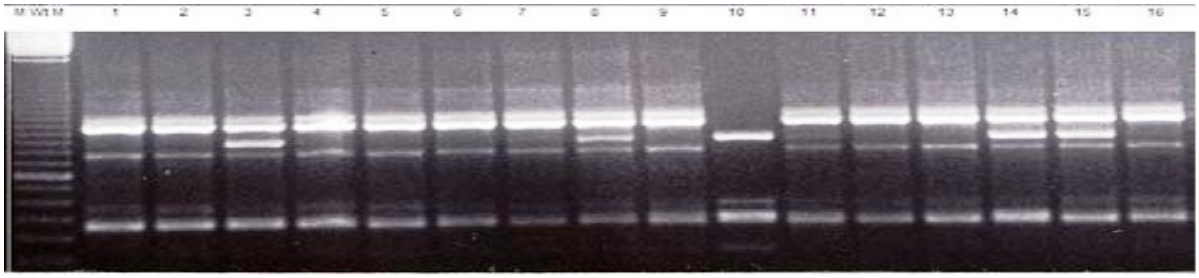


Plate 2. RAPD PCR product 1 primer Inv.27.



.

Plate 3.RAPD PCR product 2 using primer Inv 27

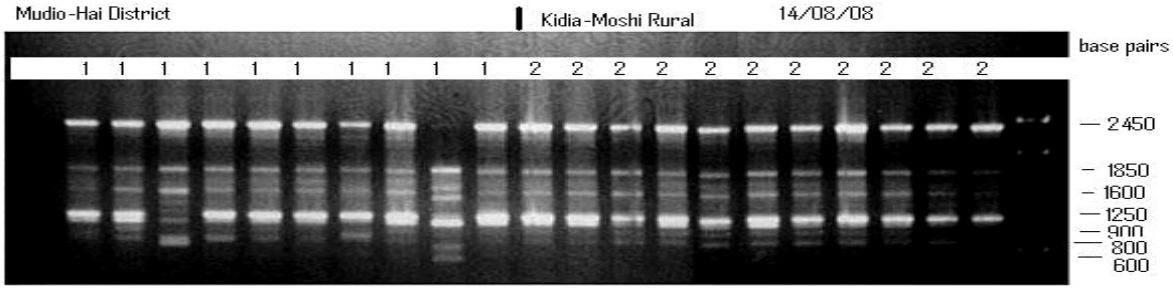


Plate 4. RAPD PCR product 3 using prime Inv 33.

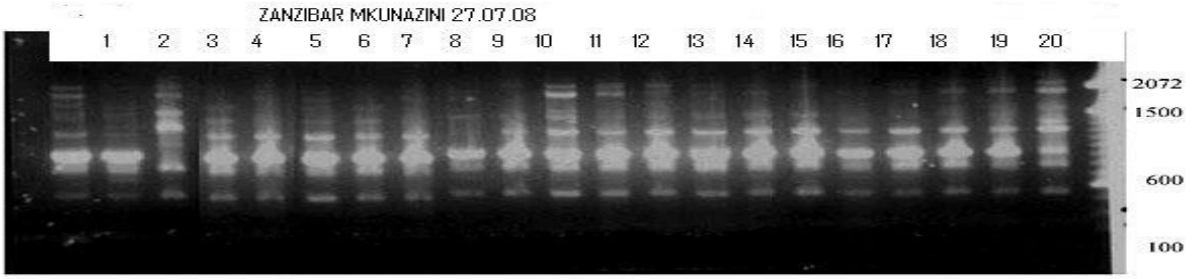


Plate 5.RAPD PCR product 4 using primer Inv 33.

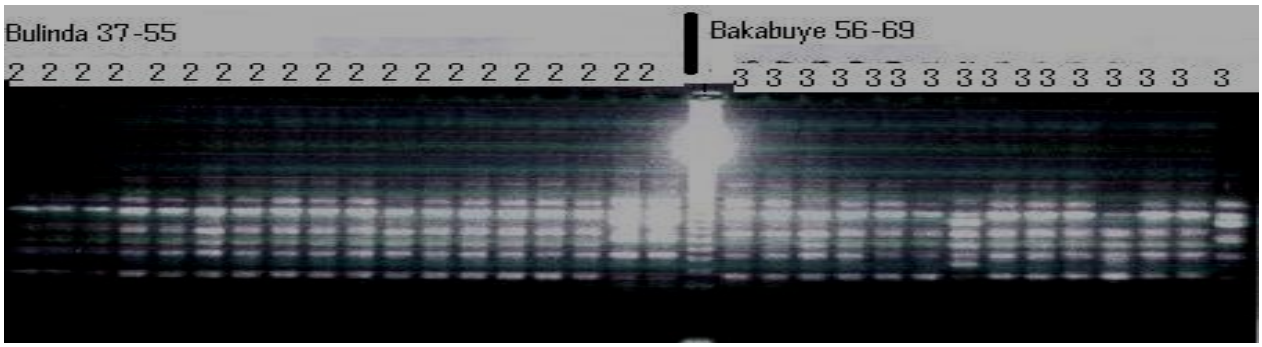


Plate 6.RAPD PCR product 4 using primer Inv 14.

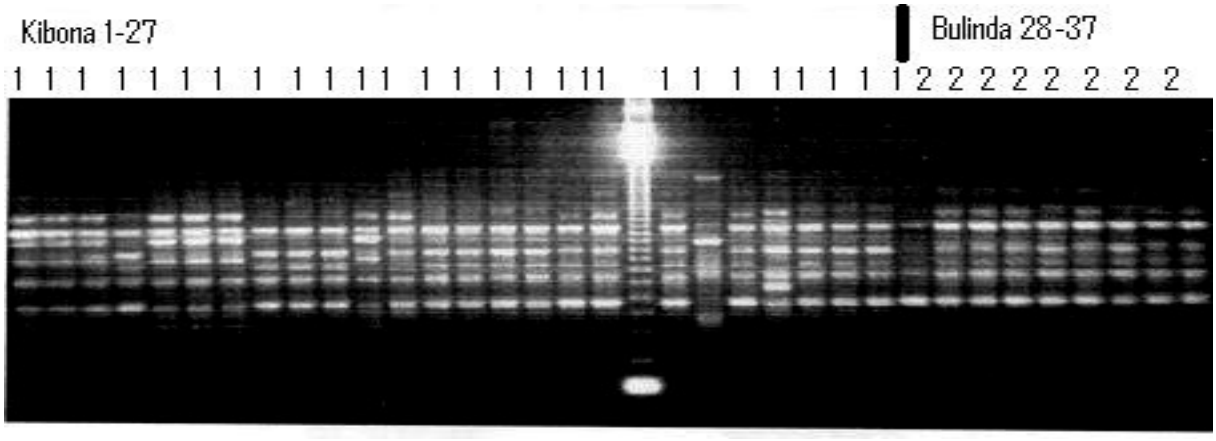


Plate 7. RAPD PCR product 5using primer Inv 14.

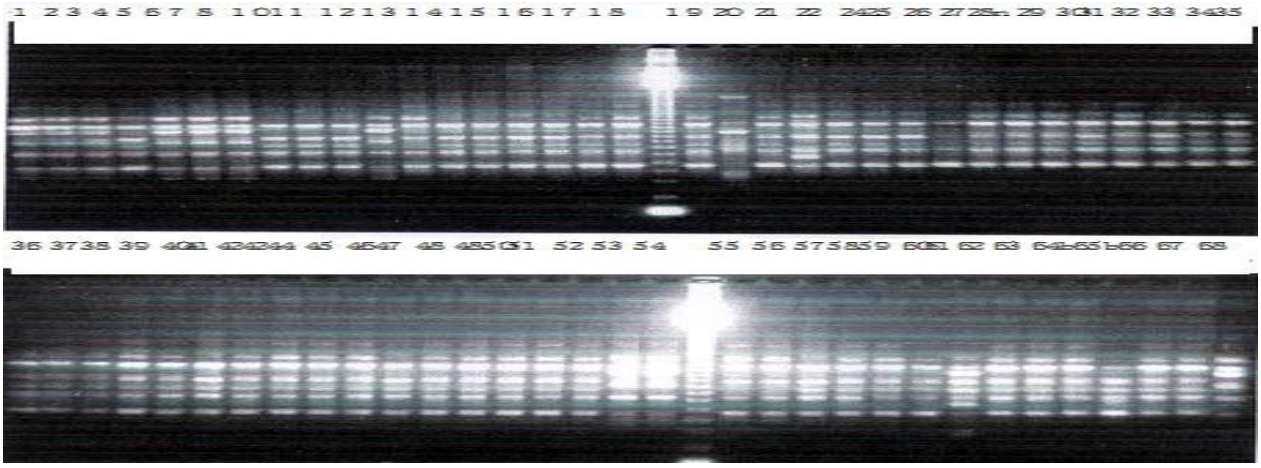


Plate 8. RAPD PCR product 6 and 7 using primer Inv 21,

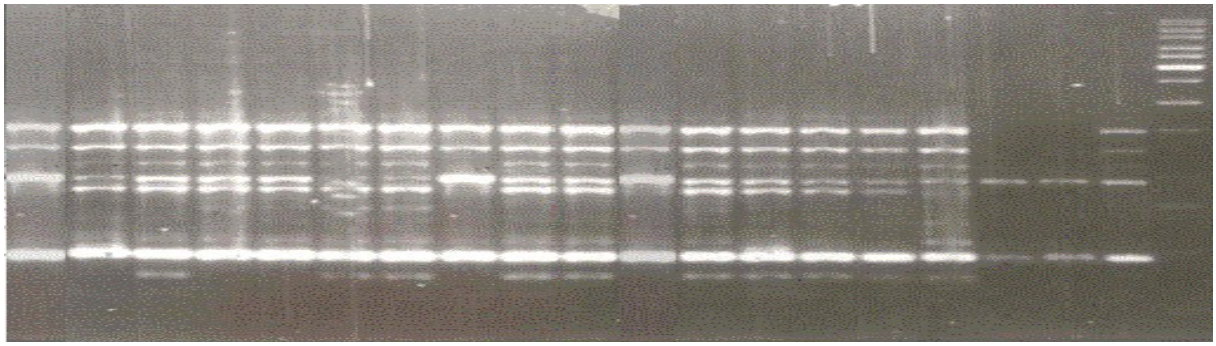
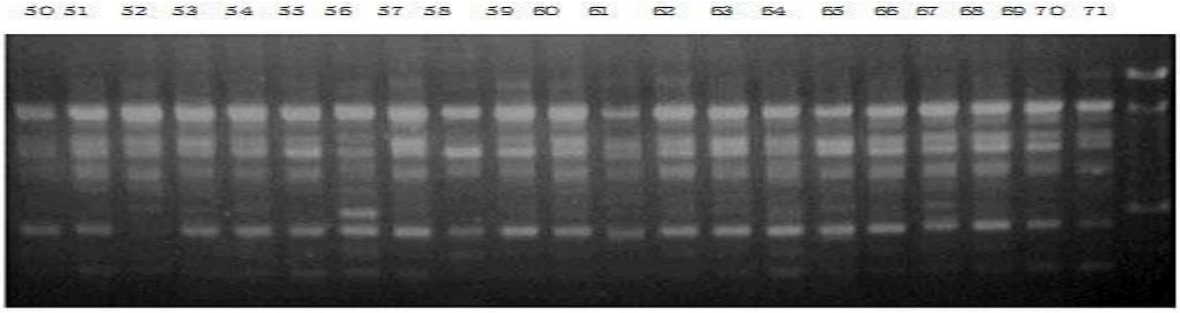


Plate 9.RAPD PCR product 8 using primer Inv 23.



.Plate 10.RAPD PCR product 9 using primer Inv 19.

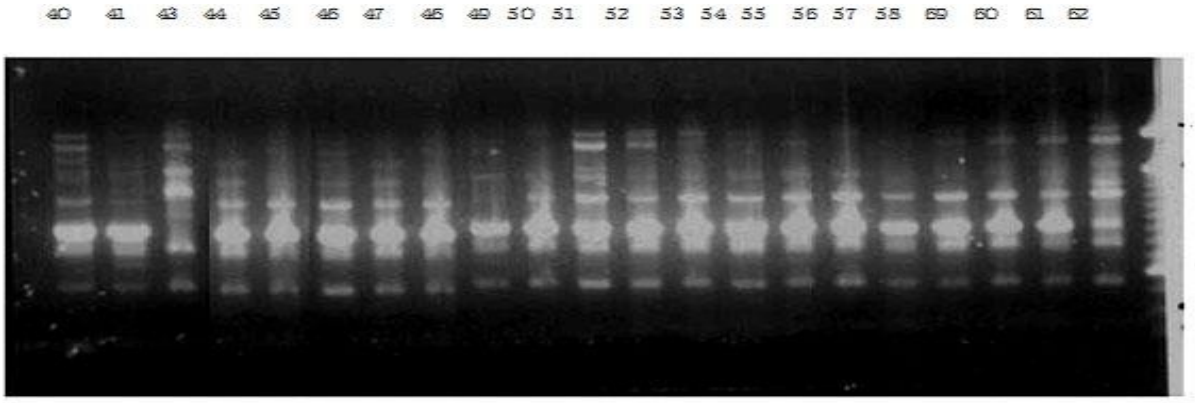


Plate 11.RAPD PCR product 10 using primer Inv 19.

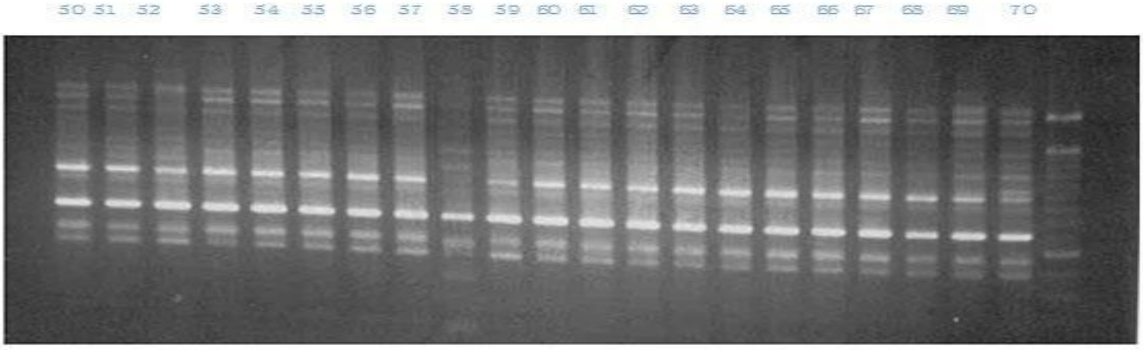


Plate 12.RAPD PCR product 11using primer Inv-21

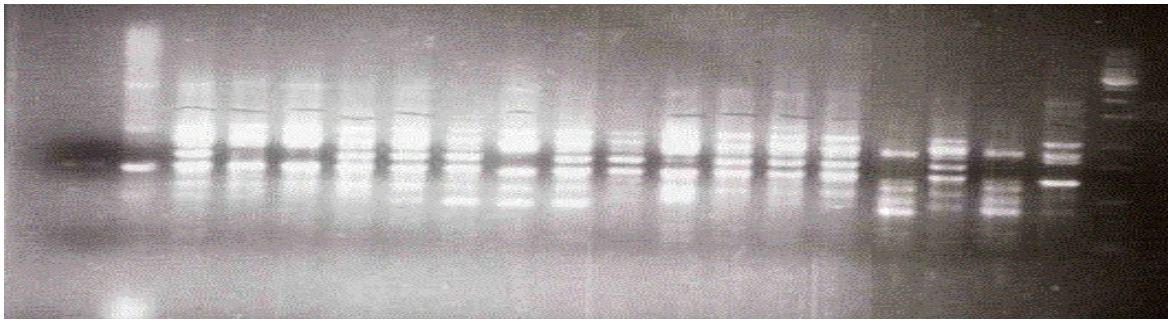


Plate13.RAPD PCR product 12 using primer Inv 20.

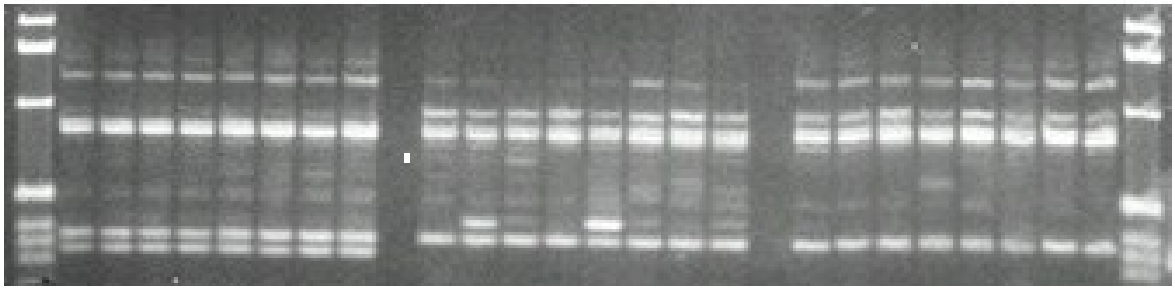


Plate 14.RAPD PCR product 13 using primer Inv 20.
.

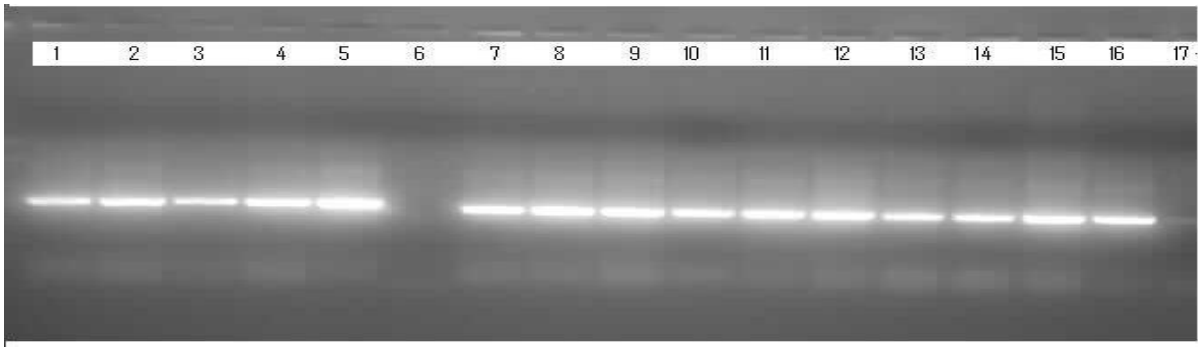


Plate 15 RAPD.PCR product 14.using primer Inv 11.

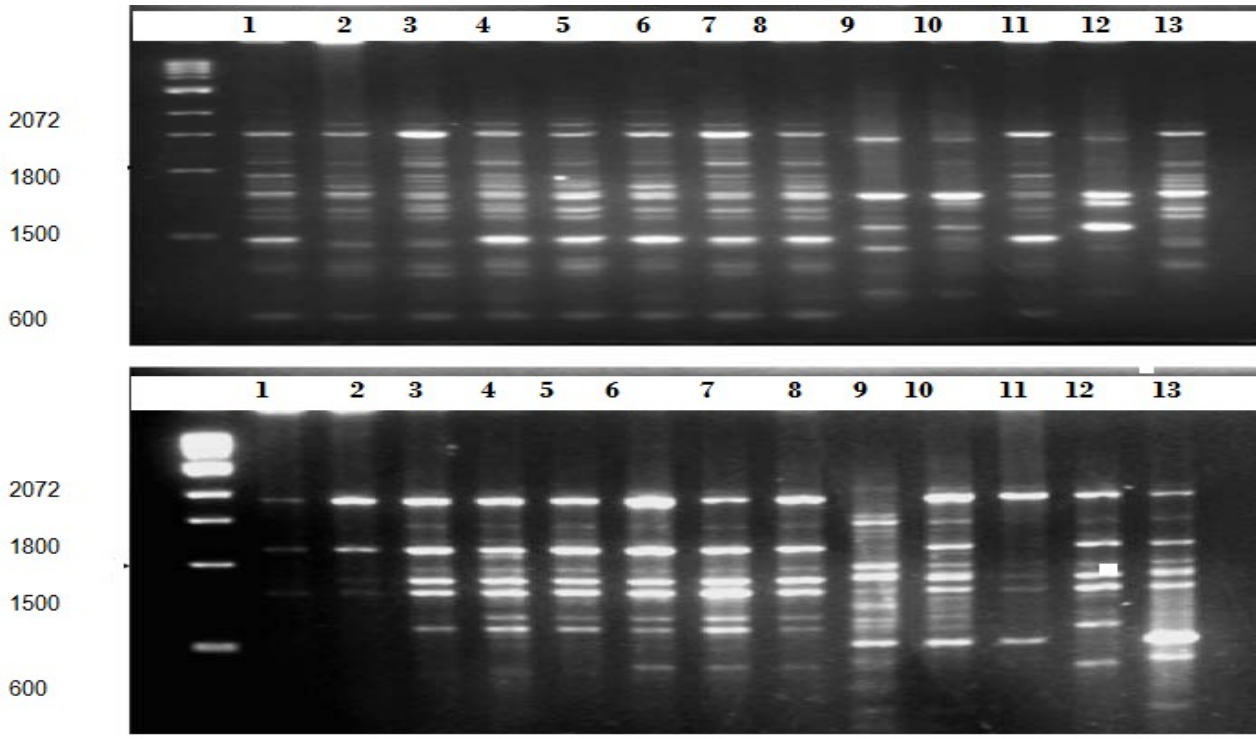


Plate 16.RAPD PCR product 16 and 17 using primer Inv 15.

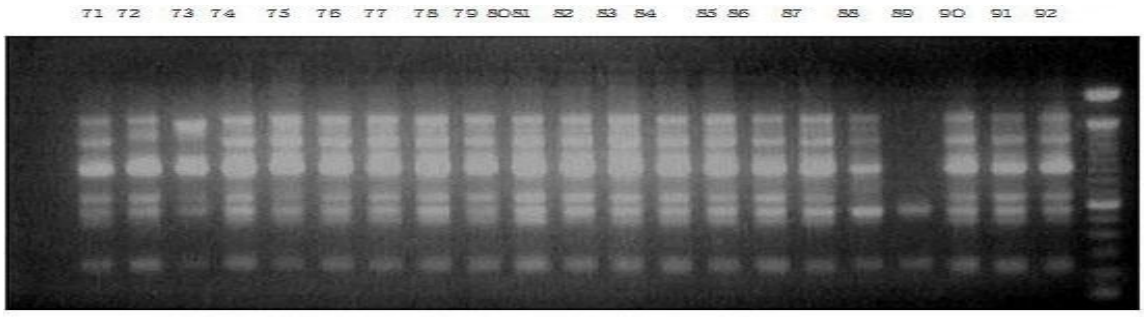


Plate 17.RAPD PCR product 16 and 17 using primer Inv 16

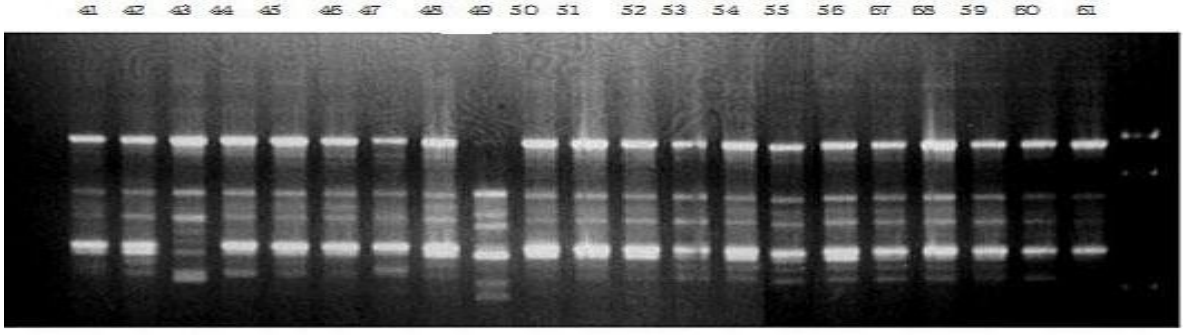


Plate 18.RAPD PCR product 19. Using primer Inv 20.

Plate 18.RAPD PCR product 16 and 17 using primer Inv 16

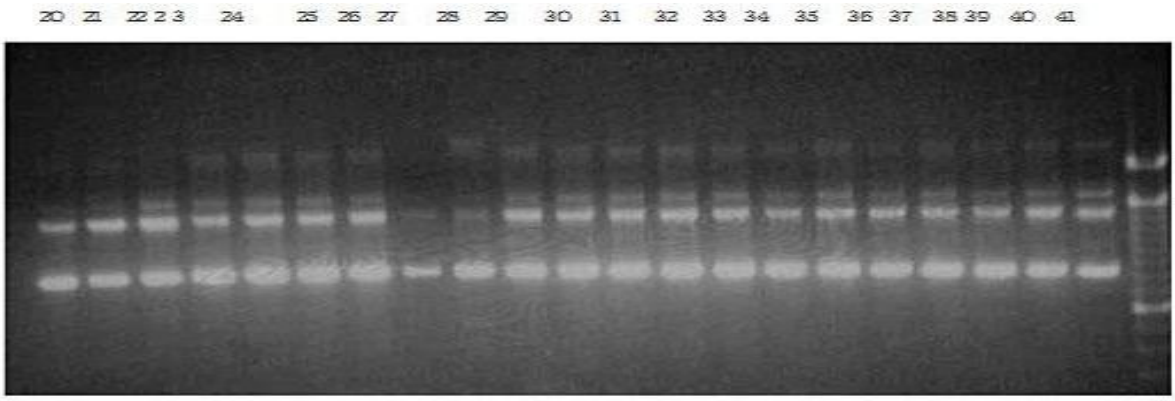


Plate 19.RAPD PCR product 20 using primer Inv 23.

Table 7. Random decamer primers used in present study; their sequence, number of

Polymorphic products and percent of polymorphic bands produced by each primer.
Primer Sequence

Primer	Sequence (5' --3'	Total number of bands	Number of polymorphic bands	PIC

		amplified		
Inv -11	CAGCACCCAC	7	1	0.13
Inv -23	GTTGCGATCC	6	3	0.34
Inv -14	CCGCATCTAC	6	1	0.12
Inv -15	TGGACCGGTG	7	4	0.52
Inv -27	TGTCTGGGTG	2	1	0.04
Inv -33	GGTGCGGGAA	5	4	0.48
Inv -21	TTATCGCCCC	6	5	0.57
Inv - 16	GGCGGCTGA	3	1	0.06
Inv -19	TGCGGCTGA	7	4	0.53
Inv -20	GAGGATCCCT	4	3	0.23
Total		53	27	0.31

.

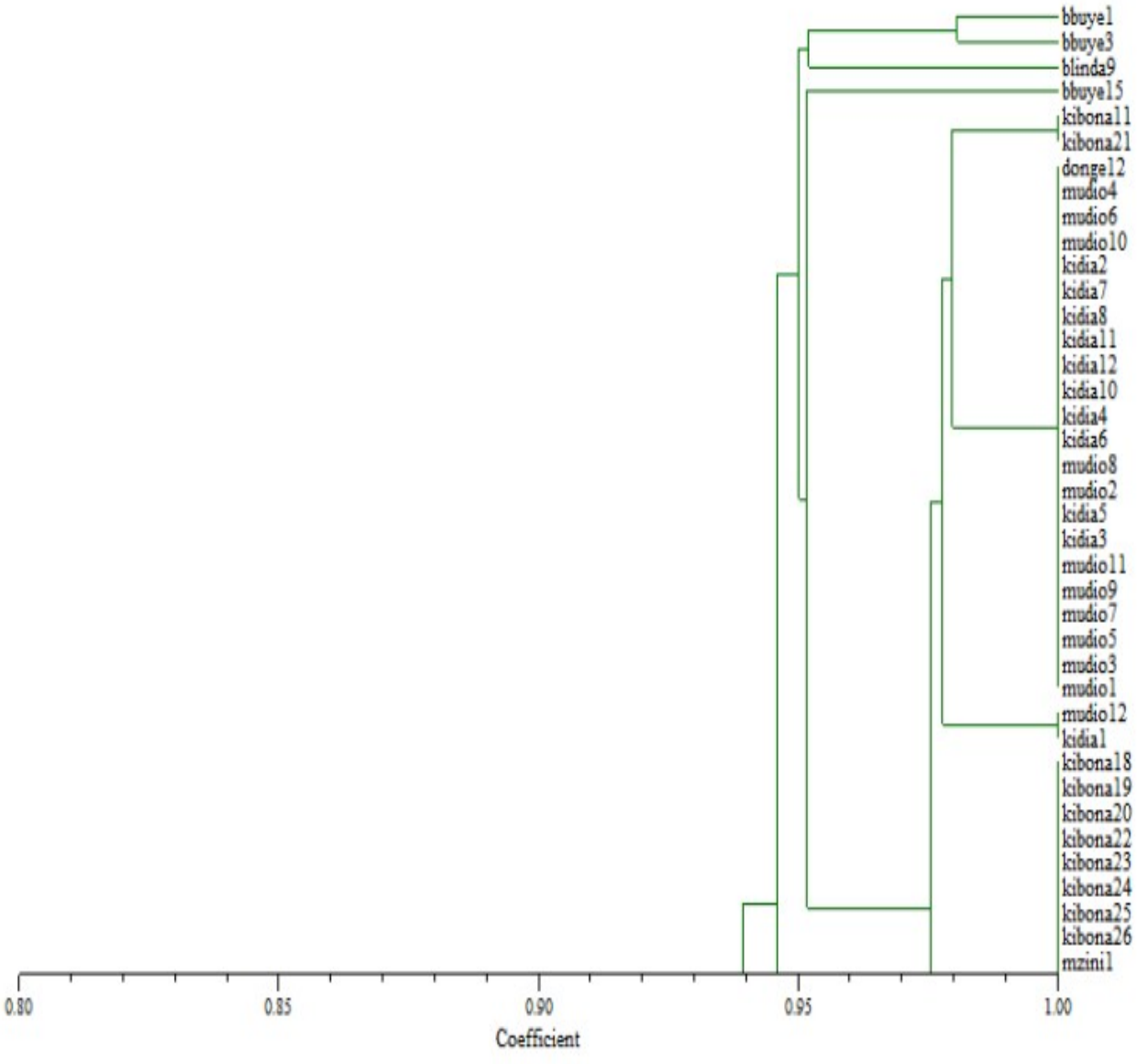
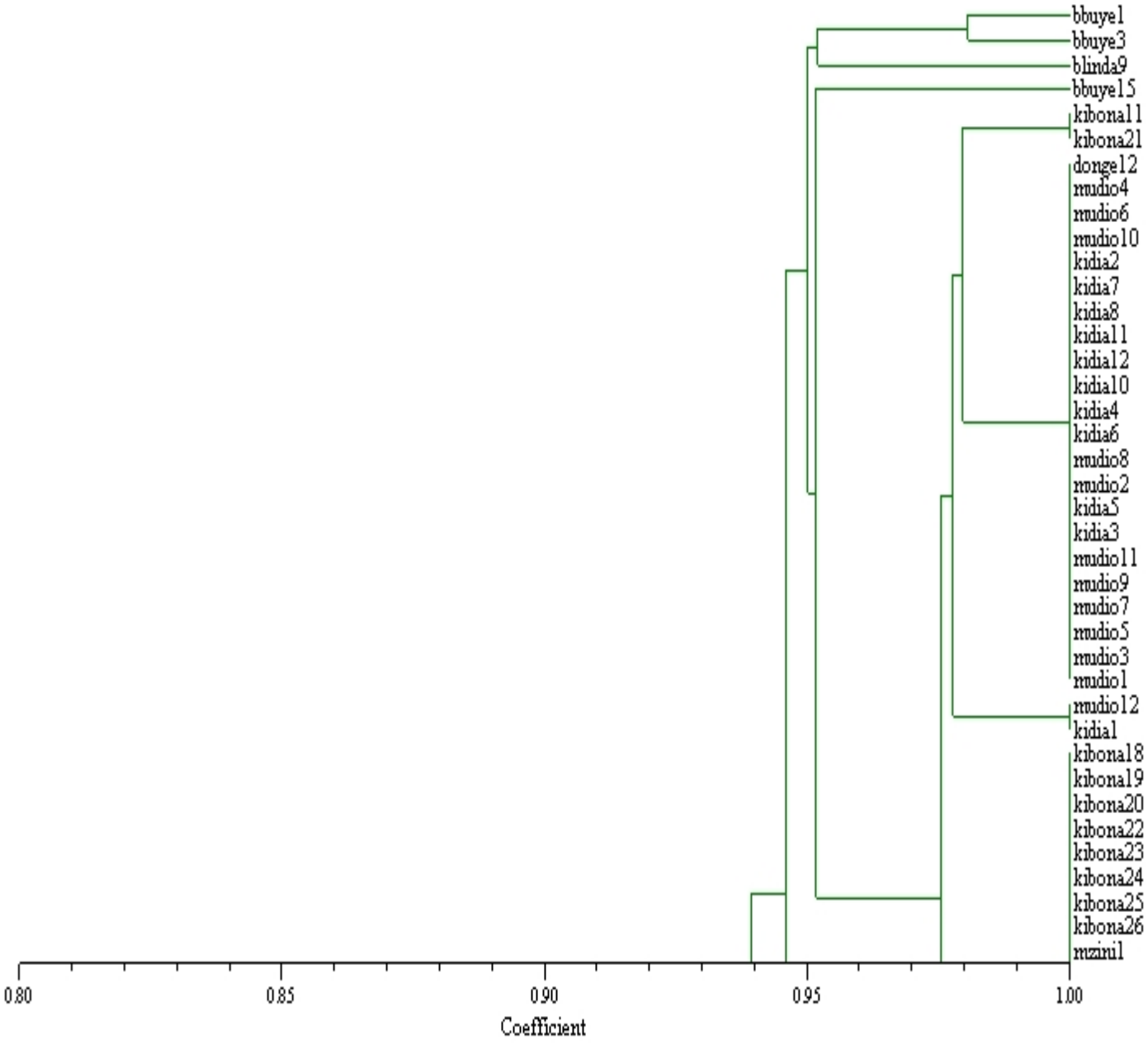


Fig.1a Dendrogram showing phylogenetic relationship among 126 Vanilla cultivars from Bukoba, Kilimanjaro and Zanzibar based on 10 RAPD primers using UPGMA



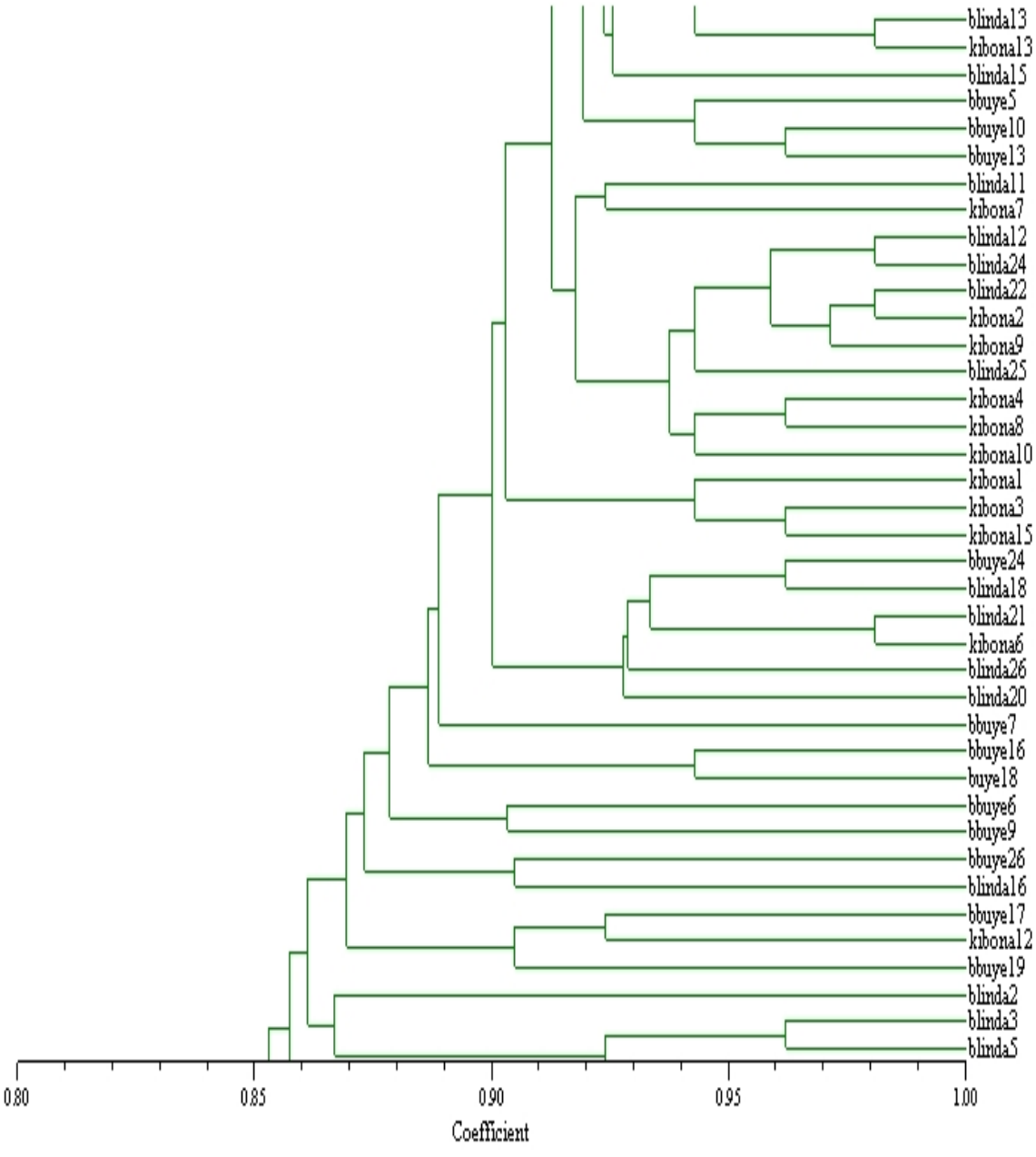

blinda13
kibona13
blinda15
bbuye5
bbuye10
bbuye13
blinda11
kibona7
blinda12
blinda24
blinda22
kibona2
kibona9
blinda25
kibona4
kibona8
kibona10
kibona1
kibona3
kibona15
bbuye24
blinda18
blinda21
kibona6
blinda26
blinda20
bbuye7
bbuye16
buye18
bbuye6
bbuye9
bbuye26
blinda16
bbuye17
kibona12
bbuye19
blinda2
blinda3
blinda5
0.80
0.85
0.90
0.95
1.00
Coefficient

Fig 1c Dendogram showing phylogenetic relationship among 126 Vanilla cultivars

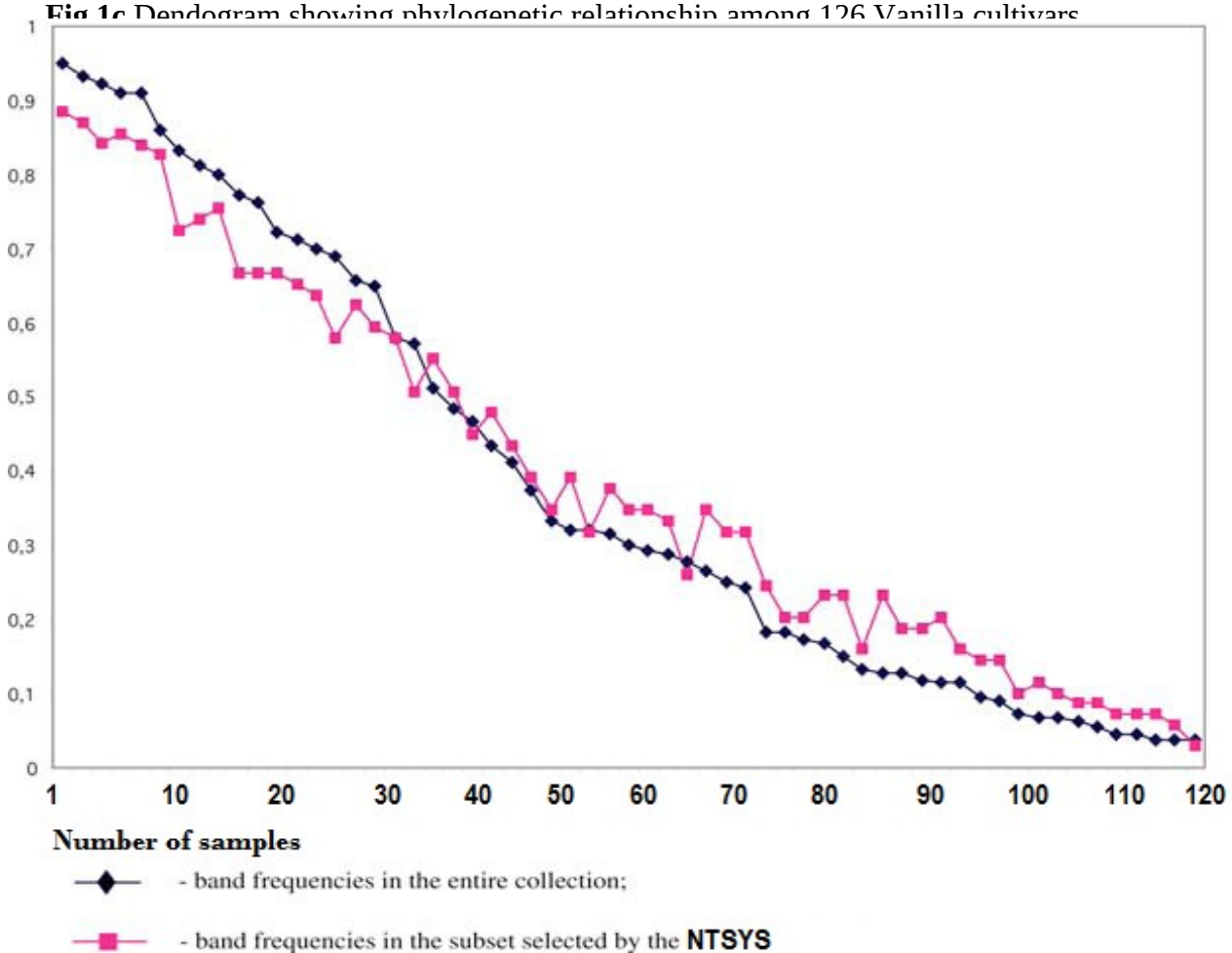


Fig.2.RAPD band frequencies in the entire collection and in the subset obtained by using the NTSYSpc program

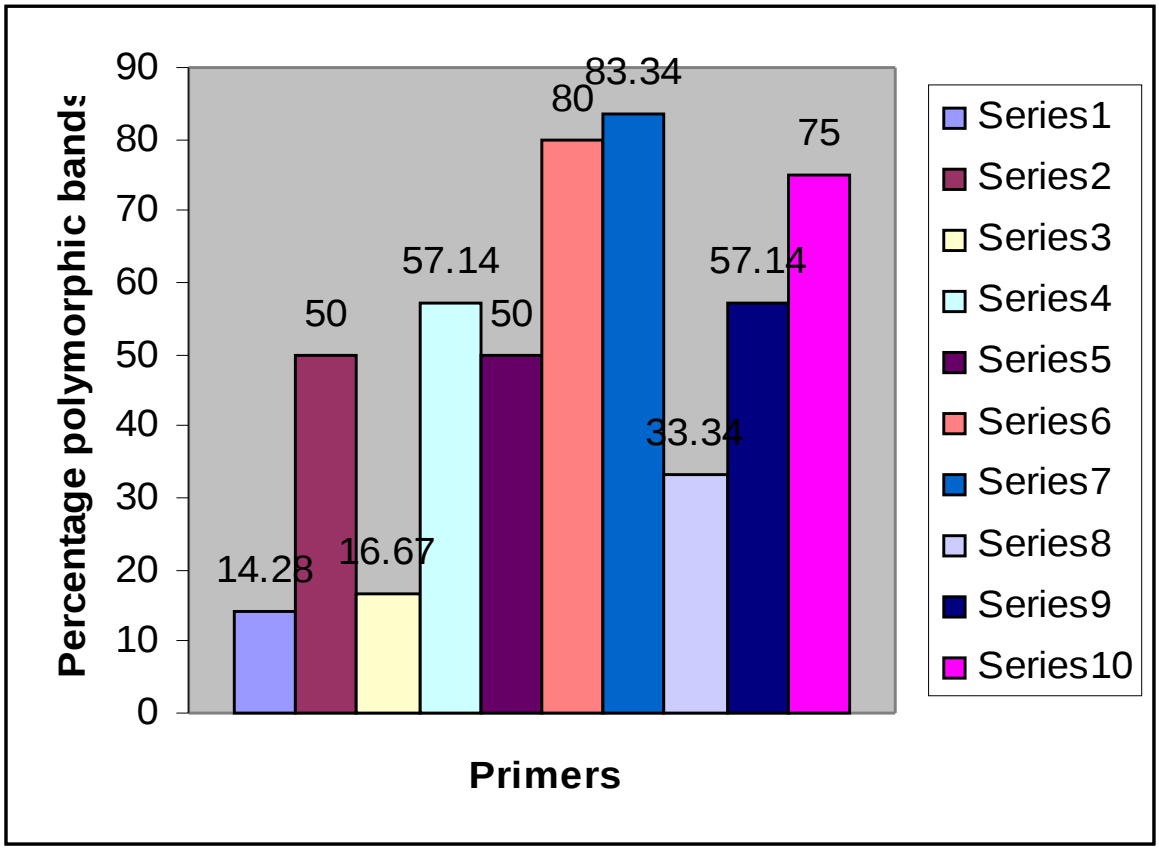


Fig 3. Histogram depicting primers used based on percentage of polymorphic bands of the RAPD data

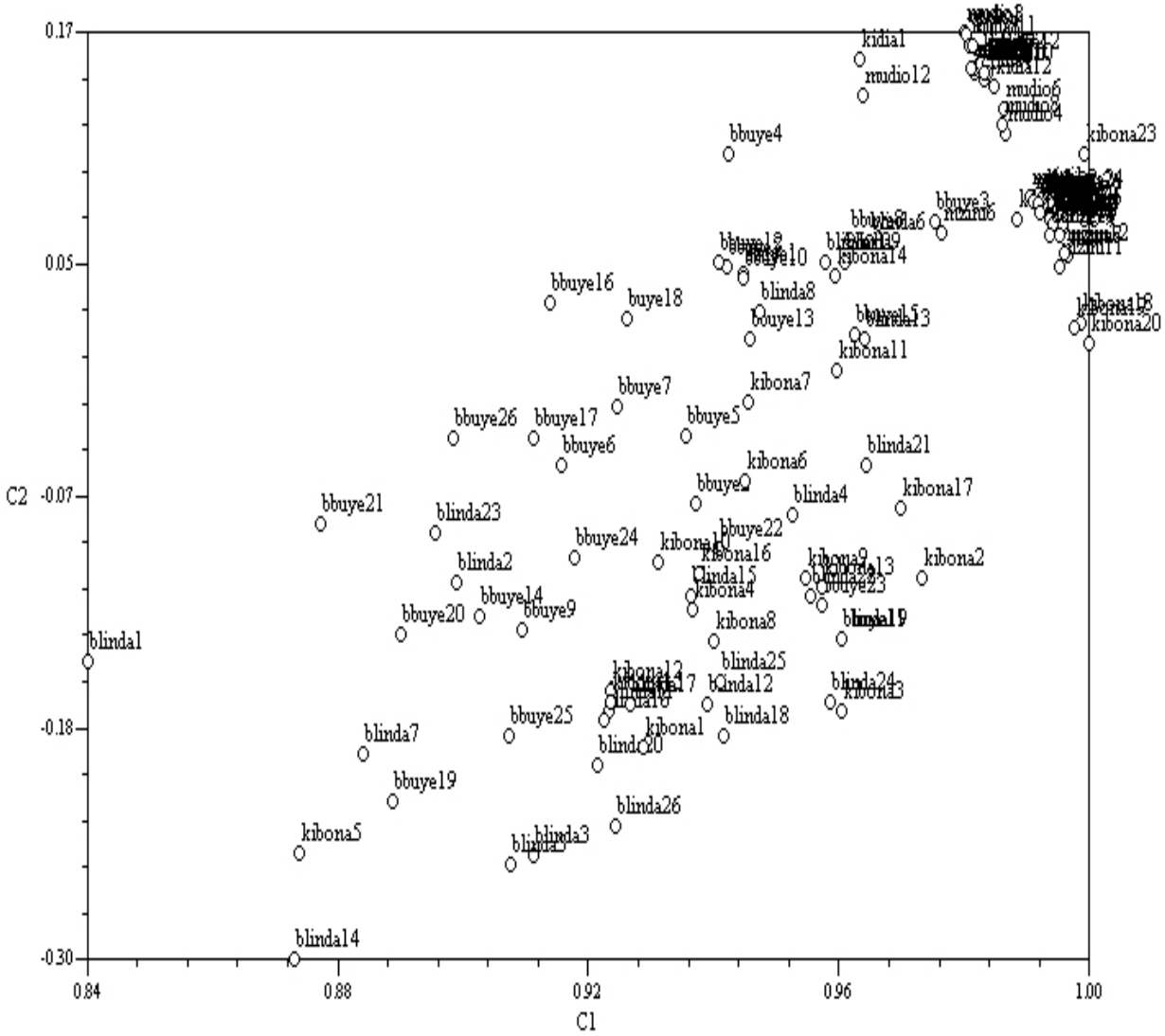


Fig 4a A 2-dimension scatter plot of the 126 vanilla genotypes based on the scores of the first three principal components using NTSYS software package.

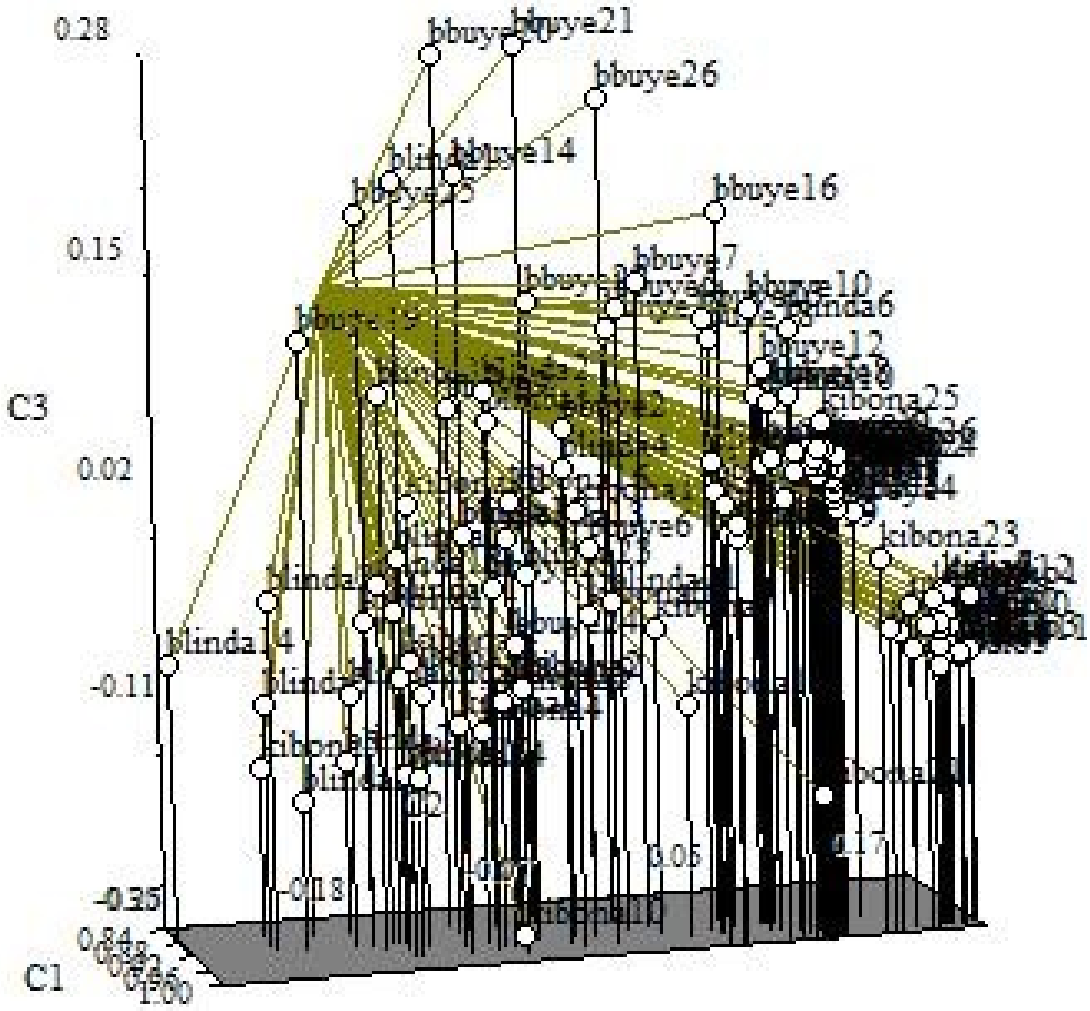


Fig 4a A 3-dimension scatter plot of the 126 vanilla genotypes based on the scores of the first three principal components using NTSYS software package.

CHAPTER FIVE

5.0 DISCUSSION

5.1 Determinations of the Extent of Intra Population Genetic Diversity

The first objective of this current research was on determination of intrapopulation genetic diversity of vanilla cultivars. The data and result so far has revealed that the genetic distance in the sample taken at the same locality (divided according to region where sampling was done) and analysed using 10 RAPD primers was determine and calculated by simple matching matrix and NTSYS programme model found a genetic distance of 0.667, 0.705 and 0.805 for Kagera, Kilimanjaro and Zanzibar respectively. Analysis of molecular variance (AMOVA) tested the result at $F< 0.001$ and 95% level of confidence and it showed the samples to be significantly similar to each other. This suggests that the samples taken from Kagera region were 99.34% similar to each other genetically, samples from Zanzibar were 99.3% similar to each other and also samples from Kilimanjaro were 99.2% to each other genetically.

This result indicates that there is a narrow genetic similar diversity among the cultivars grown in the same localities. The relatively low polymorphism and the lack of evident diversity observed among the vanilla cultivars grown in Kagera region (Bakabuye, Kibona and Bulinda), Kilimanjaro region (Mudio and Kidia) and Zanzibar (Donge and Mkunazini) could be related to the mode of introduction, propagation and maintenance of germplasm.

5.2 Determinations of the Extent of Inter Population Genetic Diversity

Similarly, one of the objectives of this research study was on determination of interpopulation genetic diversity of vanilla crop, Likewise the data and results which has obtained using 10 RAPD primers during this research work using NTSYS programme model found a similarity coefficient of 98.20 hence a genetic distance of 1.28 which indicates that vanilla crops grown in Kagera, Bukoba and Zanzibar are 98.20% similar to each other. . Analysis of molecular variance (AMOVA) tested the result at F< 0.001 and 95% level of confidence and it showed the samples to be significantly similar to each other This further indicates that there is a significant narrow genetic base of vanilla crop grown in Tanzania and this also could reflect on the mode of introduction, propagation methods and maintenance of germplasm.

Moreover, analysis of the 10 selected primers among the 126 cultivars included in this study generated 53 bands, 27 of which were polymorphic. There were 0.23 polymorphic bands per primer on average. Examples of polymorphism are shown in (Table.7) and the result showed almost all of the 126 cultivar sampled were distinguishable by similar band patterns. The PIC result also revealed the 3 primers as RAPD primers with highest PIC values that ranged from 0.52 to 0.57 with an average of 0.54 (Table7). This implies that a minimum of two highly polymorphic RAPD primers could detect polymorphism among *Vanilla planifolia* cultivars. The DNA band frequencies and Histogram of RAPD polymorphic bands percentage

(Fig2 and Fig3) also explains both DNA extractions using Dellaporta extraction protocol proved to be very useful and practical methods for the vanilla crop. The Principal coordinate analysis (fig 4a and fig 4b) also support and explain the same scenario of closeness among all the samples that were taken, hence the grouping is not scattered.

It was well noted that most of the cultivars from Zanzibar Island or Kilimanjaro region were associated with accessions already grown in Kagera region. The fact that the vanilla cultivars from Kilimanjaro and Kagera did not markedly diverge from the genetic diversity present in Zanzibar Island suggests a narrow genetic diversity of populations from which the present cultivars have been derived from common source parents and maintained over several decades. Exchange of cultivars between plantations and seedling selection may also have played a role.

5.3 Identification of Potential Parental Genotypes.

The potential parental genotypes identification which would have to be selected and used for breeding programmes as one as parents in order to permit exchange of traits which are useful such as pest and disease resistance, drought resistance and higher yield was not achieved due to the fact that this research produced such low level of polymorphism and closeness of the cultivars of vanilla cultivated in the country. This information obtained during this research that no breeding or crop improvement of vanilla crop can be done unless artificial means such as artificial mutation, genetic engineering or interspecific hybridization is applied.

These current results on genetic diversity study of *Vanilla planifolia* (Orchidaceae) in Tanzania corresponds with numerous and deliberate efforts done by other researchers worldwide to characterize vanilla biodiversity using various molecular technique such as Simple sequence repeat markers (SSR) Random amplified polymorphic DNA (RAPD) markers amplified fragment length polymorphism (AFLP) markers and Iso-enzymatic found that there is low level of genetic diversity (Cibrian, 1999; Nielsen and Sigismund 1999; Soto, 1999; Nielsen, 2000;Schuluter, 2002; Besse *et al.*, 2004; Minoo *et al*., 2006;Duval *et al* 2006;Bory, 2007; Grison *et al.*, 2007).

This study has demonstrated the ability of RAPD markers to reliably differentiate between different *Vanilla planifolia* germplasm/landraces and commercial varieties and also represents an initial but important step in using RAPD markers as a tool for the estimation of genetic diversity of Vanilla cultivars and other field and horticultural crops in Tanzania. The information about genetic similarity will be helpful to avoid the chance of use of genetically similar landrace/genotypes and will also be helpful in future breeding programme to select genetically diverse parents for vanilla breeding programme as well.

CHAPTER SIX

6.0 Conclusion and Recommendation

The present research results on the Vanilla crop grown in Tanzania have revealed that the vanilla cultivars which are grown in one location (intrapopulation) are closely related to each other. Further to that the research results on the comparison of vanilla crop cultivars grown in different locations (interpopulation) also revealed close similarity. These results suggest that the vanilla genetic variation in vanilla cultivars grown in the country have a very narrow genetic base implying that all the planting materials have come from one common source at different times.

The narrow genetic base revealed suggests that there is a high risk in case of new insect pest or disease outbreak. In addition, the close genetic base indicates difficulty in vanilla breeding or crop genetic improvement especially using conventional breeding methods, which heavily rely in presence of wide genetic variation among germplasm.

It is recommended that in order to improve the vanilla crop so as to develop high yielding cultivars, cultivars with high vanillin content, drought tolerance and disease tolerance, and fully exploit the potential that the vanilla crop as a country we have to base our breeding programmes on mutation breeding, genetic engineering and inter - specific hybridization methods to be applied. This will create artificial variations from where potential genotypes with deemed desirable traits might be selected.

7.0 REFERENCES

Altaf, M.K., Stewart, J.M.C.D, Wajahatullah, M.K, Zhang. J. (1997). Molecular and morphological genetics of a trispecies F2 population of cotton. Proc. Belt wide Cotton Conf., pp. 448-452.

Bekele, Y.D, **(**2005). Assessment of cup quality, morphological, biochemical and molecular diversity of *Coffea arabica* L. genotypes of Ethiopia. Ph.D. Thesis, University of the Free State.

Besse, P., Dasilva, D., Bory, S., Grisoni, M., Lebellec, F., Duval, M.F (2004).RAPD genetic diversity in cultivated vanilla: *Vanilla planifolia* and relationships with *V.tahitensis* and *V.pompona.* Plant Sci.167, 379-385.

Bolstein, A.H.D., Hopkins .M.S., Mitchel, S.E (1980). Core Collection: A Practical Approach to Genetic Resources Management. Genome 31:818-824.

Bory, S., Grisoni. M., Duval, M.F., Besse, P. (2007a). Biodiversity and preservation of vanilla: present state of knowledge. Genet.Resour.Crop Evol.101007, 220-236.

Bory, S., Grisoni. M., Duval, M.F., Besse, P. (2007b). Development of microsatellite markers in cultivated vanilla: polymorphism and transferability to other vanilla species, Sci.Hortc.10, 10-16.

Capdevielle, F.M. (2001). Evaluation of a discriminant analysis procedure combining agronomic and molecular marker information for germplasm improvement in rice. MS Thesis. Louisiana State University, Baton Rouge.

Cibrian.A, (1999) Variacion genética de *Vanilla planifolia* en Mexico, Ph.D. dissertation, UNAM, Mexico City.

Dean, R.E., Dahlberg J.A., M.S.Hopkins, S.E. Mitchell, S. Kresovich. (1999). Genetic redundancy and diversity among 'Orange' accessions in the US national sorghum collection as assessed with simple sequence repeat (SSR) markers. Crop Sci 39: 1215–1221.

De Bustos, C.E., Ferreira, M.E, Camargo, L.E.A, Williams, P.H. (1998). Comparison of genetic relationship estimates among cruciferous species based on RFLP and RAPD markers. *Theor. Appl. Genet.* 88: 973-980.

Dellaporta, S.L., Wood. J, and Hicks, J.B. (1983). A plant DNA minipreparation: version II. Plant Molecular Biology Reporter, 1:19-21.

Devos.K. M. and M. D. Gale **(**1992). The use of random amplified polymorphic DNA Markers in wheat. Theor.Appl.Genet. 84: 567-572.

Dignum, M.J.W., Kerler J., Verpoorte R., (2001). Vanilla production: technological, chemical, and biosynthetic aspects. Food Res. Int. 17, 199–219.

Duval,M.F.,Bory,S.,Andrzejewsk,S.,Grisoni,M.,Besse,P.,Causse,S.,Charon,C.,Dron, Odoux Wong,M.(2006).Genetic diversity of vanilla in secondary dispersion zones.Les Actes du BRG6, 181-196.

Excoffier,L., Smouse. P. E and Quattro, J. M. (1992). "Analysis of Molecular Variance Inferred from Metric Distances among DNA Haplotypes: Application to Human Mitochondrial DNA Restriction Data." *Genetics* 131. pp 479-491.

FAO (2006). Production data 2006. http://www.fao.org. Accessed 13 March 2008.

FAO (2007). Production data 2007. http://www.fao.org. Accessed 12 January 2009.

Felsenstein, J. (1985). Confidence limits of phylogenies: an approach using the Bootstrap. *Evolution* 39:783-791.

Felsenstein, J (1993) PHYLIP (Phylogeny Inference Package) version 3.5c. Department of Genetics, University of Washington, Seattle.

Fouche, J. G. and L. Jouve (1999). Vanilla planifolia: History, botany and culture in Reunion Island. Agronomie Paris. Dec. 19(8): 689-703.

Grisoni,M.,Moles,M.,Besee,P.,Bory,S.,Duval,M.F.,Kahane,R.(2007).Towards an international plant collection to maintain and characterize the endangered genetic resources of vanilla.Acta Hortic.(ISHS)760,83-91.

Guarino, P.A., Brown, S.M. (1985), Liquid chromatographic determination of vanillin and related flavour compounds in vanilla extract: cooperative study. *J. Assoc. Off. Anal. Chem.*68 pp. 1198–1201.

Joseph T and S. Bhai (2001).Sclerotium rot—a new disease of vanilla (*Vanilla planifolia Andr.)* In India. J. Spices Aromatic Crops 9 (2), 175–176.

Joseph, T., Bhai, S, S., A.K. Vijayan, A.K., (2003). Diseases and their management. In Vanilla, the Prince of Spices. Publ. Spices Board, Kochi. P .29-41.

Kumar, A. S. (2004). Vanilla cultivation: A profitable agri-based enterprise. In Kerala Calling, 26-30.

Lerceteau, S.W., Ulanch, P.E., Dong, J., Obert, J.A., Morishige,D.T., Schlueter, S.D., Childs, K.L., Ale, M., Mullet, J.E. (1997). A High-throughputAFLP-based Method for Constructing Integrated Genetic and Physical Maps: Progress toward a Sorghum Genome Map. Genome Research. 10:789-807.

Manceau, M. (2003). A *future for natural vanilla. European perspective*. First International Conference on the Future of the Vanilla Business. New Brunswick, NJ, Rutgers University. 80:27-43.

Mategu, M. (2007) Assessment of morphological characterization of *Vanilla planifolia* grown at SUA Hort.Unit special project report submitted as partial fulfilment of requirements for the degree of BSc. Horticulture. SUA. Morogoro, Tanzania.pp 35.

Markert, C.L. and Moller F. (1959). Multiple forms of enzymes: Tissues, cytogenetic and species specific patterns. *Proceedings of the National Academy of Sciences of the United States of America* 45, 753–763.

MAYAWA (2001). Kilimo bora cha vanilla Mkoani Kagera. 4th edition 233pp.

Merricum, C.M. (1991). A genetic linkage map of cowpea (*Vigna unguiculata*) developed from a cross between two inbred, domesticated lines. Theor. Appl. Genet. 95:1210–1217.

Minoo, D. (2002).Seedling and somaclonal variation and their characterization in Vanilla, University of Calcutta, Kerala, India. 320pp.

Minoo,D.,Nirmal,K.,Ravindran,P.N.,Peter,K.V. (2006).Interspecific hybridization in Vanilla and molecular characterization of hybrids and selfed progenies using RAPD and AFLP markers.Sci.Hortic.108,414-422.

Minoo,D.,Jayakumar,V.N.Veena,S.S.,Vimala,J.Saji,K.V.,Nirmal,K.Peter,K.V. (2007).Genetic variations and interrelationships in Vanilla planifolia and few related species as expresses by RAPD polymorphism.Genet.Resour.Crop Evol.Doi:10.1007/s10722-007-9252-3.

Mueller U. G. and L. L. Wolfenbarger. (1999). AFLP genotyping and fingerprinting. Trends in Ecology and Evolution 14: 389-394.

NBS (2002).Household Budget Survey, National Bureau of statistics in Tanzania. President's Office-planning and privatization, Dar es salaam.188pp.

Ngaga, Y.M., Z.S.K.Mvena, C.L.Rweyemamu, A.P.Maerere, I.LNdyetabula, S.Moshi, R.K.Kagombola. (2006). Baseline Survey report on" Broadening genetic base, increasing vines for planting, on farm processing and marketing of Vanilla in selected areas of Tanzania". PANTIL research Project.Sokoine University of Agriculture.23 pp.

Nielsen, L. R. and Siegismund，H. R. (1999). Interspecific differentiation and hybridization in vanilla species (Orchidaceae). Heredity 83, 560-567.

Nielsen, L. R. (2000). Natural hybridization between *Vanilla claviculata* (W. Wright) Sw. and *V. barbellata* Rchb. f. (Orchidaceae): genetic, morphological, and pollination experimental data. Botanical Journal of the Linnean Society 133(3), 285-302(18). doi: 10.1006/bojl.2000.0336.

Papa, R., Gepts P. 2003. A symmetry of gene flow and differential geographical structure of molecular diversity in wild and domesticated common bean (*Phaseolus vulgaris L.*) from Mesoamerica. Theor. ApplGenet., 106:239–250.genetic structure of coconut (*Cocos nucifera L.*). Genome 43: 15–21.

Paul, E.S. L, (2003). Trends in vanilla global markets: *A review of American Journal of Agricultural Research 52: 1043-1077.*

Partos, S, Y (2006), *Vanilla planifolia* crop husbandry, harvesting and curing in Mexico. University of Mexico City, Sci.Hortc.10, 13-18.

Perera, L., J.R.Provan, J. Russell, W. Powell. (2000).Use of microsatellite DNA markers to investigate the level of genetic diversity and population.

Pearson, M. N., Jackson, G. V. and Zettler, H. F. W. (1991). FAO/IBPGR technical guidelines fro the safe movement of vanilla germplasm, Food and Agriculture Organization of the United Nations, Rome/International Board for Plant Genetic Resources, Rome.

Porteres. R, (1954) Le genre *Vanilla* et ses espèces, in: G. Bourriquet (Ed.), Le vanillier et la vanille dans le monde, Paul Le chevalier, Paris, pp. 94–290.

Prenner, W., Morgante, M., Andre, M., Hanafey, M., Vogel, J., Tingey, S., and Rafalski, A. (1999). The utility of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. Mol. Breed. **2**: 225–238.

Purseglove, J.W., E.G. Brown, C.L Green, S.R.J. Robbins (1981). Spices, vol. 2. Longman Inc., New York, pp. 644–735.

Ranadive, A. S., (1994) Vanilla – Cultivation, Curing, Chemistry, Technology and Commercial Products pp 517-577 in *Spices, Herbs and Edible Fungi*, Developments in Food Science 34, Elsevier, Amsterdam.

Ranadive, A. S., (1996) Vanilla Cultivation Uganda – Current Status and Recommendations for Improvement, USAID IDEA Project Report Vol.9 pp43-66.

Rohlf, F. J. (1998). NTSYSpc Version 2.0: User Guide. Applied Biostatistics Inc. RAPD protocol.

Rweyemamu, C.L (2007) Vanilla: a potential alternative to orthodox crops, PANTIL newsletter Vol.1 No.1:13-15.

Ruan, X. Y., Chen J. B. and Zhu Y. Y. (1998). Renew on Vanilla Diseases Researches. Journal of Yunnan Agricultural University 13(1), 139-144.

Rudall, P.J., J.G. Conran, and M.W. Chase. (2000). Systematic of Ruscaceae/Convallariaceae: a combined morphological and molecular investigation. Bot. J. Linnean Soc. 134:73-92.

Schuluter, P.M., (2002).RAPD variation in *Vanilla planifolia* Jackson (Orchidaceae) and assessment of the putative hybrid *Vanilla tahitensis* Moore.Mbiochem Thesis. University of Oxford, Oxford, P.55.

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

Sengupta, R. (2004) www.kissanwatch.org Vanilla: The spice of life for Indian farmers.

Smith, S., and Helentjaris, T. (1996). DNA fingerprinting and plant variety protection. *In* Genome mapping in plants. *Edited by* A.H. Paterson. Academic Press, London. pp. 95–110.

Smulders, S., Roessli, D., and Excoffier, L. (1997). Arlequin: software for population genetics data analysis. Ver 2.000. Genetics and Biometry Lab, Dept. of Anthropology, University of Geneva.

Soto, M.A., (1999).Philogeography of genetic diversity of vanilla n México, Project J101. CONABIO, Mexico.

Soto, M.A., Pridgeon, A.M., Cribb, P.J., Chase, M.W., Rasmussen, F.N., (2003) Vanilla in Genera Orchidacearum: Orchidoideae. Oxford University Press, USA, pp. 321-334.

Stern, W. L. and W. S. Judd (1999). Comparative vegetative anatomy and systematic of Vanilla Botanical Journal of the Linnean Society. Dec. 131(4): 353.

Sun ,Q.Z, Gao, N.J, Huang .T (1998). Genetic relationships and diversity among Tibetan wheat, common wheat and European Spelt wheat revealed by RAPD markers. Euphytica. 99: 205-211.

Taasisi ya uchunguzi wa Kiswahili (2002) Toleo la pili, Dar-es salaam.

Tanksley, S.D (1983). Molecular markers in plant breeding. Plant Mol. Biol. Rep., 1: 3-8.

Templeton, A.R. (1991). Genetics and conservation biology. In: Seitz, A. and Loeschche, V. (Ed.), Species Conservation a Population biological approach. Basel, Birkhauser Verlag., 15-29.

Templeton, A. (1993).Translocation as conservation tool. In Szaro R. (Ed), Biodiversity in mangrove landscapes, theory and practice. Oxford University press.

Todd H. (1999) The North American market for vanilla. *Perfumer and Flavorist* 23: 23–25.

URT (2005).United Republic Of Tanzania. National Strategy for Growth and Reduction Of. Poverty. Vice president's office, Dar es Salaam.73pp.

Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M,Frijters A, Pot J, Peleman J, Kuiper M, Zabeau M (1995). AFLP: anew technique for DNA fingerprinting. Nucl. Acids Res. 23: 4407-4414.

Walton.N. J., Mayer, M. J. and Narbad, A. (2003). Vanillin. Phytochemistry 63, 505-515.

Wang, Z.Y and Tanksley, S.D (1989.) Restriction length polymorphism in Oryza sativa L. Genome 32, 1113–1118.

Weising, K., Nybom. H., Wolff, K. & Kahl, G. 2005. *DNA fingerprinting in plants*: *Principles, methods and applications*. Taylor & Francis/CRC press.

Welsh.J and M. McClelland**,** (1990). Fingerprinting genomes using PCR with arbitrary primers. Nucleic Acids Res. 18: 7213-7218.

Wen, Y. T and Li, M. R. (1992). Identification of pathogen of vanilla bacterial soft-rot disease. Chinese Journal of Tropical Crops 1, 101-104.

Williams, J.G.K, Kubelik AR, Livak K.J, Rafalski J.A, Tingey, S.V (1990). DNA polymorphism amplified by arbitrary primers is useful as genetic markers. Nucl. Acid. Res. 18: 6531-6535.

Xu, Y., Kang, G., Xiang, N., Hong, Y., Zhang, H. (2000). Genetic Assessment of Watermelon Germplasm using the AFLP Technique. HortScience. 38(1):81-84.

Zhang, J., N.Y. Klueva, Z. Zhang, R. Wu,T-H. David Ho, and H.T. Nguyen. (1999). Genetic engineering for a biotic stress resistance in crop plants. *In Vitro Cellular Developmental BiologyPlantTheor. Appl. Genet*. 92: 627–36.

8. APPENDICES

Appendix 1. Analysis of Molecular Variance (AMOVA**)**

The first step that was done before the analysis was convert the bands detected in the gels to binary variables with a value of either 0 or 1. Then, the sums of presences (1) were calculated and so proceeding with the sum of squares.

Calculations were first done for one population and continued for the others until we have (X...k). We have i = 126 individuals (effect b), j= 2 alleles (effect w), k = 3 populations (effect a).

Where,

X...k is the result of summing up all the band presences (1) in the individuals per population.X...k^2 is the result of squaring the number obtained above ΣΣXi...^{k2} is the result of adding up the squares of the sum of allele presences in each individual (Indiv.1 in Pop.1 will be (0 + 0)3 + Indiv.2 in Pop.1 (1 + 1)3 + Indiv. ...) ΣΣΣXijk ^{2} is the sum of each value squared SS is the sum of squares for effects a, b and w.

Calculation of SS:

SSa = ΣX...k 3/ij – X...3/ijk = [9900/ (10 x 3)] - [10060/ (10 x 3)] = 0.6 MS are the mean squares for effects a, b and w Calculation of MS: SSa/dfa = 0.6/3 = 0.2, where dfA refers to the degrees of freedom for effect a (populations)..

Table 1.Analysisi of Molecular Variance (AMOVA) for 126 samples collected from 3 different populations (Bukoba, Moshi and Zanzibar).

SV	df	SS	MS	EMS	
Populations	3	5.3	1.733	0.0198413	< 0.0001
Within population	24	19.2	0.80	0.0012698	< 0.0001
Between poplation	126	96.25	0.76	0.0086758	< 0.0001
Total	154				

Where,

SV = sources of variation

df = degrees of freedom

SS = sum of squares

MS = mean squares

σ2 = total estimated variance

EMS = expected mean squares

σw2 = 0.2222222

σb2 = (MSb – MSw)/3 = (0.26190476 – 0.22222222)/3 = 0.0198413.

σa2 = (MSa – MSb)/3 x 15 = (0.3 – 0.26190476)/2 x 15 = 0.0012698.

σa2 = (MSa – MSb)/3 x 15 = (0.3 – 0.26190476)/2 x 15 = 0.086758.

Appendix 2. Pictures of Vanilla

Appendix 3.Diagnostics in action

Appendix 4.Binary number of bands scored

	1	2	3	4	5	6	7	6	9	10
1	1 1 0 1 1 1 1 1 0 0 1 1 1 1 1 1 0	1 1 1 1 1 0	1 1 1 1 1 0	1 1 1 1 1 1 1 1 1 1 0	1 1 1 1
2	1 1 0 1 1 1 1 1 0 0 1 1 1 1 1 1 0	1 1 1 1 1 0	1 1 1 1 1 0	1 1 1 1 1 1 1 1 1 1 0	1 1 1 1
3	1 1 0 1 1 1 1 1 0 0 1 1 1 1 1 1 0	1 1 1 1 1 0	1 1 1 1 1 0	1 1 1 1 1 1 1 1 1 1 0	1 1 1 1
4	1 1 0 1 1 1 1 1 0 0 1 1 1 1 1 1 0	1 1 1 1 1 0	1 1 1 1 1 0	1 1 1 1 1 1 1 1 1 1 1	1 1 1 1
5	1 1 0 1 1 1 1 1 0 0 1 1 1 1 1 1 0	1 1 1 1 1 0	1 1 1 1 1 0	1 1 1 1 1 1 1 1 1 1 0	1 1 1 1
6	1 1 0 1 1 1 1 1 0 0 1 1 1 1 1 1 0	1 1 1 1 1 0	1 1 1 1 1 0	1 1 1 1 1 1 1 1 1 1 0	1 1 1 1
7	1 1 0 1 1 1 1 1 0 0 1 1 1 1 1 1 0	1 1 1 1 1 0	1 1 1 1 1 0	1 1 1 1 1 1 1 1 1 1 0	1 1 1 1
8	1 1 0 1 1 1 1 1 0 0 1 1 1 1 1 1 0	1 1 1 1 1 0	1 1 1 1 1 0	1 1 1 1 1 1 1 1 1 1 0	1 1 1 1
9	1 1 0 1 1 1 1 1 0 0 1 1 1 1 1 1 0	1 1 1 1 1 0	1 1 1 1 1 0	1 1 1 1 1 1 1 1 1 1 0	1 1 1 1
10	1 1 0 1 1 1 1 1 0 0 1 1 1 1 1 1 0	1 1 1 1 1 0	1 1 0 1 1 1 1 1 1 1 0	1 1 1 1
11	1 1 0 1 1 1 1 1 0 0 1 1 1 1 1 1 0	1 1 1 1 1 0	1 1 1 1 1 1 1 1 1 1 0	1 1 1 1
12	1 1 0 1 1 1 1 1 0 0 1 1 1 1 1 1 0	1 1 1 1 1 0	1 1 1 1 1 1 1 1 1 1 0	1 1 1 1
13	1 1 0 1 1 1 1 1 0 0 1 1 1 1 1 1 0	1 1 1 1 1 0	1 1 1 1 1 1 1 1 1 1 0	1 1 1 1
14	1 1 0 1 1 1 1 1 0 0 1 1 1 1 1 1 0	1 1 1 1 1 0	1 1 1 1 1 1 1 1 1 1 0	1 1 1 1
15	1 1 0 1 1 1 1 1 0 0 1 1 1 1 1 1 0	1 1 1 1 1 0	1 1 1 1 1 1 1 1 1 1 0	1 1 1 1
16	1 1 0 1 1 1 1 1 0 0 1 1 1 1 1 1 0	1 1 1 1 1 0	1 1 1 1 1 1 1 1 1 1 0	1 1 1 1
17	1 1 0 1 1 1 1 1 0 0 1 1 1 1 1 1 0	1 1 1 1 1 0	1 1 1 1 1 1 1 1 1 1 0	1 1 1 1
18	1 1 0 1 1 1 1 1 0 0 1 1 1 1 1 1 0	1 1 1 1 1 0	1 1 1 1 1 1 1 1 1 1 0	1 1 1 1
19	1 1 0 1 1 1 1 1 0 0 1 1 1 1 1 1 0	1 1 1 1 1 0	1 1 1 1 1 1 1 1 1 1 0	1 1 1 1
20	1 1 0 1 1 1 1 1 0 0 1 1 1 1 1 1 0	1 1 1 1 1 0	0 1 1 1 1 1 1 1 1 0	1 0 1 1
21	1 1 0 1 1 1 1 1 0 0 1 1 1 1 1 1 0	1 1 1 1 1 0	1 1 1 1 1 1 1 1 1 0	1 1 1 1
22	1 1 0 1 1 1 1 1 0 0 1 1 1 1 1 1 0	1 1 1 1 1 0	1 1 1 1 1 1 1 1 1 0	1 1 1 1
23	1 1 0 1 1 1 1 0 1 0 1 1 0 1 1 1 0	0 1 1 1 0	1 1 1 1 1 1 1 1 1 1	1 0 1
24	1 1 0 1 1 1 1 1 0 0 1 1 1 1 1 0 0	1 1 1 1 1 0	1 1 1 1 1 1 1 1 1 1	1 1 1 1
25	1 1 0 1 1 1 1 1 0 0 1 1 1 1 1 1 0	1 1 1 1 1 0	1 1 1 1 1 0	1 1 0 0	1 1 1 1
26	1 1 0 1 1 1 1 1 0 0 1 1 1 1 1 1 0	1 1 1 1 1 0	1 1 1 1 1 1 1 1 1 0	1 1 1 1

27 1 1 0 1 1 1 1 1 0 0 1 1 1 1 1 1 0 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 0 1 1 1 1
28 1 1 0 1 1 1 1 1 0 0 1 1 1 1 1 1 0 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 1 0 1 1 1 1
29 1 1 0 1 1 1 1 1 0 0 1 1 1 1 1 1 0 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 1 0 1 1 1 1
30 1 1 0 1 1 1 1 1 0 0 1 1 1 1 1 1 0 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 1 0 1 1 1 1
31 1 1 0 1 1 1 1 1 0 0 1 1 1 1 1 1 0 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 1 0 1 1 1 1
32 1 1 0 1 1 1 1 1 0 0 1 1 1 1 1 1 0 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 1 0 1 1 1 1
33 1 1 0 1 1 1 1 1 0 0 1 1 1 1 1 1 0 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 1 0 1 1 1 1
34 1 1 0 1 1 1 1 1 0 0 1 1 1 1 1 1 0 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 1 0 1 1 1 1
35 1 1 0 1 1 1 1 1 0 0 1 1 1 1 1 1 0 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 1 1 0 1
36 1 1 0 1 1 1 1 1 0 0 1 1 1 1 1 1 1 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 0 1 1 1 1
37 1 1 0 1 1 1 1 1 0 0 1 1 1 1 1 1 0 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 0 1 1 1 1
38 1 1 0 1 1 1 1 1 0 0 1 1 1 1 1 1 0 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 0 1 1 1 1
39 1 1 0 1 1 1 1 1 0 0 1 1 1 1 1 1 0 1 1 1 1 1 0 1 1 1 1 1 1 1 1 0 1 0 1 1 1 1
40 1 1 0 1 1 1 1 1 0 0 1 1 1 1 1 1 0 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 0 1 1 1 1
41 1 1 0 1 1 1 1 1 0 0 1 1 1 1 1 1 0 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 0 1 1 1 1
42 1 1 0 1 1 1 1 1 0 0 1 1 1 1 1 1 0 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 0 1 1 1 1
43 1 1 0 1 1 1 1 1 0 0 1 1 1 1 1 1 0 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 0 1 1 1 1
44 1 1 0 1 1 1 1 1 0 0 1 1 1 1 1 1 0 1 1 1 1 1 0 0 1 1 0 1 1 1 1 1 0 1 1 1 1
45 1 1 0 1 1 1 1 1 0 0 1 1 1 1 1 1 0 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 0 1 1 1 1
46 1 1 0 1 1 1 1 1 0 0 1 1 1 1 1 1 0 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 0 1 1 1 1
47 1 1 0 1 1 1 1 1 0 1 1 1 1 1 1 1 0 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 0 1 1 1 1
48 1 1 0 1 1 1 1 1 0 0 1 1 1 1 1 1 0 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 0 1 1 1 1
49 1 1 0 1 1 1 1 1 0 0 1 1 1 1 1 1 0 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 0 1 1 1 1
50 1 1 0 1 1 1 1 1 0 0 1 1 1 1 1 1 0 1 1 1 1 1 0 1 1 1 0 1 1 1 1 1 0 1 1 1 1
51 1 1 0 1 1 1 1 1 0 0 1 1 1 1 1 1 0 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 0 1 1 1 1
52 1 1 0 1 1 1 1 1 0 0 1 1 1 1 1 1 0 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 0 1 1 1 1
53 1 1 0 1 1 1 1 1 0 0 1 1 1 1 1 1 0 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 0 1 1 1 1
54 1 1 0 1 1 1 1 1 0 0 1 1 1 1 1 1 0 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 0 1 1 1 1
55 1 1 0 1 1 1 1 1 0 0 1 1 1 1 1 1 0 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 0 1 1 1 1
56 1 1 0 1 1 1 1 1 0 0 1 1 1 1 1 1 0 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 0 1 1 1 1
57 1 1 0 1 1 1 1 1 0 0 1 1 1 1 1 1 0 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 0 1 1 1 1
58 1 1 0 1 1 1 1 1 0 0 1 1 1 1 1 1 0 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 0 1 1 1 1
59 1 1 0 1 1 1 1 1 0 0 1 1 1 1 1 1 0 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1
60 1 1 0 1 1 1 1 1 0 0 1 1 1 1 1 1 0 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 0 1 1 1 1
61 1 1 0 1 1 1 1 1 0 0 1 1 0 1 1 1 0 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 0 1 1 1 1
62 1 1 0 1 1 1 1 1 0 0 1 1 1 1 1 0 0 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 0 1 1 1 1
63 1 1 0 1 1 1 1 1 0 0 1 1 0 1 1 1 0 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 0 1 1 1 1
64 1 1 0 1 1 1 1 1 0 0 1 1 0 1 1 1 0 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 0 1 1 1 1
65 1 1 0 1 1 1 1 1 0 0 1 1 1 1 1 1 0 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 0 1 1 1 1
66 1 1 0 1 1 1 1 1 0 0 1 1 1 1 1 1 0 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 0 1 1 1 1
67 1 1 0 1 1 1 1 1 0 0 1 1 1 1 1 1 0 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 0 1 1 1 1
68 1 1 0 1 1 1 1 1 0 0 1 1 1 1 1 1 0 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 0 1 1 1 1
69 1 1 0 1 1 1 1 1 0 0 1 1 1 1 1 1 0 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 0 1 1 1 1
70 1 1 0 1 1 1 1 1 0 0 1 1 1 1 0 1 0 1 1 1 1 0 0 1 1 1 1 1 1 1 1 1 0 0 1 1 1
71 1 1 0 1 1 1 1 1 0 0 1 1 1 1 1 1 1 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 0 1 1 1 1
72 1 1 0 1 1 1 1 1 0 0 1 1 1 1 1 0 1 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 0 1 1 1 1
73 1 1 0 1 1 1 1 1 0 0 1 1 1 1 1 0 1 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 0 1 1 1 1
74 1 1 0 1 1 1 1 1 0 0 1 1 1 1 1 0 1 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 0 1 1 1 1
75 1 1 0 1 1 1 1 1 0 0 1 1 1 1 1 1 1 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 0 1 1 1 1
76 1 1 0 1 1 1 1 1 0 0 1 1 1 1 1 1 1 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 0 1 1 1 1
77 1 1 0 1 1 1 1 1 0 0 1 1 1 1 1 1 1 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 0 1 1 1 1

78 1 1 0 1 1 1 1 1 0 0 1 1 1 1 1 1 1 1 1 1 1 0 1 1 1 1 1 1 1 1 0 1 1 1 1
79 1 1 01 1 1 1 1 0 0 1 1 1 1 1 1 1 1 1 1 0 1 0 1 1 1 1 1 1 1 1 1 0 1 1 1 1
80 1 1 0 1 1 1 1 1 0 0 1 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 1 1 1 1 0 1 1 1 1
81 1 1 0 1 1 1 1 1 0 0 1 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 1 1 1 11 0 0 1 1 1
82 1 1 0 1 1 1 1 1 0 0 1 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 1 1 1 1 0 1 1 1 1
83 1 1 0 1 1 1 1 1 0 0 1 1 1 1 1 1 1 1 1 1 1 0 1 1 0 1 1 1 1 1 1 0 1 1 0 1
84 1 1 0 1 1 1 1 1 0 0 1 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 0 1 0 1 1
85 1 1 0 1 1 1 1 1 0 1 1 1 1 1 1 1 0 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 0 1 1 1 1
86 1 1 0 1 1 1 1 1 0 1 1 1 1 1 1 1 0 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 0 1 1 1 1
87 1 1 0 1 1 1 1 1 0 1 1 1 1 1 1 1 0 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 01 1 1 1
88 1 1 0 1 1 1 1 1 0 1 1 1 1 1 1 1 0 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 0 1 1 1 1
89 1 1 0 1 1 1 1 1 0 1 1 1 1 1 1 1 0 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 0 1 1 1 1
90 1 1 0 1 1 1 1 1 0 0 1 1 1 1 1 1 1 1 0 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 0 1 1 1 1
91 1 1 0 1 1 1 1 1 0 0 1 1 1 1 1 1 1 1 0 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 0 1 1 1 1
92 1 1 0 1 1 1 1 1 0 0 1 1 1 1 1 1 1 1 0 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 0 1 1 1 1
93 1 1 0 1 1 11 0 0 0 1 1 1 1 1 1 1 1 1 1 1 1 0 1 0 1 1 1 1 1 1 1 1 1 0 1 0 1 1

94 1 1 0 1 1 1 1 1 0 0 1 1 1 1 1 1 1 0 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 0 1 1 1 1
95 1 1 0 1 1 1 1 1 0 0 1 1 1 1 1 1 1 0 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 0 1 1 1 1
96 1 1 0 1 1 1 1 1 0 0 1 1 1 1 0 1 0 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 0 1 1 1 1
97 1 1 0 1 1 1 1 1 0 0 1 1 1 1 1 1 1 0 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 0 1 1 1 1
98 1 1 0 1 1 1 1 1 0 0 1 1 1 1 1 1 1 0 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 0 1 1 1 1
99 1 1 1 1 1 1 1 1 1 1 0 1 1 1 1 1 1 1 0 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 0 1 1 1 1
100 1 1 0 1 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 0 1 1 1 0 1 0 1 1 1 1 1 1 1 1 1 0 0 1 1 1
101 1 1 1 1 1 1 1 1 1 1 0 1 1 1 1 1 1 1 0 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 0 1 1 1 1
102 1 1 1 0 1 1 1 1 1 1 0 1 1 1 1 1 1 1 0 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 0 1 1 1 1
103 1 1 0 1 1 1 1 1 1 0 1 1 1 1 1 1 1 1 0 1 1 1 0 1 0 1 1 1 1 1 1 1 1 1 0 1 1 1 1
104 1 1 1 1 1 1 1 1 1 0 0 1 1 1 1 1 1 1 0 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 0 1 1 1 1
105 1 1 1 1 1 1 1 1 1 0 0 1 1 1 1 1 1 1 0 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 0 1 1 1 1
106 1 1 1 1 1 1 1 1 1 1 0 1 1 1 1 1 1 1 0 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 0 1 1 1 1
107 1 1 1 1 1 1 1 1 1 1 0 1 1 1 1 1 1 1 0 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 0 1 1 1 1
108 1 1 1 1 1 1 1 1 1 1 0 1 1 1 1 1 1 1 0 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 0 1 1 1 1
109 1 1 0 1 1 1 1 1 1 0 0 1 1 1 1 1 1 1 0 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 0 1 1 1 1
110 1 1 0 1 1 1 1 1 1 0 0 1 1 1 1 1 1 1 0 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 0 1 1 1 1
111 1 1 0 1 1 1 1 1 1 0 0 1 1 1 1 1 1 1 0 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 0 1 1 1 1
112 1 1 0 1 1 1 1 1 1 0 0 1 1 1 1 1 1 1 0 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 0 1 1 1 1
113 1 1 0 1 1 1 1 1 1 0 0 1 1 1 1 1 1 1 0 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 0 1 1 1 1
114 1 1 0 1 1 1 1 1 1 0 1 1 1 1 1 1 1 1 0 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 0 1 1 11
115 1 1 0 1 1 1 1 1 1 0 0 1 1 1 1 1 1 1 0 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1
116 1 1 0 1 1 1 1 1 0 0 1 1 1 1 1 1 1 0 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
117 1 1 0 1 1 1 1 1 1 0 0 1 1 1 1 1 1 1 0 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1
118 1 1 0 1 1 1 1 1 1 0 0 1 1 1 1 1 1 1 0 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1
119 1 1 0 1 1 1 1 1 1 0 0 1 1 1 0 1 1 0 1 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 0 1 1 1 1
120 1 1 0 1 1 1 1 1 1 0 0 1 1 1 0 1 1 0 1 1 0 1 1 0 1 1 1 1 1 1 1 1 1 0 1 1 1 1
121 1 1 0 1 1 1 1 1 1 0 0 1 1 1 1 1 1 1 0 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 0 1 1 1 1
122 1 1 0 1 1 1 1 1 1 0 0 1 1 1 1 1 1 1 0 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 0 1 1 1 1
1**23** 1 1 0 1 1 1 1 1 1 0 0 1 1 1 1 1 1 1 0 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 0 1 1 1 1
124 1 1 0 1 1 1 1 1 1 0 0 1 1 1 1 1 1 1 0 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 0 1 1 1 1
125 1 1 0 1 1 1 1 1 1 0 0 1 1 1 1 1 1 1 0 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 0 1 1 1 1
126 1 1 0 1 1 1 1 1 1 0 0 1 1 1 1 1 1 1 0 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 0 1 1 1 1

Appendix 5. Eigen values

Input parameters
Read input from file: C:\Documents and Settings\Administrator\Desktop\vanilla\vanilla.SIM
Number of dimensions: 4
Save eigenvectors in output file: VANILLA VEC.NTS
Save eigenvalues in output file: VANILLA VAL.NTS
Scaling: SQRT(LAMBDA)

*** File: vanilla.SIM has a non-standard file extension (.SIM).
*** Assumed to actually be an NTS file.

Comments:
SIMQUAL: input=C:\Documents and Settings\Administrator\Desktop\vanilla\vanilla.NTS, coeff=SM by Cols
Matrix type = 3, size = 126 by 126, missing value code = "none" (similarity)

i	Eigenvalue	Percent	Cumulative
1	115.49871848	91.6656	91.6656
2	1.84862875	1.4672	93.1328
3	1.03570655	0.8220	93.9548
4	0.92534191	0.7344	94.6892
5	0.79767795	0.6331	95.3223
6	0.71342240	0.5662	95.8885
7	0.64451809	0.5115	96.4000
8	0.56611112	0.4493	96.8493
9	0.47291293	0.3753	97.2246
10	0.44532979	0.3534	97.5781
11	0.36122460	0.2867	97.8648
12	0.33855166	0.2687	98.1334
13	0.30414597	0.2414	98.3748
14	0.27501954	0.2183	98.5931
15	0.25296124	0.2008	98.7939
16	0.22958497	0.1822	98.9761
17	0.22178380	0.1760	99.1521
18	0.18897611	0.1500	99.3021
19	0.18531310	0.1471	99.4491
20	0.15165992	0.1204	99.5695
21	0.13682960	0.1086	99.6781

22	0.12505937	0.0993	99.7774
23	0.11046609	0.0877	99.8650
24	0.09148265	0.0726	99.9376
25	0.08919846	0.0708	> 100%
26	0.07754795	0.0615	> 100%
27	0.07427841	0.0590	> 100%
28	0.06854738	0.0544	> 100%
29	0.06662199	0.0529	> 100%
30	0.05993555	0.0476	> 100%
31	0.05309503	0.0421	> 100%
32	0.04760338	0.0378	> 100%
33	0.04413924	0.0350	> 100%
34	0.04112481	0.0326	> 100%
35	0.03446405	0.0274	> 100%
36	0.03295879	0.0262	> 100%
37	0.03058134	0.0243	> 100%
38	0.02537277	0.0201	> 100%
39	0.02057017	0.0163	> 100%
40	0.01267377	0.0101	> 100%
41	0.01195676	0.0095	> 100%
42	0.00974051	0.0077	> 100%
43	0.00087627	0.0007	> 100%
44	0.00052624	0.0004	> 100%
45	0.00032501	0.0003	> 100%
46	0.00013060	0.0001	> 100%
47	0.00003255	0.0000	> 100%
48	0.00000456	0.0000	> 100%
49	0.00000330	0.0000	> 100%
50	0.00000120	0.0000	> 100%
51	0.00000055	0.0000	> 100%
52	0.00000000	0.0000	> 100%
53	0.00000000	0.0000	> 100%
54	0.00000000	0.0000	> 100%
55	0.00000000	0.0000	> 100%
56	0.00000000	0.0000	> 100%
57	0.00000000	0.0000	> 100%
58	0.00000000	0.0000	> 100%
59	0.00000000	0.0000	> 100%
60	0.00000000	0.0000	> 100%
61	0.00000000	0.0000	> 100%
62	0.00000000	0.0000	> 100%
63	0.00000000	0.0000	> 100%
64	0.00000000	0.0000	> 100%
65	0.00000000	0.0000	> 100%
66	0.00000000	0.0000	> 100%
67	0.00000000	0.0000	> 100%
68	0.00000000	0.0000	> 100%
69	0.00000000	0.0000	> 100%
70	0.00000000	0.0000	> 100%
71	0.00000000	0.0000	> 100%
72	0.00000000	0.0000	> 100%
73	0.00000000	0.0000	> 100%
74	0.00000000	0.0000	> 100%
75	0.00000000	0.0000	> 100%
76	0.00000000	0.0000	> 100%
77	0.00000000	0.0000	> 100%
78	0.00000000	0.0000	> 100%
79	0.00000000	0.0000	> 100%
80	0.00000000	0.0000	> 100%
81	0.00000000	0.0000	> 100%
82	-0.00000418	0.0000	> 100%
83	-0.00007069	-0.0001	> 100%
84	-0.00014098	-0.0001	> 100%

85	-0.00016598	-0.0001	> 100%
86	-0.00023216	-0.0002	> 100%
87	-0.00030911	-0.0002	> 100%
88	-0.00033483	-0.0003	> 100%
89	-0.00036976	-0.0003	> 100%
90	-0.00041120	-0.0003	> 100%
91	-0.00049519	-0.0004	> 100%
92	-0.00066503	-0.0005	> 100%
93	-0.00067532	-0.0005	> 100%
94	-0.00090803	-0.0007	> 100%
95	-0.00097571	-0.0008	> 100%
96	-0.00144704	-0.0011	> 100%
97	-0.00152872	-0.0012	> 100%
98	-0.00158430	-0.0013	> 100%
99	-0.00246696	-0.0020	> 100%
100	-0.00305595	-0.0024	> 100%
101	-0.00352168	-0.0028	> 100%
102	-0.00416397	-0.0033	> 100%
103	-0.00470083	-0.0037	> 100%
104	-0.00607265	-0.0048	> 100%
105	-0.00699568	-0.0056	> 100%
106	-0.00743310	-0.0059	> 100%
107	-0.00841036	-0.0067	> 100%
108	-0.00913385	-0.0072	> 100%
109	-0.00944305	-0.0075	> 100%
110	-0.01048798	-0.0083	> 100%
111	-0.01179144	-0.0094	> 100%
112	-0.01456089	-0.0116	> 100%
113	-0.01639015	-0.0130	> 100%
114	-0.01907486	-0.0151	> 100%
115	-0.02103099	-0.0167	> 100%
116	-0.02297314	-0.0182	> 100%
117	-0.02670952	-0.0212	> 100%
118	-0.02935518	-0.0233	> 100%
119	-0.03423588	-0.0272	> 100%
120	-0.03632373	-0.0288	> 100%
121	-0.04096982	-0.0325	> 100%
122	-0.05058242	-0.0401	> 100%
123	-0.06222179	-0.0494	> 100%
124	-0.06791066	-0.0539	> 100%
125	-0.07918269	-0.0628	> 100%
126	-0.10421976	-0.0827	100.0000

Sum of eigenvalues = 126.000000
Eigenvalues (4 by 4) saved in file: VANILLA VAL.NTS
Eigenvectors matrix (126 by 4) saved in file: VANILLA VEC.NTS

Ending date & time: 6/18/2009 6:11:29 AM