UNIVERSITY OF
KWAZULU-NATAL
INYUVESI
YAKWAZULU-NATALI

Morphological and cytological diversity of some yams (*Dioscorea* spp.) in Sierra Leone

by

Prince Emmanuel Norman

BSc (Hons) Njala University College, University of Sierra Leone

Submitted in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

in the

Discipline of Plant Breeding

School of Agricultural Sciences and Agribusiness

Faculty of Science and Agriculture

University of KwaZulu-Natal

Pietermaritzburg

Republic of South Africa

December 2010

TABLE OF CONTENT

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

DECLARATION

I, PRINCE EMMANUEL NORMAN, declare that:

1. The research reported in this thesis, except where otherwise indicated, is my original work.
2. This thesis has not been submitted to any other University for any degree consideration.
3. This thesis does not contain data, graphs and pictures obtained from anyone or other sources of information.
4. This thesis does not contain other authorsô writing, unless specifically acknowledged as being sourced from other authors. Where information from other authors have been used:

    a) Their words have been re-written but the general information attributed to them has been referenced.
    b) Their exact words have been placed in italics and inside quotation marks, and referenced to them.

5. This thesis does not contain text, graphics or tables copied and pasted from the internet, unless specifically acknowledged, and the source being referenced.

Signed:

é é é é é é é é é é é é é é é é é .Dateé é é é é é é é é ..é é

Prince Emmanuel Norman (MSc Candidate)

As research supervisors we agree to submission of this dissertation for examination:

Signed: é é é é é é é é é é é é é é é Dateé é é é é é é é é é é

Prof. Pangirayi Tongoona (Supervisor)

Signed: é é é é é é é é é é é é é é é Dateé é é é é é é é é é é

Dr. Paul E. Shanahan (Co-supervisor)

ACKNOWLEDGEMENT

I would like to first and foremost register my sincere gratitude and gratefulness to the Almighty God who faithfully enabled me to successfully accomplish my dissertation write-up. To Him be all the glory for His grace and favour.

This research was made possible through the financial support provided by Forum for Agricultural Research in Africa (FARA) under the African Development Bank (AfDB) supported project ôPromotion of Science and Technology for Agricultural Development in Africa.ô I am therefore grateful to the stakeholders of this capacity strengthening vision. All views, interpretations, recommendations and conclusions expressed in this document are those of the author and not necessarily those of the supporting or cooperating organizations.

Without the insightful inputs and encouragements from mentors, colleagues, family members and friends, it would have been difficult and/or rather impossible to accomplish my dream. I am sincerely grateful and deeply indebted to my supervisors; Professor Pangirayi Tongoona and Dr. Paul Shanahan for their patient guidance and insightful inputs throughout the study. The invaluable contribution and great encouragement from Professor Pangirayi Tongoona helped to brighten my research career.

Heartfelt thanks and appreciation are extended to Dr. Alfred G. Dixon, Director General, Sierra Leone Agricultural Research Institute (SLARI), Professor Edward R. Rhodes, Deputy Director General, SLARI, Dr. Abdulai Jalloh, CORAF, Ghana and Dr. Sahr N. Fomba, Director, Njala Agricultural Research Centre (NARC), Sierra Leone, for their immense support and their recommendation, without which I would not have been in South Africa to pursue Plant Breeding career. My thanks and appreciation also go to the Government of Sierra Leone for providing me with the required study leave to pursue higher studies in South Africa.

My sincere thanks and gratefulness are due to Dr. Chrisna Durandt, Flow Cytometry Senior Sales Consultant, Beckman Coulter, South Africa, Dr. Habtom B. Tesfagiorgis, Plant Pathology, University of KwaZulu-Natal, and Dr. Samson Z. Tesfay, Horticultural Science, University of KwaZulu-Natal, for their intellectual support during cytological and flow cytometry experiments at the University of KwaZulu-Natal. I register my appreciation to Mrs. Shirley Machellar, Chief Laboratory Technician, Center for Electron Microscopy,

University of KwaZulu-Natal, for her invaluable input in accomplishing my chromosome experiment. The excellent technical support of Mr. Mathew J. Erasmus; the special efforts of Ms. Celeste Clerk, Senior Laboratory Technician, Crop Science, University of KwaZulu-Natal in ordering and cataloguing of all the chemicals and products that I used during my work. My thanks and appreciation also go to Prof. Kelvin P. Kirkman, Deputy Dean, Faculty of Science and Agriculture, University of KwaZulu-Natal; and Mr. Brendan Boyce, Manager, School of Agricultural Sciences and Agribusiness, University of KwaZulu-Natal, for helping in admistering our scholarship; and Dr. Joseph Adjety for his unwavering encouragement and concern.

Finally, I owe debts of gratitude and appreciation to all family members including my mother and father, Mr. and Mrs. Daniel J. Norman, brothers, sisters, nieces and nephews for their huge encouragement from Sierra Leone. And to my one and only beloved wife, Sylvia, I register my heartfelt love and earnest thanks for her enormous sacrifice, support and share of my life abroad.

ABSTRACT

Yam (*Dioscorea* spp.) is a major source of income and a food security crop for many households in Sierra Leone. Despite the economic importance of the yam crop its improvement has suffered from the lack of knowledge of existing germplasm and the genetic potential within the yam gene-pool. As a consequence, many species of yam are being lost to changing tastes, industrialization and urbanization. All these lead to habitat destruction.

This study assessed the extent of diversity in some yam germplasm from Sierra Leone using morphological and cytological descriptors, and ascertained the interrelationship between these two data sets. To this end, 52 genotypes comprising of forty three *D. alata*, two *D. bulbifera*, and seven *D. rotundata* sampled from the Sierra Leone germplasm were grown in a three replicate, in a randomized complete block design (RCBD) during 2010 planting season at the University of KwaZulu-Natal, Pietermaritzburg, South Africa.

Principal component analysis (PCA) of 28 morphological characters indicated that the first 10 principal components (PCs) with eigen-values greater than 0.6 explained 86.61% of the total variation. The PCs that largely contributed to the variability included number of days to shoot emergence, leaf position, leaf shape, leaf size, density of leaf, leaf vein colour; colour of leaf, petiole, petiole wing and stem, shoot growth rate, tuber shape and flesh colour of central cross section of tuber. The two-dimensional plot of the first two PCs grouped the accessions according to their species, but did not separate them into the tuber shape groups of irregular, oblong, oval-oblong, round and cylindrical.

Factor analysis (FA) grouped the morphological traits into six factors, which together explained 75% of the total phenotypic variation in the dependence structure. Factor 1 was strongly associated with absence or presence of wings, distance between lobes, leaf apex shape, leaf colour, leaf margin colour, leaf measurement length-2, leaf vein colour of upper surface, number of branches, number of stems, stem colour and tip length of mature leaf; factor 2 with leaf density, leaf measurement length-1, leaf vein colour of lower surface, petiole wing colour, tip colour, wing colour and flesh colour of central cross section of tuber; factor 3 with leaf measurement width-1; factor 4 with leaf measurement width-2; factor 5 with stem colour; and factor 6 with number of days to emergence.

The dendrogram of the cluster analysis produced six major groups supporting the PCA and FA groupings**.** Clusters A, B, C, D, E and F were formed at the dissimilarity distance = 0.90; and they consisted of two, thirty eight, one, seven, two and two genotypes respectively. Genotypes of cluster A belong to *D. bulbifera*, while genotypes of clusters B, C, E and F belong to *D. alata*, and genotypes of cluster D belong to *D. rotundata*.

The ploidy levels of the 52 genotypes were determined by flow cytometry. The various ploidy levels obtained included diploid (2*x*), triploid (3*x*), tetraploid (4*x*), pentaploid (5*x*) and hexaploid (6*x*). The estimated nuclear DNA content ranged from 1.634 pg for G_1 nuclei of diploid *Dioscorea alata* to 2.118 pg for G_1 nuclei of hexaploid *Dioscorea rotundata*. Genotypes NR 07/045 (4*x*) and NR 07/040 (5*x*), which belong to *D. bulbifera* had nuclear DNA content of 1.905 and 2.017 pg respectively. The nuclear DNA content per genome was higher in diploids compared to polyploids. The variations within the 4*x* accessions ($p<0.005$) and among the three species ($p<0.037$) were significant. Root tips of six genotypes were prepared for chromosome counting using the acetocarmine staining technique. Genotype ER 07/030 had 20 chromosomes, four genotypes including ER 07/036, NR 07/060, NR 07/071 and SR 07/072 had 40 chromosomes, and TDr 95/18544 had 60 chromosomes.

The findings agree with the hypothesis that studied germplasm from Sierra Leone were morphologically different expressing inter- and intra-group variability. Duplicate accessions were observed supporting the hypothesis that some genotypes had different names in diverse cultural setting. Flow cytometric measurements and conventional chromosome counting showed the existence of inter- and intra-group diversity in ploidy level and nuclear DNA content. A correlation was established between agro-morphological and cytological traits used in the study. This study contributes to an understanding of yam diversity in Sierra Leone which will facilitate yam genetic resource management, conservation and utilization.

LIST OF ABBREVIATIONS

AfDB	African Development Bank
CORAF	Conférence des Résponsables de la Recherche Agronomique Africains
FAO	Food and Agricultural Organization
FARA	Forum for Agricultural Research in Africa
IAR	Institute of Agricultural Research
IBPGR	International Board for Plant Genetic Resources
IITA	International Institute of Tropical Agriculture
INRAB	Institut National des Recherches Agricoles du Bénin
IPGRI	International Plant Genetic Resources Institute
NARC	Njala Agricultural Research Institute
NARS	Njala Agriculture Research Systems
SLARI	Sierra Leone Agricultural Research Institute

GENERAL INTRODUCTION

Yams are important monocotyledonous tuberous plants, which belong to the genus *Dioscorea* of the family *Dioscoreaceae,* order Dioscoreales (Ake Assi, 1998). The genus *Dioscorea* contains 600 species with more than 10 species cultivated for food and pharmaceutical use (Ake Assi, 1998). Six species are important staples including *D. rotundata* (white yam), *D. alata* (water yam), *D. cayenensis* (yellow yam), *D. bulbifera* (aerial yam), *D. dumetorum* (trifoliate yam) and *D. esculenta* (Chinese yam) (Ng and Ng, 1994). These major food species originated in three distinct regions of the world: Southeast Asia, West Africa, and Tropical America, which are also considered the main centers of yam domestication and diversity (Asiedu *et al*., 1997). Although yams are monocot, they possess some features similar to dicots such as reticulate venation, stalked net-nerving leaves, circular arrangement of vascular bundles in the stem, and the lateral position of the pistil. In relation to phylogenetic relationships, the yam is a representation of the biological link between grasses and eudicot plants (Chase, 2004). A diverse assemblage of flowering plants with an enormous range of diversity in morphology, chemistry, habitat, geographic distributions and other attributes is known as eudicot. Such a unique model crop enlightens our understanding of plant biology and evolution (Chase, 2004). They are herbaceous or woody climbing perennials with starch-rich storage organs and a distinct annual cycle of growth (Coursey, 1983). The word óyamô has its root in languages spoken by tribes of West Africa such as Mande óniamô or the Temme óen yameô. It was then adapted into Portuguese as óynhameô, Spanish as ónameô, French as óignameôand English as óyamô(Adesuyi, 1997).

Yam has played a significant role in food and agricultural system diversification, widened the food base and brought food security to about 300 million people in low income, food deficient countries (Obidiegwu *et al*., 2009). The nutrient contents of yam per 100 g fresh edible tuber include 50-84% moisture, 15-31% carbohydrate, 0.04-0.6% crude fat, 1.1-2.8% protein and 71-142 kcal energy (Asiedu *et al*., 1997; Opara, 1999). The tubers are rich in arginine, leucine, isoleucine, and valine, with substantial quantities of thiamine, riboflavin, niacin, and ascorbic acid (Eka, 1985). Yams are also a good source of calcium (5-70 mg/ 100 g edible portion of tuber), phosphorus (5-60 mg/ 100 g), and iron (0.5 mg/ 100 g) (Eka, 1985). In addition to its economic and nutritional values, the crop also plays a significant role in the cultural life of traditionalists in Africa, where more than 95% of the world yam is

produced (Zannou *et al*., 2004, 2007). The first cultivation of *D. cayenensis-D. rotundata* complex (Guinea yams) in Western and Central Africa regions dates back 7000 years, when farmersô selection of genotypes was based on their needs. The production of yam has steadily increased from 18 million metric tonnes in 1990 to more than 38 million metric tonnes in recent years (FAO, 2006). In fact, the yam belt alone in West Africa accounts for about 95% of the global annual production, which is estimated at over 51 million metric tonnes (FAO, 2007). This is as a result of increasing utilization of traditional landraces and expansion into marginal areas. Such expansion demands the provision of improved, high yielding, pest and disease resistant cultivars with tuber quality acceptable to farmers (Manyong *et al*., 2001).

In Sierra Leone, yam is a highly valued crop, which not only provides food for household consumption, but also improves many livelihoods through the sale of harvested tubers. It is considered the third most important root and tuber crop after *Manihot esculenta* (cassava) and *Ipomoea batatas* (sweet potato). However, during festive seasons, some people prefer using yams to grace their traditional meals rather than cassava and sweet potato. Also, wild types of yam are consumed by some farming communities in the rural areas especially during the mid-rainy season to overcome hunger (IAR, 2004). This emphasizes the significant role yam plays in food insecure homes in major yam producing areas in West Africa. Despite its economic importance, food yams have not been accorded the scientific attention required to investigate genetic traits that are desirable for their improvement. This has contributed to the susceptibility of many genotypes to pests and diseases and their low yields (Orkwor *et al*., 1998). The lack of improved genotypes and pest and disease free planting material are mainly responsible for the relatively small production areas and low yields worldwide.

Presently, germplasm has been collected for this study from some parts of Sierra Leone, but thorough morphological, cytogenetic and molecular classifications are yet to be conducted. Since characterization based only on morphological or agronomic characteristics masks important genetic information, complementation with cytogenetic and molecular examination of the germplasm using techniques such as isozyme analysis (Mignouna and Dansi, 2003), flow cytometry (Egesi *et al*., 2002) and marker assisted selection (Dumont *et al*., 2005), will fully reveal existing polymorphism in the various populations. Determination of the genetic diversity of yam is complicated by the fact that farmers with different ethnicity have different vernacular names possibly for the same genotypes. The various names have led to confusion

in the number of varieties of yam considered to be cultivated in the country and perhaps overestimation of the actual extent of genetic diversity.

The number and origin of chromosomes in yams may affect such factors as the type of hybrid exhibiting maximum heterosis, the amount of genetic variation lost through self-pollination and the probability of obtaining useful traits from the species (Hamon *et al*., 1995). The basic chromosome number in yams is x = 10 (Zoundjihekpon *et al*., 1990). Earlier cytogenetic work in *D. alata* showed the existence of different ploidy levels (2n = 4x, 6x, 8x) in the species (Martin and Ortiz, 1963, 1966). Cytological abnormalities associated with polyploid formation are often responsible for the erratic flowering and reproductive behaviour in yams (Egesi *et al*., 2002). Thus, the development and application of molecular cytogenetic techniques remain central in determining chromosome structure and karyotype variation of *Dioscorea* spp. (Egesi *et al*., 2002). Furthermore, nuclear genome size within populations of inter-mating individuals must be kept constant to avoid a high rate of meiotic aberrations (Egesi *et al*., 2002).

Determination of ploidy levels is traditionally done by counting the chromosomes after staining (Abraham, 1998). This method is, however, unsuitable for large-scale screening of breeding populations in yam since the chromosomes are generally small, dot-like and most often clumped together making counting difficult and laborious (Zoundjihekpon *et al*., 1990). Although stomatal size, density and pollen size have been used to determine ploidy in some species (Vandenhout *et al*., 1995; Tenkouano *et al*., 1998), these methods have been found not to be reproducible and therefore unreliable. Deoxyribonucleic acid (DNA) flow cytometry, which measures the fluorescence of a large number of stained nuclei within seconds, provides an estimate of nuclear DNA content within somatic plant tissues (Arumuganathan and Earle, 1991). The main merits of flow cytometry include its simplicity, speed, accuracy, convenience and ability to screen a large number of samples per day. The technique also circumvents a long generation interval for meiotic analysis utilizing a small amount of tissue. It is therefore non-destructive, with the possibility of analyzing large population of cells where mixoploidy or aneuploidy exists (Dolezel, 1997).

Knowledge of ploidy status facilitates breeding hybrids with higher yields, in conjunction with tuber characteristics (quality, shape, etc.) adapted to commercial production, and resistance to anthracnose, an important disease of yams (Arnau *et al*., 2007). Generally, the breeding scheme of yams begins with characterization and evaluation of germplasm received

from farmers, the National Agricultural Research Systems (NARS) collection, and other collections for field performance, morphology, tuber quality, and ploidy status culminating in the selection of parents with desirable traits for hybridization (Mignouna *et al*., 2007). Access to a wide range of genetic diversity is essential in order for the plant breeder to develop superior hybrid genotypes. From the diverse gene pool, parents with superior complementary genes are selected and crossed to produce genotypes with adaptable ecological, culinary, and pest and disease resistance traits. Such genetic diversity is fundamental to the success of the breeding programme because farmersô needs may vary in different regions depending on the socio-economic value of the crop, farming system and the desired traits needed by the end users (Zannou *et al*., 2004).

It is likely that the genotypes sampled from the Sierra Leone germplasm are genetically different, but no detailed morphological and/ or genetic study had been conducted prior to this research. The aim of this study was to record the level of morphological diversity of the samples and to determine the ploidy level using flow cytometric and conventional chromosome counting techniques. The specific objectives of this study included: (i) the determination of the level of diversity among the accessions through morphological classification; (ii) the identification of genotype duplicates having different vernacular names but exhibiting similar morphological characters; (iii) the identification of various polyploidy levels for further genetic manipulations; and (iv) the determination of the relationships between agronomic traits and variation in nuclear DNA content among species.

This study involved a number of working hypotheses which included the following:

i. The various genotypes studied were morphologically different.
ii. The same genotype was called differently by the various ethnic groups.
iii. The local accessions had wide inter- and intra-group diversity in ploidy level nuclear DNA content.
iv. There is a correlation between agro-morphological and cytological traits used in the two methods of characterization.

This thesis consists of five chapters excluding the general introduction. The review of literature is presented in Chapter one. Chapters two and three contain the methodologies, results, discussions and conclusions of morphological and cytological (using flow cytometry and conventional chromosome counting techniques) characterizations of the yam genotypes

respectively. Chapter four consists of study of the relationships between morphological and cytological traits using canonical correlation analysis. The overview of the two methods of classifications and implications for future research are presented in Chapter five.

References

Abraham, K. 1998. Occurrence of hexaploid males in *Dioscorea alata* L. *Euphytica* 99: 5-7.

Adesuyi, S.A. 1997. Curing technique for reducing the incidence of rot in yams. Technical Report No. 9. Nigerian Storage Research Product Institute. pp. 57-60.

Aké Assi, L. 1998. La diversification des utilisations des ignames: Usage pharmaceutique traditionnel. In: Berthaud, J., Bricas, N. and Marchand, J.L. (eds.), *L'igname, plante séculaire et culture d'avenir*. Actes du Séminaire International, 3-6 Juin 1997, Montpellier, France. Montpellier: Centre de Coopération Internationale en Recherche Agronomique pour le Développement (CIRAD). pp. 263-273.

Arnau, G., E. Maledon, and A. Nemorin. 2007. Genetic improvement of the greater yam *D. alata* through polyploidy breeding In: Wilfredo, C. and Wanda L.I. (eds.). *Marketing opportunities for agriculture and forestry products in the Greater Caribbean - A challenge for the 21st century*. San Juan: Caribbean Food Crops Society, Annual Meeting Caribbean Food Crops Society, 16-19, September, San José, Costa Rica. 112 pp.

Arumuganathan, K. and E.D. Earle. 1991. Estimation of nuclear DNA content of plants by flow cytometry. *Plant Molecular Biology Reporter* 9: 229-241.

Asiedu, R., N.M. Wanyera, S.Y.C. Ng, and N.Q. Ng. 1997. Yams. In: D. Fuccillo et al. (eds.) *Biodiversity in trust: Conversation and use of plant genetic resources in CGAIR centres*. Cambridge University Press, Cambridge, UK. pp. 57-66.

Chase, M.W. 2004. Monocot relationships: An overview. *American Journal of Botany* 91: 1645-1655.

Coursey, D.G. 1983. Yams. In: *Handbook of Tropical Foods*, H.C. Chan (ed.) Marcel Dekker, Inc. New York, USA. pp. 555-601.

Dolezel, J. 1997. Application of flow cytometry for the study of plant genomes. *Journal of Applied Genetics* 38: 285-302.

Dumont, R., A. Dansi, P. Vernie, and J. Zoundjihékpon. 2005. Biodiversité et domestication des ignames en Afrique de lôOuest. Pratiques traditionnelles conduisant á *Dioscorea rotundata*. CIRAD, IPGRI, Repères, France.

Egesi, C.N., M. Pillay, R. Asiedu, and J.K. Egunjobi. 2002. Ploidy analysis in water yam, *Dioscorea alata* L. germplasm. *Euphytica* 128: 225-230.

Eka, O.U. 1985. The chemical composition of yam tubers. In: *Advances in Yam Research*. G. Osuji, (ed.), Anambra State University of Technology, Enugu, Nigeria. pp 49-75.

Food and Agriculture Organisation (FAO), 2006. FAOSTAT Database FAO, Rome.

Food and Agricultural Organisation (FAO), 2007. Production Statistics. Retrieved December 24, 2007, from http://faostat.fao.org/site/567/default.aspx.

Frohne, D. and U. Jensen. 1998. Systematik des Pflanzereichs; unter besonderer Beruecksichtigung chemischer narkmale und pflanzlicher Drogen. 5. Auflage, Wissenschaftliche Verlagsgesellschaft mbH. Stuttgart, Germany.

Hamon, P., R. Dumont, J. Zoundjihekpon, B. Tio-Toure, and S. Hamon. 1995. Les ignames sauvages dôAfrique de IôOuest. Caracteristiques morphologiques. Wild yam in West Africa. Morphological characteristics. ORSTOM, Paris. 84 pp.

Institute of Agricultural Research (IAR), Annual Report, 2004. Roots, tubers and legumes crop production. IAR, Njala, Sierra Leone. 60 pp.

Manyong, V.M., R. Asiedu, and G.O. Olaniya. 2001. Farmersô perception of and actions on resource management constraints in the yam based systems of Western Nigeria. In: M.O. Akoroda and J.M. Ngeve (eds.) *Root Crops in the 21st Century, Proceedings 7th Triennial Symposium ISTRC-AB*, Cotonou, Benin Republic, 11-17 October 1998. pp. 156-167.

Martin, F.W. and S. Ortiz. 1963. Chromosome numbers and behaviour in some species of *Dioscorea. Cytologia* 28: 96-101.

Martin, F.W. and S. Ortiz. 1966. New chromosome numbers in some *Dioscorea* species. *Cytologia* 31: 105-107.

Mignouna, H.D. and A. Dansi. 2003. Yam (*Dioscorea* spp) domestication by the Nago and Fon ethnic groups in Benin. *Genetic Resources and Crop Evolution* 50: 519-528.

Mignouna, H.D., M.M. Abang, and R. Asiedu. 2007. Yams. In: *Genome Mapping and Molecular Breeding in Plants: Pulses, Sugar and Tuber Crops*. Volume 3. C. Kole (ed.) Springer- Verlag Berling Heidelberg. pp. 271-296.

Ng, N.Q. and S.Y.C. Ng. 1994. Approaches for germplasm conservation. In: Akoroda, M.O. (eds.), *Root crops for food security in Africa*. Proceedings of the fifth symposium of the International Society for Tropical Root Crops ï African Branch, Kampala, Uganda, 22-28 November 1992. International Society for Tropical Root Crops, African Branch, Kampala, Uganda. pp. 135-140.

Opara, L.U. 1999. Yam storage. In: *CGIAR Handbook of Agricultural Engineering Volume IV Agro Processing*. Bakker-Akkema (ed.). The American Society of Agricultural Engineers, St. Joseph, M.I, USA. pp. 182-214.

Orkwor, G.C., R. Asiedu, and I.J. Ekanayake. 1998. *Food Yams. Advances in Research*, IITA and NRCRI, Nigeria. 249 pp.

Tenkouano, A., J.H. Crouch, H.K. Crouch, and D. Vuylsteke. 1998. Ploidy determination in Musa germplasm using pollen and chloroplast characteristics. *Horticulture Science* 33: 889-890.

Vandenhout, H., R. Ortiz, D. Vuylsteke, R. Swenen, and K.V. Bai. 1995. Effect of ploidy on stomatal and other quantitative traits in plantain and banana hybrids. *Euphytica* 83: 117-122.

Zannou, A., A. Ahanchédé, P.C. Struik, P. Richards, J. Zoundjihékpon, R. Tossou, and S. Vodouhè. 2004. Yam and cowpea diversity management by farmers in the guinea-sudan transition zone of Benin, *NJAS – Wageningen Journal of Life Sciences* 52: 393-420.

Zannou, A., R.C. Tossou, S. Vodouhè, P. Richards, P.C. Struik, J. Zoundjihékpon, A. Ahanchédé, and V. Agbo. 2007. Socio-cultural factors influencing and maintaining yam and cowpea diversity in Benin. *International Journal of Agricultural Sustainability* 5: 140-160.

Zoundjihekpon, J., S. Essad, and B. Toure. 1990. Denombrement chromosomique dans dix groupes varietaux du complexe *Dioscorea cayenensis-rotundata. Cytologia* 55: 115-120.

CHAPTER ONE

Literature review

1.1 Taxonomy, morphology and floral biology of yam

1.1.1 Taxonomy of yam

The genus *Dioscorea* is divided into different sections, based on gross morphological traits (Burkill, 1960). A section is a group of species that is separated by peculiar characters from the others of the same genus. The five most important sections in yams are *Enantiophyllum, Lasiophyton, Combilium, Opsophyton* and *Macrogynodium* (Bai and Ekanayake, 1998). Five of the edible yam species such as *D. rotundata, D. alata, D. cayenensis* and the minor species, *D. opposite* and *D. japonica* belong to *Enantiophyllum* and are distinguished by clockwise twining on support (Bai and Ekanayake, 1998). *Dioscorea dumetorum* and *D. hispida* belong to *Lasiophyton*; *D. bulbifera* to *Opsophyton*; *D. esculenta* to *Combilium*; and *D. trifida* to *Macrogynodium* (Bai and Ekanayake, 1998). Members of these four sections twine anticlockwise (Table 1). Alexander and Coursey (1969) detailed the features that differentiate between the various sections. The features of members of the *Enantiophyllum* section include the formation of a large single tuber weighing 5 to 10 kg and 2 to 3 m in length. It is rare to have two or three tubers and extremely infrequent to have more than three tubers per season. The members of the *Lasiophyton* are distinguished by a cluster of medium sized tubers that are fused together. They have peculiar compound leaves, usually with three leaflets comparable to most *Dioscorea* leaves which are simple. *Dioscorea esculenta* is the only member of the *Combilium* section. It consists of a large number of small tubers, each weighing only a few hundred grams. The shoot system is of smaller stature compared to *Enantiophyllum* yams, while the root system resembles that of *Solanum tuberosum* (potato). The members of the *Macrogynodium* section are distinguished by a group of even smaller tubers compared to *D. esculenta*.

There is a dearth of information on yam phylogenetic relationships due to the difficulty in identifying species and the high level of polymorphism in morphological traits. For instance, a controversial relationship between *D. rotundata* and *D. cayenensis* was reported by various researchers (Burkill, 1960; Terauchi *et al*., 1992). They considered members of both species to be derived from a common ancestor. However, Ayensu (1971) argued that they are

different species based on the different anatomical structures he observed in them. Terauchi *et al*. (1992) investigating the origin and phylogeny of guinea yams using restriction fragment length polymorphism (RFLP) analysis of chloroplast DNA and nuclear ribosomal DNA suggested that *D. rotundata* was domesticated from either *D. abyssinica*, *D. liebrechtsiana*, *D. praehensilis*, or hybrids from any of the two. *Dioscorea cayenensis* was also proposed as a hybrid derived from pollination of a staminate plant of either *D. burkilliana*, *D. minutiflora*, or *D. smilacifolia* and a pistillate plant of either *D. abyssinica*, *D. rotundata*, *D. liebrechtsiana* or *D. praehensilis* (Terauchi *et al*., 1992).

Table 1. The main sections under the genus *Dioscorea* and corresponding cultivated species including their common names, origin and ploidy levels

Section*	Characteristics	Species	Common Name	Origin	Ploidy
Enantiophyllum	Vines twining to the right	*D. alata* L.	Water yam; Greater yam; winged yam	S.E Asia	2n=20,30,40, 50,60,70,80
		D. rotundata Poir.	White Guinea yam; White yam	W. Africa	2n=40;80
		D. cayenensis Lam.	Yellow Guinea yam; White yam	W. Africa	2n=36,54,60, 63,66,80,120, 140
		D. opposite Thumb.	Cinnamon yam	China	2n=40
		D. japonica Thumb.	Chinese yam	Japan	2n=40
		D. transversa R.Br.		SE Asia	-
Lasiophyton	Vines twining anticlockwise	*D. dumetorum* (Kunth.) Pax	Bitter yam Trifoliate yam; Cluster yam	Africa	2n=36,40,45, 54
		D. hispida Dennst.	Asiatic bitter yam	SE Asia India	2n=40,60
Opsophyton		*D. bulbifera* L.	Aerial yam; potato yam	Africa Tropical Asia	2n=30,40,50, 60,70,80,100
Combilium		*D. esculenta* (Lour.) Burkill	Lesser yam Asiatic yam	Indo-China Oceania	2n=30,40,60, 90,100
Macrogynodium		*D. trifida* L.f.	Cush-cush yam	Tropical America	2n=54,72,81

Sources: Coursey (1967); Alexander and Coursey, 1969; Purseglove (1972); Rehm and Espig (1991); Degras (1993); Onwueme and Charles (1994); Asiedu *et al*. (1997). *a section is a group of species separated by some distinction from others of the same genus.

Based on recent analysis of morphological and molecular data, the Dioscoreaceae family is now considered to have four distinct genera including *Dioscorea, Stenomeris, Trichopus* and *Tacca* (previously known as *Taccaceae*) (Chaddick *et al*., 2002). Various researchers have

also observed that some species of *Dioscorea* originated from a common ancestor known as the monophyletic group (Kawanabe *et al*., 1997; Wilkin *et al*., 2005). Both *D. tunuipes* and *D. tokoro*, for example, clustered in one of the monophyletic groups, while other species formed separate monophyletic groups. The different clustering patterns indicated that species in the same cluster may have evolved from a similar ancestor, whereas those obtained in different clusters were possibly from different ancestors. These observations have led to the suggestion that yams should be reclassified (Chair *et al*., 2005).

1.1.2 Morphology of yam

Yam leaves are commonly simple, cordate, or acuminate borne on long petioles, but lobed or palmate types could exist in some species with pointed tips (Okonkwo, 1985). Different leaf arrangements such as alternate, opposite or both may occur on the same stem depending on the plant species. For instance, *D. rotundata* has simple cordate leaves oppositely arranged on the nodes. *Dioscorea dumetorum* has compound leaves which are different from other species that have simple leaves with opposite or alternate leaf arrangement on the stem. Leaves consist of reticulate veins, unserrated lamina and are non-pubescent (Frageria, 1992). Yam leaf anatomy consists of stomata on the lower leaf epidermis (Okonkwo, 1985), except for the *D. bulbifera* with few stomata occurring on the upper leaf epidermis (Onwueme, 1978).

The top growth of yam consists of twining vines which may be several meters long, depending on species and growing conditions (Hahn *et al*., 1987). Vines of some species have spines which support twining and deter animals (Okonkwo, 1985). Spines are more common in wild than cultivated yams (Onwueme, 1978). Stems of most species are cylindrical, but *D. alata* comprises of stellate, rectangular or polygonal structures with angular extension of membranous wings forming a four sided cross section (Onwueme, 1978). A small number of the minor species of *Dioscorea* has dwarf plant architecture (IBPGR, 1980). Dwarf genotypes have been noted among *D. rotundata* with mean vine length of 1.4 m compared to 19.8 m non-dwarf cultivars (Abraham *et al*., 1989).

Yam plants possess two underground structures, the fibrous roots and the storage tubers in which starch accumulates. Roots arising from the tubers are normally thin and short (Okezie *et al*., 1981). The fibrous roots are generally smooth except in spinate vine cultivars. The number, shape and size of tubers are genotype and species dependent. *Dioscorea rotundata* tubers are commonly large and cylindrically shaped with white flesh consisting of ovoid,

large starch grains as a variant from other species (Okonkwo, 1985). Yam tubers have a peculiar longer dormant period of 10 to 16 weeks compared to other root and tuber crops (Orkwor *et al*., 1998).

1.1.3 Floral biology of yam

Flowering in many edible yams has been reported to be erratic, sparse or completely absent in some genotypes thereby limiting yam hybridization (Egesi *et al*., 2002). The sticky nature of yam pollen and the small openings of the female flowers limit wind pollination in yams (Sadik and Rockwood, 1975). In *D. rotundata*, Sadik and Okereke (1975a) reported the occurrence of bisexual flowers on the same spike. In *D. cayenensis*, however, only male flowers have been reported (Hahn, 1988). Observations in *D. alata* indicate the existence of few pistillate flowers and many staminate accessions used in hybridization (IITA, 1993). Sadik and Okereke (1975a) outlined the following characteristics of staminate, pistillate and complete flowers of yam. The florets of staminate flowers are 1 to 3 mm in diameter, sessile and borne on spikes subtended by small bracts. The number of florets on each raceme is variable. At least one spike is formed at a leaf axil and usually droops downwards. The perianth is slightly connate at the base and consists of three light-green sepals and a corolla of three light-yellow petals. Sepals and petals are usually similar in size and colour. The androecium consists of two whorls of each stamen. Pistillate flowers measure about 0.5 cm long, and are borne on axillary spikes. The perianth consists of three green sepals and three yellow-green petals. The sepals and petals are lobed above the ovary or otherwise they resemble those in staminate flowers. The ovary is inferior and trilocular with each locule containing two ovules. The placentation of the ovary is axial, and continues to develop into a capsule, whereas the perianth dries out during maturation (Sadik and Okereke, 1975a). The structure and shape of complete flowers are similar to pistillate flowers except for the presence of two whorls of stamens as in staminate flowers. It is presumed that complete flowers are merely advanced forms of pistillate flowers in which the staminodes develop into functional stamens (Sadik and Okereke, 1975a).

Prevalence of staminate flowers has been observed in imperfect sex separation (Akoroda, 1981). *Dioscorea rotundata* plants which originated from true seeds had a high frequency of flowering at 80% and a ratio of staminate to pistillate flowers of approximately 1:1, with 4% of the plants monoecious and presenting large number of flowers per plant (Sadik, 1975).

Yams in West Africa seldom reach more than 50% flowering, and the flowering genotypes exhibit a high staminate to pistillate ratio of 40:1 (Waitt, 1964) and monoecious to pistillate ratio of 5:1 (Sadik and Okereke, 1975b). Abortion or reduction of sex organ primordia of most species results in unisexuality (Dellaporta and Calderon-Urrea, 1993). Sexuality can, however, be reversed by hormonal treatment in certain conditions (Dellaporta and Calderon-Urrea, 1993). This erratic flowering pattern and sex ratios are influenced by ecological factors such as ratio of day- to night-length, light intensity, soil mineral balance, length of vegetative to reproductive phase and genetic factors (Degras, 1977). Flowering intensity varies among yam genotypes ranging between non flowering and profuse flowering. It is usually highest in staminate than in pistillate plants of *D. rotundata* and *D. alata* (Bai and Ekanayake, 1998).

In general, yam is a short day plant with diverse photoperiod requirements for flowering (Arnolin, 1982). Time of planting, quality and sett size of planting materials play crucial roles in flowering time. Yam setts are pieces obtained by cutting the mother- or ware-yam. An investigation into the effect of planting dates and types of setts on flowering in Ibadan, Nigeria, showed that setts of *D. rotundata* planted in January and April, flowered in early and late June, respectively (Edem, 1975). Also, setts planted in January produced more flowers and spikes than those planted in April (Edem, 1975).

1.2 Breeding scheme of yams

Breeding of elite genotypes of yams with adaptable ecological fitness of prevalent pest and disease attacks, desirable food quality traits and stable yields are needed to increase and sustain the productivity of yam cultivation in the face of a deteriorating resource base (Orkwor *et al.*, 1998). In the past, farmers based their selection of suitable genotypes on natural variation but changes in the physical and socioeconomic environments have necessitated the development of scientific breeding programmes. New yam cultivars are developed according to specific objectives with the understanding that the desirable traits vary from one species and region to another. Generally, the principal objectives of most yam improvement programmes include high and stable yield of marketable tubers; good tuber quality such as high dry mass content; culinary traits including texture, taste, dormancy period, rate of enzymatic browning; resistance to biotic stresses in the field and during post-harvest storage; tolerance to abiotic stress such as drought

and low soil fertility; and suitability of plant architecture, vigour, and maturity period to prevailing cropping systems (Orkwor *et al*., 1998).

Breeding and selection of yams include the introduction of germplasm, evaluation and phenotypic mass selection of clones. Yam improvement programmes often focus on hybridization as a short-term approach to broaden the cropôs genetic base. Cytogenetic or genetic research is often considered as a medium-term approach to increase efficiency; while biotechnological or molecular approaches are considered long-term techniques for yam improvement (Orkwor *et al*., 1998). Breeding and selection use large numbers of progeny to ensure that the few traits considered most relevant are incorporated in the local ideotype. A local ideotype is a description of idealized appearance of a plant variety. For this purpose, the selection of parents for hybridization is done such that mostly complementary, desirable genes are combined in the new hybrid cultivar. Selected parents are crossed to generate botanic seeds either using hand pollination on pairs of parents, or open pollination among clones planted in isolated fields, clonal trials or farmersô fields (Orkwor *et al*., 1998).

The selection of progeny is largely based on visual assessment developed according to the breeding objectives. The selection intensity during the first clonal generations is high in order to effectively discard undesirable genotypes (Lebot *et al*., 2005). The number of clones of each selected genotype increases in succeeding generations to facilitate precise evaluation of their agronomic performance (Figure 1). The base population includes both local and introduced germplasm. The base population is established in hybridization blocks where crossing is permitted for botanic seed production. The early clonal and preliminary yield trials are unreplicated since large numbers of clones and stands per clone are used. Clones at the uniform and advanced yield trials are established using replicated complete or incomplete block designs. The number of replicates depends on the quantity of seeds available. At the yield trial stages, genotypes can be either provided with or without stakes (Orkwor *et al*., 1998).

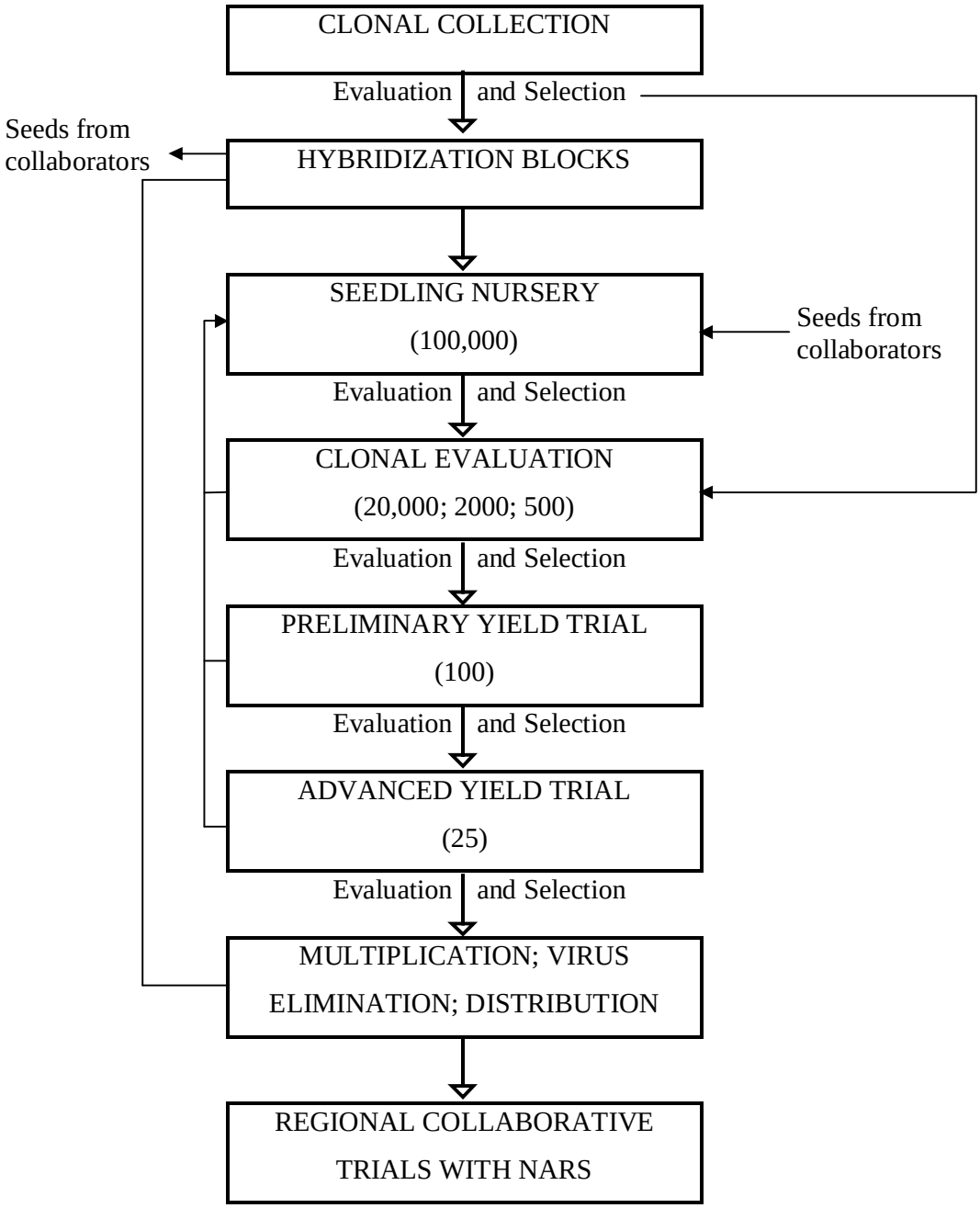


Figure 1. Yam improvement scheme showing approximate number of clones in parenthesis (Asiedu *et al*., 1998)

1.3 Genetic diversity and its significance in yams

The amount of genetic variation among individuals of a genotype, species or population, which provides adaptability to erratic environmental conditions and the potential to develop new genotypes is known as genetic diversity (Brown, 2000). Such variability among genotypes is expressed through molecular, (eg. DNA sequence), biochemical, physiological, cytogenetic and morphological traits (Ramanatha and Hodgkin, 2002). Thus, an extensive germplasm assessment and characterization involves measurement of more than one of these traits.

Increasing research into the genetic diversity of yams has contributed to an understanding of the extent and distribution of diversity present in cultivated genotypes and their wild relatives. This is due to threat of genetic erosion of valuable local and introduced genetic resources at the crop domestication and diversity centers (Lebot *et al*., 2005). Nevertheless, the yam is among cultivated staple crops considered as underutilized, minor or neglected where its conservation status and production potential is still to be unraveled in many areas (Tamiru *et al*., 2006).

The general belief is that yam diversity has been considerably maintained in traditional agro-ecosystems and through sustainable dynamic evolutionary processes (Kehlenbeck and Maass, 2004; Tamiru *et al*., 2006). Additionally, human knowledge that shaped diversity for generations is preserved (Bellon, 1991). The key players involved in crop evolution include genetic diversity, farmersô knowledge and selection, and exchange of crop varieties (Brush, 2000). Farmers often treasure diversity in crops wrought by factors of heterogeneous environment and production, risk, market demand and supply, which affect how different products are utilized (Bellon, 1996). This is evident in farmersô decisions about which cultivar to grow belonging to similar or different species. Such preferences in the development and utilisation of traditional varieties or landraces has influenced specific and intra-specific diversity in yams (Tamiru *et al*., 2006).

The concept of a landrace is complex. A landrace is defined as an integrated and adapted population that is genetically variable (Harlan, 1975). It is also referred to as a crop population in balance with its agro-ecological environment, stable over a long period of time with a potential for adaptive changes (Frankel, 1970). Landrace also refers to an early cultivated form of a crop species, evolved from a wild population (FAO, 1999). The presence

of diversity in genotypes plays a significant role in providing food for millions of people and as parent populations for breeding modern cultivars (Orkwor *et al*., 1998).

1.4 Yam genetic diversity with respect to morphological and ploidy markers

There is considerable agreement in research results that all cultivated forms of *D. cayenensis-D. rotundata* complex are products of ancient domestication of the four major wild species namely *D. abyssinica* Hochst, *D. praehensilis* Benth, *D. burkilliana* Miege, and *D. mangenotiana* Miege (Dumont and Vernier, 2000; Mignouna and Dansi, 2003; Scarcelli *et al*., 2006). But, the challenge still remains of understanding how individuals identified in the wild as *D. praehensilis* or *D. abyssinica* can directly become *D. rotundata* or *D. cayenensis* following ñdomesticationò without any genetic change (Mignouna and Dansi, 2003).

Yam shows considerable diversity both at inter- and intraspecific levels (Okoli, 1991). The diversity under cultivation is further enhanced by the ongoing domestication of wild yam in various countries (Scarcelli *et al*., 2006). Nevertheless, the extent of genetic diversity in many *Dioscorea* species and their relationships is yet to be investigated in detail. Characterization of yam germplasm based on morphological characters (Dansi *et al*., 1999), soluble protein profiles (Ikediobi and Igboanusi, 1983) and isozyme patterns (Dansi *et al*., 2000a) has revealed some degree of variability. Morphological characterization is necessary as a first step in a plant breeding programme. Many yam genotypes cultivated in the past, can nowadays only be found among a few farmers in small quantities. Moreover, the wild species are possibly experiencing severe erosion due to the disappearance of forest reserves. As a consequence, the genetic resources of yam are at risk (INRAB, 1996). The identification and conservation of new elite genotypes with ecological adaptation and resistance to pests and diseases will provide plant breeders with a wider range of diversity. This can be facilitated by establishing a molecular database of the diversity of existing traditional cultivars held by farmers (INRAB, 1996).

Chromosome counts are variable in yams, ranging from 2n = 20 to 2n = 140 in the common food species (Hahn, 1995). In yam the occurrence of extra chromosomes, B chromosomes or satellites, is common and sometimes they are as large as the chromosomes themselves (Essad, 1984). Various chromosome counts in *Dioscorea* revealed the existence of one or two chromosome base numbers, x = 9 and x = 10, with a high frequency of polyploid species

(Zoundjihekpon *et al*., 1990; Dansi *et al*., 2000b). Tetraploid genotypes are often the most frequent, compared to 2*x*, 6*x* and 8*x* genotypes. Mixoploid formation could be possible in yam, though infrequent. Among the 90 cultivars assessed for ploidy diversity in Benin, two landraces, õTam-Samô and õYoubeô, showed 4*x* and 8*x* mixoploid (Dansi *et al*., 2000b).

Different marker assisted techniques have been explored for ploidy determination in yam. In segregating populations of water yam (*D. alata*) and white yam (*D. rotundata*) using RAPD markers, Mignouna *et al*. (2002) observed disomic inheritance with 2n = 4*x* = 40, indicating that both species were allotetraploid. However, analysis using isozyme and microsatellites markers, revealed *D. rotundata* as a diploid species with 20 chromosomes (Scarcelli *et al*., 2005), whereas *D. trifida* was classified as octoploid with 80 chromosomes (Essad, 1984). In microsatellite segregation analysis, individual patterns showed a maximum of four alleles suggesting *D. trifida* to be tetraploid with 2n = 4*x* = 80 chromosomes (Hochu *et al*., 2006). Further cytogenetic investigation indicated that *D. trifida* is an autotetraploid species with a basic chromosome number of *x* = 20 (Bousalem *et al*., 2006). Segarra-Moragues *et al*. (2004) studying two species of the Bordera section, *D. pyrenaica* and *D. chonardii* (classified in the section by Chaddick *et al*. (2002)) confirmed that they are allotetraploid endemic to the Pyrenees (Spain and France). Both *D. pyrenaica* and *D. chonardii* have not been well documented in the Dioscoreaceae family. The two new basic chromosome numbers *x* = 6 and *x* = 20 also raised concerns about the validity of ploidy data in the genus *Dioscorea* (Segarra-Moragues and Catalan, 2003; Scarcelli *et al*., 2005).

1.5 Background and relevance of DNA flow cytometry

According to Robinson (2006), ñ*a flow cytometer consists of fluidics, optics and electronics, as it measures cells in suspension that flow in single-file through an illuminated volume where they scatter light and emit a fluorescence that is collected, filtered and converted to digital values for storage on a computer*.ò Deoxyribonucleic acid (DNA) flow cytometry estimates the amount of DNA in cell nuclei. This technique involves preparation of aqueous suspensions of intact nuclei with DNA stained in DNA fluorochrome and the quantification of DNA content is based on the relative florescence intensity of the samples (Dolezel and Bartos, 2005). Since sample preparation and analysis are convenient and rapid, DNA flow cytometry has gained popularity in ploidy screening, detection of mixoploidy and anueploidy and in cell cycle analysis (Dolezel and Bartos, 2005). Flow cytometry is also used in the assessment of the degree of polysomaty, the occurrence of nuclei of different ploidy levels in

the same organism, found in varying cells or tissues (Inze and De Veylder, 2006). This is also known as endopolyploidy, which occurs as a result of replication of DNA without the process of mitosis. However, there is still limited understanding regarding the extent, role and control of endopolyploidy in plants (Bennett, 2004). Through flow cytometry, the reproductive processes can be studied. For instance, the ploidy of pollen nuclei can be determined by identifying and quantifying the unreduced gametes produced (Kron *et al*., 2007). Characterization of other pollen traits such as the proportions of male and female determining pollen types in dioecious plants is made possible with flow cytometry (Stehlik *et al*., 2007). Flow cytometry may be used in screening of seeds where the ploidy levels of embryos and endosperms are separately determined. Such an exercise provides the possibility of associating particular seed traits with DNA content (Smarda and Stanciik, 2006). The estimation of absolute DNA amount or genome size is another application of flow cytometry. The genome size is the description of the DNA content in picograms per haploid genome and is often referred to as the C-value (Dolezel and Bartos, 2005). Flow cytometry can detect interspecific hybrids according to intermediate DNA values (Dolezel, 1997). This has been applied to detect hybrids in *Allium* spp (Keller *et al*., 1996). Karyological stability of somatic hybrids produced by protoplasm fusion can be assessed by flow cytometry (Binsfeld and Schnabl, 2002). Flow cytometry can be used to detect new cytotypes (Weiss *et al*., 2002); and in cell cycle kinetics (Sandoval *et al*., 2003).

Previous efforts aimed at estimating the quantity of DNA in cell nuclei preceded the discovery of DNAôs central role in heredity (Caspersson and Schultz, 1938). Afterwards, DNA amount per organism was established; and the DNA content of an unreplicated haploid chromosome complement (n) was termed the C-value (Swift, 1950). Thus, a nucleus in G_1 phase of the cell cycle with two copies of the unreplicated genome has a 2C DNA amount. Later, no correlation was observed between DNA C-values and organism complexity (Mirsky and Ris, 1951). The lack of correlation was termed the C-value paradox (Thomas, 1971).

Two methods developed to facilitate the determination of 2C DNA content in organisms include chemical analysis (Schmidt and Thannhauser, 1945) and reassociation kinetics (Britten and Kohne, 1968). The second approach (single cellular) had high reproducibility and performance, but was more complex to handle compared to the first. Research aimed at eliminating errors caused by irregularly shaped nuclei and chromosomes with non-homogeneously stained chromatin led to the development of scanning

microspectrophotometry (Deeley, 1955). The DNA image cytometry was apparently considered an electronic alternative to microspectrophotometry, since results obtained from different laboratories were accurate and reproducible (Vilhar *et al*., 2001).

Beside microspectrophotometry and image cytometry, a new technology with wider and more efficient application was developed. Flow cytometry analyses microscopic particles in suspension and compels their mobility in a single fluid stream by focusing intense light on the particles (Dolezel and Bartos, 2005). Pulses of scattered light and fluorescence are collected and converted to electric current pulses by optical sensors and are classified based on number of pulses received. The technique allows analysis of single particles at high speed, and measurement of large populations and detection of subpopulations within short periods (Shapiro, 2003). Since there is no need to use tissues with dividing cells, the ease of sample preparation, and the ability to measure DNA quickly in large populations of cells, has made flow cytometry an attractive alternative to microspectrophotometry (Bennett *et al*., 2000).

The first flow cytometers quantified DNA in human cells by measuring absorbance of UV light (Kamentsky *et al*., 1965). This method was replaced by florescence technology, and nowadays DNA is determined indirectly by measuring fluorescence emission (Dolezel and Bartos, 2005). The result of the analysis is usually displayed in the form of a histogram of relative fluorescence intensity (RFI), from which DNA content is estimated (Dolezel and Bartos, 2005). Because large populations of cells may be measured in a short time, DNA flow cytometry has been used extensively in the monitoring of cell cycle kinetics and its perturbations (Rabinovitch, 1994), biomedical research to detect anueploidy (Kawara *et al*., 1999), and apoptosis (Vermes *et al*., 2000). Apoptosis is a process of cell death, which occurs naturally during normal development, maintenance and renewal of tissue in an organism (FAO, 1999).

The application of flow cytometry in plants was hampered by difficulties in the preparation of suspensions of intact cells and nuclei suitability for the technique. The first breakthrough occurred when Heller (1973) prepared a suspension of field bean nuclei from alcohol acetic acid-fixed root tips after enzymatic treatment with pectinase and pepsin. Nuclear DNA was stained with ethidium bromide and the analysis of relative fluorescence intensity indicated a potential for analysis of cell cycle kinetics. Hellerôs (1973) application was not actively pursued by other plant scientists for about a decade, either due to its expensiveness or its applications largely restricted to biomedical research.

The possibility of nuclear DNA content quantification within intact plant cells started gaining attention during the early 80s (Puite and Ten Broeke, 1983). However, the presence of a rigid cell wall, which is autofluorescent and confers an irregular cell shape that disturbs the fluid stream, makes isolated plant cells unsuitable for estimation of DNA content using flow cytometry. Removal of the cell wall using cellulases or pectinases (hydrolytic enzymes) in the presence of an inert osmoticum converts cells to protoplasts, which are spherical and behave regularly within the flow stream. Although nuclear DNA could be stained in plant protoplasts (Puite and Ten Broeke, 1983), the histograms of fluorescence intensity could not be interpreted in terms of cell cycle distribution. This may be due to the effect of cytoplasmic autofluorescence and low permeability of the plasma membrane. Fixation with ethanol-acetic acid permeates cell membrane and decreases the autofluorescence. Notwithstanding, the quality of resulting histograms is poor, possibly as a result of misalignment of the instrument (Galbraith, 1990). A more successful method depends on intact nuclei analysis, which may be released from protoplasts by lysis either in the presence of a detergent or a hypotonic medium, leading to very good histograms of DNA content (Ulrich *et al*., 1988).

1.5.1 In-field application of DNA flow cytometry

The successful application of flow cytometry and its increasing utilization in plant taxonomy, systematics and ecology may present some interesting challenges. Like most analytical methods, the materials used by flow cytometry analysis are sampled and then dispatched to the laboratory (Dolezel *et al*., 1998). Cultivating plants within reasonable distance to the laboratory may reduce the deterioation of plant samples. However, difficulties set in where materials have to be transported over great distances and/or maintained or preserved for any length of time (Dolezel and Bartos, 2005). Leaf samples, the most popular tissue for DNA flow cytometry, may be transported in humid paper tissue and kept at low temperature. However, dispatching is not feasible for leaf samples of some species that deteriorate rapidly (Dolezel and Bartos, 2005).

Application of ñwheel-barrowò or ñbushò flow cytometry in plant analysis was first proposed by Brown (1993). At the time it seemed like an impractical proposition, but todayôs compact and portable flow cytometers operating off a single 12 V car battery have made the establishment of a field laboratory for on-site sample preparation and flow cytometry analysis a reality. Marine biologists analyzing phytoplankton usually have portable flow cytometry laboratories aboard research vessels (Sosik and Olson, 2002) or make use of cytometers that

free float in the ocean (Dubelaar and Gerritzen, 2000). These developments have not solved all the difficulties associated with field DNA flow cytometry. The cost of transporting or establishing the laboratory may prohibit the application of flow cytometery in certain areas. Costing involves growing or transporting a series of plant standards, running preliminary experiments to identify an optimal sample preparation protocol, and testing for interference of cytosolic compounds with DNA staining. Despite these challenges, bush flow cytometry is a very attractive tool for ploidy screening on-site, where a similar protocol could be applied to screening hundreds or thousands of accessions representing a number of species (Dolezel and Bartos, 2005).

1.5.2 Estimation of nuclear genome size

The genome size of an unknown sample is estimated after comparison with the nuclei of a reference standard with known genome size. This can be obtained through flow cytometry, which analyses the relative florescence intensity (RFI), and hence relative DNA content (Dolezel and Bartos, 2005). The measurements of RFI of stained nuclei are determined on a linear scale with analysis of 5000-20000 nuclei per sample. Estimation of absolute DNA amount of a sample is based on the values of the G_1 peak means (Galbraith *et al*., 1998). These are means obtained at the G_1 phase, the initial proliferative phase of the cell cycle (FAO, 1999). The absolute DNA content of a sample is conventionally estimated in pg DNA based on the values of the G_1 peak means as:

Sample 2C DNA content
= [(sample G_1 peak mean)/ (standard G_1 peak mean)] x standard 2C DNA content (pg DNA).

Since the advent of molecular biology and progress in genome sequencing, DNA amount has been reported in number of base pairs (bp) using the term genome size (Greilhuber *et al*., 2005). Genome size lacks a precise definition; it has either been used to describe the DNA amount in G_1 phase nucleus or unreplicated haploid chromosome set (n). This problem is exacerbated in polyploids, where genome size is used to describe the haploid (n) and monoploid (x) chromosome set(s) (Greilhuber *et al*., 2005). Some authors were previously using 0.965×10^9 bp to 0.980×10^9 bp: 1 pg DNA for the estimation of the mean relative mass of a nucleotide pair (Straus, 1971; Cavalier-Smith, 1985). Recently, a 1:1 ratio of AT: GC pairs (1 pg DNA = 0.978×10^9 bp) ignoring the presence of modified nucleotides in the DNA molecule and maintaining errors <1% has been reported (Dolezel *et al*., 2003).

Estimation of the genome size of a species involves random selection of plants with each analysed several times (Suda *et al*., 2003). Carrying out replicate measurements on a plant would enhance detection of diversity in the procedure, while analyzing several plants permits monitoring of intraspecific variation. The number of plants and replicated measurements differ among various studies. With smaller number of replicates in large-scale screening experiments, at least three plants should be analysed, and each thrice when intraspecific genome size diversity is studied (Suda *et al*., 2003). To determine accurate and reliable genome size, the nuclei must be isolated in sufficient quantity; the DNA staining must be specific and stoichiometric for both the target and standard nuclei; and the genome size of the reference standard must be known (Dolezel and Bartos, 2005).

1.5.3 Preparation of nuclei suspensions and optimisation of DNA content histograms

Preparation of intact nuclei suspensions for the estimation of absolute DNA amounts has been almost globally carried out following the approach of Galbraith *et al*. (1983). This technique involves the release of nuclei into a nuclei isolation buffer by mechanical homogenization of a small sample of fresh plant tissue (Galbraith *et al*., 1983). The isolation buffer composition is crucial in enhancing the release of sufficient nuclei free of cytoplasm and maintaining the integrity of isolated nuclei; protecting the nucleic DNA against endonucleases; and facilitating DNA staining (Dolezel and Bartos, 2005).

The Otto (1990) procedure was applied to plant flow cytometry and adjusted for use with non-fixed nuclei by Ulrich and Ulrich (1991) and Doleel and Gohde (1995), respectively. In this method, nuclei are released into the Otto I buffer, where they are fixed by citric acid. Staining is done in a mixture of Otto I and Otto II buffers (1:4), both comprising a phosphate/ citric acid buffer of pH 7.3. Most plant species may yield unsurpassed resolution in DNA content histograms probably due to the citric acid reaction, which improves chromatin accessibility and homogenizes chromatin structure. This may result in the canceling of variations in staining intensity among populations of nuclei with similar DNA content but different chromatin states (Dolezel and Bartos, 2005).

The quality of DNA content histograms depends on sample preparation, instrument alignment and the data analysis. In order to minimise the amount of debris and clumps during analysis,

samples should be prepared with the intention of obtaining single cells (Dolezel and Bartos, 2005). It is also important to check the flow cytometer for linearity using flow check fluorospheres. This helps in the proper alignment of the instrument and in setting the coefficient of variation (CV) values to as low as possible before analysing the samples. If the cells or nuclei concentration in the prepared sample is high, a sufficient dye should be added to facilitate stoichiometric or proportionate binding. The CV is estimated as the ratio of the standard deviation to the peak mean expressed as a percentage; where the peak mean represents the mean channel number of the peak. The smaller the CV of the peaks in the DNA histogram, the more reliable is the estimation of ploidy level, and the better is the estimate of the percent of cells present in the various stages of the cell cycle. Doleel and Gohde (1995) showed that histograms with peak CVs <1% may be obtained under specific conditions such as leaving fixed cells overnight at 4°C, and allowing sufficient time for RNase to remove all the double stranded RNA from fixed cells stained with propidium iodide. In practice, CVs <3% are acceptable for most crops; however, in recalcitrant species where such precision is hindered, CVs <5% are acceptable (Galbraith *et al*., 1998).

Previous DNA content estimation work using flow cytometry was done by utilization of different fluorescent dyes to stain nuclear DNA such as ethidium bromide, mithramycin and Hoechst dyes (Puite and Ten Broeke, 1983). Lack of knowledge of AT:GC ratio of the standard and sample DNA resulted in inadequate estimates of absolute DNA amounts. Mithramycin has not been frequently utilized (Galbraith *et al*., 1983). However, 4ϕ6-Diamidino-2-phenylindole (DAPI), which binds to AT-rich regions, has been widely used due to two reasons. Firstly, DAPI is specific for double-stranded DNA and its binding to DNA is not influenced by chromatin structure, which results in low peak CVs similar to that obtained with Hoechst dye (Cowdon and Curtis, 1981). Secondly, many plant scientists preferred using arc-lamp-based flow cytometers because DAPI fluorescence was particularly easy to excite and measure (Dolezel and Bartos, 2005). The DNA binding properties of propidium iodide (PI) stained samples revealed similar results to Fuelgen microspectrophotometry, but were contrary to DAPI stained samples (Dolezel *et al*., 1998). However, optimum dye concentration for specific species, shorter staining period (2 to 20 min), isolation of nuclei from tissues of similar metabolic and developmental state remains crucial to good results (Galbraith *et al*., 1998).

1.6 Polyploid induction and expression in yams

Polyploidy involves the duplication of single genome (to produce autoploids) or the combination of at least two different genomes (to form alloploids) (Grant, 1981). Polyploidy normally occurs in cells of organisms when there are more than two sets of homologous chromosomes. Most organisms are diploid, having two sets of chromosomes, with each set inherited from each parent. A somatic chromosome doubling event in the zygote, developing seedlings, or active apical meristematic tissues is one of the mechanisms of ploidization. It is well noted that the union of two unreduced gametes as a result of meiotic mishaps would immediately produce polyploids whereas somatic chromosome doubling is rare (Ramsey and Schemske, 1998; Nasrallah *et al*., 2000; Grant, 2002). It is also believed that other ploidy levels such as triploid and pentaploid are produced by the union between reduced (n) and unreduced (2n and 4n) parental gametes respectively (Ramsey and Schemske, 1998). Husband (2004) reported that back-crossing of eggs from viable and fertile triploids to a normal diploid may result in tetraploid formation.

Polyploids arise from two main mechanisms: asexual or somatic polyploidization and sexual polyploidization (Carputo *et al*., 2003). In the asexual type, chromosome restitution occurs during mitosis such that one daughter nucleus comprises all the chromosomes of a somatic cell thereby yielding a cell with doubled number of chromosomes. However, scanty information is known about the natural frequency of somatic chromosome doubling in plants and their effects on interspecific hybridization (Carputo *et al*., 2003). As a result, somatic doubling was considered nonsignificant in the production of polyploids (Ramsey and Schemske, 1998). In sexual polyploidization, two 2n gametes or 1n and 2n gametes unite to produce polyploid embryos depending on the genotype and the prevailing environmental conditions. There is more evidence on the genetic and environmental factors affecting the frequency of 2n gamete production and their contribution to polyploidy evolution than asexual polyploidization (Carputo *et al*., 2003). It is noted that somatic doubling does not alter the genotype except for allelic redundancy and diplod plants with doubled chromosomes become inbred. Sexual polyploidy results in new allelic variation and gene combinations without the occurrence of inbreeding. Knowledge about the rate of fitness of sexual polyploids versus diploid progenitors in natural habitats could be intriguing. Such information may provide a better understanding of the role of polyploid formation in adaptation. Otto and Whitton (2000) assumed that both sexual and asexual polyploids adapt

faster than their diploid counterparts in the presence of beneficial mutations in small populations with a partial dominant effect on fitness.

Polyploids are also believed to arise from genomic duplication or hybridization of at least two different genomes to produce auto- or allotetraploids, respectively (Grant, 1981). Altogether, four types of polyploidy have been identified viz: strict autopolyploidy (from self replicating diploids), strict allopolyploidy (from hybridization between different diploids), segmental polyploidy (intermediate between auto- and allopolyploid) and autoallopolyploidy (Wendel and Doyle, 2005). Allopolyploidy has been observed to be more prevalent in nature than autopolyploidy (Wendel and Doyle, 2005). Both allopolyploidy and autopolyploidy are frequent in plants, including yam. These forms are also important to human nutrition, as are the intermediate types such as segmental allopolyploids (Hilu, 1993).

Polyploidization has had a tremendous influence on the evolution of plants as depicted by replicated genome, genomic sequencing and the extensive expressed sequence tag (EST) (Seoighe, 2003). Expressed sequence tags (ESTs) are short (200-500 nucleotides) DNA sequences or tiny portion of a gene that can be used in the identification of unknown genes and mapping their positions within the genome (NCBI, 2004). The inception of polyploidization also influences the transfer of genes or genetic materials between genomes and the loss of gene due to migration (Levy and Feldman, 2004). Such processes lead to the induction of epigenetic gene silencing with a significant impact on gene expression (Liu and Wendel, 2003). Genetic drift accompanying polyploid formation causes deviation in collinearity within similar plant species (Paterson *et al*., 2003).

De Wet (1980) and Ramsey and Schemske (1998) advanced factors that could influence polyploid formation and establishment via unreduced gamete production including adverse growing conditions, genotypes of varying populations and hybridization. The evolution of higher polyploidy via a triploid bridge is a significant challenge because triploidy results in reduced fertility due to meiotic irregularities yielding anueploid gametes. Polyploidy is also hampered by the triploid block, which is as a result of partial or complete failure of the endosperm tissue following certain interploidal or interspecific crosses (Bretagnolle and Thompson, 1995). Despite the potential for triploid block, the triploid bridge remains central for polyploidy formation (Ramsey and Schemske, 1998).

In polyploid plants, desirable traits such as disease could be found in wild ecotypes, a population or strain of an organism that is adapted to specific habitat (FAO, 1999). Heteroploid crosses of parents with different ploidy levels are normally done to enhance introgression of desirable genes into the cultivated types. However, heteroploid hybridization is frequently complicated by the generation of progeny of varying ploidy status (Costich *et al*., 1993). Crop species that require laborious somatic chromosome counting and have a long generation interval for meiotic analysis, require a quick and fast method of ploidy level establishment at the seedling stage. Determination of the phylogenetic nature of the variation in genomic DNA content among these species may serve as a guide to their prospective utility in breeding programmes (Costich *et al*., 1993).

A hybrid derived from two diploid species can be induced to form a stable allotetraploid via spontaneous chromosome doubling or by colchicine treatment. Alternatively, allotetraploid formation can be obtained by fusing two unreduced gametes from two diploids or by hybridization of two autotetraploids (Chen, 2007). Polyploid induction through colchicine applications also led to the emergence of octoploids and 16-ploids of *D. floribunda* (Martin and Ortiz, 1966). Martin and Ortiz (1966) further discovered that hexaploids derived from crosses between tetraploids and octoploids, as well as plants with 54 chromosomes including *D. composita* and *D. polygonoides* exhibited bivalent and trivalent chromosome pairings. Martin and Ortiz (1963) hypothesized, based on quadrivalent pairing of 54 chromosome lines of *D. floribunda*, that a backcross of an autotetraploid to one of its diplod progenitors accompanied by chromosomal doubling, further suggesting that Central American species are allotetraploids. Most of the hybrids of obtained from the cross between *D. floribunda* and diploid parent had normal chromosome pairing indicating genome homology of the species; and that speciation involved little karyotype differentiation. On the basis of high trivalent (17.269 III + 2.925 II + 2.342 I at M-1) and chiasmata frequencies in a triploid clone of *D. hispida* (2n = 60), Jos *et al*. (1977) concluded that the species was autotetraploid in nature; and considering the autosyndetic pairing nature suggested the basic chromosome number to be as low as eight (Jos *et al*., 1980). Bai and Jos (1978) studying the type of polyploidy in *D. esculenta* var. spinosa (2n = 90) recorded the chromosomal association, which consisted of hexavalents, quadrivalents, trivalents, bivalents and univalents with mean chromosomal association = 0.3 VI + 0.15 IV + 4.35 III + 32.65 II + 9.25 I at M-1. The occurrence of both low multivalent frequency and low chiasmata frequency led to the conclusion that *D. esculenta* is an autoploid, which is in the process of converting into an allopolyploid (Bai and

Jos, 1978). It is therefore apparent that in yams polyploidy often occurs due to abnormal meiotic division during M-1.

1.7 Role of polyploidy in yam improvement

The genus *Dioscorea* offers an attractive model for polyploid investigation and chromosome evolution in both wild and cultivated species, particularly regarding its vegetative propagation and the process of domestication (Bousalem *et al*., 2006). Several factors have played key roles in the duplication of genes and genomes. Among these included genome buffering, increased allelic diversity, increased or fixed heterozygosity and the possibility of novel phenotypic variation in duplicated genes (Udall and Wendel, 2006).

The emergence of multiple genomes in polyploid species was considered to affect mutation and recombination, with a resultant retardation of natural or artificial selection (Stebbins, 1971). This report assumed equality in mutation rates of diploids and polyploids and tetrasomic segregation ratios. Recent trends, however, suggest buffering of some mutational characters by genomic redundancy (Yamamori *et al*., 2000). Genetic or genomic redundancy occurs where two or more genes are performing similar function, whereby activiation of one gene will result in little or no effect on phenotypic expression. Genome redundancy is more common in higher than lower organisms (Nowak *et al*., 1997). Gene redundancy makes the genome vulnerable to accumulation of deleterious mutations (Nowak *et al*., 1997). On the contrary, instantaneous individual gene duplication, which follows polyploid formation produce dosage effects which are likely to impact phenotypic expression (Guo *et al*., 1996).

Genome doubling is also considered to increase allele variation and/or maintain heterozygosity. It is established that blocks of genes from different allelic backgrounds (intergenomic heterozygosity) could be hybridized to increase allele variation in polyploids (Udall and Wendel, 2006). Allopolyploidy is a major cause of increase allelic diversity (Udall and Wendel, 2006).

A distinctly peculiar phenotypic diversity is also a product of polyploid formation (Soltis *et al*., 2004). Despite the fact that most agro-morphological diversity in root and tuber crops including yams, is found within cultivars, most allelic variation is obtained in the wild gene pool (Lebot *et al*., 2005). However, these wild morphotypes have limited phenotypic diversity. The genetic relationships among 269 cultivars of *D. alata* originating from South

Pacific, Asia, Africa and the Caribean were studied (Lebot *et al*., 1998). A limited allelic diversity as indicative of the phenotypes was observed among these cultivars. The large similarity on one hand could be due partly to the evolution of many genotypes from a narrow genetic base, while the different genotypes were as a result of earlier sexual recombinations from different parents.

1.8 Molecular cytogenetics

The genus *Dioscorea* is made up of Old and New World species. The Old World species have the basic chromosome number of 10, while the New World species have the basic chromosome number of nine (Orkwor *et al*., 1998). A third basic number of 12 was proposed based on the occurrence of two, 24 chromosome species in the Pyrenees regions (Miege, 1954).

The problem of sparse, irregular or no flowering of some cultivated species of yam and the small chromosomal size has restricted cytogenetic analysis in *Dioscorea* species. As a consequence only little cytogenetic research has been done. In the available literature, Asian species are well documented (Raghavan, 1959; Martin and Ortiz, 1963; Ramachandran, 1968) compared to American species (Martin and Ortiz, 1966) and African species (Zoundjikhekpon *et al*., 1990; Orkwor *et al*., 1998).

Although polyploidy has been observed in most *Dioscorea* species, studies have extensively been restricted to determine chromosome numbers from somatic cells due partly to sparse, irregular or no flowering in some genotypes. Determination of the type of ploidy prevalent in the genus in relation to its evolution and distribution are yet to be done in many regions (Orkwor *et al*., 1998). Various workers have reported different races in *D. bulbifera* with: 2n = 36 and 54 (Miege, 1954); 2n = 40, 60, 80 or 100 chromosomes (Martin, 1974b); 10 African varieties with 2n = 36, 40, 54 and 60 chromosomes. Different races have also been reported in *D. esculenta* with: 2n = 40 (Martin, 1974a); 90 (Bai and Jos, 1978), 90 and 100 (Raghavan, 1958); in *D. alata* with 2n = 30, 50 and 70 (Martin, 1976); 52, 55, 66, 81 and 88 (Sharma and De, 1956); and *D. cayenensis* with 2n = 36, 54 and 140 chromosomes (Miege, 1954) (Table 1).

The chromosomes of *Dioscorea* species are too small ranging in size between 0.5 and 2.7 µ to conduct karyotype analysis since they are dot-like and the location of centromeric regions

is difficult to determine and view chromosomes. The size of the somatic chromosomes of the New World species was observed to range from 0.5-2.0 µ (Martin and Ortiz, 1963). Comparing the somatic chromosomes of *D. esculenta* and *D. hispida*, Ramachandran (1968) reported that the largest chromosome of the former measured 2.4 µ with the remaining found within the chromosome range of 0.75-1.9 µ, while the chromosome range of the latter was 0.9-2.7 µ. It was clear from this study that the highest chromosome number and the smallest chromosome sizes occurred in the tropical *Dioscorea* species; whereas the smallest chromosome number and largest sizes occurred in the temperate species (Orkwor *et al*., 1998).

Knowledge of the ploidy level of yam genotypes is imperative for yam breeders, especially new introductions, before their utilization in breeding programme, to enhance matching of ploidy levels and to facilitate ploidy manipulations in intraspecific crosses (Dessauw, 1988). The existence of various ploidy levels and the lack of diploid relatives in cultivated polyploid yams have contributed to making the genetic studies of the crop complex. Dessauw (1988) reported that variations in ploidy levels in yams are not reflected in the morphological traits compared to other plant species. Phenotypic variations were also noted to be higher within than between ploidy levels similar to other plant species. Thus, cytological aberrations causing erratic flowering and reproductive behavior are obvious (Dessauw, 1988).

Genetic improvement research was formerly, entirely devoted to the development of diploid varieties. Triploid (4*x* x 2*x*) and tetraploid (4*x* x 4*x*) hybrids were found to be more vigorous and to have a higher potential yield than diploid varieties (Arnau *et al*., 2007). The existence of octoploidy has been widely noted in both *D. alata* and *D. rotundata* (Hamon *et al*., 1992; Gamiette *et al*., 1999; Dansi *et al*., 2001), and mixoploidy in the *D. cayenensis*-*D. rotundata* complex where two cultivars were mixoploid with each exhibiting 4*x* and 8*x* ploidy levels (Dansi *et al*., 2000b). Further, Egesi *et al*. (2002) observed to a greater extent (84.9%) hexaploid and a lesser percent (15.1%) tetraploid in *D. alata*. Diverse research findings showed that males were usually tetraploids whilst females were mostly hexaploids (Abraham and Nair, 1991; Abraham, 1998). Contrarily, higher ploidy levels in yam were not found to be always directly correlated with erratic anthesis or female sex, but rather poor viability (Egesi *et al*., 2002).

A dearth of information exists on yam genome size determination. Chromosome number has been reported to be significantly correlated with nuclear DNA content (Costich *et al*., 1993).

Dansi *et al*. (2001) also observed that the relative nuclear DNA content in arbitrary units (AU), expressed as channel numbers varied from 43.4-52.2 AU for tetraploids; 65.4-76.9 AU for hexaploids and 85.5-100.3 AU for octoploids. Channel numbers are measured values of a parameter representing the signal intensity of an event after amplification. These values obtained at the various channels (0 to 1024) as relative fluorescence intensity, are arbitrary since they can be transformed into some other units such as picograms. Estimates of *D. alata* genome size was at 550 Mbp/1C (in million base pairs per haploid genome) and *D. rotundata* at 800 Mbp/1C with basic chromosome number, n = 10 (Mignouna *et al*., 2007). A recent report indicated mean 2C nuclear DNA contents ranged from 0.702 $\pm$ 0.004 pg for G_1 nuclei of diploid *D. dumetorum* to 2.573 $\pm$ 0.020 pg for G_1 nuclei of octoploid *D. cayenensis* (Obidiegwu *et al*., 2009).

1.9 Multivariate statistical techniques

Multivariate data can be obtained in all branches of science. For instance, a marketing researcher might be interested in identifying determinants of demand and supply of a specific product. A yam breeder might not only breed for high yielding genotypes, but other traits such as resistance to local pests and diseases, earliness and desired culinary qualities. A social scientist might be interested in studying relationships between the courtship conducts of teenage girls and their fathersô attitudes (Njuho, 2002).

The objects of applying multivariate techniques in scientific investigations include (i) data reduction or structural simplification; (ii) sorting and grouping; (iii) investigation of the dependence among variables; (iv) prediction and (v) hypothesis construction and testing (Manly, 1994). Multivariate techniques are applied where more than one variable is measured in an experimental unit. Since these variables could be correlated, the use of univariate analysis is inappropriate to extract relevant information (Njuho, 2002). Multivariate techniques are classified into two categories, namely variable-directed and individual or experimental unit directed. The variable directed techniques include principal component analysis (PCA), factor analysis (FA), canonical correlation analysis (CCA) and multiple regression analysis (MRA). The individual directed techniques include discriminant analysis (DA), cluster analysis (CA) and multivariate analysis of variance (MANOVA) (Njuho, 2002).

1.9.1 Principal component analysis

Principal Components Analysis (PCA) involves the identification of patterns in data, and expression of traits in data that highlight their similarities and variances (Smith, 2002). Principal component analysis technique is useful in discovering dimensionality of the data, data screening, checking clusters and finding abnormalities (Smith, 2002). In this technique, variables that are highly correlated are grouped together. The variables within a group are highly correlated whereas those between groups are uncorrelated. New variables are expressed as linear combination of the p original variables X_1, X_2, é ., X_p to produce uncorrelated indices Z_1, Z_2, é ., Z_p known as principal components (Manly, 1994). The first principal component contributes most to the variation in the data set and is related to the other components as: var $(Z_1) \geq$ var $(Z_2) \geq$ é $\geq (Z_p)$. The general PCA model is given as:
$Z_i = a_{i1}X_1 + a_{i2}X_2 +$ é .. $+ a_{ip}X_p$ where a_i = eigenvectors and X_p are the vaues of the different characters measured (Manly, 1994). The correlation between PCs and X-variables is given by $P(Y_i, X_k) = \frac{a_{ik}\sqrt{\lambda_i}}{\sqrt{\sigma_k^2}}$ where a_{ik} is the k element of a_i and $\mathring{u}^2$ is the standard deviation of X_k.

Principal component scores are relevant inputs in other multivariate analysis techniques (Njuho, 2002). Multicollinearity is among major problems encountered in multiple regression analysis, due to correlation between predictor variables. However, this can be avoided in PCA by using selected PC scores as regressors (Njuho, 2002). Plots of the first PC scores help to identify outliers and clusters that may be associated with the data (Njuho, 2002).

1.9.2 Factor analysis

Factor analysis follows the same principle of PCA. The main difference being that the former has distributional properties whereas the later does not. A few factors do explain the original variables without loss of information. When the new factors cannot be explained, rotation techniques, some which are orthogonal, are applied. The PCs selected using PCA can be used as the new factors (Manly, 1994).

The study of Charles Spearman, which explained the performance of students in relation to various school subjects, led to the development of the general factor analysis model as:
$X_i = a_{i1}F_1 + a_{i2}F_2 +$ ..... $+ a_{im}F_m + e_i$ (Manly, 1994). Where X_i is the i^{th} test score with mean zero and unit variance; a_{i1}, a_{i2}, ..., a_{im} are the factor loadings for the i^{th} test; F_1, F_2, ..., F_m are

m uncorrelated common factors, each with mean zero and unit variance; and e_i is a specific factor only to the i^{th} test, with mean zero, and is uncorrelated with the common factors. The model can be written as:

var (X_i) = 1= a_{i1}^2 var (F_1) + a_{i2}^2 var (F_2) + ... + a_{im}^2 var (F_m) + var (e_i) (Manly, 1994). Where a_{i1}^2, a_{i2}^2, é .., a_{im}^2 are the common factors whereas var (e_i) is called the specific factor of X_i. The variance of (X_i) = Communality *i* + Specificity *i*.

The correlation between X_i and X_j is $r_{ij} = a_{i1}a_{j1} + a_{i2}a_{j2} + ... + a_{im}a_{jm}$. Thus, two traits can only be highly correlated if they have high loadings on the same factors. The range of the common factor is $-1 \leq a_{ij} \leq +1$ and cannot exceed one (Manly, 1994).

The use of the factor analysis model requires the satisfaction of various assumptions which include: (i) the common factors (F_i, F_j) are assumed to be independent of each other. That is cov (F_i, F_j) equals zero for all iÍ j; (ii) specific factors (e_i,e_j) are assumed to be independent of each other. That is cov (e_i,e_j) equals zero for all iÍ j. The specific factors are also independent of the common factors. That is cov (e_i,F_j) equals zero for all iÍ j; (iii) the *X* variables and common factors have zero mean and unit variance; (iv) the variance of the specific factors ($e_i$ôs) may vary; and (v) the *X* variables, $F_i$ôs and $e_i$ôs are assumed to be multivariate X = عF + e, which gives the matrix form of the factor analysis model (Manly, 1994).

The main purpose of factor analysis is to find if there are some underlying factors affecting our original *X*-random variables. It is desired that each *X*-random variable has as much loading on one factor as possible so that all the random *X*-variables have one dominant factor. The most common method is the verimax rotation which is an algorithmic method used to obtain ófavourableô factor loadings. Favourable factor loading implies that a particular variable is dominated by such a factor. For instance, variable X_1 is possibly predominantly affected by factor 2 or variable X_2 is dominated by factor 1 and so on (Manly, 1994).

1.9.3 Cluster Analysis

Cluster Analysis is a way of grouping cases of data based on the similarity of responses to several variables (Manly, 1994). Cluster analysis is different from discriminant analysis where the researcher initially knows how many distinct groups are existing (Manly, 1994). Cluster analysis is used to determine the similarity of an accession relative to other accessions

so that all accessions in the population can be assigned to their specific group of similarity (Peeters and Martinelli, 1989).

Field (2000) reported possible factors that may limit clustering patterns as follows: (i) The different methods of clustering usually give very different results. This occurs because of the different criterion for merging clusters (including cases). It is important to think carefully about which method is best for what you are interested in looking at. (ii) With the exception of simple linkage, the results will be affected by the way in which the variables are ordered. (iii) The analysis is not stable when cases are dropped. This occurs because selection of a case (or merger of clusters) depends on similarity of one case to the cluster. Dropping one case can drastically affect the course in which the analysis progresses. (iv) The hierarchical nature of the analysis means that early ôbad judgementsô cannot be rectified.

Euclidean distance is the geometric distance between two objects (Manly, 1994). With Euclidean distances the smaller the distance, the more similar the cases. However, this measure is heavily affected by variables with large size or dispersion differences. So, if cases are being compared across variables that have very different variances, then the Euclidean distances will be inaccurate (Manly, 1994). As such it is important to standardise scores before proceeding with the analysis. Standardising scores is especially important if variables have been measured on different scales (Manly, 1994).

The data for cluster analysis consists of values f p variables X_1, X_2,é é ..X_p for n objects. For hierarchical algorithms these values are then used to produce an array of distances between the individuals. The Euclidean distance function $d_{ij} = \sqrt{\{\sum_{k=i}^{p} (Xik - Xjk)^2\}}$ is mostly used for quantitative variables, where X_{ik} is the value of the variable X_k for individual i and X_{jk} is the value of the same variable for individual j. In some cluster analysis algorithms we begin with principal components analysis to reduce large number of original variables to a smaller number of principal components. However, when the first principal components account for a higher percentage of variation in the data a plot of individuals against these components is useful in cluster analysis (Manly, 1994).

1.9.4 Canonical correlation analysis

Canonical correlation analysis (CCA) can be defined as a way of measuring the linear relationship between two sets of basis vectors, one for *x* and the other for *y*, such that the correlations between the projections of the variables to these basis vectors are mutually maximized (Akbas and Takma, 2005). It finds two bases, one for each trait, that are optimal with respect to correlations and, at the same time, it finds the corresponding correlations (Manly, 1994). The dimensionality of these new bases is equal to or less than the smallest dimensionality of the two traits (Manly, 1994). Canonical correlation analysis is related to other multivariate techniques. It is similar to regression, where it quantifies the strength of the relationship between the two sets of traits. It creates composites of variables similar to factor analysis; and resembles discriminant analysis in the determination of independent dimensions for each variable set (Hair *et al*., 1998).

An important feature of canonical correlation analysis is that it is invariant with respect to affine transformations of the traits (Borga, 2001). This is the most important difference between CCA and ordinary correlation analysis which highly depend on the basis in which the traits are described. Ordinary correlation analysis is dependent on the coordinate system in which the variables are described. This means that CCA maximizes the estimate of correlation between linear combinations of traits in the two sets, but does not maximize the amount of variance accounted for in one set of traits by the other set of traits (Akbas and Takma, 2005). This technique therefore limits the probability of committing type I error (Akbas and Takma, 2005).

The maximum correlations between two sets of traits in CCA is done using two linear combinations as shown below:

$W_i = a_{i1}X_1 + a_{i2}X_2 + \text{--------} + a_{ip}X_p$ é é é é é é é é é é é é é é é é é é ..é é ..(Eqn 1)

$V_i = b_{i1}Y_1 + b_{i2}Y_2 + \text{--------} + b_{iq}Y_q$ é é é é é é é é é .é é é é é é é é é é é (Eqn 2)

The symbols W and V represents canonical variates; a and b are canonical coefficients of the X and Y trait sets; and p and q are the number of traits in the X and Y trait sets, (Tabachnick and Fidell, 2001).

The canonical correlation for both sets of traits are generated from the following relationships:

var (W) = a'Cov (X) a.é é é é é é é é é ....é é é é é .é é é é é é é é é .....(Eqn 3)

var (V) = b'Cov (Y) bé é é é é é é é é ...é é é é é é é é é ..é é é é ..(Eqn 4)

$$C_{WV} = \frac{b'Cov(Y)\,b}{\sqrt{(a'Cov(X)\,a)(b'Cov(Y)b)}}$$ é é é é é é ...é é é é é é é é é é é é é é ...(Eqn 5)

where var (W) represents variance of the canonical variate W; var (V) is the variance of the covariate V; C_{wv} is the canonical correlation between the X and Y trait sets; Cov (Y) and Cov (X) are the covariances of the traits in the X and Y trait sets, respectively (Keskin and Yasar, 2007).

The relationship of a set of canonical variate is maximized when the correlation (r-value) of the p and q is small. The first set of canonical variate (W_1 and V_1) gives the highest correlation and is considered the most important. The correlation between W_2 and V_2 is only maximized where the traits measured are uncorrelated to W_1 and V_1. Similarly, the correlation between W_3 and V_3 is maximized if traits are not correlated with W_1, V_1, W_2 and V_2 (Manly, 1994).

The squared canonical correlation which represents the amount of variance in one canonical variate accounted for by the other canonical variate (Hair *et al*., 1998). The standardized coefficients are similar to the standardized regression coefficients in multiple regression, which gives an indication of the relative importance of the independent traits in determining the value of dependent traits.

Sharma (1996) suggested the estimation of the redundancy measure (RM) for each canonical correlation for the determination of the amount of variance in one set of traits that is accounted for by another set of traits. The redundancy measure (RM) estimate is given as:

$RM_{vi/wi}$ = AV (Y/V_i) x C_i^2é ..é ..(Eqn 6)

AV (Y/V_i) = $[\times^{q} LY_{ij}^2/q]$ é ....(Eqn 7)

where AV (Y/V_i) = the averaged variance in Y traits that is accounted for by the canonical variate V_i; LY_{ij}^2 = the loading of the j^{th} Y trait on the i^{th} canonical variate V_i; q = the number of traits in canonical variates; C_i^2 = the shared variance between V_i and W_i; W_i and V_i are canonical variates of Y and X trait sets, respectively.

1.10 References

Abraham, K. 1998. Occurrence of hexaploid males in *Dioscorea alata* L. *Euphytica* 99: 5-7.

Abraham, K. and P.G. Nair. 1991. Polyploidy and sterility in relation to sex in *Dioscorea alata* L. *Genetica* 83: 93-97.

Abraham, K., S.G. Nair, and M.T. Sreekumari. 1989. White yam, *Dioscorea rotundata* Poir, dwarf type. *Tropical Agriculture* (Trinidad) 66: 184-186.

Akbas, Y. And C. Takma. 2005. Canonical correlation analysis for studying the relationship between egg production traits and body weight, egg weight and age at sexual maturity in layers. *Czech Journal of Animal Science* 50: 163-168.

Akoroda, M.O. 1981. Studies on the genetics and floral biology of yams (*Dioscorea rotundata* Poir and *D. cayenensis* Lam.). Ph.D. Thesis, University of Ibadan, Nigeria. 398 pp.

Alexander, J. and D.G. Coursey. 1969. In: *The domestication and exploitation of plants and animals*. P.J. Ukcho and G.W. Dimpleby (eds.) Pergamon Press, London. pp. 405-425.

Arnau, G. E. Maledon, and A. Nemorin. 2007. Genetic improvement of the greater yam *D. alata* through polyploidy breeding In: Wilfredo, C. and Wanda L.I. (eds.). *Marketing opportunities for agriculture and forestry products in the Greater Caribbean - A challenge for the 21st century*. San Juan: Caribbean Food Crops Society, Annual Meeting Caribbean Food Crops Society, 16-19, September, San José, Costa Rica. 112 pp.

Arnolin, R. 1982. Vegetable cycle of the yam *D. alata* cv. Tahiti and Belep: Influence of spaced planting. In: *Proceedings 18th Annual Meeting, Caribbean Food Crops Society*, Mayaguez, Puerto Rico University, Puerto Rico. pp. 146-169.

Asiedu, R., N.M. Wanyera, S.Y.C. Ng, and N.Q. Ng. 1997. Yams. In: D. Fuccillo et al. (eds.) *Biodiversity in trust: Conversation and use of plant genetic resources in CGAIR centres*. Cambridge University Press, Cambridge, UK. pp. 57-66.

Asiedu, R., S.Y.C. Ng, K.V. Bai, I.J. Ekanayake, and N.M.W. Wanyera. 1998. Genetic Improvement. In: Orkwor GC, Asiedu R, Ekanayake IJ (eds.) *Food yams: Advances in research*, Ibadan, Nigeria: IITA and NRCRI. pp. 63-104.

Ayensu, E.S. 1971. Comparative anatomy of *Dioscorea rotundata* and *Dioscorea cayenensis*. *Journal of Linean Society* London (Botany) 63 (Supplementary): 127-136.

Bai, K.V. and J.S. Jos. 1978. Polyploidy in the Spinosa variety of *Dioscorea esculenta* (Lour.) Burk. *Journal of Root Crops* 4: 11-14.

Bai, K.V. and I.J. Ekanayake. 1998. Taxonomy, morphology and floral biology. In: G.C. Orkwor, R. Asiedu and I.J. Ekanayake, (eds.), *Food Yams: Advances in Research*, NRCRI and IITA, Ibadan, Nigeria. pp. 13-37.

Bellon, M.R. 1991. The ethnoecology of maize variety management: a case study from Mexico. *Human Ecology* 19: 389-418.

Bellon, M.R. 1996. The dynamics of crop intraspecific diversity: A conceptual framework at the farmer level. *Economic Botany* 50: 26-39.

Bennett, M.D. 2004. Perspectives on polyploidy in plants ï ancient and neo. *Biological Journal of the Linean Society* 82: 411-423.

Bennett, M.D., P. Bhandol, and I. Leitch. 2000. Nuclear DNA amounts in angiosperms and their modern uses: 807 new estimates. *Annals of Botany* 86: 859-909.

Binsfeld, P. C. and H. Schnabl. 2002. Molecular and cytogenetic constitution of plants obtained via two different somatic hybridization methods. *Plant Cell Reports* 21:58-62.

Borga, M. 2001. Canonical correlation tutorial. Available at http://people.imt.liu.se/magnus/cca/. 12 pp.

Bousalem, M., G. Arnou, I. Houchu, R. Arnolin, V. Viader, S. Santoni, and J. David. 2006. Microsatellite segregation analysis and cytogenetic evidence for tetrasomic inheritance in the American yam *Dioscorea trifida* and a new basic chromosome number in the Dioscoreae. *Theoretical Applied Genetics* 113: 439-451.

Bretagnolle, F. and J.D. Thompson. 1995. Gametes with the somatic chromosome number: mechanisms of their formation and role in the evolution of autopolyploid plants. *New Phytologia* 129: 1-22.

Britten, R.J. and D.E. Kohne. 1968. Repeated sequences in DNA. *Science* 161: 529-540.

Brown, S.C. 1993. Cytometric *tout terrain* or bush DNA cytometry. In: Sablon A, (ed.) *Flow cytometry, Jaquemin-NATO ASI Series, Volume H67.* Heidelberg: Springer-Verlag. pp. 227-241.

Brown, A.H.D. 2000. The genetic structure of crop landraces and the challenge to conserve them *in situ* on farms. In: *Genes in the Field: On-Farm Conservation of Crop Diversity* (S.B. Brush, ed.). Lewis Publishers, Boca Raton, FL, USA. pp. 29-48.

Brush, S. B. 2000. Genes in the Field: On-Farm Conservation of Crop Diversity. Lewis Publishers. Boca Raton, Rome, Italy. 288 pp.

Burkill, I.H. 1960. The organography and the evolution of the Dioscoreaceae, the family of the yams. *Journal of Linnean Society (Botany)* 56: 319-412.

Carputo, D., L. Frusciante, and S.J. Peloquin. 2003. The role of 2n gametes and endosperm balance number in the origin and evolution of polyploids in the tuber-bearing solanums. *Genetics* 163: 287-294.

Caspersson, T. and J. Schultz. 1938. Nucleic acid metabolism of the chromosomes in relation to gene reproduction. *Nature* 142: 294.

Cavalier-Smith, T. 1985. Cell volume and the evolution of eukaryote genome size. In: Cavalier-Smith, T. (ed.) *The evolution of genome size*. John Wiley, Chichester. pp. 105-184.

Chaddick, L.R., P.J. Rudall, P. Wilkin, A.J. Hedderson, and M.W. Chase. 2002. Phylogenetics of Dioscoreales based on combined analysis of morphological and molecular data. *Journal of Linnean Society (Botany)* 138: 123-144.

Chair, H., X. Perrier, C. Agbangla, J.L. Marchand, O. Dainou, and J.L. Noyer. 2005. Use of cpSSRs for the characterization of yam phylogeny in Benin. *Genome* 48: 674-684.

Chen, Z.J. 2007. Genetic and epigenetic mechanisms for gene expression and phenotypic variation in plant polyploids. *Annual Review of Plant Biology*. 58: 377-406.

Costich, D.E., R. Ortiz, T.R. Meagher, L.P. Bruederle, and N. Vorsa. 1993. Determination of ploidy level and nuclear DNA content in blueberry by flow cytometry. *Theoretical Applied Genetics* 86: 1001-1006.

Coursey, D.G. 1967. *Yams*, Longman, London. 230 pp.

Cowdon, R.R. and S.K. Curtis. 1981. Microfluorometeric investigations of chromatin structure I. Evaluation of nine DNA-specific fluorochromes as probes of chromation organization. *Histochemistry* 72: 11-23.

Dansi, A., H.D. Mignouna, M. Pillay, and S. Zok. 2001. Ploidy variation in the cultivated yams (*Dioscorea cayenensis-Dioscorea rotundata* complex) from Cameroon as determined by Flow cytometry. *Euphytica* 119: 301-307.

Dansi, A., H.D. Mignouna, J. Zoundjihekpon, A. Sangare, R. Asiedu, and N. Ahoussou. 2000a. Using isozyme polymorphism for identifying and assessing genetic variation in cultivated yam (*Dioscorea cayenensis/ Dioscorea rotundata* complex) of Benin. *Genetic Resources and Crop Evolution* 47: 371-383.

Dansi, A., H.D. Mignouna, J. Zoundjihekpon, A. Sangare, N. Ahoussou, and R. Asiedu. 2000b. Identification of some Benin Republicôs Guinea yam (*Dioscorea cayenensis-Dioscorea rotundata* complex) using Random Amplified Polymorphic DNA. *Genetic Resources and Crop Evolution* 47: 619-625.

Dansi, A., H.D. Mignouna, J. Zoundjihekpon, A. Sangare, R. Asiedu, and F.M. Quin. 1999. Morphological diversity, cultivar groups and possible descent in the cultivated yams (*Dioscorea cayenensis-Dioscorea rotundata* complex) of Benin Republic. *Genetic Resources and Crop Evolution* 46: 371-388.

De Wet, J.M.J. 1980. Origins of polyploid. In: Lewis W.H. (ed.) *Polyploidy: Biological Relevance.* New York, Plenum Press, pp. 5-15.

Deeley, E.M. 1955. An integrating micodensitometer for biological cells. *Journal of Cell Instrumentation* 32: 263-267.

Degras, L. 1977. Vegetative and sexual management in food yam improvement. In: *Proceedings Fourth Symposium of International Society of Tropical Root Crops*. J. Cock, R. MacIntyre and M. Graham, eds. CIAT, Columbia. pp. 58-63.

Degras, L. 1993. *The Yam*: *A tropical root crop*. The Technical Centre for Agricultural and Rural Cooperation (CTA). The Macmillan Press, London. 408 pp.

Dellaporta, S.L. and A. Calderon-Urrea. 1993. Sex determination in flowering plants. *The Plant Cells* 5: 1241-1251.

Dessauw, D. 1988. Etude des facteurs de la sterilite du bananier (*Musa* spp.) et des relations cytotaxonomiques entre *M. acuminata* et *M. balbisiana* Colla. *Fruits* 43: 539-700.

Dolezel, J. 1997. Application of flow cytometry for the study of plant genomes. *Journal of Applied Genetics* 38: 285-302.

Dolezel, J. and J. Bartos. 2005. Plant DNA flow cytometry and estimation of nuclear genome size. *Annals of Botany* 95: 99-110.

Dolezel, J. and W. Gohde. 1995. Sex determination in dioecious plants *Melandrium album* and *M. robrum* using high resolution flow cytometry. *Cytometry* 19: 103-106.

Dolezel, J., J. Barto, H. Voglmayr, and J. Greilhuber. 2003. Nuclear DNA content and genome size of trout and human. *Cytometry* 51: 127-128.

Dolezel, J., J. Greilhuber, S. Lucretti, A. Meister, M.A. Lysak, L. Nardi, and R. Obermayer. 1998. Plant genome size estimation by flow cytometry: inter-laboratory comparison. *Annals of Botany* 82 (Supplementary A): 72-26.

Dolezel J., P. Binarová and S. Lucretti. 1989. Analysis of nuclear DNA content in plant cells by flow cytometry. ï *Biologia Plantarum* 31: 113-120.

Dubelaar, G.B.J. and P.I. Gerritzen. 2000. CytoBuoy: A step forward towards using flow cytometry in operational oceanography. *Scientia Marina* 64: 255-265.

Dumont, R. and P. Vernier. 2000. Domestication of yams (*Dioscorea cayenensis -rotundata*) within the Bariba ethnic group in Benin. *Outlook on Agriculture* 29: 137-142.

Edem, E.U. 1975. Preliminary investigations into the effect of planting dates and types of sett on the flowering of white yam (*Dioscorea rotundata* Poir) varieties. Federal Department of Agriculture, Ibadan, Nigeria. Memo 115: 6.

Egesi, C.N., M. Pillay, R. Asiedu, and J.K. Egunjobi. 2002. Ploidy analysis in water yam, *Dioscorea alata* L. germplasm. *Euphytica* 128: 225-230.

Essad, S. 1984. Variation geographique des nombres chromosomiques de base et polyploidie dans le genre *Dioscorea*, a propos du denombrement des especes *transversa* Brown, *pilosiuscula* bert. et *trifida* L. *Agronomie* 4: 611-617.

Field, A. 2000. Postgraduate statistics: Cluster analysis. www.statisticshell.com/cluster. 10 pp.

Food and Agricultural Organisation (FAO), 1999. Glosary of biotechnology and genetic engineering. A. Zaid, H.G. Hughes, E. Porceddu, F.W. Nicholas (eds.) FAO Research and Technology Paper 7. 250 pp.

Frageria, N.K. 1992. *Maximizing Crop Yields*. Marcel Decker Inc., New York. 274 pp.

Frankel, O.H. 1970. Genetic conservation in perspective. In: *Genetic Resources in Plants - their Exploration and Conservation* O.H. Frankel and E. Bennett (eds.). IBP Handbook Nº 11. Blackwell, Oxford (U.K.).

Galbraith, D.W. 1990. Flow cytometric analysis of plant genomes. *Methods in Cell Biology* 33: 549-562.

Galbraith, D.W., G.M. Lambert, J. Macas, and J. Doleel. 1998. Analysis of nuclear DNA content in higher plants. In: Robinson, J.P., Darzynkiewicz, Z., Dean, P.N., Dressler, L.G., Orfao, A., Robinovitch, P.S., Stewart, C.C., Tanke, H.J., Wheeless, L.L. (eds.) *Current Protocols in Cytometry*. New York: John Wiley & Sons. pp. 7.6.1-7.6.22.

Galbraith, D.W., K.R. Harkins, J.M. Maddox, N.M. Ayres, D.P. Sharma, and E. Firoozabady. 1983. Rapid flow cytometric analysis of the cell cycle in intact plant tissues. *Science* 220: 1049-1051.

Gamiette, F., F. Bakry, and G. Ano. 1999. Ploidy determination of some yam species (*Dioscorea* spp.) by flow cytometry and conventional chromosomes counting. *Genetic Resources and Crop Evolution* 46: 19-27.

Grant, V. 1981. *Plant Speciation*, 2nd edition. Columbia University Press, New York. 368 pp.

Grant, V. 2002. Frequency of spontaneous amphiploids in *Gilia* (Polemoniaceae) hybrids. *America Journal of Botany* 89: 1197-1202.

Greilhuber, J., J. Doleel, M. Lysak, and M.D. Bennett. 2005. The origin, evolution and proposed stabilization of the terms ógenome sizeô and óC-valueô to describe nuclear DNA contents. *Annals of Botany* 95: 255-260.

Guo, M., D. Davis, and J.A. Birchler. 1996. Dosage effects on gene expression in a maize ploidy series. *Genetics* 142: 1349-1355.

Hahn, S.K. 1995. Yams: *Dioscorea* spp. (Dioscoreaceae). In: J. Smartt and N.W. Simmonds (Eds.), *Evolution of Crop Plants*, Longman Scientific and Technical, UK. pp. 112-120.

Hahn, S.K. 1988. Constraints in cassava and yam research in Africa. In: *The use of Biotechnology for the Improvement of Cassava, Yams and Plantains in Africa*. IITA Meeting Report Series. IITA, Ibadan, Nigeria 1998/2. pp. 10-15.

Hahn, S.K., D.S.O. Osiru, M.O. Akoroda, and J.A. Otoo. 1987. Yam production and its future prospects. *Outlook on Agriculture* 16: 105-110.

Hair, J.F., R.E. Anderson, R.H. Tatham, and C.B. Williams. 1998. *Multivariate Data Analysis*. 5th Edition. Prentice Hall, Inc. pp. 442-462.

Hamon, P., J. Brizard, J., Zoundjihékpon, C. Duperray, and A. Borge. 1992. Etude des index dôADN de huit espèces dôignames (*Dioscorea* spp.) par cytometrie en flux. *Canada Journal of Botany* 70: 996-1000.

Harlan, J.R. 1975. Our vanishing resources. *Science* 188: 618-621.

Heller, F.O. 1973. DNS-Bestimmung an Keimwurzeln von *Vicia faba* L. mit Hilfe der impulscytophotometrie. *Bericht der Deutschen Botanischen Gesellschaft* 86: 437-441.

Hilu, K.W. 1993. Polyploidy and the evolution of domesticated plants. *American Journal of Botany* 80: 1494-1499.

Hochu, I., S. Santoni, and M. Bousalem. 2006. Isolation, characterization and cross-species amplification of microsatellite DNA loci in the tropical American yam *Dioscorea trifida*. *Molecular Ecology Notes* 6: 137-140.

Husband, B.C. 2004. The role of triploid hybrids in the evolutionary dynamics of mixed ï ploidy populations. *Biology Journal of Linnean Society* 82: 537-546.

Ikediobi, C.O. and L.C. Igboanusi. 1983. Identification of yam, *Dioscorea* spp. cultivars by use of electrophoretic patterns of soluble tuber proteins. *Biotropica* 15: 65-67.

Institut National des Recherches Agricoles du Bénin (INRAB), 1996. Plan Directeur de la Recherche Agricole du Bénin. Volume II: Deuxième partie. Plan de développement à long terme. Programmes sectoriels. Ministère du éveloppement Rural; République du Bénin. pp. 155-312.

International Board for Plant Genetic Resources (IBPGR), 1980. International Board for Plant Genetic Resources. Regional Committee for Southeast Asia. Descriptors for yam (*Dioscorea* spp.) IBPGR Secretariat, Rome. 19 pp.

International Institute of Tropical Agriculture (IITA), 1993. Genetic resources Unit *Annual Report* 1992. Crop Improvement Division. IITA, Ibadan, Nigeria. 38 pp.

Inze, D. and L. De Veylder. 2006. Cell cycle regulation in plant development. *Annual Review of Genetics* 40: 77-105.

Jos, J.S., K.V. Bai, and N. Hrishi. 1977. Meiosis in a triploid intoxicating yam. *Journal of Root Crops* 3: 17-20.

Kamentsky, L.A., M.R. Melamed, and H. Derman. 1965. Spectophotometer: new instrument for ultrarapid cell analysis. *Science* 150: 630-631.

Kawara, S., M. Takata, and K. Takehara. 1999. High frequency of DNA aneuploidy detected by DNA flow cytometry in Bowenôs disease. *Journal of Dermatological Science* 21: 23-26.

Kawanabe, A., N.T. Miyashita, and R. Terauchi. 1997. Phylogenetic relationship among section Stenophora in the genus *Dioscorea* based on the analysis of nucleotide sequence variation in the phosphoglucose isomerase (Pgi) locus. *Genes Genetics Systems* 72: 253-262.

Kehlenbeck, K. and B.L. Maass. 2004. Crop diversity and classification of homegardens in Central Sulawesi, Indonesia. *Agroforestry Systems* 63: 53-62.

Keller, E. R. J., I. Schubert, J. Fuchs, and A. Meister. 1996. Interspecific crosses of onion with distant *Allium* species and characterization of the presumed hybrids by means of flow cytometry, karyotype analysis and genomic in situ hybridization. *Theoretical Applied Genetics* 92: 417-424.

Keskin, S. And F. Yasar. 2007. Use of canonical correlation analysis for determination of relationships among several traits in egg plants (*Solanum melongana* L.) under salt stress. *Pakistan Journal of Botany* 39: 1547-1552.

Kron, R., J. Suda, and B.C. Husband. 2007. Applications of flow cytometry to evolutionary and population biology. *AREES* 38: 847-876.

Lebot, V., A. Ivancic, and K. Ibrahim. 2005. The geographical distribution of allelic diversity, a practical means of preserving and using minor root crop genetic resources. *Experimental Agriculture* 41: 475-489.

Lebot, V., B. Trilles, J. Noyer, and J. Modesto. 1998. Genetic relationship between *Dioscorea alata* L. cultivars. *Journal of Genetic Resources and Crop Evolution* 45: 499-509.

Levy, A.A. and M. Feldman. 2004. Genetic and epigenetic reprogramming of the wheat genome upon allopolyploidization. *Biology Journal Linnean Society* 82: 607-613.

Liu, B. and J.F. Wendel. 2003. Epigenetic phenomena and the evolution of plant allopolyploids. *Molecular Phylogenetic Evolution* 29: 365-379.

Manly, B.F.J. 1994. Multivariate statistical methods. A Primer. Second edition. Chapman and Hall, 2-6 Boundary Row, London SE1 8HN, UK. 214 pp.

Martin, F.W. 1974a. Tropical yams and their potential. Series ï Part 1. *Dioscorea esculenta*. USDA Agriculture Handbook No. 457. ARS, USDA. 18 pp.

Martin, F.W. 1974b. Tropical yams and their potential. Series ï Part 2. *Dioscorea bulbifera*. USDA Agriculture Handbook No. 466. ARS, USDA. 20 pp.

Martin, F.W. 1976. Tropical yams and their potential. Series ï Part 3. *Dioscorea alata*. USDA Agriculture Handbook No. 495 ARS, USDA. 40 pp.

Martin, F.W. and S. Ortiz. 1963. Chromosome numbers and behavior in some species of *Dioscorea. Cytologia* 28: 96-101.

Martin, F.W. and S. Ortiz. 1966. New chromosome numbers in some *Dioscorea* species. *Cytologia* 31: 105-107.

Miège, J. 1954. Nombre chromosomiques et répartition géographique de quelques plantes tropicales et équatoriales. *Revista Cytologia* Paris 15: 314-348.

Mignouna, H.D. and A. Dansi. 2003. Yam (*Dioscorea* spp) domestication by the Nago and Fon ethnic groups in Benin. *Genetic Resources and Crop Evolution* 50: 519-528.

Mignouna, H.D., M.M. Abang, and R. Asiedu. 2007. Yams. In: *Genome Mapping and Molecular Breeding in Plants: Pulses, Sugar and Tuber Crops*. Volume 3. C. Kole (ed.) ©Springer- Verlag Berling Heidelberg. pp. 271-296.

Mignouna, H.D., M.M. Abang, A. Onasanya, and R. Asiedu. 2002. Identification and application of RAPD markers for anthracnose resistance in water yam (*Dioscorea alata*). *Annual Applied Biology* 141: 61-66.

Mirsky, A.E. and H. Ris. 1951. The deoxyribonucleic acid content of animal cells and its evolutionary significance. *Journal of General Physiology* 34: 451-462.

Nasrallah, M.E., K. Yogeeswaran, S. Snyder, and J.B. Nasrallah. 2000. Arabidopsis species hybrids in the study of species differences and evolution of amphiploidy in plants. *Plant Physiology* 124: 1605-1614.

National Center for Biotechnology Information (NCBI), 2004. Just the facts: A basic introduction to the science underlying NCBI Resouces. Available at www.ncbi.nlm.nih.gov/About/primer/est.html (modified 29 March 2004; accessed 24 November 2010).

Njoho, P.M. 2002. EDEL 812 Statistical Analysis in Research Module, University of KwaZulu-Natal Pietermatritzburg Campus South Africa. Available at www.stat.unp.ac.za. 115 pp.

Nowak, M.A., M.C. Boerlijst, J. Cooke. And J.M. Smith. 1997. Evolution of genetic redundancy. Nature 388: 167-171.

Obidiegwu, J.E., E. Rodriguez, E.E. Ene-Obong, J. Loureiro, C.O. Muoneke, C. Santos, M. Kolesnikova-Allen, and R. Asiedu. 2009. Estimation of the nuclear DNA content in some representative of genus *Dioscorea*. *Scientific Research and Essay* 4: 448-452.

Okezie, C.E.A., S.N.C. Okonkwo, and F.I. Nwoke. 1981. Growth pattern and growth analysis of the white guinea yam raised from seed. In: *Proceedings; First International Symposium of Tropical Root Crops – African Branch*. pp. 180-194.

Okoli, O.O. 1991. Yam germplasm diversity, uses and prospects for crop improvement in Africa. In: Ng, N.Q., Perrino, P., Attere F.and Zedan H. (eds.), *Crop genetic resources of Africa, Vol. II*. Co-publication IITA/IBPGR/UNEP/CNR, Rome, and IITA, Ibadan. pp. 109-117.

Okonkwo, S.N.C. 1985. The botany of the yam plant and its exploitation in enhanced productivity of the crop. In: *Advances in Yam Research. The biochemistry and technology of the yam tuber.* G. Osuji, (ed.) Biochemical Society of Nigeria and Anambra State University of Technology. pp. 3-29.

Onwueme, I.C. 1978. *The Tropical Tuber Crops: Yams, Cassava, Sweet potato, Cocoyam.* John Wiley and Sons, New York. 234 pp.

Onwueme, I.C. and W.B. Charles. 1994. *Tropical root and tuber crops: production, perspectives, and future prospects*. FAO plant production and protection papers ; 126. Rome: FAO, 1994. xii. 228 pp.

Orkwor, G.C., R. Asiedu, and I.J. Ekanayake. (eds.) 1998. *Food Yams. Advances in Research*, IITA and NRCRI, Nigeria. 249 pp.

Otto, F.J. 1990. DAPI staining of fixed cells for high resolution flow cytometry of nuclear DNA. In: Darzynkiewickz, Z. and Crissman, H.A. (eds.), *Methods in Cell Biology, Volume 33*. San Diego: Academic Press. pp. 105-110.

Otto, S.P. and J. Whitton. 2000. Polyploid incidence and evolution. *Annual Review Genetics* 34: 401-437.

Paterson, A.H., J.E. Bowers, D.G. Peterson, J.C. Estill, and B.A. Chapman. 2003. Structure and evolution of cereal genomes. *Current Opinion Genetics Development* 13: 644-650.

Peeters, J.P. and J.A. Martinelli. 1989. Hierarchical cluster analysis as a tool to manage variation in germplasm collections. *Theoretical Applied Genetics* 78: 42-48.

Puite, K.J. and W.R.R. Ten Broeke. 1983. DNA staining of fixed and non-fixed plant protoplasts for flow cytometry with Hoechst 33342. *Plant Science Letters* 32: 79-88.

Purseglove, J.W. 1972. *Tropical crops. 1. Monocotyledons*. Longman, London. Pp. 39-62.

Rabinovitch, P.S. 1994. DNA content histogram and cell-cycle analysis. In: Darzynkiewicz, Z., Robinson, J.P. and Crissman, H.A. (eds.), *Methods in Cell Biology: Flow Cytometry, Volume 41.* San Diego: Academic Press. pp. 263-296.

Raghavan, R.S. 1958. A chromosome survey of Indian *Dioscorea* spp. *Proceedings, Indian Academy Science* 48: 59-63.

Raghavan, R.S. 1959. A note on some South Indian species of the genus *Dioscorea. Current Science* 28: 337-338.

Ramachandran, K. 1968. Cytological studies in *Dioscoreacea. Cytologia* 33: 401-410.

Ramanatha Rao, V.R. and T. Hodgkin. 2002. Genetic diversity and conservation and utilization of plant genetic resources. *Plant Cell, Tissue and Organ Culture* 68: 1-19.

Ramsey, J. and D.W. Schemske. 1998. Pathways, mechanisms, and rates of polyploid formation in flowering plants. *Annual Review of Ecology Systems* 29: 467-501.

Rehm, S. and G. Espig. 1991. The cultivated plants of the tropics and sub-tropics (translated by G. McNamara & C. Ernsting). Verlay Josef Margraf, Germany, viii + 552 pp. ISBN 3-8236-1169-0.

Robinosn, J.P. 2006. Introduction to flow cytometry. Flow cytometry talks. Purdue University Cytometry Laboratories. Accessible at http://www.cyto.purdue.edu/flowcyt/educate/pptslide.htm

Sadik, S. 1975. Root and tuber physiology with emphasis on yam improvement. In: *Proceedings Physiology Program Formulation Workshop*. IITA, Ibadan, Nigeria. April 1975. pp. 35-40.

Sadik, S. and N.G. Rockwood. 1975. Introducing the hybrid yam. *Span* 18: 67-68.

Sadik, S. and O.U. Okereke. 1975a. Flowering, pollen grain germination, fruiting, seed germination and seedling development of white yam, *Dioscorea rotundata* Poir. *Annals of Botany* 39: 597-604.

Sadik, S. and O.U. Okereke. 1975b. A new approach to the improvement of yam, *Dioscorea rotundata. Nature* 254: 134-135.

Sandoval A., V. Hocker and J.L. Verdeil. 2003. Flow cytometric analysis of the cell cycle in different coconut palm (*Cocos nucifera* L.) tissues cultured in vitro. *Plant Cell Reports* 22: 25-31.

Scarcelli, N., O. Dainou, C. Agbangla, S. Tostain, and J.L. Pham. 2005. Segregation patterns of isozyme loci and microsatellite markers show the diploidy of African yam *Dioscorea rotundata* (2n=40). *Theoretical Applied Genetics* 111: 226-232.

Scarcelli, N., S. Tostain, C. Mariac, C. Agbangla, O. Da, J. Berthaud, and J.L. Pham. 2006. Genetic nature of yams (*Dioscorea* spp.) domesticated by farmers in Benin (West Africa). *Genetic Resource Crop Evolution* 53: 121-130.

Schmidt, G. and S.J. Thannhauser. 1945. A method for the determination of deoxyribonucleic acid, ribonucleic acid, and phosphoproteins in animal tissues. *Journal of Biological Chemistry* 161: 83-89.

Segarra-Moragues, J.G. and P. Catalan. 2003. Life history variation between species of relictual genus *Borderea* (Dioscoreaceae): Phylogeography, genetic diversity and population genetic structures assessed by RAPD markers. *Biology Journal of Linnean Society* 80: 483-498.

Segarra-Moragues, J.G., M. Palop-Esteban, F. Gonzalez-Caudelas, and P. Catalan. 2004. Characterization of seven $(CTT)_n$ microsatellite loci in the Pyrenean endemic *Borderea pyreniaca* (Dioscoreaceae): Remarks on ploidy level and hybrid origin assessed through allozymes and microsatellite analyses. *Journal of Heredity* 95: 177-183.

Seoighe, C. 2003. Turning the clock back on ancient genome duplication. *Current Opinion Genetics Development* 2003, 13: 636-643.

Shapiro, H.M. 2003. *Practical Flow Cytometry*. 4th edition. Wiley-Liss, New York. (ISBN 0-471-41125-6)
http://probes.invitrogen.com/products/flowcytometry/practicalflowcytometry.html.

Sharma, S. 1996. *Applied Multivariate Techniques*. John Willey and Sons, Inc., Canada. pp. 391-404.

Sharma, A.K. and N. De. 1956. Polyploidy in *Dioscorea. Genetica* 2B: 112-120.

Smarda, P. And D. Stanciik. 2006. Ploidy level variability in South American fescues (Festuca L., Poaceae): Use of flow cytometry in up to five and half year old caryopses and herbarium specimens. *Plant Biology* 8: 73-80.

Smith, L.I. 2002. A tutorial on principal component analysis. Available at www.cs.otago.ac.nz/cosc453/. 27 pp.

Soltis, D.E., P.S. Soltis, J.C. Pires, A. Kovarik, J.A. Tate, and E. Mavrodiev. 2004. Recent and recurrent polyploidy in Tragopogon (Asteraceae): Cytogenetic, genomic and genetic comparisons. *Biology Journal of Linn. Society* 82: 485-501.

Sosik, H.M. and R.J. Olson. 2002. Phytoplankton and iron limitation of photosynthetic efficiency in the Southern Ocean during late summer. *Deep Sea Research. Part I. Oceanographic Research Papers* 49: 1195-1216.

Stebbins, G.L. 1971. Chromosomal evolution of higher plants. Edward Anold, London. pp. 85-86.

Stehlik, I., P. Kron, S.C.H. Barrett, and B.C. Husband. 2007. Sexing pollen reveals female bias in a dioecious plant. *New Phytology* 175: 185-194.

Straus, N.A. 1971. Comparative DNA renaturation kinetics in amphibians. *Proceedings of the National Academy of Sciences of the USA* 68: 799-802.

Suda, J., T. Kyncl, and R. Freiova. 2003. Nuclear DNA amounts in Macronesian angiosperms. *Annals of Botany* 92: 153-164.

Swift, H. 1950. The constancy of deoxyribose nucleic acid in plant nuclei. *Proceedings of the National Academy of Sciences of the USA* 36: 643-654.

Tabachnick, B. and L.S. Fidell. 2001. *Using Multivariate Statistics*. A Pearson Education Company, Needham Heights, USA. 966 pp.

Tamiru, M., H.C. Becker, and B.L. Maass. 2006. How farmers select and manage yam (*Dioscorea* spp.) diversity in Wolayita, Southern Ethiopia. Paper presented at the *47th Annual Meeting of the Society for Economic Botany "Folk Botanical Wisdom: Towards Global Markets*, 4-9 June 2006, Chiang Mai, Thailand. Book of Abstracts. pp. 90.

Terauchi, R., V.A. Chikaleke, G. Thottappilly, and S.K. Hahn. 1992. Origin and phylogeny of Guinea yam as revealed by RLFP analysis of chloroplast DNA and nuclear ribosomal DNA. *Theoretical Applied Genetics* 83: 743-751.

Thomas, C.A. 1971. The genetic organization of chromosomes. *Annual Review of Genetics* 5: 237-256.

Udall, J.A. and J.F. Wendel. 2006. Polyploidy and crop improvement. *The Plant Genome (A Supplementary to Crop Science)* 46: 3-14.

Ulrich, I. and W. Ulrich. 1991. High-resolution flow cytometry of nuclear DNA in higher plants. *Protoplasma* 165: 212-215.

Ulrich, I., B. Fritz, and W. Ulrich. 1988. Application of DNA fluorochromes for flow cytometric DNA analysis of plant protoplast. *Plant Science* 55: 151-158.

Vermes, I., C. Haanen, and C. Reutelingsperger. 2000. Flow cytometry of apoptotic cell death. *Journal of Immunological Methods* 243: 167-190.

Vilhar, B., J. Greilhuber, J. Dolenc Koce, E.M. Temsch, and M. Dermastia. 2001. Plant genome size measurement with DNA image cytometry. *Annals of Botany* 87: 719-728.

Waitt, A.W. 1964. Annual Report for 1960-1961. Federal Department of Agricultural Research. Lagos, Nigeria. Federal Ministry of Information, Printing Division, Nigeria.

Weiss H., C. Dobeġ, G. M. Schneeweiss. and J. Greimler. 2002. Occurrence of tetraploid and hexaploid cytotypes between and within populations in *Dianthus* sect. *Plumaria* (Caryophyllaceae). ï *New Phytologia* 156: 85-94.

Wendel, J.F. and J. Doyle. 2005. Polyploidy and evolution in plants. *In*: R.J. Henry (ed.) *Plant diversity and evolution*: Genotypic and phenotypic variation in higher plants. CAB International. pp. 97-117.

Wilkin, P., P. Schols, M.W. Chase, K. Chayamarit, C.A. Furness, S. Huysmans, F. Rakotonasolo, E. Smets, and C. Thapyai. 2005. A plastid gene phylogeny of the yam genus, *Dioscorea*: roots, fruits and madagascar. *Systematic Botany* 30: 736-749.

Yamamori, M., S. Fujita, K. Hayakawa, J. Matsuki, and T. Yasui. 2000. Genetic elimination of a starch granule protein, SGP-1, of wheat generates an altered starch with apparent high amylase. *Theoretical Applied Genetics* 101: 21-29.

Zoundjihekpon, J., S. Essad, and B. Toure. 1990. Denombrement chromosomique dans dix groupes varietaux du complexe *Dioscorea cayenensis-rotundata. Cytologia* 55: 115-120.

CHAPTER TWO

Diversity of morphological traits of yam (*Dioscorea* spp.) genotypes from Sierra Leone

Abstract

Various vernacular names which may refer to the same genotype have been used by farmers and consumers. Morphological characterization of the variation among cultivated yams is essential for improved management and efficient utilization of yam genetic resources. A total of 52 yam genotypes from Sierra Leone were grown in a randomized complete block design with three replications during 2010 at the University of KwaZulu-Natal, Pietermaritzburg, South Africa. The aim of the study was to explore the existing morphological variability within this germplasm. Twenty-eight morphological traits measured from the genotypes were analysed using principal component analysis (PCA), factor analysis (FA) and cluster analysis (CA). The first 10 principal components (PCs), which had eigen-values >0.6 explained 86.61% of the total variability. The PCA results indicated a number of traits that largely contributed to the variability within and between the species and they included: number of days to shoot emergence, shoot traits (position, shape, size, density, vein colour and measurements of leaves; shoot growth rate) and tuber traits (tuber shape and flesh colour of cross section of tuber). The two-dimensional plot of the first two PCs grouped the accessions according to their species, whereas some of the genotypes within species were grouped according to the various tuber shapes: irregular, oblong, oval-oblong, round and cylindrical. Genotypes WR 07/024, SR 07/075, 07/073, ER 07/032 and NR 07/042 overlapped in sub-groups B_1 and B_2; whereas genotypes WR07/010, NR 07/041, ER 07/038 and NR 07/067 overlapped in sub-groups B_1 and B_3. This indicated the possibility of duplicate genotypes in the germplasm. Factor analysis had six factors, which explained 75% of the total genetic variation in the dependence structure. Factor 1 was strongly associated with absence or presence of wings, distance between lobes, leaf apex shape, leaf colour, leaf density, leaf margin colour, leaf length-2 leaf vein colour of the upper surface, number of branch, number of stem, stem colour and tip length of mature leaf; factor 2 with leaf density, leaf length-1, leaf vein colour of lower surface, petiole wing colour, tip colour, wing colour and flesh colour of central cross section of tuber; factor 3 with absence or presence of wings and leaf width-1; factor 4 with leaf width-2; factor 5 with stem colour; and factor 6 with number of

days to emergence. Other factors (≥7) explained the rest of the genetic variation and may not be important in yam breeding programmes. The dendrogram of the CA, based on the morphological characters, showed six major groups, which also supported groupings in the PCA and FA. This study demonstrated that local accessions from Sierra Leone have wide inter- and intra-group variability.

2.1 Introduction

Yams (*Dioscorea* spp.) are food security crops that sustain many livelihoods in the tropics and subtropics especially in countries in West Africa where large commercial scale production is practiced (Mwiringi *et al*., 2009). The crop serves as a source of food, medicine and income for many small scale farmers in Africa. Despite itôs importance, efforts of breeding and selection of yam genotypes with improved traits are currently inhibited by the lack of adequately characterized native genotypes at the morphological and molecular level (Asiedu *et al*., 1998). This is due to the fact that the distribution of genotypes and their characteristics are not well documented, which constraints the efficient conservation of these genetic resources thereby limiting their use in breeding programmes. This dearth of knowledge of existing germplasm in some of the countries where yams are cultivated has significantly contributed to genetic erosion of yams (Dansi *et al*., 1997). In Sierra Leone for instance, despite the importance of yams, many farmers and scientists do not know the existing level of diversity among the various species or varieties within species under cultivation. Also, pests and diseases are among major factors responsible for genetic erosion in yams (IAR, 2004). Genetic erosion can be overcome by collecting, characterizing and conserving existing germplasm for diversity studies and breeding work (Mignouna *et al*., 2002).

Several morphological diversity studies have been carried out between and within yam populations to catalogue existing diversity (Sastrapradja, 1982; Velayudhan *et al*., 1989; Asiedu *et al*., 1997; Lebot *et al*., 1998; Dansi *et al*., 2001; Hasan *et al*., 2008). However, yams are heterogeneous perennials with many overlapping morphological, physiological and chemical attributes. The efficient utilization of large genetic variability can be optimized when it has been systematically evaluated, quantified and characterized (Amurrio *et al*., 1995). The use of one or more of these systematic methods to determine the extent of variability present in yam germplasm has provided better understanding in major yam

producing countries like Cote dôIvoire (Hamon, 1987), Benin (Dansi *et al*., 1999), Cameroon (Mignouna *et al*., 2002) and Malaysia (Hasan *et al*., 2008), but the method is yet to be fully explored in Sierra Leone. Proper characterization of genotypes should include a morphological description, as well as determination of biochemical and molecular markers for genetic evaluation (Mignouna *et al*., 2003; Dumont *et al.*, 2005).

In experiments where large amounts of data are obtained such as diversity studies, data mining (knowledge discovery) is relevant to resolve difficulties in interpreting results. Principal component analysis is used to reduce dimensionality in data by performing a covariance analysis between factors. Factor analysis is used to describe variability among observed variables in terms of a lower number of unobserved variables known as factors. Factor analysis is related to PCA, but the two are different. Principal component analysis performs a variance-maximizing rotation of the variable space, taking into account all variability present in the traits. Conversely, FA estimates how much of the variability is due to common factors (communality) and specific factors (specificity). The two techniques may only be on equal terms if the error in the FA model (the variability not explained by the communality) is assumed to have constant variance. Cluster analysis aims at sorting different objects into groups based on their degree of associations and may not necessarily have similar results as those of PCA and FA (Manly, 1994). In this study, the multivariate techniques of principal components, factor and cluster analyses were used to determine the levels of phenotypic diversity/ similarity in 52 yam accessions. The objectives of this study were to determine the relationships between the accessions, and to identify duplicates and groupings of genotypes in the accessions of *Dioscorea* spp. obtained from Sierra Leone. The results of the research would contribute to conservation planning, and genetic improvement of yams in Sierra Leone.

2.2 Materials and methods

2.2.1 Plant material

A total of 52 genotypes which included 50 landraces collected in various locations within four regions (southern, northern, eastern and western) of Sierra Leone, and two improved checks of *D. rotundata* from the International Institute of Tropical Agriculture (IITA) were morphologically characterized (Table 2.1). The collections were made during the 2007 harvest season (November to December). From two to five tubers of each genotype were

collected and assigned an accession number. The accessions were maintained in experimental plots at the Njala Agricultural Research Centre (NARC), Sierra Leone. In early January 2010, about three minitubers of each genotype collected were imported to South Africa.

Minisetts each weighing 50 g were established in 25 cm (diameter) x 20 cm (height) pots in a green-house at the University of KwaZulu-Natal, Pietermaritzburg, South Africa in mid January 2010. The environmental temperature and relative humidity of the green-house ranged between 20 and 33°C, and between 60 and 85% respectively, typical of the Sierra Leone weather (Table 2.2). The pots were filled with composted seedling mix, and water was supplied by drip irrigation. The pots were arranged in a three replicate, randomized complete block design. The planting distance between pots was 0.25 m. Each pot was fertigated at the rate of 200 kg ha^{-1} of NPK (40:40:60) daily throughout the growing period. Hand weeding was done as necessary. At one month after sprouting, staking of vines was done using the trellis method. Harvesting of fresh storge tubers was done at seven months after planting.

2.2.2 Morphological characterization

Morphological characterization was conducted by measuring 28 agro-morphological characters from at least three healthy plants (Table 2.3). The traits measured and data collection procedure used were based on those presented in the International Plant Genetic Resources Instituteôs descriptor list for *Dioscorea* spp. (yam) (IPGRI/ IITA, 1997) with slight modifications (Appendix 1). Only those descriptors or traits that discriminated between genotypes were used in this study. Data were the averages measurements of at least five different healthy plants per genotype. Measurements of the quantitative characters of each accession were made using a meter rule for petiole, vine, leaf length and width and an electronic balance for weight characters. Qualitative traits were visually scored on scales as indicated (Table 2.3). Morphological data collected during the experiment were: number of days to emergence (number of days between planting and emergence), number of stems per plant, number of internode to first branching, stem colour, internode length, absence or presence of wings, wing colour, position of leaves, leaf density, leaf lobation, leaf colour, leaf margin colour, leaf vein colour of upper and lower surfaces, leaf shape, leaf apex shape, distance between lobes, leaf length and width measurements, tip length of mature leaf, tip colour, petiole length of mature leaf, petiole colour, petiole wing colour, tuber shape and flesh colour of central cross section of tuber.

2.2.3 Correlation and principal component analyses

Multivariate analysis of the 52 x 28 data matrix comprising of correlation and PCA was performed in Genstat 12.1 (Payne *et al*., 2009) for Windows statistical software package. Correlation analysis was done in order to determine the interrelationship of the metric traits which are essential for designing breeding strategy. In the PCA, eigenvalues and load coefficient values were generated from the data set. The PCs that had eigenvalues > 0.7 were selected, and those traits that had load coefficient values > 0.25 were considered as relevant scores for the PC, which significantly contributed to distinguish between the genotypes (Jeffers, 1967). The first two PCs which accounted for the higher proportion of the total variation were used to present a two-dimensional scatter plot of the groupings of the accessions.

For the purpose of graphing principal components, the data of the 28 agro-morphological traits were standardised using the formula:

$$\text{Std}(X_{ij}) = (X_{ij} - \bar{x}_{ij})/(\text{variance}(X_{ij})^2 \quad \text{(Eqn 1)}$$

Standardization was done in order to ensure that all the variables have equal weighting in the analysis. The standardized X-variables were then used to compute the principal component scores using the formula:

$$Y_i = \times a_i X_i \quad \text{(Eqn 2);}$$

where Y_i = the principal components, a_i = the eigenvectors, X_i = the values of the various traits and i = 1, 2, ..., n (Manly, 1994).

Table 2.1. Accession numbers, collection sites and main tuber traits of 52 yam (*Dioscorea* spp.) accessions

Species	Accession no.	Local name	Source (village, district/ division, province)	Main tuber traits
D. alata	WR 07/001	Water yam	Waterloo, Koya, Western Rural	Oval-white
	WR 07/004	Yamsiguwi	Waterloo, Koya, Western Rural	Oval-white
	WR 07/007	Water yam	Waterloo, Koya, Western Rural	Round-white
	WR 07/008	Water yam	Waterloo, Koya, Western Rural	Irregular-white
	WR 07/010	White yam	Waterloo, Koya, Western Rural	Irregular-white
	WR 07/013	Water yam	Waterloo, Koya, Western Rural	Oblong-white
	WR 07/014	Water yam	Waterloo, Koya, Western Rural	Oval-oblong-white
	WR 07/015	Water yam	Waterloo, Koya, Western Rural	Oval-oblong-white
	WR 07/016	Water yam	Waterloo, Koya, Western Rural	Oval-oblong-white
	WR 07/020	Yamsiguwi	Waterloo, Koya, Western Rural	Oval-white
	WR 07/022	Unknown	Waterloo, Koya, Western Rural	Round-white
	WR 07/024	Water yam	Waterloo, Koya, Western Rural	Oblong-white
	WR 07/025	Water yam	Waterloo, Koya, Western Rural	Oblong-white
	WR 07/028	Water yam	Waterloo, Koya, Western Rural	Oblong-white
	ER 07/029	Yamsiegbamie	Blama, Small Bo, Kenema	Round-white
	ER 07/030	Gbuheyamsie	Blama, Small Bo, Kenema	Oval-white
	ER 07/031	Gbogboi	Gofor, Dama, Kenema	Round-white
	ER 07/032	Yamsigbamie	Gofor, Dama, Kenema	Oval-white
	ER 07/033	Water yam	Levuma, Kando Leppeama, Kenema	Oval-white
	ER 07/034	Mende yamsie	Levuma, Kando Leppeama, Kenema	Oval-white
	ER 07/036	Yamsieguwi	Nganyagwehun, Nongowa, Kenema	Oval-white
	ER 07/037	Yamsiegbamie	Kenema, Nongowa Kenema	Oval-white
	ER 07/038	Yamsieguwi	Kenema, Nongowa Kenema	Round-white
	ER 07/039	Njayamsi	Kenema, Nongowa Kenema	Oval-oblong-white
	NR 07/041	Mowonmiferra	Rokupr, Magbema, Kambia	Round-white
	NR 07/042	Mowomiferra	Masorie, Magbema, Kambia	Round-white
	NR 07/043	Mabonk	Makassa, Magbema, Kambia	Round-light purple
	NR 07/047	Mawonmiyalla	Ro-thain, Magbema, Kambia	Oblong-white
	NR 07/054	Eneyi	Makoloh, Pakimasabong, Makeni	Oblong-white
	NR 07/057	Anayeyim	Makoloh, Pakimasabong, Makeni	Oval-white
	NR 07/059	Anayeyim	Mangay Loko, Makari Gbanti, Makeni	Round-white
	NR 07/067	White yam	Makeni, Bombali Shebora, Makeni	Round-white
	NR 07/068	Mawonmiyim	Makeni, Bombali Shebora, Makeni	Oblong-white
	NR 07/069	Mawonmiyim	Makeni, Bombali Shebora, Makeni	Round-white
	SR 07/073	Yamsiegboi	Nguala, Kaiyamba, Moyamba	Oval-white
	SR 07/075	Jakenakie	Nguabu, Kaiyamba, Moyamba	Round-white
	SR 07/076	Agabi	Yayema, Kaiyamba, Moyamba	Cylindrical-light purple
	SR 07/079	Njayamsie	Lungi, Kaiyamba, Moyamba	Round-white
	SR 07/080	Njamagha	Pelewahun, Kamajeh, Moyamba	Round-white
	SR 07/081	Njayamsie	Pelewahun, Kamajeh, Moyamba	Oblong-white
	SR 07/082	Njayamsie	Mosongo, Kori, Moyamba	Round-white
	SR 07/084	Njamagha	Mokonde, Kori, Moyamba	Round-white
	SR 07/085	Darvie	Mokonde, Kori, Moyamba	Round-light purple
D. bulbifera	NR 07/045	Mowonmiferra	Kalangba, Magbema, Kambia	Round-yellow
	NR 07/040	Mowonmiyim	Ro-Bolie, Magbema, Kambia	Round-yellow
D. rotundata	NR 07/052	Mabonk	Simbeck, Magbema, Kambia	Cylindrical-white
	NR 07/060	Mawonmiyella	Mangay Loko, Makari Gbanti, Makeni	Cylindrical-white
	NR 07/071	White yams	Makeni, Bombali Shebora, Makeni	Cylindrical-white
	SR 07/072	Agbanie	Nguabu, Kaiyamba, Moyamba	Cylindrical-white
	SR 07/074	Yamsieguwi	Nguala, Kaiyamba, Moyamba	Cylindrical-white
	TDr 95/00005	Improved check	IITA, Nigeria	Cylindrical-white
	TDr 95/18544	Improved check	IITA, Nigeria	Cylindrical-white

Key: WR=Western Region, ER=Eastern Region, NR=Northern Region, SR=Southern Region and TDr=Tropical *Dioscorea rotundata*.

Table 2.2. Mean monthly temperature, rainfall and relative humidity of germplasm collection regions of Sierra Leone (Sierra Leone Met. Sta., 2010)

Year	Site* / Region	Temperature (°C)		Rainfall (mm)	Relative humidity (%)	
		Min.	Max.		0900 am	1500 pm
2006	Waterloo, West	23.3	31.1	232.3	81.5	70.5
	Kenema, East	20.5	31.4	249.7	83.7	70.7
	Kambia, North	19.7	30.6	237.6	80.2	67.2
	Moyamba, South	21.2	32.4	242.2	82.5	65.3
2007	Waterloo, West	23.5	31.3	264.6	82.3	69.9
	Kenema, East	20.6	31.0	247.5	85.4	71.0
	Kambia, North	19.0	30.5	233.5	83.2	68.4
	Moyamba, South	21.7	31.3	232.8	85.1	75.2
2008	Waterloo, West	23.3	31.4	249.2	84.6	69.9
	Kenema, East	20.5	31.5	236.1	83.3	68.9
	Kambia, North	19.8	30.4	245.6	85.9	68.7
	Moyamba, South	21.6	32.9	207.2	82.6	63.4
2009	Waterloo, West	23.5	31.5	291.5	82.5	69.1
	Kenema, East	20.7	31.6	232.4	82.4	65.7
	Kambia, North	20.4	31.7	222.5	82.2	73.1
	Moyamba, South	21.4	32.3	220.5	83.4	62.8
2010	Waterloo, West	23.7	31.7	260.3	81.3	68.7
	Kenema, East	21.0	31.9	225.6	82.7	69.8
	Kambia, North	20.9	32.4	209.9	83.7	73.4
	Moyamba, South	21.7	32.8	204.3	83.1	62.6

*Weather data for five years were only available for four germplasm collection sites

Figure 2.1. Regional map of Sierra Leone showing germplasm collection districts

2.2.4 Factor analysis

For factor analysis (FA), the general model formula was used:

Var (X_i) = Var ($a_i F + e_i$) = Var ($a_i F$) + Var (e_i) = a_i^2Var (F) + Var (e_i) é é é é é é (Eqn 3);

where F and e_i are independent, and the variance of F and X_i are assumed to be unity (Manly, 1994). Thus, $1 = a_i^2 + \text{Var}(e_i)$.

The communality (variance due to common factors) and specificity (variance due to specific factors) were estimated from the relationship:

specificity = 1 ï communality é (Eqn 4);

and their respective percentages were estimated:

$\%F_1 = a_1^2 \times 100\%$ or $\%F_2 = a_2^2 \times 100\%$ .....(Eqn 5);

and % specificity = (1 ï communality) x 100% .é é é é é é é é é é é é é é é (Eqn 6);

(Manly, 1994).

2.2.5 Cluster analysis

For cluster analysis (CA), the standardized data matrix was used to generate pair-wise genetic similarity values among accessions, i.e. the Euclidean dissimilarity coefficient, and then used to generate a hierarchical dendogram through an unweighted pair-group method average (UPGMA) (Sokal and Michener, 1958) using Genstat 12.1 (Payne *et al*., 2009). This analysis was used to study patterns of variance and relationships among accessions, where accessions with close genetic distances were placed in close proximity in the dendrogram.

Table 2.3. Morphological traits measured in 52 yam (*Dioscorea* spp.) accessions. The traits and measurement methods were based on the International Plant Genetic Resources Institute descriptor list (IPGRI/ IITA, 1997)

IPGRI code	Trait acronym*	Trait/ descriptor	Score code ï descriptor state
		Qualitative traits	
7.1.18	SC	Stem colour	1 ï green; 2 ï purplish green; 3 ï brownish green; 4 ï purple
7.1.25	APW	Absence or presence of wings	0 ï absent; 1 ï present
7.1.27	WC	Wing colour	1 ï green; 2 ï green with purple edge; 3 ï purple
7.2.9	PL	Position of leaves (mature leaves)	1 ï alternate, 2 ï opposite, 3 ï alternate at base/ opposite above
7.2.12.2	LL	Leaf lobation	1 ï shallowly lobed; 2 ï deeply lobed
7.2.15	LC	Leaf colour	1 ï yellowish; 2 ï pale green; 3 ï dark green; 4 ï purplish green; 5 ï purple
7.2.16	LVCUS	Leaf vein colour (upper surface)	1 ï yellowish; 2 ï green; 3 ï pale purple; 4 ï purple
7.2.17	LVCLS	Leaf vein colour (lower surface)	1 ï yellowish; 2 ï green; 3 ï pale purple; 4 ï purple
7.2.18	LMC	Leaf margin colour	1 ï green; 2 ï purple
7.2.22	LS	Leaf shape	1ï ovate; 2ï cordate; 3ï cordate long; 4ï cordate broad; 5ï sagittate long; 6ï saggitate broad; 7ï hastate
7.2.23	LAS	Leaf apex shape	1 ï obtuse; 2 ï acute; 3 ï emarginated; 4 ï acuminate; 5 ï aristate; 6 ï caudate; 7 ï cuspidate
7.2.33	TC	Tip colour	1 ï light green; 2 ï dark green; 3 ï purple/ green; 4 ï red; 5 ï yellowish green; 6 ï greenish yellow; 7 ï greenish purple
7.2.37	PC	Petiole colour	1 ï green with purple base; 2 ï green with purple leaf junction; 3 ï green with purple with purple at both ends; 4 ï purplish green with base; 5 ï purplish green with purple leaf junction; 6 ï purplish green with purple at both ends; 7 ï green; 8 ï purple
7.2.38	PWC	Petiole wing colour	1 ï green; 2 ï green with purple; 3 ï purple
7.6.14	TS	Tuber shape	1ï round; 2ï oval; 3ï oval oblong; 4ï cylindrical; 5ï flattened; 6 ï irregular
7.6.30	FCCCS	Flesh colour at central cross section of tuber	1 ï white; 2 ï yellow; 3 ï light purple
		Quantitative traits	
7.1.1	DE	Number of days to emergence	Direct measurement
7.1.17	NS	Number of stems per plant	1 ï 1 stem; 3 ï 3 stem; 5 ï 5 stem; 7 ï 7 stem
7.1.19	NB	Number of internodes to first banching	Direct measurement
7.1.23	IL	Internode length	1 - ≤2.9 cm; 2 ï 3-6.9 cm; 3 ï 7-10.9 cm; 4 ï 11-14.9 cm; 5 - ≥15 cm
7.2.10	LD	Number of leaves (density) per plant	3 ï low; 5 ï intermediate; 7 ï high
7.2.25	DBL	Distance between lobes	1 ï no distance; 5 ï medium; 9 ï very distant
7.2.30.1	LL1	Leaf length-1	1 - ≤5 cm; 2 ï 5.1-8 cm; 3 ï 8.1-11 cm; 4 ï 11.1-14 cm; 5 ï 14.1-18 cm
7.2.30.2	LL2	Leaf length-2	1 ï ≤2 cm; 2 ï 2.1-4 cm; 3 ï 4.1-6 cm; 4 ï 6.1-8 cm
7.2.30.3	LW1	Leaf width-1	1 - ≤5 cm; 2 ï 5.1-8 cm; 3 ï 8.1-11 cm; 4 ï 11.1-14 cm; 5 ï 14.1-18 cm
7.2.30.4	LW2	Leaf width-2	1 ï ≤2 cm; 2 ï 2.1-4 cm; 3 ï 4.1-6 cm; 4 ï 6.1-8 cm; 5 ï 8.1-10 cm
7.2.32	TLM	Tip length of mature leaves	1 ï ≤4 mm; 2 ï 5-9 mm; 3 ï 10-14 mm; 4 ï 15-19 mm; 5 ï 20 mm
7.2.34	PLM	Petiole length of mature leaves	1 - ≤2.9 cm; 2 ï 3-6.9 cm; 3 ï 7-10.9 cm; 4 ï 11-14.9 cm; 5 - ≥15 cm

**Dioscorea alata* genotypes are identified by presence of wings on stem, while *D. bulbifera* and *D. rotundata* are wingless. This trait was visually assessed. Original IPGRI descriptor is presented in Appendix 1

2.3 Results

2.3.1 Correlation analysis

2.3.1.1 Correlation between absence or presence of wings and associated morphological traits

Generally, all the traits whose correlations were greater than or equal to 0.5 significantly ($p<0.05$) influenced the phenotypic expression of the various genotypes (Appendix 3). The correlation between absence or presence of wings and leaf colour was negative (r = -0.517) (Appendix 2). Similar negative correlations were also observed between absence or presence of wings and leaf margin colour (r = -0.562), between absence or presence of wings and number of branches (r = -0.684) and between absence or presence of wings and stem colour (r = -0.556). On the other hand, absence or presence of wings was positively correlated with leaf lobation (r = 0.612), petiole wing colour (r = 0.628) and wing colour of stems (r = 0.714). Wingless genotypes had mostly shallow leaf lobation, which contrasts with *D. alata* genotypes with predominantly deep lobation and winged stems and petioles. The correlations between the absence or presence of wings and the remaining traits were either weak, negative ($-0.014 \leq r \leq -0.481$) or weak, positive ($0.114 \leq r \leq 0.491$).

2.3.1.2 Correlation between distance between lobes, leaf lobation, internode length, petiole length and associated morphological traits

Distance between lobes (DBL) was negatively correlated (r= -0.549) with leaf apex shape (Appendix 2). The distance between lobes in *D. bulbifera* cultivars was so small that the lobes of most leaves overlapped. Genotypes of this species had a peculiar cuspidate leaf apex shape. Also, the correlations of distance between lobes and leaf length-2 (r = -0.601) and distance between lobes and tip colour of mature leaves (r = -0.503) were negative. This contrasts with the positive correlations of distance between lobes and leaf shape (r = 0.613) and distance between lobes and stem colour (r = 0.578). The expansive lobation expressed by most *D. rotundata* genotypes gave them a peculiar saggitate broad leaf shape and purplish to brownish-green vine colour. The correlations of distance between lobes and the remaining traits were either weak, negative ($-0.043 \leq r \leq -0.433$) or weak, positive ($0.074 \leq r \leq 0.42$).

The correlation between leaf lobation (LL) and petiole wing colour was positive (r = 0.520). Some *D. alata* cultivars with purple petiole wing colour were mostly deeply lobed. The correlations between leaf lobation and the other traits were either weak, negative (-0.036 ≤ r ≤ -0.474) or weak, positive (0.056 ≤ r ≤ 0.434).

The correlation between internode length (IL) and leaf length-1 (LL1) was positive (r = 0.634). The correlations between internode length and the other traits were either weak, negative (-0.049 ≤ r ≤ -0.339) or weak, positive (0.046 ≤ r ≤ 0.499).

Petiole length of mature leaf (PLM) had a positive (r = 0.502) correlation with tip length of mature leaves (TLM). The correlations between petiole length of mature leaves and the other traits were either weak, negative (-0.040 ≤ r ≤ -0.154) or weak, positive (0.167 ≤ r ≤ 0.446).

2.3.1.3 Correlation between leaf colour, leaf margin colour, leaf vein colour of upper surface, leaf vein colour of lower surface, leaf shape and associated morphological traits

Leaf colour (LC) had a strong, positive (r = 0.872) correlation with leaf margin colour, and a positive correlation with: leaf density (r = 0.558); leaf vein colour of upper leaf surface (r = 0.630); number of branches per plant (r = 0.510); number of stems per plant (r = 0.593); stem colour (r = 0.627); and flesh colour of central cross section of tuber (r = 0.520) (Appendix 2). The correlations between leaf colour and the other traits was either weak, negative (-0.047 ≤ r ≤ -0.324) or weak, positive (0.035 ≤ r ≤ 0.374).

The correlations between leaf margin colour (LMC) and leaf vein colour of upper leaf surface (r = 0.673), between leaf margin colour and number of stems per plant (r = 0.628), between leaf margin colour and stem colour (r = 0.718) and between leaf margin colour and number of branches per plant (r = 0.503) were positive. The correlations between leaf margin colour and the other traits were either weak, negative (-0.023 ≤ r ≤ -0.435) or weak, positive (0.110 ≤ r ≤ 0.423).

The correlations between leaf vein colour of upper leaf surface (LVCUS) and number of branches per plant (r = 0.520) and between leaf vein colour of upper leaf surface and number of stems per plant (r = 0.544) were positive. Most of the genotypes that produced stem with one or few branches had green upper surface leaf venation, whereas pale purple to purple

venation was common among profuse branching genotypes. The correlations between leaf vein colour of upper leaf surface and other traits were either weak, negative (-0.064 ≤ r ≤ -0.333) or weak, positive (0.138 ≤ r ≤ 0.413).

The correlations between leaf vein colour of lower leaf surface (LVCLS) and flesh colour of the central cross section of tuber (r = 0.644) were positive. The correlations between leaf vein colour of lower leaf surface and the other traits were either weak, negative (-0.010 ≤ r ≤ -0.146) or weak, positive (0.053 ≤ r ≤ 0.411).

Leaf apex shape (LAS) was negatively correlated (r = -0.551) with leaf margin colour, and leaf vein colour of upper surface (r = -0.605). However, the correlations between leaf apex shape and the other traits were either weak, negative (-0.117 ≤ r ≤ -0.497) or weak, positive (0.137 ≤ r ≤ 0.445).

2.3.1.4 Correlation between leaf length and width and associated morphological traits

The correlations between: leaf length-1 and leaf length-2 (r = 0.667); leaf length-1 and leaf width-1 (r = 0.684); leaf length-1 and tip length of mature leaves (r = 0.620); leaf length-1 and leaf width-2 (r = 0.562); and leaf length-1 and petiole length of mature leaves (r = 0.535) were all positive (Appendix 2). In the main, genotypes with larger leaf length-2, leaf width-1, leaf width-2, and tip length of mature leaves, also exhibited larger leaf length-1. However, the correlations between leaf length-1 and other traits were either weak, negative (-0.076 ≤ r ≤ -0.211) or weak, positive (0.065 ≤ r ≤ 0.427).

The correlations between: leaf length-2 (LL2) and leaf width-1 (r = 0.533); between leaf length-2 and petiole wing colour (r = 0.504); and between leaf length-2 and tip length of mature leaves (r = 0.599) were positive. The correlation between leaf length-2 and stem colour (r = -0.502) was negative. It was evident that genotypes which exhibited larger leaf width-1 and tip length of mature leaves with characteristic green with purple petiole wing, also had larger leaf length-2. The correlations between leaf length-2 and the other traits were either weak, negative (-0.163 ≤ r ≤ -0.470) or weak, positive (0.099 ≤ r ≤ 0.463).

Both leaf width-2 (r = 0.612) and petiole length of mature leaves (r = 0.718) were positively correlated with leaf width-1 (LW1). Genotypes which had larger leaf width-2 and petiole

length of mature leaves, also exhibited larger leaf width-1. The correlations between leaf width-1 and the other traits were either weak, negative ($-0.072 \leq r \leq -0.213$) or weak, positive ($0.006) \leq r \leq 0.471$).

Petiole length of mature leaf had a positive ($r = 0.636$) correlation with leaf width-2 (LW2). Genotypes with larger petiole length of mature leaves also had larger leaf width-2. However, the corrleation between leaf width-2 and other traits was either weak, negative ($-0.016 \leq r \leq -0.274$) or weak, positive ($0.022 \leq r \leq 0.455$).

2.3.1.5 Correlation between number of branches, leaf density and associated morphological variables or traits

Number of branches per plant (NB) was positively correlated ($r = 0.572$) with number of stems per plant (Appendix 2). Apparently, some genotypes with high number of stems (NS) also branched profusely. On the contrary, wing colour (WC) was negatively correlated ($r = -0.509$) with number of branches per plant. Among genotypes of *D. alata*, profuse branching was most common in genotypes with wing colour ranging from green with purple edge to purple. The correlations between number of branches per plant and the other traits was either weak, negative ($-0.045 \leq r \leq -0.389$) or weak, positive ($0.019 \leq r \leq 0.330$).

The correlations between: leaf density (LD) and number of branches per plant ($r = 0.685$); leaf density and number of stems per plant ($r = 0.693$); leaf density and leaf margin colour ($r = 0.602$); and between leaf density and flesh colour of the central cross section of tuber ($r = 0.571$) were all positive. The correlations between leaf density and the remaining traits was either weak, negative ($-0.020 \leq r \leq -0.366$) or weak, positive ($0.016 \leq r \leq 0.489$).

2.3.1.6 Correlation between petiole wing colour, tip colour, tip length of mature leaf and associated morphological variables or traits

Petiole wing colour (PWC) had a positive correlations with tip length of mature leaves ($r = 0.565$) and wing colour ($r = 0.708$) (Appendix 2). Among *D. alata* genotypes, many which had tip length of mature leaves ≥ 1.0 cm and green wing with purple edge also exhibited green with purple edge petiole wing colour. The correlations between petiole wing colour and the remaining traits was either weak, negative ($-0.090 \leq r \leq -0.280$) or weak, positive ($0.345 \leq r \leq 0.379$).

Tip colour (TC) had a positive correlation (r = 0.500) with wing colour, but a weak, negative correlation (r = -0.207) with tuber shape. Similarly, the correlation between tip length of mature leaf and wing colour of stem was positive (r = 0.532); but the correlation between tip length of mature leaf and tuber shape was weak, negative (r = -0.331).

2.3.2 Phenotypic variation among genotypes

The variation in leaf, tuber morphology and flesh colour of central cross section of tuber are represented in Figures 2.2, 2.3 and 2.4, respectively. Of the 52 genotypes, 43 belonged to *D. alata*, two genotypes to *D. bulbifera* and seven genotypes to *D. rotundata* (Table 2.1). The 43 genotypes of *D. alata* exhibited different leaf traits ranging from saggitate long green leaf to cordate long dark green leaf. Of the 43 gentotypes, 17 had round, 11 oval, seven oblong, five oval-oblong, two irregular and one cylindrical tuber shape; while the flesh colour of central cross section of tuber of 40 genotypes was white, three exhibited light purple colour.

Genotypes of *D. bulbifera* exhibited cordate light green leaf and cuspidate leaf apex shape. The tuber shape of both genotypes was roun. The bulbils of NR 07/045 were larger than those of NR 07/040. The flesh colour of central cross section of tuber of both genotypes was yellow.

Members of *D. rotundata* exhibited mainly saggitate light green leaf, cordate green purple leaf and saggitate long green leaf. The tuber shape of all genotypes was cylindrical possessing white flesh colour of central cross section of tuber.

a) SR 07/074: Saggitate light green

b) NR 07/052: Cordate green-purple leaf

c) ER 07/038: Saggitate long green leaf

d) WR 07/013: Cordate long green leaf

e) WR 07/025: Saggitate long light green leaf

f) SR 07/085: Cordate long dark green leaf

g) NR 07/041: Cordate long dark green leaf

h) NR 07/040: Cordate broad light green leaf

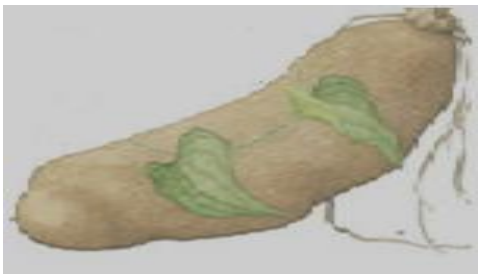
Figure 2.2. Variation in leaf colour, type and shape among yam (*Dioscorea* spp.) germplasm with a-b, and c-g, and f representing accessions of *D rotundata*, *D. alata*, and *D. bulfifera* respectively

a) WR 07/010: Irregular

b) WR 07/024: Oblong shape

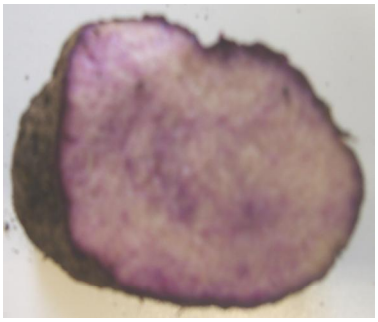
c) SR 07/079: Round shape

d) ER 07/039: Oval-oblong shape

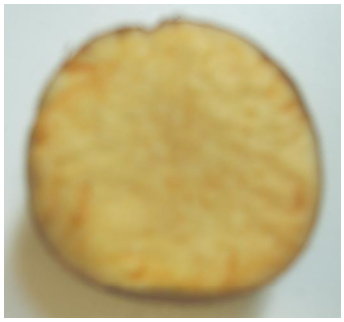


e) SR 07/072: Cylindrical shape

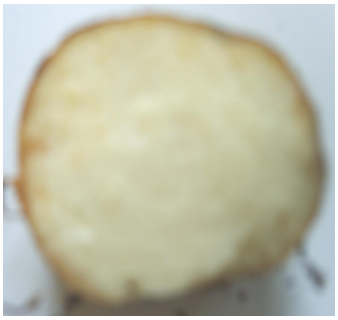
Figure 2.3. Variation in tuber shape among yam (*Dioscorea* spp.) germplasm with a-d, and e representing accessions of *D. alata*, and *D rotundata*, respectively



a) NR 07/043: light purple flesh



b) NR 07/045: yellow flesh



c) SR 07/072: white flesh

Figure 2.4. Variation in flesh colour of central cross section of tuber among yam (*Dioscorea* spp.) germplasm with a, b and c representing accessions of *D. alata*, *D. bulbifera* and *D rotundata*, respectively

2.3.3 Principal component analysis

The principal component scores (PC1 ï PC10) are the eigenvectors (latent vectors) for each of the 28 morphological traits analysed (Table 2.4; Appendix 6). The multivariate analysis based on the 28 morphological traits revealed considerable diversity among the 52 accessions of *Dioscorea* spp. (*D. alata*, *D. bulbifera* and *D. rotundata*) evaluated in this study. Each of the first 10 principal components had eigen-value greater than 0.6 and together explained 86.61% [(7.672 + 5.025 + 3.419 + 2.161 + 1.636 + 1.067 + 0.992 + 0.869 + 0.711 + 0.701)/ 28.00 x 100%] of the total variance in the data set (Table 2.4). An eigen-value is a quantitative assessment of how much a component represents the data. The higher the eigen-value of a component, the more representative it is of the data. The percent of variance explained is dependent on how well all the components summarize the data.

Scores on PC1, which accounted for 27.40% [7.672/ 28.00 x 100%] of the total variation, were correlated (r > 0.25) with traits related to the shoot (absence or presence of wings, distance between lobes, leaf margin colour, leaf length-2 and tip length of mature leaf) (Table 2.4). Scores of PC2, which explained 17.95% [5.025/ 28.00 x 100%] of the total variation, were correlated (r > 0.25) with shoot traits such as leaf colour, leaf density per plant, leaf length-1, leaf leaf vein colour of lower surface and number of stems per plant; and flesh colour of central cross section of tuber. The scores of PC3, which explained 12.21% [3.419/ 28.00 x 100%] of the total variation, were correlated (r > 0.25) with days to emergence; leaf and petiole traits (leaf width-1, leaf width-2, leaf shape, leaf vein colour of lower surface and petiole length of mature leaf); and flesh colour of central cross section of tuber. The scores of PC4, which explained 7.72% [2.161/ 28.00 x 100%] of the variation, were mainly correlated (r > 0.26) with shoot traits (distance between lobes, leaf shape, petiole colour, stem colour and wing colour) and the tuber traits of tuber shape. The scores of PC5, which explained 5.84% [1.636/ 28.00 x 100%] of the total variation, were correlated (r > 0.29) with days to emergence, internode length, leaf apex shape, petiole length, stem colour and tip colour. The scores of PC6, which explained 3.81% [1.067/ 28.00 x 100%] of the total variation, were correlated (r > 0.25) with the shoot traits: internode length, petiole length and below ground trait, tuber shape. The scores of PC7, which explained 3.54% [0.992/ 28.00 x 100%] of the total variation, were correlated (r > 0.25) with days to emergence, leaf apex shape, leaf shape, petiole and petiole wing colour.

Table 2.4. First 10 principal components (PCs) scores of 28 trait means across 52 yam genotypes

No.	Traits[+]	PC-1	PC-2	PC-3	PC-4	PC-5	PC-6	PC-7	PC-8	PC-9	PC-10
1	APW	**0.274**	-0.093	-0.205	0.213	0.127	0.099	-0.152	0.057	0.044	-0.071
2	DBL	**-0.251**	-0.149	0.071	**0.301**	0.043	-0.039	0.113	-0.125	0.024	-0.086
3	DE	-0.055	0.048	**0.252**	0.060	**0.431**	0.089	**0.284**	**-0.435**	-0.084	**0.357**
4	IL	0.130	0.248	0.142	-0.036	**0.326**	**0.258**	0.072	0.055	**-0.264**	-0.214
5	LAS	0.239	-0.055	0.067	-0.246	**-0.293**	0.007	**0.401**	-0.044	-0.051	-0.028
6	LC	-0.217	**0.273**	-0.035	0.191	-0.102	0.007	0.204	-0.036	0.043	0.002
7	LD	-0.138	**0.333**	0.002	-0.182	-0.002	0.135	0.021	0.150	-0.134	0.084
8	LL	0.214	-0.024	-0.213	0.052	0.109	0.086	0.065	0.121	0.007	**0.653**
9	LMC	**-0.264**	0.226	-0.110	0.157	-0.068	-0.086	0.182	-0.064	0.064	0.033
10	LL1	0.182	**0.258**	0.226	0.105	0.105	0.074	-0.017	-0.040	-0.034	-0.069
11	LL2	**0.271**	0.169	0.061	-0.071	0.119	0.045	-0.154	-0.088	0.193	0.081
12	LW1	0.129	0.226	**0.329**	0.007	-0.198	0.022	-0.169	-0.025	0.089	0.214
13	LW2	0.090	0.091	**0.425**	0.093	0.025	-0.116	-0.015	0.082	**0.280**	-0.223
14	LS	-0.139	-0.163	**0.259**	**0.266**	-0.045	0.248	**0.251**	0.148	-0.209	-0.172
15	LVCLS	-0.001	**0.282**	**-0.295**	0.078	0.102	-0.076	-0.161	-0.040	0.063	-0.148
16	LVCUS	-0.211	0.225	-0.076	0.019	0.183	-0.192	-0.236	-0.138	0.239	-0.071
17	NB	-0.216	0.211	0.172	-0.205	0.062	0.129	-0.198	0.109	0.013	0.013
18	NS	-0.192	**0.265**	0.005	0.027	0.023	0.032	-0.116	0.204	**-0.265**	0.107
19	PC	0.110	-0.075	0.022	**0.446**	-0.008	0.010	**-0.427**	-0.155	**-0.487**	0.028
20	PL	-0.078	-0.048	0.014	0.170	**0.334**	**-0.474**	0.169	**0.623**	-0.038	0.118
21	PLM	0.166	0.109	**0.313**	0.160	-0.229	-0.067	-0.088	0.225	0.244	0.145
22	PWC	0.214	0.156	-0.178	0.208	0.197	0.004	**0.259**	-0.132	**0.274**	0.009
23	SC	-0.238	0.075	0.048	**0.264**	**-0.295**	-0.193	0.039	**-0.274**	0.072	0.158
24	TC	0.171	0.237	-0.025	0.038	**-0.345**	-0.236	0.060	0.014	**-0.336**	0.182
25	TLM	**0.274**	0.108	0.105	0.176	0.065	-0.151	0.211	0.001	-0.089	-0.181
26	TS	-0.131	-0.002	-0.082	**0.278**	-0.128	**0.620**	0.016	**0.272**	**0.258**	0.126
27	WC	0.232	0.119	-0.225	**0.272**	-0.155	0.040	0.034	0.029	0.085	-0.190
28	FCCCS	0.019	**0.328**	**-0.266**	-0.086	-0.090	0.080	0.226	0.029	-0.145	-0.160

[+]Details of each trait acronym are provided in Table 2. Values in bold indicate the most important traits (>0.25) that had large contributions to the total variance of a particular principal component

The scores of PC8, which explained 3.10% [0.869/ 28.00 x 100%] of the total variation, were correlated (r > 0.27) with shoot traits (days to emergence, petiole length, stem colour and tuber shape) and the below ground trait, tuber shape. The scores of PC9, which explained 2.54% [0.711/ 28.00 x 100%] of the total variation, were correlated (r > 0.25) with shoot traits (internode length, leaf width-2, number of stems, petiole colour, petiole wing colour, tip colour and tuber shape). The scores of PC10, which explained 2.50% [0.701/ 28.00 x 100%] of the total variation, were correlated (r > 0.35) with internode length and leaf lobation. The PC11 and subsequent PCs were considered less significant since their percentage contribution to the total variation were small. Therefore they are not discussed (Table 2.5).

Table 2.5. Eigen-value, percentage variation and accumulated variation explained by each component of, the first 10 principal components (PCs)

Principal component (PC)	Eigen-values*	Variation of each component (%)	Accumulated variation (%)
1	7.672	27.40	27.40
2	5.025	17.95	45.35
3	3.419	12.21	57.56
4	2.161	7.72	65.28
5	1.636	5.84	71.12
6	1.067	3.81	74.93
7	0.992	3.54	78.47
8	0.869	3.10	81.57
9	0.711	2.54	84.11
10	0.701	2.50	86.61

*[The total of all the variances of the PCs is known as trace. Trace (عم) = $\sum_{i=1}^{p} \mathbf{Var\ (Yi) = 28.00}$]. Each eigen-value (ٯ$_1$) was divided by the trace to estimate percent variability and then accumulated variation

2.3.3.1 Principal component scores

The latent vectors (loadings) (Table 2.4, Appendix 6) are important components used to estimate the first 10 principal components for each of the 52 genotypes presented in Table 2.6. The PC equations are depicted below (Manly, 1994):

PC1 = 0.274 APW ï 0.251 DBL ï 0.055 DE + ---------- + 0.019 FCCCS

PC2 = -0.093 APW ï 0.149 DBL + 0.048 DE + ---------- + 0.328 FCCCS

PC3 = -0.205 APW + 0.071 DBL + 0.252 DE + ---------- - 0.266 FCCCS

PC4 = 0.213 APW + 0.301 DBL + 0.060 DE + ---------- - 0.086 FCCCS

PC5 = 0.127 APW + 0.043 DBL + 0.431 DE + ---------- - 0.090 FCCCS

PC6 = 0.099 APW ï 0.039 DBL + 0.089 DE + ---------- + 0.080 FCCCS

PC7 = -0.152 APW + 0.113 DBL + 0.284 DE + ---------- + 0.226 FCCCS

PC8 = 0.057 APW ï 0.125 DBL ï 0.435 DE + ---------- + 0.029 FCCCS

PC9 = 0.044 APW + 0.024 DBL ï 0.084 DE + ---------- - 0.145 FCCCS

PC10 = -0.071 APW ï 0.086 DBL + 0.357 DE + ---------- - 0.160 FCCCS

Where APW denotes absence or presence of wings, DBL is the distance between lobes, DE is the number of days to emergence (number of days between planting and emergence) and FCCCS is the flesh colour of central cross section of tuber. Other traits represented by dash lines were internode length (IL), leaf apex shape (LAS), leaf colour (LC), leaf density (LD), leaf lobation (LL), leaf margin colour (LMC), leaf length and width measurements, leaf shape (LS), leaf vein colour of upper leaf surface (LVCUS) and leaf vein colour of lower surface (LVCLS), number of internode to first branching (NB), number of stems per plant (NS), petiole colour (PC), position of leaves (PL), petiole length of mature leaf (PLM), petiole wing colour (PWC), stem colour (SC), tip colour (TC), tip length of mature leaf (TLM), tuber shape and wing colour (WC).

Table 2.6. Principal component scores (PC_n) of 28 traits of 52 yam (*Dioscorea* spp.) accessions

No.	Genotype	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10
1	WR 07/001	0.447	3.584	3.419	-5.421	-1.543	-0.236	0.043	0.582	0.063	-1.121
2	WR 07/004	-0.493	3.368	2.537	-6.001	-0.377	0.703	0.344	-0.894	-0.963	0.197
3	WR 07/007	3.226	2.085	3.073	1.252	1.018	-0.594	0.344	0.638	0.206	0.479
4	WR 07/008	2.691	0.884	1.089	1.271	-0.035	0.359	0.411	1.512	0.046	0.228
5	WR 07/010	0.547	5.121	-5.191	-1.628	-1.157	0.631	1.089	-1.137	1.063	0.884
6	WR 07/013	2.596	0.596	1.417	-0.801	0.431	-1.216	-0.744	0.687	-0.526	0.222
7	WR 07/014	1.919	1.147	1.743	-0.119	0.918	0.196	0.008	1.104	-0.985	0.677
8	WR 07/015	-6.512	-0.044	1.417	0.256	0.410	-0.160	0.869	-0.575	0.541	0.493
9	WR 07/016	0.604	-1.255	-0.398	0.476	1.638	0.416	0.679	0.299	-0.978	0.114
10	WR 07/020	0.424	4.281	-1.886	1.270	-0.828	1.801	1.984	-0.370	-0.832	-0.678
11	WR 07/022	3.607	0.663	2.295	1.358	1.181	-0.054	0.906	-0.551	0.502	1.203
12	WR 07/024	2.996	-0.024	1.774	0.038	0.262	-0.115	0.274	-1.682	-0.094	0.332
13	WR 07/025	0.215	1.170	1.745	0.177	1.635	1.215	-1.815	1.988	0.302	0.171
14	WR 07/028	-1.502	-0.243	-0.795	0.097	1.913	1.539	-1.707	-1.776	0.056	-0.836
15	ER 07/029	-0.701	6.456	-3.506	1.744	1.041	-0.076	-0.022	0.959	-1.545	-0.637
16	ER 07/030	-1.041	-0.068	1.256	0.443	0.706	0.608	-1.577	1.670	0.099	-1.432
17	ER 07/031	2.749	1.148	2.548	1.224	1.146	-1.472	0.666	-1.083	-0.214	-1.377
18	ER 07/032	-0.507	0.808	0.907	-0.055	1.813	4.024	-1.820	-0.268	0.273	0.938
19	ER 07/033	-6.532	0.543	1.777	1.799	-0.152	-0.475	0.284	-1.530	-0.967	0.877
20	ER 07/034	3.040	1.605	2.683	1.869	1.449	-0.268	0.942	-0.614	0.964	0.528
21	ER 07/036	1.076	6.533	-4.474	0.862	0.395	-1.894	-1.406	0.118	1.471	-0.516
22	ER 07/037	3.213	0.367	1.929	0.919	0.228	-0.364	0.867	-2.416	-0.085	0.470
23	ER 07/038	1.200	-0.617	0.896	0.677	-2.464	-0.396	-0.891	-0.670	0.770	-1.573
24	ER 07/039	1.080	-1.782	-1.133	0.254	-2.469	0.326	-0.449	-0.363	-0.096	0.413
25	NR 07/040	1.039	-0.586	-0.451	0.485	-1.795	0.077	-0.693	-0.677	-0.259	0.042
26	NR 07/041	1.427	-0.922	-0.249	0.303	-2.195	0.071	-1.089	-0.671	0.846	0.498
27	NR 07/042	1.117	-0.165	0.932	1.047	-2.123	1.704	-1.100	-0.186	1.382	-1.142
28	NR 07/043	1.703	-0.945	-0.083	0.855	-1.760	1.001	-0.044	-0.195	-0.084	-0.134
29	NR 07/045	1.000	-2.180	-1.453	-0.139	-1.806	0.467	-0.361	-0.894	-0.553	0.370
30	NR 07/047	1.275	-0.887	0.992	1.010	-1.036	-0.842	0.219	0.704	-0.096	-1.531
31	NR 07/052	1.503	-1.545	0.388	0.720	-1.707	-1.018	-0.289	0.942	0.294	0.715
32	NR 07/054	1.380	-1.024	0.420	0.718	-1.455	-0.279	-0.322	0.819	0.352	1.347
33	NR 07/057	0.902	-0.095	-2.150	0.549	-0.792	0.647	1.669	0.987	-1.326	-0.352
34	NR 07/059	2.384	-1.106	-0.317	-0.244	-1.053	0.591	-0.173	1.512	-0.512	0.842
35	NR 07/060	-0.904	-3.000	-1.245	-0.287	1.349	0.025	0.118	-0.305	-0.712	-1.921
36	NR 07/067	0.022	-2.943	-1.529	-0.763	0.775	0.340	-0.223	-0.529	-0.049	-0.446
37	NR 07/068	-0.633	-0.763	-1.903	-1.565	0.817	-1.916	-2.436	-0.191	-0.216	1.651
38	NR 07/069	-0.132	-2.872	-1.820	-0.444	1.599	1.037	-0.574	-0.362	-0.906	-0.283
39	NR 07/071	0.468	-3.191	-1.215	-1.216	0.917	-0.410	0.135	-0.960	-0.785	-0.440
40	SR 07/072	0.285	-0.466	-1.590	-1.192	2.021	-2.366	-1.442	-0.476	0.800	0.176
41	SR 07/073	0.930	-2.206	-0.884	-1.367	0.800	0.126	1.305	0.905	1.797	-0.195
42	SR 07/074	-0.882	-3.211	-1.848	-1.420	0.207	0.439	1.940	1.292	1.185	0.313
43	SR 07/075	-6.409	-0.133	0.947	0.145	-0.397	-0.327	0.339	0.239	0.698	-0.174
44	SR 07/076	-4.972	-0.156	0.021	2.009	0.150	0.045	-0.278	-0.124	0.475	-0.411
45	SR 07/079	1.414	-1.765	-0.047	-0.387	-0.082	-1.458	0.185	0.142	-1.309	-1.753
46	SR 07/080	0.809	-2.268	-1.332	0.257	-0.917	-0.546	0.112	0.324	-0.665	0.757
47	SR 07/081	-6.420	-0.123	0.997	0.157	-0.309	-0.309	0.396	0.151	0.681	-0.102
48	SR 07/082	-6.409	-0.133	0.947	0.145	-0.397	-0.327	0.339	0.239	0.698	-0.174
49	SR 07/084	0.512	-1.859	-1.184	-1.197	2.012	0.095	2.053	0.442	2.047	-0.206
50	SR 07/085	0.320	-1.698	-1.020	-0.586	0.162	-0.323	-0.010	0.347	-0.828	1.103
51	TDr 95/00005	0.613	-1.673	-1.454	0.063	0.960	-0.197	0.749	0.029	-0.701	0.470
52	TDr 95/18544	-6.682	1.592	1.010	0.386	-1.105	-0.842	0.196	0.866	-1.327	0.923

2.3.3.2 Graphical presentation of principle component analysis

The standardized data (Eqn 1) were used to compute the principal component scores using the PC formulae (Eqn 2; Appendix 6). The first two most important principal components, (PC1 and PC2), which contributed 45.35% of the total variance in the data set, were plotted in a graph (Figure 2.5).

The 28 morphological traits classified the 52 genotypes into five main clusters, namely: groups A, B, C, D and E. While most of the genotypes clustered around the center of the graph (Figure 2.5), others were widely scattered along both PC axes. Despite the small amount of overlap between sub-groups B_1 and B_2, the dispersion pattern generally separated the species based on the measured morphological traits (Figure 2.5). Of the 52 genotypes studied, five (WR 07/024, SR 07/075, 07/073, ER 07/032 and NR 07/042) overlapped in sub-groups B_1 and B_2, whereas four genotypes (WR07/010, NR 07/041, ER 07/038 and NR 07/067) overlapped in sub-groups B_1 and B_3 indicating the possibility of duplicate genotypes in the germplasm.

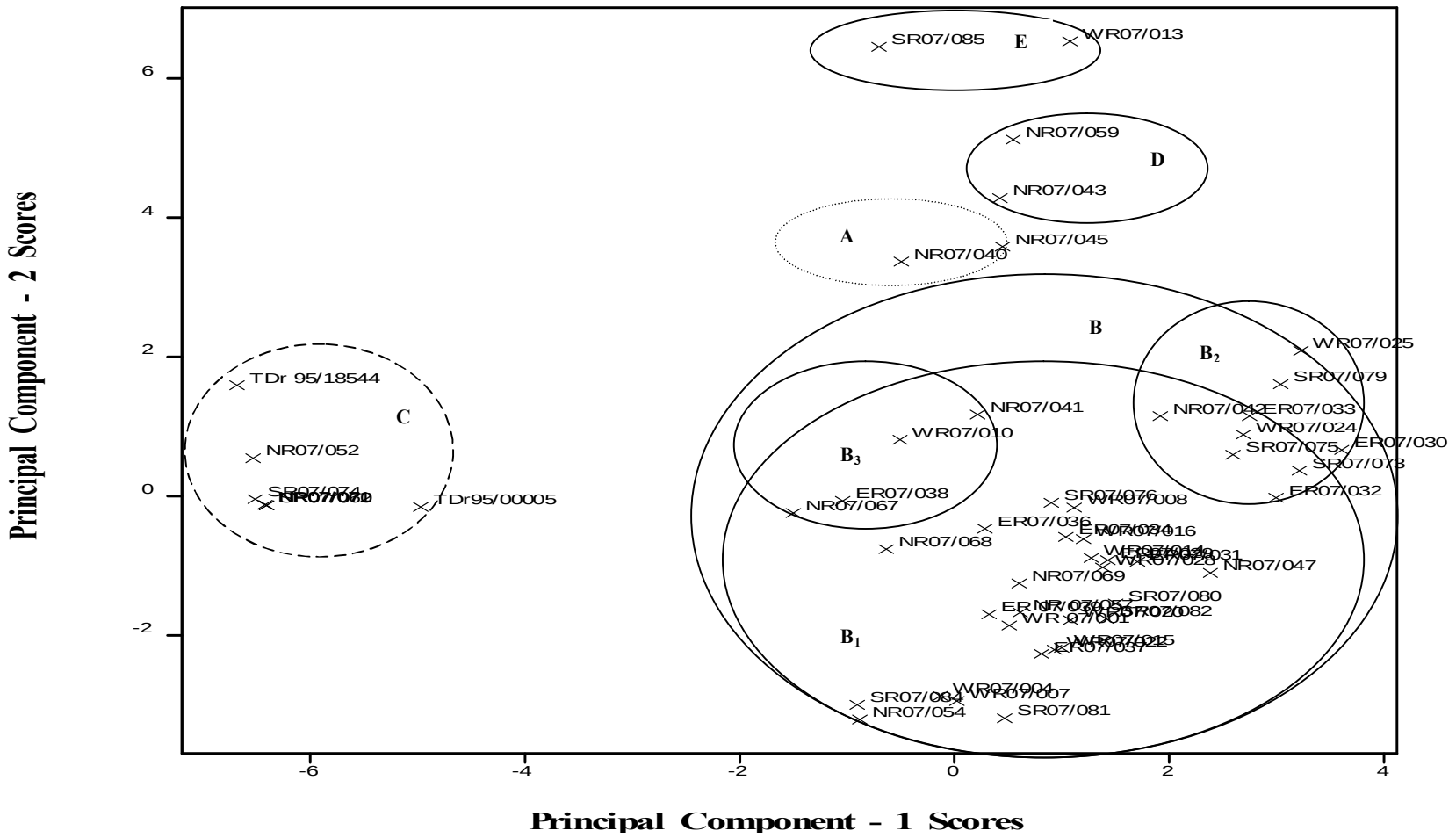


Figure 2.5. Two-dimensional plot of the first two principal components (PC-1 and PC-2). Accessions that are encircled by the dotted (group A = *D. bulbifera*), dashed (group C = *D. rotundata*), and solid (groups B, D and E = *D. alata*) lines

2.3.4 Factor analysis

The six principal component eigen-values that were greater than 1.0 (Table 2.7; Appendix 7) suggest the use of six factors in the factor analysis (Biabani and Pakniyat, 2008). Factor loadings with coefficients greater than or equal to 0.5 (ignoring the sign) were considered important and emboldened. These large and moderate loadings indicated how the traits were related to the factors (Manly, 1994). The contributions by the communalities were fairly high with 24 traits exhibiting higher communality over the specificity (four traits) (Table 7). Factor 1 was heavily loaded with: absence or presence of wings (0.7010); distance between lobes (-0.6071); leaf apex shape (0.7589); leaf colour (-0.7714); leaf density (0.5178); leaf margin colour (-0.8797); leaf length-2; leaf vein colour of upper surface (0.7061); number of branches per plant (-0.6407); number of stems (-0.6472); stem colour (-0.7118); and tip length of mature leaves (0.6008). Factor 2 was loaded with: leaf density (0.5270); leaf length-1 (0.6210); leaf vein colour of lower surface (0.6135); petiole wing colour (0.5824); tip colour (0.6545); wing colour (0.5780); and flesh colour of central cross section of tuber (0.7674). Factor 3 was loaded with: absence or presence of wings (0.6355); and leaf width-1 (-0.5539). Factor 4 was loaded with leaf width-2 (-0.5799). Factor 5 was loded with stem colour (0.5070); and factor 6 with number of days to emergence (-0.4793). Most of the variation in the traits was accounted for by factor 1, with moderate (-0.5178) to large (-0.8797) loadings compared to the traits loaded in the other factors. This makes rotating the factors to further explore the variables unnecessary.

The variation in absence or presence of wings (APW) was strongly influenced by communality (92.56%) compared with the specificity (7.44%). Factor 1 (50.14%) contributed most of the variation in the communality compared to factors 2 (0.92%), 3 (40.39%), 4 (0.80%), 5 (0.30%) and 6 (0.01%).

The variation in distance between lobes (DBL) was largely due to communality (80.51%) compared to specificity (19.49%). Factor 1 (36.86%) contributed most to the variation in the communality compared to 21.15, 1.19, 0.75, 17.25 and 3.31% contributions by factors 2, 3, 4, 5 and 6, respectively.

Table 2.7. Loadings of common and specific factors of 28 traits of 52 yam (*Dioscorea* spp.) accessions analyzed by factor analysis

	Factor loadings								% variation						
Traits	F1	F2	F3	F4	F5	F6	Communality	Specificity	F1	F2	F3	F4	F5	F6	Specificity
APW	**0.7081**	0.0959	**0.6355**	-0.0895	0.0544	0.0116	**0.9256**	0.0744	**50.14**	0.92	**40.39**	0.80	0.30	0.01	7.44
DBL	**-0.6071**	-0.4599	0.1090	-0.0871	0.4153	-0.1819	**0.8051**	0.1949	**36.86**	21.15	1.19	0.75	17.25	3.31	19.49
DE	-0.1707	-0.0351	-0.2377	-0.3440	0.0036	-0.4793	0.4350	**0.5650**	2.91	0.12	5.65	11.83	0.00	**22.97**	56.50
IL	0.2065	0.4953	-0.1399	-0.3618	-0.3238	-0.3424	**0.6605**	0.3395	4.26	**24.53**	1.96	13.09	10.49	11.72	33.95
LAS	**0.7589**	0.1253	-0.4187	0.4185	0.1088	-0.0552	**0.9570**	0.0430	**57.59**	1.57	17.53	17.51	1.18	0.30	4.30
LC	**-0.7714**	0.4614	-0.0352	0.0539	0.1519	-0.0603	**0.8387**	0.1613	**59.51**	21.29	0.12	0.29	2.31	0.36	16.13
LD	**-0.5178**	**0.5270**	-0.2606	0.0425	-0.3422	0.0143	**0.7329**	0.2671	26.81	**27.77**	6.79	0.18	11.71	0.02	26.71
LL	0.4938	0.1671	0.3904	0.0571	-0.0900	-0.0084	0.4355	**0.5645**	**24.38**	2.79	15.24	0.32	0.81	0.01	56.45
LMC	**-0.8797**	0.3536	0.0550	0.1795	0.1409	-0.0545	**0.9569**	0.0431	**77.39**	12.50	0.30	3.22	1.98	0.30	4.31
LL1	0.2932	**0.6210**	-0.2106	-0.4911	0.0313	-0.1619	**0.7844**	0.2156	8.60	**38.56**	4.44	24.12	0.10	2.62	21.56
LL2	**0.5929**	0.4898	-0.0479	-0.3006	-0.1985	-0.0060	**0.7235**	0.2765	**35.15**	23.99	0.23	9.04	3.94	0.00	27.65
LW1	0.2297	0.4974	**-0.5539**	-0.4343	0.1314	0.2173	**0.8601**	0.1399	5.28	24.74	**30.68**	18.86	1.73	4.72	13.99
LW2	0.1835	0.1602	-0.4571	**-0.5799**	0.2321	-0.1628	**0.6849**	0.3151	3.37	2.56	20.89	**33.63**	5.39	2.65	31.51
LS	-0.2516	-0.4091	-0.1940	-0.1737	0.4500	-0.2334	**0.5554**	0.4446	6.33	16.73	3.76	3.02	20.25	5.45	44.46
LVCLS	-0.2240	**0.6135**	0.4010	0.0706	-0.2413	0.1101	**0.6627**	0.3373	5.02	**37.64**	16.08	0.50	5.82	1.21	33.73
LVCUS	**-0.7061**	0.2404	0.0803	-0.1019	-0.2164	-0.0011	**0.6201**	0.3799	**49.86**	5.78	0.65	1.04	4.68	0.00	37.99
NB	**-0.6407**	0.1444	-0.4200	-0.2204	-0.3513	0.0745	**0.7854**	0.2146	**41.05**	2.09	17.64	4.86	12.34	0.56	21.46
NS	**-0.6472**	0.3306	-0.0796	-0.1323	-0.1660	0.0726	**0.5848**	0.4152	**41.89**	10.93	0.63	1.75	2.75	0.53	41.52
PC	0.2376	-0.0312	0.3677	-0.3925	0.3608	0.1730	**0.5068**	0.4932	5.65	0.10	13.52	15.40	13.02	2.99	49.32
PL	-0.2157	-0.1560	0.1515	-0.1395	0.0482	-0.2189	0.1635	**0.8365**	4.65	2.43	2.30	1.95	0.23	4.79	83.65
PLM	0.3678	0.3192	-0.3449	-0.4519	0.3484	0.1847	**0.7158**	0.2842	13.53	10.19	11.89	20.42	12.14	3.41	28.42
PWC	0.3863	**0.5824**	0.4338	-0.0107	0.0818	-0.3023	**0.7748**	0.2252	14.92	**33.92**	18.82	0.01	0.67	9.14	22.52
SC	**-0.7118**	0.0528	-0.1052	0.0524	**0.5070**	0.1731	**0.8103**	0.1897	**50.67**	0.28	1.11	0.28	25.70	2.99	18.97
TC	0.2988	**0.6545**	-0.1452	0.1123	0.1465	0.2340	**0.6276**	0.3724	8.93	**42.84**	2.11	1.26	2.15	5.47	37.24
TLM	**0.6008**	0.4542	-0.0002	-0.2410	0.2048	-0.2553	**0.7325**	0.2675	**36.10**	20.63	0.00	5.81	4.19	6.52	26.75
TS	-0.3410	-0.0226	0.1976	0.0232	0.1983	0.0809	0.2023	**0.7977**	11.63	0.05	3.90	0.05	3.93	0.65	79.77
WC	0.4650	**0.5780**	0.4900	0.0734	0.2476	0.1236	**0.8723**	0.1277	21.62	**33.41**	24.01	0.54	6.13	1.52	12.77
FCCCS	-0.1542	**0.7674**	0.1247	0.3606	-0.2693	0.0013	**0.8308**	0.1692	2.38	**58.89**	1.56	13.00	7.25	0.00	16.92

[+]Details of each acronym/ trait are provided in Table 2. Values in bold indicate the most important traits (>0.25) that contributed much to the total variance of the particular component

The variation in leaf apex shape (LAS) was strongly influenced by the communality (95.70%) compared to the specificity (4.30%). Factor 1 (57.59%%) accounted for the largest variation in the communality compared to factors 2 (1.57%), 3 (17.53%), 4 (17.51%), 5 (1.18%) and 6 (0.30%).

The variation in leaf colour (LC) was largely due to the communality (83.87%) compared with the specificity (16.13%). Factor 1 (59.51%) contributed most of the variation in the communality compared to factors 2 (21.29%), 3 (0.12%), 4 (0.29%), 5 (2.31%) and 6 (0.36%).

The variation in leaf density (LD) was more influenced by the communality (73.29%) compared to the specificity (26.71%). Factor 2 (27.77%) contributed most of the variation in the communality compared to factors 1 (26.81%), 3 (6.79%), 4 (0.18%) 5 (11.71%) and 6 (0.02%).

The variation in leaf margin colour (LMC) was strongly influenced by the communality (95.69%) of which factor 1 (77.39%) contributed the most compared with factors 2 (12.50%), 3 (0.30%), 4 (3.22%), 5 (1.98%) and 6 (0.30%). The specificity contributed 4.31%.

The variation in leaf length-1 (LL1) was explained by 78.44% contribution from communality of which factor 2 (36.97%) contributed most compared to factors 1 (8.60%), 3 (4.44%), 4 (24.12), 5 (0.10%) and 6 (2.62%). The specificity accounted for 21.56%.

The variation in leaf length-2 (LL2) was largely due to communality (72.35%) of which factor 1 (35.15%) contributed highest compared to 23.99, 0.23, 9.04, 3.94 and 0.00% contributions by factors 2, 3, 4, 5 and 6 respectively. The specificity accounted for 27.65%.

The variation in leaf width-1 (LW1) was strongly influenced by the communality (86.01%) compared with the specificity (13.99%). Factor 3 (30.68%) contributed most to the variation in the communality compared with factors 1 (5.28%), 2 (24.74%), 4 (18.86%), 5 (1.73%) and 6 (4.72%).

The variation in leaf width-2 (LW2) was explained by 68.49% contribution from communality of which factor 4 (33.63%) contributed most compared to factors 1 (3.37%), 2 (2.56%), 3 (20.89%), 5 (5.39%) and 6 (2.65%). The specificity contributed 31.51%.

The variation in leaf shape (LS) was explained by 55.54% contribution from communality of which factor 5 (20.25%) contributed highest than factors 1 (6.33%), 2 (16.73%), 3 (3.76%), 4 (3.02%) and 6 (5.45%). The specificity accounted for 44.46%.

The variation in leaf vein colour lower surface (LVCLS) was largely due to communality (66.27%) than the specificity (33.73%). Factor 2 (34.27%) contributed highest compared to 5.02, 16.08, 0.50, 5.82 and 1.21% inputs by factors 1, 3, 4, 5 and 6 respectively.

The variation in leaf vein colour upper surface (LVCUS) was explained by 62.01% contribution from communality of which factor 1 (49.86%) contributed most compared with factors 2 (7.58%), 3 (0.65%), 4 (1.04%), 5 (4.68%) and 6 (0.00%). The specificity accounted for 37.99%.

The variation in number of branches (NB) was more influenced by communality (78.54%) compared with the specificity (21.46%). Factor 1 (41.05%) contributed the highest to the variation in the communality compared to 2.09, 17.64, 4.86, 12.34 and 0.56% inputs by factors 2, 3, 4, 5 and 6 respectively.

The variation in number of stems (NS) was explained by 58.48% contribution from communality of which factor 1 (41.89%) contributed most compared with factors 2 (10.93%), 3 (0.63%), 4 (1.75%), 5 (2.75%) and 6 (0.53%). The specificity accounted for 41.52%.

The variation in petiole length of mature leaf (PLM) was explained by 71.58% contribution from communality of which factor 4 (20.42%) contributed most compared with factors 1 (13.53%), 2 (10.19%), 3 (11.89%), 5 (12.14%) and 6 (3.41%). The specificity accounted for 28.42%.

The variation in petiole wing colour (PWC) was strongly influenced by 77.48% contribution from communality compared to 22.52% input by the specificity. Factor 2 (33.92%) contributed most compared with factors 1 (14.92%), 3 (18.82%), 4 (0.01%), 5 (0.67%) and 6 (9.14%).

The variation in stem colour (SC) was largely due to communality (81.03%) than the specificity (18.97%). Factor 1 (50.67%) contributed most to the variation in the communality compared with 0.28, 1.11, 0.28, 25.70, and 2.99% contributions by factors 1, 3, 4, 5 and 6 respectively.

The variation in tip colour (TC) was explained by 62.76% contribution from communality of which factor 2 (42.84%) contributed most compared with factors 1 (8.93%), 3 (2.11%), 4 (1.26%), 5 (2.15%) and 6 (5.47%). The specificity accounted for 37.24%.

The variation in tip length of mature leaf (TLM) was explained by 73.25% contribution from communality of which factor 1 (36.10%) contributed most compared with factors 2 (20.63%), 3 (0.00%), 4 (5.81%), 5 (4.19%) and 6 (6.52%). The specificity accounted for 26.75%.

The variation in wing colour (WC) was strongly influenced by the common factors (87.23%) compared to the specificity (12.77%). Factor 2 (33.41%) accounted for most of the variation in the communality compared to factors 1 (21.62%), 3 (24.01%), 4 (0.54%) 5 (6.13%) and 6 (1.52%).

The variation in flesh colour of central cross section of tuber (FCCCS) was explained by 83.08% contribution from communality of which factor 2 (58.89%) contributed most compared with factors 1 (1.38%), 3 (1.56%), 4 (13.00%), 5 (7.25%) with no contribution by factor 6. The specificity accounted for 16.92%.

Generally, all the traits measured were useful in determining variability in the 52 accessions. The six factors considered distinguished the traits into groups. The highest weightings by factor 1 were given to leaf colour (-0.7719) and leaf margin colour (-0.8797). These traits were not only highly positively correlated (r = 0.872) (Appendix 2), but also served as important components in distinguishing between the accessions.

2.3.5 Cluster analysis

The dendrogram of the hierarchical cluster analysis (HCA) separated the 52 genotypes into different clusters with Euclidean distance dissimilarities ranging between 0.8 and 1.0 (Figure 2.6; Appendix 8). At the dissimilarity distance of 0.90, the dendrogram identified six main clusters, A, B, C, D, E and F. Clusters A, E and F had two genotypes each, cluster B consisted of 38 genotypes and cluster C consisted of one genotype and cluster D had seven genotypes. Genotypes of cluster A belong to *D. bulbifera*, while genotypes of clusters B, C, E and F belong to *D. alata*, and genotypes of cluster D belong to *D. rotundata*. At the 0.95 dissimilarity distance, cluster B was further divided into three sub-clusters: B_1, B_2 and B_3, each consisting of 5, 30 and 3 genotypes, respectively. The dendogram of the hierarchical

cluster analysis (Figure 2.6) produced a similar grouping of genotypes as did the PCA scatter plot (Figure 2.5). The clustering patterns of the various genotypes in the dendrogram revealed the proximity of their genetic distance.

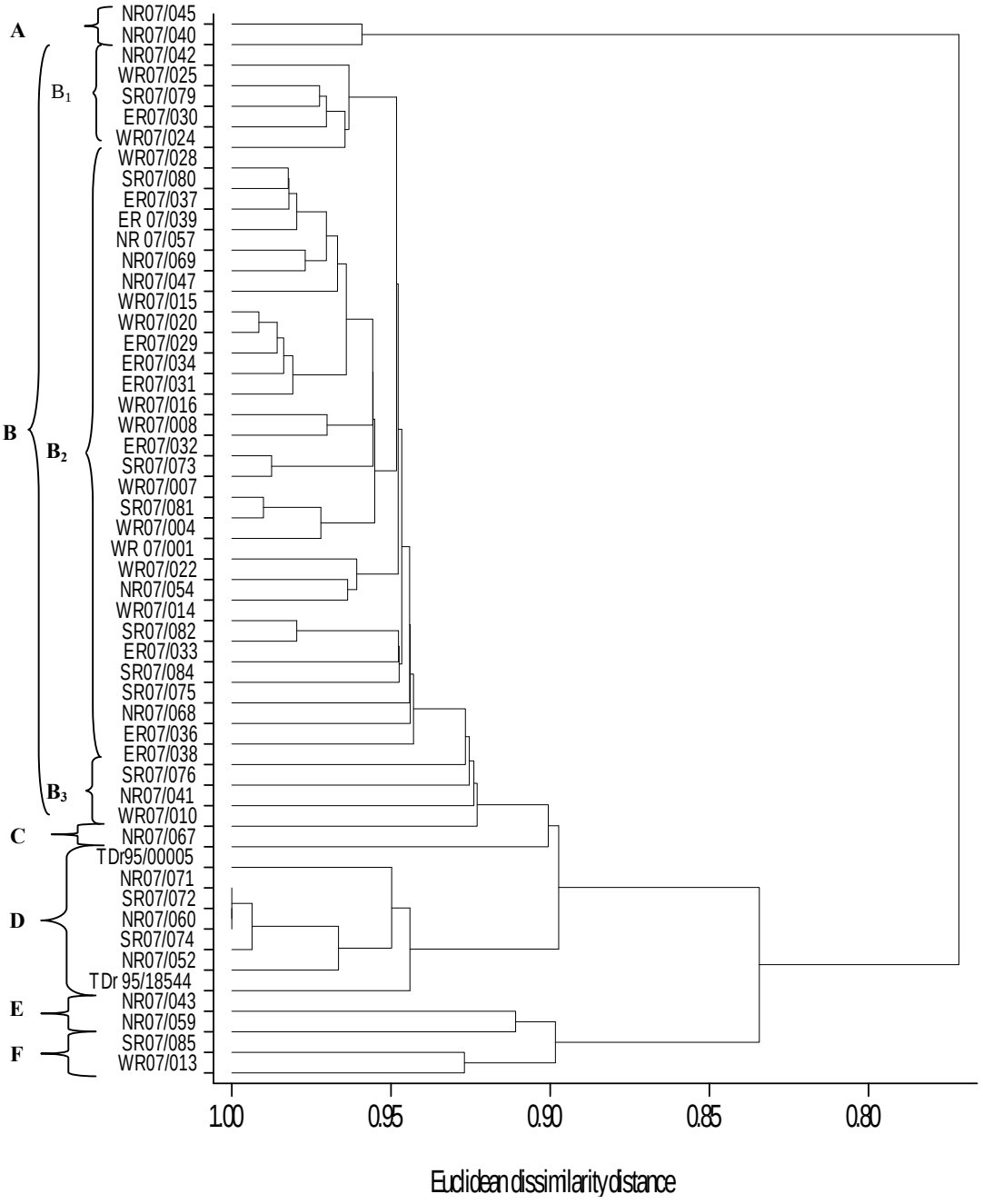


Figure 2.6. Dendrogram showing genetic diversity among 52 yam accessions (43 each of *D. alata*, two each of *D. bulbifera* and seven each of *D. rotundata*) based on morphological traits

2.4 Discussion

The standard yam descriptor list (IPGRI/ IITA, 1997) was a useful tool for assessing the available variation among Sierra Leone accessions. The polymorphism showed for 16 qualitative descriptors and 12 quantitative traits confirm that the selected descriptors are appropriate for appraising yam diversity. A better understanding of the existing traditional yam cultivars in Sierra Leone is one of the prerequisites for breeding new cultivars with novel or improved characteristics.

A correlation coefficient quantifies the degree to which the variation in one variable (or trait) is mirrored by or ñaffectsò variation in another i.e. it provides a measure of the intensity of the biological or otherwise association between the two variables. The sign of the correlation coefficient provides an indication of either a positive or negative association between two variables. Correlation coefficients provide guidance with regard to the execution of direct or indirect selection of traits and the consequences thereof for other traits. In this study, it was considered that quantification of the biological associations between morphological traits in yams would provide invaluable information to current and future breeding programmes. Pearson product correlation coefficients revealed significant associations among most of the morphological traits studied (Appendix 3). Other associations with winglessness in genotypes were profuse branching pattern, purple leaf margin and purple-green stem.

The interrelationships between internode length, leaf length-1, petiole length and tip length were particularly significant in the classification of the genotypes. For instance, the higher internode length noted in some accessions was associated with a corresponding increase in leaf length-1. It appeared that as the leaves were more well spaced apart on vines, thereby improving the harnessing of solar radiation for photosynthesis, there was an associated increase in leaf length-1. Similarly, some genotypes which had mature leaves with larger tip length also had larger petiole length of the mature leaves.

The positive association revealed by Pearsonôs correlation for the morphological traits: leaf colour, leaf margin colour, leaf vein colour of upper surface and leaf shape was the probable cause of the unique colour venation in the leaves of some genotypes. Another important observation was that some of the genotypes that had profuse branching, purplish-green stem and pale purple to purple leaf vein colour of upper leaf surface also had purple leaf margin colour. These findings suggest that variability in leaf apex shape was partly due to the

influence by leaf margin colour and leaf vein colour of upper surface, with the latter stronger than leaf margin colour. Genotypes WR 07/013 and SR 07/085, which exhibited pale purple to purple venation, also had purple leaf margin colour and acute leaf apex shape.

The variation in morphological traits within and between landraces of *D. alata, D. bulbifera* and *D. rotundata* is likely due to initial sexual recombination and possibly mutation. This is often followed by intensive selection by isolated human communities in diverse environments (Martin, 1976). Yams are dioceous implying that spontaneous hybridization may have contributed to the ancestory of some of the accessions, and improvement may have been far more often by selection of somatic mutants. Also, Velayudhan *et al*. (1989) suggested that continuous vegetative propagation and selection within germplasm may contribute to phenotypic variation in the species. In the present study, however, only two genotypes of *D. bulbifera*, NR 07/040 and NR 07/045, flowered. Thus, the interspecific variation across species level was possibly due to the fact that *D. alata* and *D. rotundata* form part of the section Enanthiophylum while *D. bulbifera* belongs to Opsophyton.

The traits identified in this study should be useful as markers for the classification and genetic improvement of the genotypes. Those that discriminated the most between the accessions were: the number of days to emergence, shoot traits (absence or presence of wings, leaf colour, density, lobation, position, shape and size of leaf, number of stems and branches) and below ground traits (tuber shape and flesh colour of central cross section of tuber). Martin and Rhodes (1973, 1977), Martin (1976), Onwueme (1978), Sastrapradja (1982), and Hasan *et al*. (2008) also noted that leaf and other shoot growth and tuber traits are the most effective morphological traits to classify yams.

The FA indicated significant contributions in the factor loadings of the 28 traits which underpins their relevance in determining the variability among the 52 accessions. Six factors which had eigen-values greater than 1.0 were retained (c.f. Manly, 1994; Biabani and Pakniyat, 2008). These factors accounted for 75% of the total genetic variability. Factor 3 had the highest negative associations (19 traits) whereas factor 4 had the least (10 traits). The sign on the loadings indicates the direction of the relationship between the factor and the trait measured (Biabani and Pakniyat, 2008). Two traits with high weighting in the same factor are expected to be highly correlated. This suggests that these traits could be probably influenced by similar gene(s) and may be used to identify variation among accessions (Biabani and

Pakniyat, 2008). Other factors (7, 8, 9 and 10) explained 25% of the genetic variation, and were considered to be not as important in characterizing the yam accessions.

Factor 1 had moderate, positive loading for leaf length-2, tip length of mature leaf, absence or presence of wing and leaf apex shape on one hand; and moderate (leaf density, distance between lobes, number of branches, number of stems, leaf vein colour, stem colour and leaf colour) to high (leaf margin colour) negative influence on characterization of the accessions. It, therefore, measured the importance of leaf shape and size attributes against shoot growth and colour traits in distinguishing the accessions. Factor 2 (leaf density, wing colour, petiole wing colour, leaf length-1, tip colour and flesh colour of central cross section of tuber) had a moderate, positive influence in the classification of the accessions. Factor 3 had a moderate, positive loading for absence or presence of wing, and a moderate, negative loading for leaf width-1. It measured the constrast between wing production ability of the various genotypes and leaf width-1. Factor 4 had a moderate, negative loading for leaf width-2. It measured the contribution of leaf growth parameter to genotype classification. Factor 5 had a moderate, positive loading for stem colour, whereas factor 6 exhibited low, negative loading for days to emergence. Days to emergence contributed the highest weighting in factor 6 compared to the other characters. Among traits that heavily loaded as specificity were days to emergence and tuber shape. The significance of these traits in yam breeding programme is crucial. For instance, the development of early maturing genotype may require the reduction in the number of days to emergence. Early emergence enhances the full utilization of the active growth period, which in turn provides tubers with the opportunity to attain their normal size and shape. The longer the number of days to emergence, the shorter the active growth period. Additionally, infertile and poorly irrigated soils on one hand, coupled with diseased planting material on the other, could affect tuber shape. Yam tuber shape is one of the the desirable traits in market-oriented breeding.

Accessions placed in group A (NR 07/045 and NR 07/040), based on PCA, belong to *D. bulbifera* and were characterized by wingless stems and petioles, and sharp angled bulbils with depressions containing preformed buds. Accessions placed in groups B, D and E belong to *D. alata*. Accessions of group B belonging to *D. alata* and were highly variable with irregular, oblong, oval-oblong and round tuber shapes, with flesh colour of central cross section of tubers ranging from light purple and white. This suggests that tuber shape alone is not sufficient to define taxonomic units in *D. alata*. Also, accessions of three sub-groups

within group B, namely: B_1, B_2 and B_3 overlapped in the PC1 versus PC2 graph (Figure 2.5). Overlap between species for morphological traits generally make characterization difficult (MacLean *et al*., 1993). The overlapping among genotypes for morphological traits increase the taxonomic complexity, which conflicts classification (MacLean *et al*., 1993). This suggests the use of molecular techniques to augment morphological classification to resolve issues of overlap and confirm morphological associations. However, the overlap among the sub-groups of the B group was within the same species rather than different species. Accessions of group D (NR 07/043 and NR 07/059) had purple wings, purplish-green young leaves, intermediate lobes and cylindrical and branched tubers. Accessions of group E (WR 07/013 and SR 07/085) sprouted in a period of a month, had purple leaf margins, purplish-green petioles with purple at both ends. Whereas WR 07/013 had an oblong shaped tuber with white flesh colour, SR 07/085 had round tubers with light purple flesh colour of central cross section of tubers.

Accessions of group C (NR 07/052, NR 07/060, NR 07/071, SR 07/072, SR 07/074, TDr 95/00005 and TDr 95/18544) belong to *D. rotundata*. Except for TDr 95/00005, rest were characterized by wingless vines. Most exhibited saggitate broad leaf shape, purplish green stems, cylindrically shaped tubers, with white flesh colour of central cross section of tubers. They had delayed sprouting, but produced fairly intermediate leaf density due to their profuse branching habit. The delayed sprouting was probably due in part to inherent genetic variation.

The classification of germplasm based on regional distribution revealed that 42.8% of the *D. rotundata* genotypes in group D came from the south of Sierra Leone, 28.6% from the north and 28.6% (the improved checks) were from IITA. Two genotypes of *D. bulbifera* in group A, NR 07/040 and NR 07/045 were from the north. Genotypes of groups B, C, D, E and F belong to *D. alata* with 27.3% from the north, 18.2% from the south, 22.7% from the east and 31.8% from the west. Since cluster analysis categorises accessions based on genetic similarity, it does not necessarily group accessions with the same geographic origin. Mwiringi *et al*. (2009) also noted the lack of association between morphological traits and their geographic origin. The high inter-mix of genotypes and the presence of overlap in the principal component graph (Figure 2.5) suggest, the possibility of duplicate genotypes. This may have occurred by introduction of similar genotypes into different regions either by exchange of germplasm between farmers, release of new genotypes, etc.

Based on the 28 morphological traits the multivariate techniques separated the genotypes into six major groups with varying tuber shapes (round, irregular, oblong, oval-oblong and cylindrical). Also, high intra- and inter-group diversity was discovered among the species (Figures 2.5 and 2.6). These results agree with Lebot *et al*. (1998) and Hasan *et al*. (2008) who obtained four major tuber groups in *D. alata* morphotypes alone.

2.5 Conclusions

A detailed chracterization of the genotypic diversity within the three *Dioscorea* species (*D. alata, D. bulbifera* and *D. rotundata*) evaluated should contribute to effective conservation and utilization of the yam genetic resources available in Sierra Leone. Knowledge of the genetic resources may also facilitate the exploration of other potential uses of these species. The overlap in sub-groups B_1 and B_2 of genotypes WR 07/024, SR 07/075, 07/073, ER 07/032 and NR 07/042, and between sub-groups B_1 and B_3 of genotypes NR 07/041, WR 07/010, ER 07/038 and NR 07/067, indicated the possibility of duplicate genotypes in the germplasm collection (Figure 2.4). Four traits loaded heavily as specificity in the factor analysis. Of the four, number of days to emergence and tuber shape are crucial in breeding for market-oriented traits. Cluster analysis which classified the genotypes based on genetic distance identified six distinct groups in the germplasm.

A major breeding objective is the development of early establishment in yams through a reduction in the number of days to emergence. Early establishment will enable a greater realization of yield potential of genotypes by maximising the active growing period. The characterization of the accessions will facilitate the identification and genetic relationships of parental genotypes in order to attain the apex breeding objectives of developing high yielding yam genotypes with desirable tuber size, shape, culinary quality, pest and disease resistance, good storability and also characters that confer lower labor requirements.

2.6 References

Amurrio, J.M., A. Ron de, and A.C. Zeven. 1995. Numerical taxonomy of Liberian pea landraces based on quantitative and qualitative characters. *Euphytica* 82: 195-205.

Asiedu, R., N.M. Wanyera, S.Y.C. Ng, and N.Q. Ng. 1997. Yams. In: D. Fuccillo et al. (eds.) *Biodiversity in trust: Conversation and use of plant genetic resources in CGAIR centres*. Cambridge University Press, Cambridge, UK. pp. 57-66.

Asiedu, R., S.Y.C. Ng, K.V. Bai, I.J. Ekanayake, and N.M.W. Wanyera. 1998. Genetic Improvement. In: Orkwor GC, Asiedu R, Ekanayake IJ (eds.) *Food yams: Advances in research*, Ibadan, Nigeria: IITA and NRCRI. pp. 63-104.

Biabani, A.R. and H. Pakniyat. 2008. Evaluation of seed yield-related characters in sesame (*Sesamum indicum* L.) using Factor and Path Analysis. *Pakistan Journal of Biological sciences* 11: 1157-1160.

Dansi A., J. Zoundjihekpon, H.D. Mignouna, and F.M. Quin. 1997. Collecte dôignames cultivées du Complexe *Dioscorea cayenensis - rotundata* au Benin. *Plant Genetic Resources Newsletter* 112: 81-85.

Dansi, A., H.D. Mignouna, J. Zoundjihekpon, A. Sangare, R. Asiedu, and F.M. Quin. 1999. Morphological diversity, cultivar groups and possible descent in the cultivated yams (*Dioscorea cayenensis-Dioscorea rotundata* complex) of Benin Republic. *Genetic Resources and Crop Evolution* 46: 371-388.

Dansi, A., H.D. Mignouna, M. Pillay, and S. Zok. 2001. Ploidy variation in the cultivated yams (*Dioscorea cayenensis-Dioscorea rotundata* complex) from Cameroon as determined by Flow cytometry. *Euphytica* 119: 301-307.

Dumont, R., A. Dansi, P. Vernie, and J. Zoundjihékpon. 2005. Biodiversité et domestication des ignames en Afrique de lôOuest. Pratiques traditionnelles conduisant á *Dioscorea rotundata*. CIRAD, IPGRI, Repères, France.

Hamon, P. 1987. Structure, origine genetique des Ignames en Afrique de IôOuest. These de Doctorat es-Sciences. Université Paris XI, Centre dôOrsay. 223 pp.

Hasan, S.M.Z., A.A. Ngadin, R.M. Shah, and N. Mohamad. 2008. Morphological variability of greater yam (*Dioscorea alata* L.) in Malaysia. *Plant Genetic Resources*: *Characterization and Utilization* 6: 52-61.

Institute of Agricultural Research (IAR), 2004. Roots, tubers and legumes crop production. IAR Annual Report, Njala, Sierra Leone. 60 pp.

International Plant Genetic Resources Institute (IPGRI/ IITA), 1997. *Descriptors for yam* (*Dioscorea* spp.) IPGRI, Rome, Italy: International Plant Genetic Resources Institute (IPGRI)/ International Institute of Tropical Agriculture (IITA), Ibadan Nigeria. 61 pp.

Jeffers, J.N.R. 1967. Two case studies in the application of principal component analysis. *Applied Statistics* 16: 225-236.

Lebot, V., B. Trilles, J. Noyer, and J. Modesto. 1998. Genetic relationship between *Dioscorea alata* L. cultivars. *Journal of Genetic Resources and Crop Evolution* 45: 499-509.

MacLean, D.J., K.S. Bronthwaite, J.M. Manners and J.A.G. Irwing. 1993. How do we identify and classify fungal pathogens in the era of DNA analysis? *Advances in Plant Pathology* 10: 207-244.

Manly, B.F.J. 1994. Multivariate statistical methods. A Primer. Second edition. Chapman & Hall, 2-6 Boundary Row, London SE1 8HN, UK. 214 pp.

Martin, F.W. 1976. Tropical yams and their potential. Series ï Part 3. *Dioscorea alata*. USDA Agriculture Handbook No. 495 ARS, USDA. 40 pp.

Martin F.W. and A.M. Rhodes. 1973. Correlations among greater yam (*Dioscorea alata* L.) cultivars. *Journal of Tropical Agriculture* 50: 183-192.

Martin F.W. and A.M. Rhodes. 1977. Intra-specific classification of *Dioscorea alata*. *Journal of Tropical Agriculture* 54: 1-13.

Mignouna, H.D., A. Dansi, and S. Zoc. 2002. Morphological and isozymic diversity of the cultivated yams (*Dioscorea cayenensis*/ *Dioscorea rotundata* complex) of Cameroon. *Genetic Resources and Crop Evolution* 49: 21-29.

Mignouna, H.D., M.M. Abang, and R. Asiedu. 2003. Harnessing modern biotechnology for tropical root crop improvement: Yam (*Dioscorea* spp.) molecular breeding. *African Journal of Biotechnology* 2: 478-485.

Mwiringi, P.N., E.M. Kahangi, A.B. Ngende, and E.G. Mamati. 2009. Morphological variability within the Kenyan yam (*Dioscorea* spp.). *Journal of Applied Biosciences* 16: 894-901.

Onwueme, I.C. 1978. *The Tropical Tuber Crops*: Yams, Cassava, Sweet potato, Cocoyam. John Wiley and Sons, New York. 234 pp.

Payne, R.W., D.A. Murray, S.A. Harding, D.B. Baird and D.M. Souter. 2009. Genstat for Windows (12th Edition) Introduction VSN International, Hemel Hempstead.

Sastrapradja, S. 1982. *Dioscorea alata*: Its variation and importance in Java, Indonesia. In: Miege J and Lyonga SN (eds.) *Yams Ignames*. Oxford: Oxford University Press. pp. 44-49.

Sokal, R.R. and C.P. Michener. 1958. A statistical method for evaluating systematic relationships. *University of Kansas Science Bulletin* 38: 1409-1438.

Sierra Leone Meteorological Station (Sierra Leone Met. Sta.), 2010.

Velayudhan, K.C., V.K. Muralidharan, V.A. Amalraj, T.A. Thomas, and P. Soudhamini. 1989. Studies on the morphotypic variability, distribution and genetic divergence in an indigenous collection of greater yam (*Dioscorea alata* L.). *Journal of Root Crops* 15: 79-89.

CHAPTER THREE

Diversity in ploidy level and nuclear DNA content (pg) of some yam (*Dioscorea* spp.) genotypes in Sierra Leone as determined by flow cytometry and chromosome counting

Abstract

The techniques of flow cytometry (FCM) and chromosome counting were used to determine the ploidy levels of 52 genotypes from Sierra Leone. For FCM, *Lycopersicum esculentum* (garden tomato) with known ploidy level and genome size was used as a standard. Nuclei were isolated from freshly harvested young leaves, squashed in lysis buffer and stained with propidium iodide. For chromosome counting, root tips were prepared using the acetocarmine staining technique and chromosomes counted using a light microscope at 1000x magnification. The various ploidy levels observed among the genotypes included diploid (2*x*), triploid (3*x*), tetraploid (4*x*), pentaploid (5*x*) and hexaploid (6*x*). The estimated nuclear DNA content ranged from 1.668 $\pm$ 0.017 pg for G_1 nuclei of diploid *D. alata* to 2.118 pg for G_1 nuclei of hexaploid *D. rotundata*. A one-way ANOVA of 4*x* accessions for DNA content, relative florescence intensity and ploidy levels indicated significant variation among species ($p<0.005$) and within genotypes of the various species ($p<0.037$). Genotype ER 07/030 had 20 chromosomes, four genotypes viz ER 07/036, NR 07/060, NR 07/071 and SR 07/072 had 40 chromosomes, and TDr 95/18544 had 60 chromosomes. The results from FCM agree with the hypothesis that chromosome size decreases with higher ploidy level. The results also suggest the possibility of the existence of polymorphism within the chromosomes of *Dioscorea*. Chromosome counts agree with the ploidy results obtained from the FCM, which indicated that FCM was a reliable technique for the rapid determination of ploidy level in yams.

3.1 Introduction

Tropical root and tuber crops are a subsidiary staple to over 20% of the worldôs population occupying an important position after cereals and grain legumes (Orkwor *et al*., 1998). *Dioscorea cayenesis* (yellow yam), *D. alata* (water yam), *D. rotundata* (white yam), and *D. bulbifera* (aerial yam) are among the most important tuber crops consumed in West Africa (Orkwor *et al*., 1998).

Despite its economic importance, yam has not been accorded the keen attention of researchers in many areas in Africa, especially in Sierra Leone. In order to develop new elite genotypes for ecological adaptation and reasonable tolerance to local pests and diseases, plant breeders require wide genetic diversity (Dansi *et al*., 2000b). Since yams are largely polyploid, knowledge of the ploidy state of existing cultivars will be helpful to breed new varieties. Phenotypic variation within ploidy level in yams is higher than between ploidy levels as also noted in other plants (Dessauw, 1988).

Among the many constraints limiting conventional breeding of *Dioscorea* spp., ranging from flowering to seedling development, are: flowering expression, pollen viability or egg receptivity, gametogenesis, pollination, fecundation, embryogenesis and seed set. These constraints encountered in sexual recombination of yams are due to the complex speciation in the crop (Obidiegwu *et al*., 2009). Although scanty information on yam phylogeny exists, many taxonomic ambiguities associated with cytological irregularities still remain unresolved. In addition, various cellular parameters including cell and nuclear volume and chromosome size, and developmental parameters such as minimum generation time or duration of meiosis, among others, are influenced by the C-value of an organism (Swift, 1950). Therefore, genome size normally determines the breeding system (Govindaraju and Cullis, 1991).

Feulgen densitometry, image cytometry, and flow cytometry (FCM) are among the cytometric techniques which have played a significant role in plant taxonomy, biosystematics, and ecology in determining chromosomal and ploidy level data (Suda *et al*., 2006). The merits of FCM lie in its simplicity and speed, the small amount of tissue sample required, the use of various types of plant tissues: leaves, stems, roots, sepals, petals and seeds in FCM assays. This provides the possibility of extensively exploring rare and endangered plant species with no risk of population destruction (Sgorbati *et al*., 2004).

Through FCM, ploidy level at various spatial scales, interactions among cytotypes, and evolutionary processes in diploid-polyploid sympatric populations can also be reliably assessed (Baack, 2004; Husband and Sabara, 2004). Moreover, FCM holds great potential in reshaping former taxonomic concepts and facilitating robust classification based on cytotype characteristics (Bures *et al*., 2003; Rosenbaumova *et al*., 2004). Thus, the application of molecular cytogenetics to the species of *Dioscorea* under study will greatly improve an understanding of chromosome structure and karyotype variation within the species.

Chromosome observation is necessary to clarify the structure, function, organisation and evolution of yam genomes. However, the determination of ploidy level in yam somatic cells by chromosome counting is limited by the polyploid nature of the crop, dot-like nature of chromosomes and small volume of mitotic cells. These characteristics hinder the preparation of a distinct and well-spread chromosomes visible in a single focal plane (Staudt, 1989). A simple, rapid and reliable procedure is needed to determine the chromosome number of meristematic regions of yam root tips (Dansi *et al*., 2001). Furthermore, an understanding of the ploidy and chromosome status in plants generated from anther, ovary and callus cultures, or cell fusion for the identification of haploids, heterokaryons or doubled haploid genotypes is imperative in augmenting plant breeding efforts to develop new genotypes. The aim of this study was to investigate the ploidy levels and nuclear DNA contents of *Dioscorea* species using flow cytometry. A conventional chromosome counting technique was also employed to confirm ploidy results. The hypothesis tested was that local accessions had wide inter- and intra-group diversity.

3.2 Materials and methods

3.2.1 Flow cytometry technique

3.2.1.1 Plant materials and growth conditions

A total of 52 genotypes (50 landraces grown in Sierra Leone and two improved lines from the International Institute of Tropical Agriculture, Ibadan, Nigeria) representing three *Dioscorea* spp. (Table 2.1) were established in 25 cm (diameter) x 20 cm (height) pots in a greenhouse at the controlled environment research unit (CERU) of the University of KwaZulu-Natal, South Africa. Minisetts each weighing 50 g were established in 25 cm (diameter) x 20 cm (height) pots in a green-house at the University of KwaZulu-Natal, Pietermaritzburg, South

Africa in January 2010. The pots were filled with composted seedling mix, and water was supplied by drip irrigation. The pots were arranged in a three replicate, randomized complete block design. The planting distance between pots was 0.25 m. Each pot was fertigated at the rate of 200 kg ha^{-1} of NPK (40:40:60) daily throughout the growing period. Hand weeding was done as necessary.

3.2.1.2 Solution and reagents

The preparation of nuclei from young leaf tissue was based on a modified protocol by Galbraith *et al*. (1983). Two buffer solutions were prepared prior to the extraction of nuclei. The first solution, buffer A, also known as LB01 Lysis buffer or nuclei isolation buffer, consisted of the following reagents: 5 mM TRIS, 2 mM Na_2EDTA, 0.5 mM spermine.4HCl, 80 mM KCl, 20 mM NaCl, 15 mM ϐ-mercaptoethanol, 0.1% (v/v) Triton X-100, and the pH adjusted to 7.5 (Dolezel *et al*., 1989).

The second solution, buffer B, also known as nuclei staining buffer, consisted of 10 mg ml^{-1} propidium iodide (PI, Fluka, Buchs, Switzerland) and 10 mg ml^{-1} RNase A (Fluka, DNase-free) prepared on ice just prior to use. Since staining of nuclei within a crude homogenate is influenced by the composition of the nuclei isolation buffer and phenolic compounds present in the cytosol, the use of ϐ-mercaptoethanol in the isolation buffer was to alleviate the influence of phenolic compounds on staining.

3.2.1.3 Sample preparation for ploidy analysis

Healthy young leaves were collected from individual plants in the screen house, bagged, transported on ice and kept in the refrigerator at a temperature of 4°C for a period of four days until analysis. Young leaves were used in order to avoid the high concentration of starch, polysaccharides, calcium oxalate, and other metabolites, which decrease the purity of intact nuclei found in old tissues. Nuclei from each accession were carefully isolated and suspended by slicing approximately 50 mg of sample material into thin strips less than 0.5 mm wide with a sharp double-edged razor blade in a glass petri dish containing 1 ml LB01 lysis buffer. This was done to eliminate the occurrence of contaminants that accelerate the degradation of nuclei, increase the viscosity of the sample and/ or block the fluidics system of the flow cytometer.

The extract of nuclei was filtered into a 15 ml Falcon tube using a 50 εm pore size nylon mesh. The nuclear DNA was stained with 10 mg ml^{-1} of propidium iodide (PI); and also 10 mg ml^{-1} of RNase was added to avoid staining of double-stranded RNA by PI. About 600 µl of buffer B was added to each sample. After 20 min incubation period on ice, samples were ready for flow cytometric analyses. *Solanum lycopersicum* L. (garden tomato) (1C Genome size = 958 Mbp; 2C = 1.96 pg DNA; Dolezel *et al*., 1992) was used as a reference standard because of its close but non-overlapping genome size.

3.2.1.4 Flow cytometric analysis

Beckman Coulter EPICS-XL flow cytometer (Beckman Coulter®, Hialeah, FL, USA) equipped with an air cooled argon-iron laser regulated at 15 mW and operating at 488 nm was used for the analysis of the samples. Propidium iodide (PI) florescence was collected through a 645 nm dichroic long pass filter and a 260 nm band pass filter. Before sample analysis, the instrument was checked for linearity with Flow Check fluorospheres daily (Beckman Coulter®). Counts were obtained using the SYSTEM II software version 3.0 (Coulter Electronics). The amplifier system was adjusted so that the G_0/G_1 peak of nuclei isolated from diploid individuals appeared at channel 200 in a scale with 1,024 channels and the flow was given a stop time of 300 s. These settings were kept constant throughout the entire experimentation.

Initially, ploidy level was determined by comparing the relative fluorescence intensity (RFI) of sample nuclei with the RFI of the reference standard. Afterwards, the reference standard was included during sample preparation. In the later experiment, when deviation greater than 10% was observed, the ploidy level was estimated by preparing a new sample with both test material and the reference standard. This was done by chopping 20 mg of reference plant tissue together with 50 mg of sample tissue. The samples were analyzed in random sequence on the flow cytometer in order to statistically account for laser drift and other sources of machine error. The total amount of DNA content present was calculated as:

Sample 2C DNA content = {(sample G_1 peak mean)/ (*S. lycopersicum* G_1 peak mean)} x *S. lycopersicum* 2C DNA content (pg DNA) (Obidiegwu *et al*., 2009).

3.2.1.5 Statistical analysis

The statistical relationship between DNA content and ploidy level was determined through regression analysis. The amount of the total variation in ploidy level explained by DNA content was evaluated through the coefficient of determination (R^2) (Steel and Torrie, 1980).

Variation in DNA content among and within tetraploid species of *Dioscorea* was evaluated using a one-way analysis of variance (ANOVA). The analysis was performed in Genstat 12.1 (Payne *et al*., 2009); error mean squares for successively higher levels in the one-way ANOVA were determined by evaluation of significance at lower levels in the analysis.

3.2.2 Chromosome counting technique

3.2.2.1 Plant materials and growth conditions

An experiment to investigate and confirm the ploidy status of six yam genotypes using chromosome counting technique was conducted at the University of KwaZulu-Natal, Pietermaritzburg, South Africa during the early summer months of September to November, 2010. Yam minisetts of six genotypes, ER 07/030, ER 07/036, NR 07/060, NR 07/071, SR 07/072 and TDr 95/ 18544 each weighing 20 g were grown in 250 mL pots containing a vermiculite medium. Plants were grown for about 4 weeks at 25°C day/night and 70% relative humidity. The moisture level of the growing medium was well managed to prevent limited oxygen availability and reduced root growth by over-wet or over-dry medium, respectively.

3.2.2.2 Pre-treatment, fixation, staining method and chromosome counts

A protocol slightly modified from Fukui and Nakayama (1996) was used for sample preparation and analyis. Root tips each 6 mm long were collected using forceps at 06h00 and treated in small vial of Carnoyôs (farmers) fixative (three parts 95% alcohol: 1 part glacial acetic acid) for 48 h at room temperature. Samples were rinsed thoroughly in 70% alcohol to remove acetic acid, which could hinder the staining of the chromosomes by acetocarmine. Samples were then stored in 70% ethanol at 4°C in a refrigerator until examined. For mitotic analysis, root tips were hydrolysed in 1N HCl for 90 mins at room temperature and then

washed in distilled water before staining. The outer (1 to 2 mm deep) layer of cells of root tips were excised using a razor blade and stained in a drop of acetocarmine for 15 mins on microscope (glass) slide. The cells were covered with a cover slip, carefully mopping excess stain using paper towel. The slide was slighly heated over a spirit bunsen burner without boiling to prevent damage to the cells.

The chromosomes were counted in three to five cells per slide in five to 10 root tips per genotype using a light microscope (AX70; Olympus Optical Co. Ltd., Japan) at the 100x objective magnification (1000x total magnification). The well spread chromosomes at the metaphase stage were digitally photographed and stored using a camera (CC12; Olympus Optical Co. Ltd., Japan) connected to a personal computer equipped with image filing software (Soft Imaging System (SIS) analysis® 3.0 Co. Ltd., Japan).

3.3 Results

3.3.1 Flow cytometry technique

There was a highly significant ($F_{1,50}$ = 77.6, $p<0.001$) linear relationship between nuclear DNA content and ploidy level among yam genotypes (Table 3.1; Appendix 9). Nuclear DNA content accounted for 60% of the total variation observed in ploidy level. The regression equation for DNA content indicated that for every picogram increase in nuclear DNA content of yams, ploidy level increases by 5.52 unit. However, nuclear DNA content was strongly negatively correlated (r = -0.999) with ploidy level (Appendix 9). This implied that as ploidy level increases, the DNA content per genome decreases.

Overall, five ploidy levels: diploid (2*x*), triploid (3*x*), tetraploid (4*x*), pentaploid (5*x*) and hexaploid (6*x*) were detected among the 52 genotypes studied (Table 3.1). Three of the *D. alata* genotypes were diploid (Figure 3.1a) and one was triploid (Figure 3.1b). A total of 45 genotypes including 38 *D. alata,* six *D. rotundata* and one *D. bulbifera* were tetraploid (Figure 3.1c); one genotype, NR 07/045 was pentaploid (Figure 3.1d), and two genotypes, SR 07/084 and TDr 95/18544 were hexaploid (Figure 3.1e).

The nuclear DNA content varied from 1.588 to 1.718 pg (1.668 $\pm$ 0.017 pg) for the diploids, 1.750 pg for the triploid, from 1.772 to 1.937 pg for the tetraploids, 1.998 pg for the pentaploid, and from 2.102 to 2.118 pg for the hexaploids (Table 3.1).

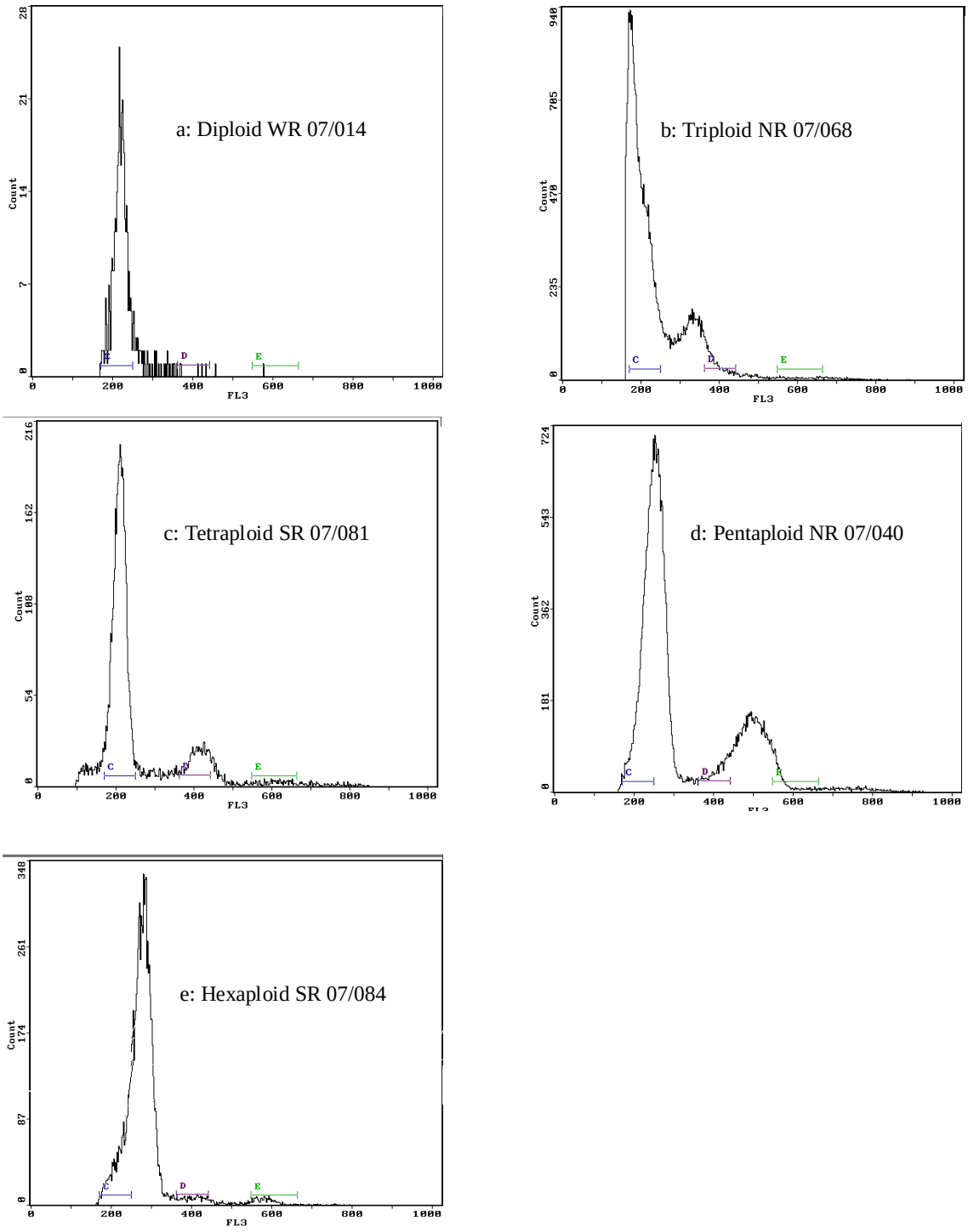


Figure 3.1. Histograms of fluorescent intensity of nuclei for G1 peaks of: (a) diploid *D. alata* (2n = 2*x* = 20, 2C = 1.689 pg); (b) triploid *D. alata* (2n = 3*x* = 30, 2C = 1.750 pg); (c) tetraploid *D. alata* (2n = 4*x* = 40, 2C = 1.781 pg); (d) pentaploid *D. bulbifera* (2n = 5*x* = 50, 2C = 2.017 pg); and (e) hexaploid *D. rotundata* (2n = 6*x* = 60, 2C = 2.102 pg)

Table 3.1. Flow cytometric measurements of relative fluorescence intensity (RFI (log)), 2C nuclear DNA content and ploidy level of 52 yams accessions from Sierra Leone

Species	Accession Number[+]	RFI	2C Nucler DNA content (pg)	Mean ± SE[*]	Ploidy Level
D. alata	WR 07/014	5.197	1.689	1.668 ± 0.017	2*x*
	WR 07/016	5.193	1.681	CV = 1.78%	2*x*
	ER 07/030	5.165	1.634		2*x*
	NR 07/068	5.233	1.750		3*x*
	WR 07/001	5.291	1.854		4*x*
	WR 07/004	5.301	1.873		4*x*
	WR 07/007	5.301	1.874		4*x*
	WR 07/008	5.315	1.899		4*x*
	WR 07/010	5.272	1.820		4*x*
	WR 07/013	5.266	1.809		4*x*
	WR 07/015	5.330	1.928		4*x*
	WR 07/020	5.284	1.842		4*x*
	WR 07/022	5.248	1.776		4*x*
	WR 07/024	5.250	1.779		4*x*
	WR 07/025	5.326	1.921		4*x*
	WR 07/028	5.268	1.812		4*x*
	ER 07/029	5.278	1.830		4*x*
	ER 07/031	5.309	1.888		4*x*
	ER 07/032	5.310	1.890	1.846 ± 0.007	4*x*
	ER 07/033	5.288	1.848	CV = 2.47%	4*x*
	ER 07/034	5.282	1.838		4*x*
	ER 07/036	5.283	1.840		4*x*
	ER 07/037	5.329	1.927		4*x*
	ER 07/038	5.261	1.800		4*x*
	ER 07/039	5.308	1.886		4*x*
	NR 07/041	5.272	1.820		4*x*
	NR 07/042	5.325	1.918		4*x*
	NR 07/043	5.271	1.818		4*x*
	NR 07/047	5.305	1.880		4*x*
	NR 07/054	5.245	1.772		4*x*
	NR 07/057	5.293	1.859		4*x*
	NR 07/059	5.279	1.832		4*x*
	NR 07/067	5.276	1.826		4*x*
	NR 07/069	5.276	1.827		4*x*
	SR 07/073	5.314	1.897		4*x*
	SR 07/075	5.260	1.798		4*x*
	SR 07/076	5.314	1.897		4*x*
	SR 07/079	5.302	1.875		4*x*
	SR 07/080	5.254	1.788		4*x*
	SR 07/081	5.251	1.781		4*x*
	SR 07/082	5.256	1.790		4*x*
	SR 07/085	5.278	1.831		4*x*
	SR 07/084	5.416	2.102	2.102	6*x*
D. bulbifera	NR 07/045	5.326	1.905	1.905	4*x*
	NR 07/040	5.375	2.017	2.017	5*x*
D. rotundata	NR 07/052	5.278	1.831		4*x*
	NR 07/060	5.335	1.938		4*x*
	NR 07/071	5.305	1.880	1.897 ± 0.017	4*x*
	SR 07/072	5.327	1.923	CV = 2.26%	4*x*
	SR 07/074	5.300	1.872		4*x*
	TDr 95/00005	5.335	1.937		4*x*
	TDr 95/18544	5.423	2.118	2.118	6*x*

[+]WR = western region of Sierra Leone; ER = eastern region; NR = northern region, SR = southern region *SE = standard error; and CV = coefficient of variation within each ploidy level

Exploration of the nuclear DNA content of the 4*x* ploidy level of *D. alata*, *D. bulbifera* and *D. rotundata* revealed significant diversity both among species ($p<0.005$) and within accessions ($p<0.037$) (Tables 3.2 and 3.3; Appendices 10 and 11). Although the variations within ploidy levels were small compared to differences between ploidy levels, this implied the probable existence of polymorphism within yam chromosomes.

Table 3.2. ANOVA in DNA content among three species of yam (*D. alata*, *D. bulbifera* and *D. rotundata*) with 4*x* genotypes

Source of variation	Df	SS	MS	F ratio	F prob
Replication	2	0.502386	0.251193	51.60	
Species	2	0.054730	0.027365	5.62	0.005
Residual	130	0.632886	0.004868		
Total	134	1.190001			

Table 3.3. ANOVA in DNA content of 45, 4*x* genotypes

Source of variation	Df	SS	MS	F ratio	F prob
Replication	2	0.502386	0.251193	57.41	
Genotype	44	0.302567	0.006877	1.57	0.037
Residual	88	0.385048	0.004376		
Total	134	1.190001			

3.3.2 Chromosome counting

The six genotypes evaluated had chromosome numbers ranging from 20 t0 60. One genotype, ER 07/030, had 20 chromosomes; four, ER 07/036, NR 07/060, NR 07/071 and SR 07/072, had 40 chromosomes; and TDr 95/18544 had 60 chromosomes. The genomic number of chromosomes of the six genotypes are a multiple of the basic chromosome number, 10 (Figures 3.2 and 3.3).

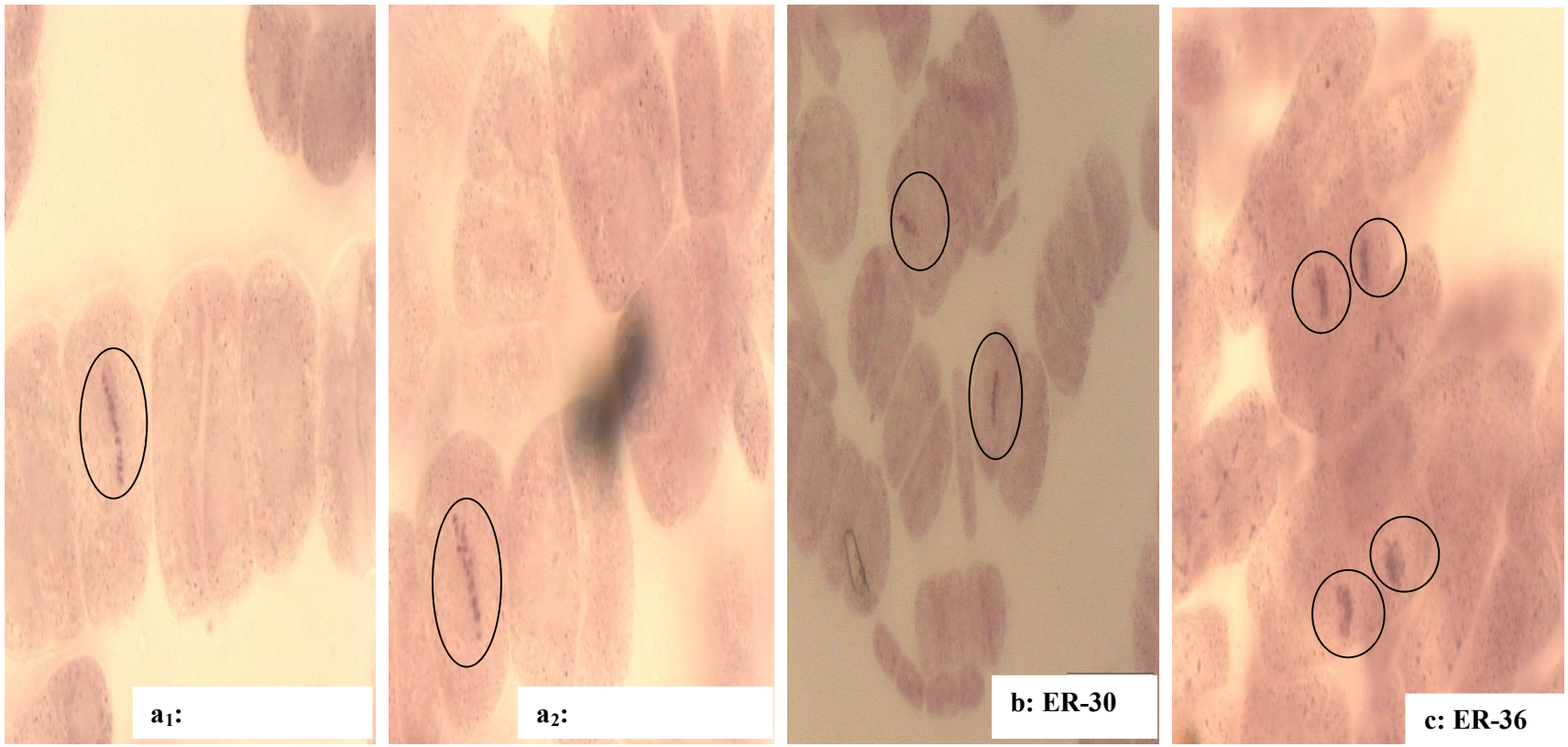


Figure 3.2. Mitotic chromosomes in root tip cells of the basic number, diploid and tetraploid yams stained with acetocarmine. Metaphase chromosomes in root tip cells of: (a_1 and a_2) examples of countable resolution with the chromosome number x = 10, (b) ER 07/030 ($2n = 2x = 20$), (c) ER 07/036 ($2n = 4x = 40$). Circled regions indicate metaphase chromosome

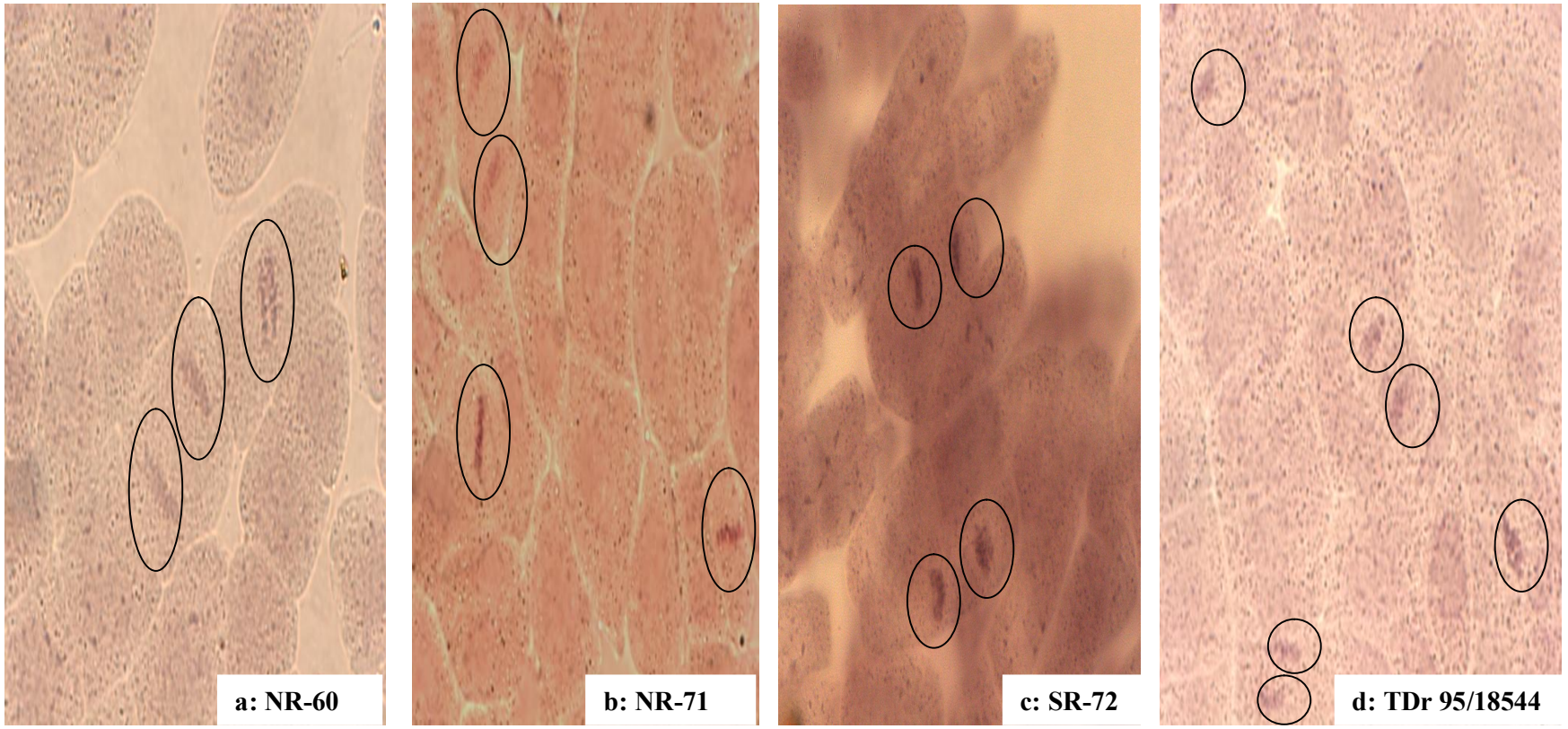


Figure 3.3. Mitotic chromosomes in root tip cells of tetraploid and hexaploid yam genotypes stained with acetocarmine. Metaphase chromosomes in root tip cells of (a) NR 07/060 (2n = 4*x* = 40), (b) NR 07/071 (2n = 4*x* = 40), (c) SR 07/072 (2n = 4*x* = 40), (d) TDr 95/18544 (2n=6*x*=60). Circled regions indicate metaphase chromosomes

3.4 Discussion

3.4.1 Flow cytometry technique

The results generally indicated that DNA flow cytometry is a useful technique to determine ploidy level of healthy young leaves of yams. The technique is also very useful for plants derived from heteroploid crosses in which parent (s) produce 2n gametes. This saves time and resources through the maintenance of plants generated from sexual polyploidisation.

Of the 52 genotypes, three (5.8%) were diploid, one (1.9%) was triploid, 45 (86.5%) were tetraploid, one (1.9%) was pentaploid and two (3.9%) were hexaploid. These findings are in agreement with those obtained by Essad (1984), Zoundjihekpon *et al*. (1990), Hamon *et al*. (1992) and Dansi *et al*. (2000a) who noted that tetraploids are the most frequent group in the *Dioscorea* species. The high number of tetraploids and the low number of hexaploids suggests that they may have evolved either by somatic doubling or sexual polyploidization. However, the presence of triploid (3*x*) and pentaploid (5*x*) individuals suggests that polyploidization by the fusion of reduced (n) and unreduced (2n) gametes may have occurred. The occurrence of triploidy and pentaploidy in *Dioscorea* was earlier reported by Sharma and De (1956) and Martin (1976).

The 2C nuclear DNA content of *D. alata* ranged from 1.668 $\pm$ 0.017 pg (diploid), 1.750 pg (one triploid), 1.846 $\pm$ 0.007 pg (tetraploid), and 2.102 pg (one hexaploid). Since the DNA content did not increase in multiples of ploidy level an obvious interpretation was an apparent decrease in chromosome size with increasing ploidy level. These data support Sharma and Senôs (2002) hypothesis that the size of chromosomes tends to diminish with polyploidization. It is possible that such a ñcompensationò mechanism serves as a defensive strategy against the increased probability of mutations that often accompany polyploidy (Sharma and Sen, 2002). Moreover, Chenuil *et al*. (1997) demonstrated that shortening of microsatellites and reducing their number could be one of the molecular mechanisms employed to eliminate excessive DNA in organisms of higher ploidy levels.

In the *D. bulbifera*, tetraploid and pentaploid accessions, 2C nuclear DNA contents were estimated at 1.905 and 2.017 pg, respectively. Six of the *D. rotundata* accessions were tetraploid with a mean of 1.897 pg, while one was hexaploid with 2C = 2.118 pg. The results are in concurrence with those obtained by Gamiette *et al*. (1999), Dansi *et al*. (2000a, 2001)

and Obidiegwu *et al*. (2009) who noted tetraploids outnumbering other ploidy levels in *D. rotundata*. A similar trend of decrease in chromosome size with higher ploidy in *D. alata* was also observed in both *D. bulbifera* and *D. rotundata*. However, the present analysis did not show the occurrence of either octoploidy or mixoploidy in genotypes of *D. alata* and *D. rotundata* as reported by Hamon *et al*. (1992), Gamiette *et al*. (1999), and Dansi *et al*. (2000a, 2001). Polyploidy has been noted as one of the main sources of an increase in DNA content in plants. Leitch and Bennett (2004) noted that an increase in ploidy level caused an average increase in the total DNA. However, the DNA content of each genome (i.e. the nuclear DNA content divided by the ploidy level) decreased in polyploid nucleus (Leitch and Bennett, 2004). The variations in nuclear DNA content within diploid plants have been associated with differences in transposon copy number (Bennetzen, 2002) and intron size (Petrov, 2001). Findings of the present study are consistent with those of Lietch and Bennett (2004) in which an increase in ploidy number led to a decrease in DNA content per genome in polyploids. Transposoon copy number and intron size influence on the nuclear DNA content of diploids was not investigated in this study.

The variation in nuclear DNA content of 4*x* genotypes among the *Dioscorea* species (*D. alata*, *D. bulbifera* and *D. rotundata*) reflects interspecific variation in either genomic DNA content or composition. The higher variations observed between species compared to within species variation were possibly due to increasing ploidy level. This suggests that these species may have evolved from different ancestors. The variation noted among genotypes within species also supports the evolution of chromosome composition hypothesis in plants (Costich *et al*., 1993). The hypothesis assumes that adaptive differentiation of a group of related species is followed by a gradual decrease in genome size as species become more specialized (Price, 1976). Polyploidy in yams may have arisen from multiplication of the basic chromosome number with the chromosome size and DNA content per chromosome not increasing in direct proportion with ploidy level. The phenotypic variations observed among the polyploid (triploid, tetraploid, pentaploid and hexaploid) genotypes was likely due to increase in the number of loci present. This is in agreement with earlier findings that polyploidy increases the number of loci, potential number of alleles each locus contains and the dosage effect of genes is altered (Udall and Wendel, 2006). The variation observed among the 2*x D. alata* genotypes was likely due to allelic differences at homologous loci.

Determination of ploidy status of genotypes, especially new introductions, before their utilization in breeding programme is crucial. Such data enables matching of ploidy levels as well as in the enhancement of ploidy manipulations in inter- and intra-specific crosses. Through DNA flow cytometry, chromosome differentiation in yams can be investigated. Knowledge of the magnitude of such differentiation facilitates an understanding of the lack of chromosome pairing in hybrids. The 52 genotypes comprising five ploidy levels (2*x*, 3*x*, 4*x*, 5*x* and 6*x*), will be evaluated, multiplied and used for either breeding purposes or other genetic investigations. Through these initiatives superior genotypes with desirable traits could be developed and released as new cultivars.

3.4.2 Conventional chromosome counting technique

Based on conventional chromosome the chromosome was determined as $x = 10$ for the various genotypes studied (Figure 3.2). A basic chromosome number, $x = 10$ was also reported by Zoundjihekpon *et al*. (1990) and Dansi *et al*. (2000a). Generally, the dot-like and clumping nature of the chromosomes made counting difficult. In yams, the occurrence of one or two extra chromosomes in cells of individual genotypes is not rare (Zoundjihekpon *et al*., 1990; Gamiette *et al*., 1999; Dansi *et al*., 2000b). However, the presence of the extra chromosomes is often attributed to the B-chromosomes or satellites which are sometimes as large as the chromosomes themselves as opposed to aneuploidy (Essad, 1984). The B-chromosomes, which may be involved in directing non-disjunction of chromatids during cell division are dispensable and extra to the basic A-chromosome set (Hasterock *et al*., 2002). Langdon *et al*. (2000) also noted that B-chromosome-specific region is possibly occupied by a block of hetero-chromatin at the distal end of the long arm in *Secale cereale* (rye). The results were also in agreement with the ploidy results obtained from FCM, which indicated that FCM was a reliable technique for rapid determination of ploidy level in yams.

This is the first report of cytogenetic work of yam genotypes from Sierra Leone. It is believed that the information generated from this study would provide guidance in a yam improvement programme both in terms of selection of initial breeding material and choice of breeding methods.

3.5 Conclusions

An adequate knowledge of the chromosome/ploidy constitution of yam genotypes is a prerequisite for their effective and efficient utilization in a breeding programme. Relative to chromosome counting, FCM is the technique of choice for the determination of ploidy levels in large yam populations as it is an easier and quicker technique provided the necessary equipment is available.

Of the 52 genotypes, three (5.8%) were diploids, one (1.9%) was triploid, 45 (86.5%) were tetraploids, one (1.9%) was pentaploid and two (3.9%) were hexaploids. The accessions of *D. alata* had diploid, triploid, tetraploid and hexaploid ploidy levels, which were more diverse compared to the accessions of *D. bulbifera* which were tetraploid and pentaploid, and *D. rotundata* which were tetraploid and hexaploid.

The nuclear DNA content varied from 1.588 to 1.718 pg for the diploids, 1.750 pg for the triploid, 1.772 to 1.937 pg for the tetraploids, 1.998 pg for the pentaploid, and from 2.102 to 2.118 pg for the hexaploids. Both the genome size variation within individuals of the tetraploid population and those involving all ploidy populations among the three species was significant. The DNA amount per genome decreased with increasing ploidy level.

The 52 genotypes from which five ploidy levels (2*x*, 3*x*, 4*x*, 5*x* and 6*x*) were obtained, will be evaluated, multiplied and used for either breeding or other genetic investigations. The yam genotypes evaluated in this study will further be used in breeding and genetic investigations to develop new genotypes with superior combinations of traits.

3.6 References

Baack, E.J. 2004. Cytotype segregation on regional and microgeographic scales in snow buttercups (*Ranunculus adoneus*: Ranunculaceae). *America Journal of Botany* 91: 1783-1788.

Bennetzen, J.L. 2002. Mechanisms and rates of genome expansion and contraction in flowering plants. *Genetica* 115:29-36.

Bures, P., L. Ticky, W.F. Wang, and J. Bartos. 2003. Occurrence of *Polypodium* x *mantoniae* and new localities for *P. interjectum* in the Czech Republic confirmed using flow cytometry. *Preslia* 75: 293-310.

Chenuil, A., E. Desmarias, L. Pouyaud, and P. Berrebi. 1997. Does polyploidy lead to fewer and shorter microsatellites in Barbus (Teleostei: Cyprinidae)? *Molecular Ecology* 6: 169-178.

Costich, D.E., R. Ortiz, T.R. Meagher, L.P. Bruederle, and N. Vorsa. 1993. Determination of ploidy level and nuclear DNA content in blueberry by flow cytometry. *Theoretical Applied Genetics* 86: 1001-1006.

Dansi, A., M. Pillay, H.D. Mignouna, O. Dainou, F. Mondeil, and K. Moutairou. 2000a. Ploidy level of the cultivated yams (*Dioscorea cayenensis/ D. rotundata* complex) from Benin Republic as determined by chromosome counting and flow cytometry. *Journal of African Crop Science* 8: 355-364.

Dansi, A., H.D. Mignouna, J. Zoundjihekpon, A. Sangare, N. Ahoussou, and R. Asiedu. 2000b. Identification of some Benin Republicôs Guinea yam (*Dioscorea cayenensis-Dioscorea rotundata* complex) using Random Amplified Polymorphic DNA. *Genetic Resources and Crop Evolution* 47: 619-625.

Dansi, A., H.D. Mignouna, M. Pillay, and S. Zok. 2001. Ploidy variation in the cultivated yams (*Dioscorea cayenensis-Dioscorea rotundata* complex) from Cameroon as determined by Flow cytometry. *Euphytica* 119: 301-307.

Dessauw, D. 1988. Etude des facteurs de la sterilite du bananier (*Musa* spp.) et des relations cytotaxonomiques entre *M. acuminata* et *M. balbisiana* Colla. *Fruits* 43: 539-700.

Dolezel, J., S. Sgorbati, and S. Lucretti. 1992. Comparison of three DNA fluorochromes for flow cytometric estimation of nuclear DNA content in plants. *Physiology of Plant* 85: 625-631.

Essad, S. 1984. Variation geographique des nombres chromosomiques de base et polyploidie dans le genre *Dioscorea*, a propos du denombrement des especes *transversa* Brown, *pilosiuscula* bert. et *trifida* L. *Agronomie* 4: 611-617.

Fukui, K. and S. Nakayama, (eds.) 1996. *Plant Chromosomes: Laboratory Methods*. 3rd Edition, CRC Press, Boka Raton, Florida, USA. 274 pp.

Galbraith, D.W., K.R. Harkins, J.M. Maddox, N.M. Ayres, D.P. Sharma, and E. Firoozabady. 1983. Rapid flow cytometric analysis of the cell cycle in intact plant tissues. *Science* 220: 1049-1051.

Gamiette, F., F. Bakry, and G. Ano. 1999. Ploidy determination of some yam species (*Dioscorea* spp.) by flow cytometry and conventional chromosomes counting. *Genetic Resources and Crop Evolution* 46: 19-27.

Govindaraju, D.R. and C.A. Cullis. 1991. Modulation of genome size in plants: the influence of breeding systems and neighbourhood size. *Evolution Trends Plants* 5: 43-51.

Hamon, P., J. Brizard, J. Zoundjihékpon, C. Duperray, and A. Borge. 1992. Etude des index dôAND de huit espèces dôignames (*Dioscorea* spp.) par cytometrie en flux. *Canadan Journal of Botany* 70: 996-1000.

Hasterock, R., G. Jenkins, T. Langdon, and R.N. Jones. 2002. The nature and density of translocated B-chromosome-specific satellite DNA of rye. *Chromosome Research* 10: 83-86.

Husband, B.C. and H.A. Sabara. 2004. Reproductive isolation between autotetraploids and their diploid progenitors in fireweed. *Chamerion angustifolium* (Onagraceae). *New Phytology* 161: 701-711.

Langdon, T., C. Seago, R.N. Jones, H. Ougham, H. Thomas, J.W. Forster, and G. Jenkins. 2000. De novo evolution of satellite DNA on rye B-chromosome. *Genetics* 154: 869-884.

Leitch, I. and M. Bennett. 2004. Genome downsizing in polyploid plants. *Biological Journal of the Linnean Society.* 82: 651-663.

Martin, F.W. 1976. Tropical yams and their potential. Series ï Part 3. *Dioscorea alata*. USDA Agriculture Handbook No. 495 ARS, USDA. 40 pp.

Obidiegwu, J.E., E. Rodriguez, E.E. Ene-Obong, J. Loureiro, C.O. Muoneke, C. Santos, M. Kolesnikova-Allen, and R. Asiedu. 2009. Estimation of the nuclear DNA content in some representative of genus *Dioscorea*. *Scientific Research and Essay* 4: 448-452.

Orkwor, G.C., R. Asiedu, and I.J. Ekanayake. (eds.) 1998. *Food Yams. Advances in Research*, IITA and NRCRI, Nigeria. 249 pp.

Payne, R.W., D.A. Murray, S.A. Harding, D.B. Baird, and D.M. Souter. 2009. Genstat for Windows (12th Edition) Introduction VSN INterbational, Hemel Hempstead.

Petrov, D. 2001. Evolution of genome size: new approaches to an old problem. *Trends in Genetics* 17: 23-28.

Price, H.J. 1976. Evolution of DNA content in flowering plants. *Botany Revolution* 42: 27-52.

Rosenbaumova, R., I. Plackova, and J. Suda. 2004. Variation in *Lamium* subg *Galeobdolon* (Lamiaceae) ï insights from ploidy levels, morphology and isozymes. *Plant Systematic Evolution* 244: 219-244.

Sgorbati, S., M. Labra, E. Grugni, G. Barcaccia, G. Galasso, L. Boni, M. Mucciarelli, S. Citterio, A.B. Iramategui, L.Y. Gonzales, and S. Scannerini. 2004. A survey of genetic diversity and reproductive biology of *Puya raimondii* (Bromeliaceae), the endangered queen of Aades. *Plant Biology* 6: 222-230.

Sharma, A.K. and N. De. 1956. Polyploidy in *Dioscorea. Genetica* 2: 112-120.

Sharma, A.K. and S. Sen. 2002. Chromosome Botany. ï Science Publishers, Inc., Enfield, USA. 155 pp.

Staudt, G. 1989. The species of Frageria, their taxonomy and geographical distribution. *Acta Horticulture* 256: 23-33.

Steel, R.G.D. and J.H. Torrie. 1980. Principles and procedures of statistics ï a biometrical approach. 2nd edn. McGraw Hill Book Company, New York. pp. 477-579.

Suda, J., A. Krahuleova, P. Travnieek, and F. Krahulee. 2006. Ploidy level versus DNA ploidy level: an appeal for consistent terminology. *Taxonomy* 55: 447-450.

Swift, H. 1950. The constancy of deoxyribose nucleic acid in plant nuclei. *Proceedings of the National Academy of Sciences of the USA* 36: 643-654.

Udall, J.A. and J.F. Wendel. 2006. Polyploidy and crop improvement. *The Plant Genome (A Supplementary to Crop Science)* 46: 3-14.

Zoundjihekpon, J., S. Essad, and B. Toure. 1990. Denombrement chromosomique dans dix groupes varietaux du complexe *Dioscorea cayenensis-rotundata. Cytologia* 55: 115-120.

CHAPTER FOUR

Determination of associations between three morphological and two cytological traits of yams (*Dioscorea* spp.) using canonical correlation analysis

Abstract

Agro-morphological traits of plants may directly or indirectly depend on cytological traits. Thus, the determination of associations between morphological traits (absence or presence of wings, number of stems per plant and wing colour of stem) and cytological traits (DNA content and ploidy level) of yams were investigated using canonical correlation analysis. This multivariate technique is used in wide fields of study to quantify the mathematical relationships between multiple sets of independent and dependent traits or properties. Canonical weights and loadings indicated that DNA content (pg) had the highest contribution to the variation of the morphological traits (presence of wings, number of stems per plant and wing colour) compared with ploidy level. It was found that cytological traits accounted for 0.09 to 0.17% of the variation in the selected morphological traits. The first and second canonical correlations exhibited 60.91 and 39.09% overlapping variance of the canonical variate sets respectively. The first and second canonical variates extracted 0.57 and 4.43% of the total variance in the cytological trait set. The study demonstrated the successful determination of complex inter-relationships between morphological and cytological traits.

4.1 Introduction

Canonical correlation analysis (CCA) is one of several multivariate analysis techniques used to determine the overall correlation between two sets of traits (X and Y). Canonical correlation is a generalization of multiple regression analysis with more than one trait in the independent and dependent trait sets. The basic principle of the technique is to determine how much variance in one set of traits is accounted for by the other set along one or more axes (Tabachnick and Fidell, 2001). In contrast to many other techniques, any of the two sets of traits is a potential candidate to be used as dependent or independent traits. Canonical correlation makes possible several combinations of two trait sets. The number of

combinations depends on the number of traits in the smaller trait set (Tabachnick and Fidell, 2001; Keskin and Yasar, 2007).

Canonical correlation analysis has been widely applied in various fields such as the plant sciences, biology, chemistry, social and management sciences. However, there is scant information available on the interrelationship between morphological and cytological traits of yams. The main aim of this study was to determine the level of association between morphological and cytological traits of yams using canonical correlation analysis. The hypothesis being tested was that correlation exists between the agro-morphological and cytological traits used in the two methods of classification.

4.2 Materials and methods

A total of five traits including three morphological (absence or presence of wings, number of stems per plant and wing colour of stem) and two cytological (DNA content and ploidy level) traits were used. The morphological traits were considered as the dependent Y-trait set, whereas the cytological traits were taken as the independent X-trait set. To obtain the maximum correlations between two sets of traits, two linear combinations were designed as shown below:

$W_i = a_{i1}X_1 + a_{i2}X_2 + \text{--------} + a_{ip}X_p$ é é é é é é é é é é é é é é é é é ..é é . (Eqn 9)

$V_i = b_{i1}Y_1 + b_{i2}Y_2 + \text{--------} + b_{iq}Y_q$ é é é é é é é é é .é é é é é é é é é é ..(Eqn 10)

The symbols W and V represents canonical variates; a and b are canonical coefficients of the X and Y trait sets; and p (two traits) and q (three traits) are the number of traits in the X and Y trait sets, respectively. The estimation of the vector coefficients, a and b, was done according to Tabachnick and Fidell (2001).

To generate the canonical correlation for both sets of traits, the following formulae were used:

var (W) = aqCov (X) a.é é é é é é é é é ...é é é é é é é é é é é .é é é é (Eqn 11)

var (V) = bqCov (Y) bé é é é é é é é é é ...é é é é é é é é é é ..é é é é (Eqn 12)

$$Cwv = \frac{b'Cov(Y)\,b}{\sqrt{(a'Cov(X)\,a)(b'Cov(Y)b)}}$$ é é é é é é ...é é é é é é é é é é é é é é é (Eqn 13)

where var (W) represents variance of the canonical variate W; var (V) is the variance of the covariate V; C_{wv} is the canonical correlation between the X and Y trait sets; Cov (Y) and Cov (X) are the covariances of the traits in the X and Y trait sets, respectively (Keskin and Yasar, 2007).

The relationship of a set of canonical variate is maximized when the correlation (r-value) of the p and q is small. The first set of canonical variate (W_1 and V_1) gives the highest correlation and is considered the most important. The correlation between W_2 and V_2 is only maximized where the traits measured are uncorrelated to W_1 and V_1. Similarly, the correlation between W_3 and V_3 is maximized if traits are not correlated with W_1, V_1, W_2 and V_2 (Manly, 1994).

The canonical correlation analysis procedure (CANCORRELATION procedure) in Genstat Version 12.1 was used to generate the relationships between sets of traits (Payne *et al*., 2009). The squared canonical correlation (also known as canonical roots or eigen-values) represents the amount of variance in one canonical variate accounted for by the other canonical variate (Hair *et al*., 1998). The standardized coefficients are similar to the standardized regression coefficients in multiple regression, which gives an indication of the relative importance of the independent traits in determining the value of dependent traits.

In order to determine the amount of variance in one set of traits that is accounted for by another set of traits, Sharma (1996) suggested the estimation of the redundancy measure (RM) for each canonical correlation. The equation for the RM is shown below:

$RM_{vi/wi} = AV(Y/V_i) \times C_i^2$ é ..é (Eqn 14)

$AV(Y/V_i) = [\times {}^{q}LY_{ij}^2/q]$ é ..(Eqn 15)

where $AV(Y/V_i)$ = the averaged variance in Y traits that is accounted for by the canonical variate V_i.

LY_{ij}^2 = the loading of the j^{th} Y trait on the i^{th} canonical variate V_i.

q = the number of traits in canonical variates.

C_i^2 = the shared variance between V_i and W_i.

W_i and V_i are canonical variates of Y and X trait sets, respectively.

This estimate is necessary because a large canonical correlation does not always imply powerful relationship between two sets of traits. Canonical correlation maximizes the estimate of correlation between linear combinations of traits in the two sets, but does not maximize the amount of variance accounted for in one set of traits by the other set of traits. Thus, the variance in one set of traits accounted for by the other set is obtained through the RM (Akbas and Takma, 2005).

To determine the level of significance between morphological and cytological traits, each morphological trait was regressed on the cytological traits in the X set (Appendix 14).

4.3 Results

The descriptive statistics and Pearson correlation coefficient (r) for the six traits are presented in Tables 4.1 and 4.2, respectively.

Table 4.1. Descriptive statistics of the cytological and morphological traits

Traits	Mean $\pm$ SE	Minimum	Maximum
DNA (pg)	1.858 $\pm$ 0.013	1.588	2.118
Ploidy	3.962 $\pm$ 0.091	2.000	6.000
APW	0.846 $\pm$ 0.051	0.000	1.000
NS	1.846 $\pm$ 0.108	1.000	5.000
WC	1.404 $\pm$ 0.117	0.000	3.000

DNA: Deoxyribonucleic acid content (pg); Ploidy level; APW: Absence or presence of wings; NS: Number of stems per plant; and WC: Wing colour

The Pearson correlation coefficients for the traits ranged between for the traits ranged between -0.3928 and 0.7798 and were statistically significant ($p<0.05$), except for the association between APW and Ploidy (r = -0.2715) and between WC and NS (r = 0.1755) (Table 4.2; Appendix 12). Although the correlations between morphological and cytological traits were generally weak, the statistical significance of the correlations except between

APW and ploidy, indicated that the morphological traits were influenced by the cytological traits.

Table 4.2. Pearson correlation coefficients between cytological and morphological traits

	DNA	Ploidy	APW	NS
Ploidy	0.7798**			
APW	-0.3928**	-0.2715ns		
NS	0.2882*	0.3468**	-0.4727**	
WC	-0.3578**	-0.2895*	0.7143**	0.1755ns

*: p< 0.05, **: p< 0.01, ns: not significant. DNA: Deoxyribonucleic acid content (pg); Ploidy level; APW: Absence or presence of wings; NS: Number of stems per plant; and WC: Wing colour

In this study, the X trait set comprised of two traits: p = 2; and the Y trait set comprised of three traits: q = 3. Thus, two pairs of canonical variates, W_1V_1 and W_2V_2 were formed based on the set with the smaller number. The canonical correlations between these variates are presented in Table 13. The first canonical correlation (W_1V_1) was 0.4441, which represents 69.91% [$\frac{0.4441}{0.4441 + 0.2850} \times 100$] of overlapping variance of the first canonical variate. The second canonical correlation W_2V_2, which exhibited 0.2850, represents 39.9% [$\frac{0.2850}{0.4441 + 0.2850} \times 100$] overlapping variance of the second canonical variate set (Table 4.3).

Table 4.3. Canonical correlations between canonical variates

Canonical variates	Canonical correlations	Squared canonical correlation	% correlation	Cumulative % correlation
W_1V_1	0.4441	0.1972	69.91	69.91
W_2V_2	0.2850	0.0812	39.09	100.0

The coefficients of canonical variates from the original data are presented in Table 4.4 These coefficients of canonical equations are not unique since the DNA coefficient value is more than 1.0. Therefore the coefficients were standardized to give canonical variates with zero mean and unit variance. The standardized canonical coefficients for the X and Y trait sets are

presented in Table 4.5. The magnitude of each canonical coefficient represents the relative contribution of each trait to its respective canonical variate.

Table 4.4. Non-standardized coefficients of the respective traits of the canonical variates

Traits	W_1	W_2	Traits	V_1	V_2
DNA	0.9605	-2.2161	APW	-0.0436	0.5411
Ploidy	0.0909	0.3209	NS	0.1046	0.1724
			WC	-0.0999	-0.0931

DNA: Deoxyribonucleic acid content (pg); Ploidy level; APW: Absence or presence of wings; NS: Number of stems per plant; and WC: Wing colour

From equations 9 and 10, the following canonical variates can be obtained from the standardized coefficients (Table 4.5):

$W_1 = 0.0890$ DNA + 0.0591 Ploidy

$V_1 = -0.0159$ APW + 0.0815 NS ï 0.0845 WC

$W_2 = -0.2052$ DNA + 0.2157 Ploidy

$V_2 = 0.1971$ APW + 0.1344 NS ï 0.0787 WC

From the equations above, W_1 estimates the additive effect between DNA amount and ploidy level; whereas V_I estimates the contrast between number of stems per plant on one hand and the other traits (wing colour and absence or presence of wing). This indicates that the large variation in morphological traits (wing colour and absence or presence of wing) compared to number of stems per plant was possibly due to the additive influence between the cytological traits, like DNA amount and ploidy level. However, the second canonical variate, W_2 estimates a contrast between DNA amount and ploidy level; whereas V_2, measures the difference between wing colour and the other traits (number of stem per plant and absence or presence of wing). This indicates that the variation in morphological traits (absence or presence of wing and number of stem per plant) compared to wing colour was possibly due to the influence of the cytological traits (DNA amount and ploidy level). Of the three morphological traits used, number of stems produced per plant and wing colour were more

stable (their signs did not change in both canonical variates) compared to absence or presence of wing (Table 4.5; Appendix 13). Ploidy level was also more stable compared to DNA content. This implies that a group of genotypes with similar ploidy level may not necessarily contain the same DNA content and consequently, the number of loci will vary.

Table 4.5. Standardized coefficients of the respective traits of the canonical variates

Traits	W_1	W_2	Traits	V_1	V_2
DNA	0.0890	-0.2052	APW	-0.0159	0.1971
Ploidy	0.0591	0.2157	NS	0.0815	0.1344
			WC	-0.0845	-0.0787

DNA: Deoxyribonucleic acid content (pg); Ploidy level; APW: Absence or presence of wings; NS: Number of stems per plant; and WC: Wing colour

The proportion of the total variance extracted from a set of traits by a canonical variate of that set is equal to the quotient of the sum of square of loadings and the number of traits in the set. Thus, the first canonical variates, W_1 in the X trait set; and V_1 in the Y trait set, were estimated as 0.0057 [$(0.0890^2 + 0.0591^2)/ 2$] and 0.0047 [$(-0.0159^2 + 0.0815^2 + (-0.0845^2)/ 3$] respectively. Therefore, the first canonical variate (W_1) extracted 0.57% in the X trait set and 0.47% in the Y trait set. The second canonical variates (W_2) in the X trait set and (V_2) in the Y trait set were estimated as 0.0443 [$(-0.2052^2 + 0.2157^2)/ 2$] and 0.0210 [$(0.1971^2 + 0.1344^2 + (-0.0787^2)/ 3$], respectively.

The redundancy index in a canonical variate is expressed as the percentage of variance it extracts from its own set of traits. Thus, the first canonical variate (V_1) extracted 0.09% [0.0047×0.4441^2] of the variance in the X trait set; whereas the second variate (V_2) extracted 0.17% [0.0210×0.2850^2] of the variance in the X trait set. The results suggest that traits in the Y trait set (APW, NS and WC) are influenced by those in the X trait set (DNA and ploidy).

The regression of each morphological trait on the two cytological traits revealed that cytological traits significantly ($p<0.05$) influenced the phenotypic expression of the morphological traits (Appendix 12).

4.4 Discussion

The first pair of canonical variates (W_1V_1) had the highest (0.4441) estimated canonical correlation compared to the second pair of canonical variates (W_2V_2 (0.2850)). The correlation between the first pair of canonical variate indicates that morphological traits: absence or presence of wing, number of stems per plant and wing colur are associated with cytological traits: DNA and Ploidy level.

The signs of the standardized coefficients reflect the effects of DNA and ploidy on absence or presence of wing, number of stems and wing colour. Wright *et al*. (2008) suggested that the total amount of DNA in the genome (genome size) roughly reflects an estimate of the number of genes within a genome. Thus, an understanding of allelic diversity within germplasm is relevant in association with observed phenotypic variation.

The redundancy estimates for the first and second canonical correlation suggested that 0.09 and 0.17% of the variance in the Y trait set (APW, NS and WC) was accounted by the X trait set (DNA and ploidy). Although the percentages were small, the variation in each of the morphological traits showed a significant ($p<0.05$) input by cytological traits. It is possible that some morphological traits were more influenced by cytological traits than others. A future challenge would be to investigate the specific traits that are influenced by specific genes.

4.5 Conclusion

The associations between morphological and cytological traits of yams were investigated using canonical correlation analysis. The phenotypic expression of morphological traits was apparently influenced by cytological traits. The first and second canonical correlations exhibited 69.91 and 39.09% overlapping variance of the canonical variate sets, respectively. The first canonical variate extracted 0.57% of the total variance in the X trait set and 0.47% in the Y trait set.

This study demonstrated the complex inter-relationships between morphological and cytological traits. This could be relevant as a pre-breeding guide to ploidy manipulation; and also for future investigation of the effect of specific genes on the phenotypic expression of genotypes.

4.6 References

Akbas, Y. And C. Takma. 2005. Canonical correlation analysis for studying the relationship between egg production traits and body weight, egg weight and age at sexual maturity in layers. *Czech Journal of Animal Science* 50: 163-168.

Hair, J.F., R.E. Anderson, R.H. Tatham, and C.B. Williams. 1998. *Multivariate Data Analysis*. 5th Edition. Prentice Hall, Inc. pp. 442-462.

Keskin, S. And F. Yasar. 2007. Use of canonical correlation analysis for determination of relationships among several traits in egg plants (*Solanum melongana* L.) under salt stress. *Pakistan Journal of Botany* 39: 1547-1552.

Manly, B.F.J. 1994. *Multivariate statistical methods*. A Primer. Second edition. Chapman & Hall, 2-6 Boundary Row, London SE1 8HN, UK. 214 pp.

Payne, R.W., D.A. Murray, S.A. Harding, D.B. Baird, and D.M. Souter. 2009. Genstat for Windows (12th Edition) Introduction VSN International, Hemel Hempstead.

Sharma, S. 1996. *Applied Multivariate Techniques*. John Willey and Sons, Inc., Canada. pp. 391-404.

Tabachnick, B. and L.S. Fidell. 2001. *Using Multivariate Statistics*. A Pearson Education Company, Needham Heights, USA. 966 pp.

Wright, S.I., R.W. Ness, J.P. Foxe, and S.C.H. Barette. 2008. Genomic consequences of out-crossing and selfing in plants. *International Journal of Plant Science* 169: 105-118.

CHAPTER FIVE

Research overview

5.1 Introduction

The focus of this study was to investigate the level of diversity within some germplasm from Sierra Leone using agro-morphological and cytological traits. Morphological characterization is often considered as a first step in the description and classification of yam germplasm before more in-depth biochemical and molecular studies are undertaken, since yam breeding heavily relies on the magnitude of genetic variability (Smith *et al*., 1995). An understanding of the genetic architecture and knowledge of existing variation in different traits of yams are important for successful yam breeding. Inter- and intra-population diversity of crop species is important for the analysis and monitoring of germplasm during the maintenance phase, and for predicting potential genetic gain in a breeding programme (Hayward and Breese, 1993). Against this background the following hypotheses were formulated and tested:

i. The various genotypes studied were morphologically different.
ii. The same genotype was called differently by the various ethnic groups.
iii. The local accessions had wide inter- and intra-group diversity in ploidy level and nuclear DNA content.
iv. There is a correlation between agro-morphological and cytological traits used in the two methods of characterization.

5.2 Summary of main findings

5.2.1 Genetic diversity of some morphological traits in yam (*Dioscorea* spp.) genotypes from Sierra Leone

The objective of this study was to determine the relationships among accessions and characterise them based on these relationships. Synopsis of findings obtained in this study were:

i. Of the 52 genotypes studied, 43 genotypes belong to *D. alata*, two belong to *D. bulbifera* and seven belong to *D. rotundata*.

ii. Diversity was observed for many morphological traits including the number of days to shoot emergence, shoot traits (position, shape, size, density, vein colour and measurements of leaves; shoot growth rate) and root traits (tuber shape and flesh colour of tuber). The exploitation of these traits would enable the development of elite genotypes in a yam breeding programme.

iii. Both graphical principal component and cluster analyses did not group all the genotypes according to their geographical origins. Whereas groups A, C and E were from the north, groups B and D had mixed origins.

iv. Genotypes WR 07/024, SR 07/075, 07/073, ER 07/032 and NR 07/042, overlapped in sub-groups B_1 and B_2 of the graphical principal component analysis which indicated the possibility of duplicate genotypes in the germplasm. This suggests the use of molecular techniques in future research to confirm the level of diversity obtained and explore the possibility of duplicate genotypes.

v. Genotypes of groups B, C, D and F belong to *D. alata* with 27.3% from the north, 18.2% from the south, 22.7% from the east and 31.8% from the west.

vi. Genotypes WR 07/013 and Sr 07/085, which exhibited pale purple to purple venation, also had purple leaf margin and light purple flesh colour of tuber.

vii. Most of the variation in the morphological traits measured was accounted for by factor 1 with moderate (-0.5178) to large (-0.8797) loadings compared to the loadings of few traits in the other factors.

viii. The highest weightings by factor 1 were given to leaf colour (-0.7710) and leaf margin colour (-0.8797). These traits were also highly positively correlated ($r = 0.872^{**}$) and served as an important descriptors in distinguishing between the accessions.

5.2.2 Diversity in ploidy level and nuclear DNA content (pg) of some yam (*Dioscorea* spp.) genotypes in Sierra Leone as determined by flow cytometry and chromosome counting

i. Determination of chromosome numbers and ploidy levels of six genotypes of yams using chromosome counts supported the ploidy level(s) obtained by FCM.

ii. Of the 52 genotypes, three (5.8%) were diploids, one (1.9%) was triploid, 45 (86.5%) were tetraploids, one (1.9%) was pentaploid and two (3.9%) were hexaploids.

iii. The genotypes of *D. alata* had diploid, triploid, tetraploid and hexaploid ploidy levels, which were more diverse compared to the genotypes of *D. bulbifera* and *D. rotundata* both of which had tetraploid and pentaploid accessions.

iv. The nuclear DNA content varied from 1.588 to 1.718 pg for the diploids, 1.750 pg for the triploid, 1.772 to 1.937 pg for the tetraploids, 1.998 pg for the pentaploid and from 2.102 to 2.118 pg for the hexaploids.

v. Both the DNA content variation within individuals of the tetraploid population and those involving all ploidy populations among the three species was significant.

vi. The DNA amount per genome decreased with increasing ploidy level.

vii. Polyploidy in yams may have arisen from multiplication of the basic chromosome number.

viii. The conventional chromosome counting technique revealed the basic chromosome number as $x = 10$.

5.2.3 Determination of associations between three morphological and two cytological traits of yams (*Dioscorea* spp.) using canonical correlation analysis

The associations between morphological and cytological traits of yams were:

i. The morphological (phenotypic) expression of the various genotypes studied was influenced by cytological traits.

ii. The first and second canonical correlations exhibited 69.91 and 39.09% overlapping variance of the canonical variate sets, respectively.

iii. The first canonical variate extracted 0.57% of the total variance in the X trait set and 0.47% in the Y trait set.

iv. The redundancy estimates for the first and second canonical correlation suggested that 0.09 and 0.17% of the variance in the Y trait set (APW, NS and WC) was accounted by the X trait set (DNA and ploidy).

5.3 Implications of findings for future research

Based on results obtained from morphological and cytological characterization of yams, the 52 genotypes were found to belong to three species (*D. alata*, *D. bulbifera* and *D. rotundata*) and five ploidy levels (2*x*, 3*x*, 4*x*, 5*x* and 6*x*). However, this is just the first step in determining the magnitude of existing morphological diversity within the germplasm from Sierra Leone. The presence of overlap between groups in the graphical principal component analysis indicates the possibility of duplicate genotypes.

The factor analysis generally identified tuber shape, petiole length, leaf lobation and days to emergence as specific factors. However, days to emergence and tuber shape are more relevant for the economic yielding ability of the crop. The wide range and high variance among genotypes for number of days to sprouting could be attributed to inter- and intra-species diversity (Appendix 4). The maximization of effective growth period, estimated as the period between shoot emergence and senescence of leaves, depends on the genotype, species, type of planting materials and environmental factors such as variability in soil fertility, weather, climate and altitude at various locations. Akoroda (1993) suggested that inherent yam tuber dormancy and the degree of sett maturity are among the major factors that determine pre-emergence duration in yams. Manipulation of tuber dormancy and sprouting period depends on the objectives of the breeding programme. For instance, prolonging tuber dormancy may increase the post-harvest shelf life of healthy yam tubers. But the duration of dormancy required may differ among farmers and sometimes not correspond with the breedersô perspective thereof. One group of farmers may need early maturing tubers with rapid, uniform and vigorous sprouting seedlings in order to plan other management operations such as staking, weeding and fertilizer application (Godwin-Egein and Igwilo, 2005). Another group of farmers on the other hand, may prefer increased shelf life for marketing flexibility. A yam researcher on the other hand, may manipulate tuber dormancy in order to synchronize the time of male and female flowering for hybridization. Delayed sprouting does not only affect flowering, but also tuber shape and size, which are determined by the amount of assimilates translocated from the source to the sink during the effective growth period. Against this complex background, it is clearly essential that the researcher conduct a thorough needs assessment towards implementing the correct strategies of a market-oriented breeding programme.

In an effort to maximize uniform sprouting and mass propagation of yams, various techniques such as: pre-sprouting setts or minisetts (cutting the mother or bigger yam tuber into smaller sizes) in different growing media; and tissue culture to eliminate disease in addition to rapid multiplication using meristem or shoot tip and nodal cultures may be employed (Orkwor *et al*., 1998). The use of tissue culture techniques could have been explored in this study to enhance rapid multiplication. However, potentially that could have caused unwanted complexity in the diversity study, since the technique is associated with somaclonal variation. Somaclonal variation can be both genotypic or phenotypic with the former the consequence of alterations in chromosome numbers, chromosome structure and DNA sequence. Although the development of such genotypic variability is crucial to the breeder in that it provides the opportunity to exploit a wider genetic-base, the primary objective of this study was to obtain a measure of the existing diversity within the sampled germplasm before inducing new genetic variation.

Morphological characterization is often considered first step in diversity study before more in-depth biochemical or molecular studies are attempted. The statistical techniques of principal components, factor and cluster analyses are important in the characterization of germplasm as they enable the identification of the specific traits that effectively contributed towards the diversity observed. Principal components analysis is useful in the identification of meaningful descriptors that effectively account for most of the diversity observed, saving time and effort for future characterization. Notwithstanding, a large number of polymorphic markers are required to measure genetic relationships and genetic diversity in a reliable manner (Tatineni *et al.,* 1996). This limits the use of morphological characters and isozymes, which are few or lack adequate levels of polymorphism in yams.

Cluster analysis of the germplasm based on regional distribution classified 42.8% of the *D. rotundata* genotypes in group D came from the south of Sierra Leone, 28.6% from the north and 28.6%, the improved checks, were from IITA. The two genotypes of *D. bulbifera* in group A, NR 07/040 and NR 07/045, were from the north. Genotypes of groups B, C, D, E and F belong to *D. alata* with 27.3% from the north, 18.2% from the south, 22.7% from the east and 31.8% from the west. Since cluster analysis categorises accessions based on the similarity of the morphological traits or other markers, it obvioiusly does not necessarily group accessions that had the same geographical origin.

Both the genome size variation within individuals of the tetraploid population and those involving populations of all ploidy levels was sigmificant. These results are in agreement with the proposition that the plant kingdom is characterized by: (i) large divergence in genome size, even among closely related taxa with the same number of chromosomes; (ii) high incidence of polyploidy; and (iii) frequently non-proportional changes of nuclear DNA amount with respect to ploidy level (Suda *et al.*, 2006).

Morphological and flow cytometric characterization of yams proved to be efficient approaches to distinguishing between the various genotypes. However, this is just the first step in determining the magnitude of existing diversity within germplasm. Marker based selections are presently regarded as efficient and effective ways to exploiting the existing level of variation in a germplasm collection. Thus, complementary markers to morphological and cytological markers based on biochemical and molecular analyses are necessary to further identify and categorize major yam cultivars grown in Sierra Leone.

The declining genetic variability in yam populations could possibly be due to anthropogenic interferences and the associated massive destruction of habitats. Ecogeographical conditions and habitat distribution patterns are also influential factors which affect genetic diversity within populations. Therefore, consideration of these factors and their interactions prior to sampling accessions for *ex situ* and *in situ* conservation programmes is extremely important.

This investigation on the correlation between morphological and cytological traits is the first report on the level of associations between these two sets of variables. The study on both methods of characterization of yams from Sierra Leone provided better understanding of existing diversity that would help to optimize the efficiency and efficacy of breeding programmes in that country and elsewhere.

5.4 References

Akoroda, M.O. 1993. Yams. *Dioscorea* spp. In: *Genetic Improvement of Vegetable Crops*. G. Kalloo and B.O. Berah, eds. Pergamon Press, New York. pp. 717-733.

Godwin-Egein, M.I. and N.H. Igwilo. 2005. Effect of sprouting medium on the survival of yam peelsetts. *International Journal of Agriculture and Biology* 2: 315-317.

Hayward, M.D. and E.L. Breese. 1993. Population structure and variability. In: *Plant Breeding: Principles and Prospects*, MD Hayward, NO Bosemark and I Romayosa (Eds.), Chapman and Hall, London. pp. 17-29.

Mwiringi, P.N., E.M. Kahangi, A.B. Ngende, and E.G. Mamati. 2009. Morphological variability within the Kenyan yam (*Dioscorea* spp.). *Journal of Applied Biosciences* 16: 894-901.

Orkwor, G.C., R. Asiedu, and I.J. Ekanayake. 1998. *Food Yams. Advances in Research*, IITA and NRCRI, Nigeria. 249 pp.

Smith, S.E., L. Guarino, A. Al. Doss, and D.M. Conta. 1995. Morphological and agronomic affinities among Middlea Eastern alfalfas accessions from Omen and Yemen. ~~*Journal of*~~ *Crop Science* 35: 1118-1194.

Suda, J., A. Krahuleova, P. Travnieek, and F. Krahulee. 2006. Ploidy level versus DNA ploidy level: an appeal for consistent terminology. *Taxonomy* 55: 447-450.

Tatineni, V., R.G. Cantrell, and D.D. Davis. 1996. Genetic diversity in elite cotton germplasm determined by morphological characters and RAPDs. *Journal of Crop Science* 36: 186-192.

APPENDICES

Appendix 1. List of morphological traits measured in 52 yam (*Dioscorea* spp.) genotypes according to the original International Plant Genetic Resources Institute descriptor list (IPGRI/ IITA, 1997)

IPGRI code	Traits acronym	Characters/ descriptors	Score code ï descriptor state
		Shoot traits	
7.1.1	DE	Number of days to emergence	Direct measurement
7.1.17	NS	Number of stems per plant	
7.1.18	SC	Stem colour	1 ï Green; 2 ï Purplish green; 3 ï Brownish green; 4 ï Dark brown; 5 ï Purple; Other (specify in descriptor 7.7 Notes)
7.1.19	NB	Number of internodes to first banching	Direct measurement
7.1.23	IL	Internode length (cm)	Recorded at 1 m height. Average of five plants
7.1.25	APW	Absence or presence of wings	0 ï absent; 1 ï present
7.1.27	WC	Wing colour	1 ï Green; 2 ï Green with purple edge; 3 ï Purple; Other (specify in descriptor 7.7 Notes)
7.2.9	PL	Position of leaves (mature leaves)	1 ï Alternate, 2 ï Opposite, 3 ï Alternate at base/ opposite above; Other (specify in descriptor 7.7 Notes)
7.2.10	LD	Leaf density	3 ï Low; 5 ï Intermediate; 7 ï High
7.2.12.2	LL	Leaf lobation	1 ï Shallowly lobed; 2 ï Deeply lobed
7.2.15	LC	Leaf colour	1 ï Yellowish; 2 ï Pale green; 3 ï Dark green; 4 ï Purplish green; 5 ï Purple; Other (specify in descriptor 7.7 Notes)
7.2.16	LVCUS	Leaf vein colour (upper surface)	1 ï Yellowish; 2 ï Green; 3 ï Pale purple; 4 ï Purple; Other (specify in descriptor 7.7 Notes)
7.2.17	LVCLS	Leaf vein colour (lower surface)	1 ï Yellowish; 2 ï Green; 3 ï Pale purple; 4 ï Purple; Other (specify in descriptor 7.7 Notes)
7.2.18	LMC	Leaf margin colour	1 ï Green; 2 ï Purple; Other (specify in descriptor 7.7)
7.2.22	LS	Leaf shape	1ï Ovate; 2ï Cordate; 3ï Cordate long; 4ï Cordate broad; 5ï Sagittate long; 6ï Saggitate broad; 7ï Hastate; Other (specify in descriptor 7.7 Notes)
7.2.23	LAS	Leaf apex shape	1 ï Obtuse; 2 ï Acute; 3 ï Emarginated; 99 ï Other (specify in descriptor 7.7 Notes)
7.2.25	DBL	Distance between lobes	1 ï No measureable distance; 5 ï Intermediate; 9 ï Very distant
7.2.30.1	LML1	Leaf measurement L1	
7.2.30.2	LML2	Leaf measurement L2	The various parts of leaves measured are shown in the figure below.
7.2.30.3	LMW1	Leaf measurement W1	
7.2.30.4	LMW2	Leaf measurement W2	
7.2.32	TLM	Tip length	1 ï >2 mm; 2 ï 2-5 mm; 3 ï >5 mm
7.2.33	TC	Tip colour	1 ï Light green; 2 ï Dark green; 3 ï Purple/ green; 4 ï Red; Other (specify in descriptor 7.7 Notes)
7.2.34	PLM	Petiole length of mature leaves	1 - ≤5 cm; 2 ï 6-9 cm; 3 ï ≥10 cm
7.2.37	PC	Petiole colour	1 ï All green with purple base; 2 ï All green with purple leaf junction; 3 ï All green with purple with purple at both ends; 4 ï All purplish green with purple base; 5 ï All purplish green with purple leaf junction; 6 ï All purplish green with purple at both ends; 7 ï Green; 8 ï Purple, 9 = Brownish green; 10 ï Brown; 11 ï Dark brown; 99 ï Other (specify in descriptor 7.7 Notes)
7.2.38	PWC	Petiole wing colour	1 ï Green; 2 ï Green with purple; 3 ï Purple; 99 ï (specify in descriptor 7.7 Notes)

Appendix 1. Continued

		Underground tuber traits	
7.6.14	TS	Tuber shape	1ï Round; 2ï Oval; 3ï Oval oblong; 4ï Cylindrical; 5ï Flattened; 6 ï Irregular; Other (spsecify in descriptor 7.7 Notes)
7.6.30	FCCCS	Flesh colour at central transverse cross section	1 ï White; 2 ï Yellowish white or off-white; 3 ï Yellow; 4 ï Orange; 5 - Light purple; 6 ï Purple; 7 ï Purple with white; 8 ï White with purple; 9 ï Outer purple/inner yellowish; 99 ï Other (specify in descriptor 7.7)

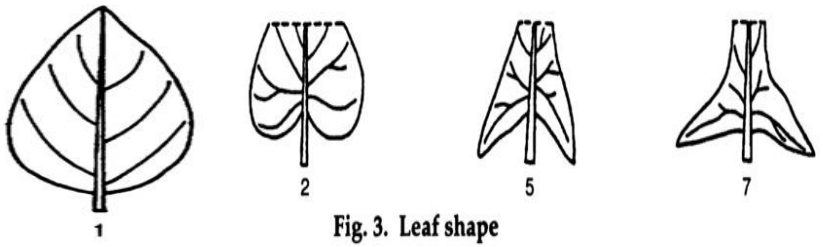


Fig. 3. Leaf shape

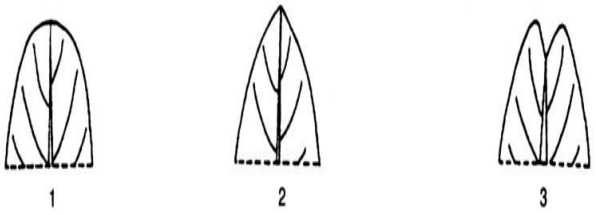


Fig. 4. Leaf apex shape

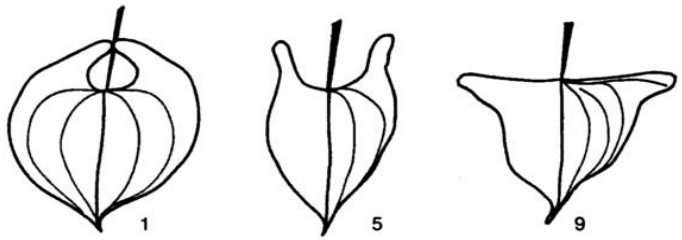


Fig. 5. Distance between lobes

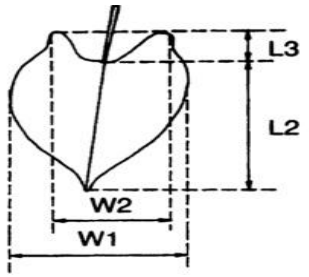


Fig. 6. Leaf measurement

Appendix 2. Correlation matrix of 28 trait means across 52 yam genotypes used in the principal component, factor and cluster analyses

	APW	DBL	DE	IL	LAS	LC	LD	LL	LMC	LML1	LML2	LMW1	LMW2	LS	LVCLS	LVCUS	NB	NS	PC	PL	PLM	PWC	SC	TC	TLM	TS	WC	Tfleshc
APW	-																											
DBL	-0.382	-																										
DE	-0.257	0.217	-																									
IL	0.114	-0.420	0.326	-																								
LAS	0.255	**-0.549**	-0.147	0.108	-																							
LC	**-0.517**	0.330	0.138	0.046	-0.476	-																						
LD	-0.481	-0.081	0.092	0.327	-0.230	**0.558**	-																					
LL	**0.612**	-0.433	-0.134	0.058	0.234	-0.324	-0.216	-																				
LMC	**-0.562**	0.420	0.091	-0.131	**-0.551**	**0.872**	**0.602**	-0.342	-																			
LL1	0.174	-0.420	0.230	**0.634**	0.194	0.035	0.205	0.123	-0.116	-																		
LL2	0.438	**-0.601**	0.044	0.499	0.383	-0.291	-0.020	0.407	-0.435	**0.667**	-																	
LW1	-0.092	-0.375	0.176	0.367	0.294	0.053	0.228	-0.036	-0.126	**0.684**	**0.533**	-																
LW2	-0.093	-0.043	0.229	0.361	0.137	-0.054	0.016	-0.186	-0.193	**0.562**	0.334	**0.612**	-															
LS	-0.288	**0.613**	0.246	-0.170	-0.162	0.124	-0.180	-0.355	0.110	-0.101	-0.470	-0.085	0.204	-														
LVCLS	0.116	-0.264	-0.119	0.184	-0.268	0.374	0.342	0.093	0.415	0.173	0.218	0.008	-0.274	-0.434	-													
LVCUS	-0.426	0.220	0.141	0.071	**-0.605**	**0.630**	0.489	-0.362	**0.673**	-0.076	-0.191	-0.072	-0.073	-0.168	0.464	-												
NB	**-0.684**	0.122	0.200	0.190	-0.426	**0.510**	**0.685**	-0.474	**0.503**	0.065	-0.163	0.243	0.132	0.100	0.086	**0.520**	-											
NS	-0.473	0.208	0.116	0.122	-0.497	**0.593**	**0.693**	-0.275	**0.628**	0.102	-0.196	0.104	-0.016	0.073	0.405	**0.544**	**0.572**	-										
PC	0.491	0.074	-0.028	0.051	-0.117	-0.176	-0.366	0.184	-0.238	0.171	0.099	0.126	0.068	0.134	-0.044	-0.205	-0.351	-0.212	-									
PL	-0.064	0.264	0.121	-0.049	-0.296	0.080	-0.048	-0.036	0.156	-0.122	-0.227	-0.213	0.022	0.165	-0.010	0.141	0.019	0.087	0.004	-								
PLM	0.146	-0.296	0.031	0.281	0.300	-0.051	-0.045	0.056	-0.277	**0.535**	0.427	**0.718**	**0.636**	0.058	-0.146	-0.237	-0.045	-0.090	0.190	-0.052	-							
PWC	**0.628**	-0.363	0.040	0.370	0.212	-0.047	-0.036	**0.520**	-0.065	0.427	**0.504**	0.105	0.065	-0.405	0.384	-0.064	-0.389	-0.160	0.137	-0.044	0.167	-						
SC	**-0.556**	**0.578**	0.098	-0.339	-0.422	**0.627**	0.228	-0.425	**0.718**	-0.184	**-0.502**	0.006	-0.027	0.274	0.053	0.413	0.330	0.357	0.032	0.057	-0.040	-0.280	-					
TC	0.155	**-0.503**	-0.201	0.297	0.419	0.105	0.168	0.238	-0.023	0.416	0.446	0.471	0.126	-0.380	0.264	-0.067	-0.182	0.133	0.159	-0.176	0.446	0.345	-0.011	-				
TLM	0.481	-0.425	0.044	0.480	0.445	-0.182	-0.227	0.367	-0.375	**0.620**	**0.599**	0.425	0.455	-0.145	0.097	-0.333	-0.389	-0.256	0.295	-0.027	**0.502**	**0.565**	-0.377	0.478	-			
TS	-0.068	0.323	-0.060	-0.094	-0.308	0.363	0.083	-0.068	0.324	-0.211	-0.298	-0.123	-0.200	0.328	0.062	0.138	0.097	0.194	0.050	-0.002	-0.074	-0.090	0.287	-0.207	-0.331	-		
WC	**0.714**	-0.402	-0.352	0.172	0.270	-0.096	-0.121	0.434	-0.137	0.357	0.463	0.143	-0.031	-0.337	0.411	-0.260	**-0.509**	-0.176	0.325	-0.170	0.274	**0.708**	-0.198	**0.500**	**0.532**	0.015	-	
Tfleshc	-0.014	-0.367	-0.207	0.339	0.045	**0.520**	**0.571**	0.130	0.423	0.207	0.216	0.082	-0.239	-0.460	**0.644**	0.264	0.185	0.318	-0.225	-0.154	-0.154	0.379	-0.003	0.459	0.135	0.034	0.421	-

Appendix 3. Spearmanôs rank correlation coefficient of the 28 trait means across 52 genotypes used in the principal component, factor and cluster analyses

	APW	DBL	DE	IL	LAS	LC	LD	LL	LMC	LML1	LML2	LMW1	LMW2	LS	LVCLS	LVCUS	NB	NS	PC	PL	PLM	PWC	SC	TC	TLM	TS	WC	Tfleshc
APW	*																											
DBL	**0.004**	*																										
DE	**0.044**	0.066	*																									
IL	0.353	**0.002**	0.064	*																								
LAS	0.096	**0.000**	0.106	0.552	*																							
LC	**0.000**	0.062	0.190	0.324	**0.002**	*																						
LD	**0.000**	0.896	0.136	0.086	**0.022**	**0.000**	*																					
LL	**0.000**	**0.001**	0.429	0.695	0.123	**0.023**	0.082	*																				
LMC	**0.000**	**0.002**	0.229	0.323	**0.000**	**0.000**	**0.000**	**0.013**	*																			
LL1	0.199	**0.001**	0.549	**0.000**	0.169	0.331	0.216	0.363	0.330	*																		
LL2	**0.004**	**0.000**	0.820	**0.000**	**0.009**	0.389	0.732	**0.006**	**0.005**	**0.000**	*																	
LW1	0.712	**0.005**	0.914	**0.015**	**0.028**	0.256	0.180	0.927	0.325	**0.000**	**0.000**	*																
LW2	0.577	0.744	0.335	**0.018**	0.529	0.486	0.713	0.161	0.252	**0.000**	**0.009**	**0.000**	*															
LS	**0.003**	**0.000**	0.052	**0.019**	**0.001**	**0.018**	0.648	**0.001**	**0.002**	0.108	**0.000**	0.204	0.466	*														
LVCLS	0.385	**0.046**	0.744	0.205	0.208	**0.020**	**0.046**	0.708	**0.005**	0.573	0.125	0.921	**0.022**	**0.023**	*													
LVCUS	**0.001**	0.071	0.058	0.550	**0.000**	**0.000**	**0.000**	**0.003**	**0.000**	0.490	0.238	0.666	0.744	0.257	**0.002**	*												
NB	**0.000**	0.313	0.068	0.252	**0.000**	**0.000**	**0.000**	**0.001**	**0.000**	0.603	0.387	0.201	0.637	0.091	0.136	**0.000**	*											
NS	**0.000**	0.310	0.124	0.229	**0.000**	**0.000**	**0.000**	**0.024**	**0.000**	0.288	0.307	0.183	0.685	0.169	**0.007**	**0.000**	**0.000**	*										
PC	**0.000**	0.840	0.610	0.679	0.778	**0.028**	**0.003**	0.276	**0.001**	0.453	0.703	0.204	0.430	0.472	0.102	**0.009**	**0.007**	**0.016**	*									
PL	0.610	0.059	0.145	0.823	**0.043**	0.780	0.662	0.721	0.241	0.379	0.129	0.187	0.923	**0.036**	0.493	0.352	0.738	0.780	0.787	*								
PLM	0.211	**0.023**	0.325	0.053	**0.015**	0.702	0.520	0.683	**0.018**	**0.000**	**0.001**	0.000	0.000	0.749	0.196	**0.044**	0.429	0.404	**0.046**	0.790	*							
PWC	**0.000**	**0.008**	0.768	**0.005**	0.120	0.533	0.244	**0.000**	0.237	**0.001**	**0.000**	0.365	0.496	**0.003**	**0.014**	0.202	**0.004**	0.124	0.305	0.832	**0.124**	*						
SC	**0.002**	**0.000**	0.562	**0.003**	**0.014**	**0.000**	0.107	**0.010**	**0.000**	0.334	**0.006**	0.634	1.000	**0.002**	0.542	**0.008**	**0.028**	**0.029**	0.528	0.935	0.959	0.095	*					
TC	0.528	**0.000**	**0.032**	**0.032**	**0.006**	0.157	0.234	0.150	0.749	**0.003**	**0.001**	**0.000**	0.368	**0.001**	**0.049**	0.708	0.624	0.693	0.741	0.294	**0.005**	**0.025**	0.420	*				
TLM	**0.000**	**0.001**	0.414	**0.000**	**0.004**	0.634	**0.049**	**0.007**	**0.007**	**0.000**	**0.000**	**0.002**	**0.002**	**0.008**	0.587	**0.011**	**0.004**	**0.040**	0.061	0.781	**0.000**	**0.000**	**0.031**	**0.002**	*			
TS	0.199	**0.008**	0.996	0.480	**0.015**	**0.001**	0.156	0.378	**0.002**	0.112	**0.008**	0.364	0.157	**0.001**	0.483	0.143	0.102	**0.040**	0.673	0.726	0.448	0.142	**0.025**	0.192	**0.015**	*		
WC	**0.000**	**0.006**	**0.001**	0.251	**0.036**	0.558	0.143	**0.005**	0.293	**0.007**	**0.002**	0.143	0.789	**0.017**	**0.016**	**0.016**	**0.001**	0.112	0.129	0.258	**0.019**	**0.000**	0.918	**0.001**	**0.000**	0.533	*	
tfleshc	0.421	**0.002**	0.202	**0.006**	0.375	**0.001**	**0.000**	0.685	**0.007**	0.109	0.063	0.354	0.212	**0.001**	**0.000**	0.108	**0.023**	**0.009**	**0.001**	0.162	0.257	0.143	0.943	**0.000**	0.305	0.987	**0.039**	*

Appendix 4. Summary statistics for 28 morphological trait means across 52 genotypes

Trait	Mean ± SD	Minimum	Maximum	Range	Variance
APW	0.846 ± 0.364	0	1	1	0.133
DBL	5.077 ± 2.168	1	9	8	4.700
DE	45.58 ± 19.78	13	91	78	391.5
IL	2.827 ± 0.760	2	4	2	0.577
LAS	4.154 ± 1.526	2	7	5	2.329
LC	2.212 ± 1.073	1	4	3	1.150
LD	4.000 ± 1.343	3	7	4	1.804
LL	1.673 ± 0.474	1	2	1	0.224
LMC	1.212 ± 0.412	1	2	1	0.170
LL1	3.750 ± 0.926	2	6	4	0.858
LL2	2.250 ± 0.738	1	4	3	0.544
LW1	3.308 ± 0.897	2	5	3	0.805
LW2	3.058 ± 0.895	1	5	4	0.801
LS	4.904 ± 0.891	2	6	4	0.794
LVCLS	2.115 ± 0.427	2	4	2	0.183
LVCUS	1.923 ± 0.837	1	4	3	0.700
NB	2.385 ± 1.962	1	7	6	3.849
NS	1.519 ± 0.779	1	5	4	0.607
PC	5.981 ± 1.955	1	8	7	3.823
PL	2.365 ± 0.908	1	3	2	0.825
PLM	3.346 ± 0.653	2	5	3	0.427
PWC	1.192 ± 0.817	0	3	3	0.668
SC	1.673 ± 0.834	1	5	4	0.695
TC	4.000 ± 2.086	1	7	6	4.353
TLM	3.327 ± 0.964	2	5	3	0.930
TS	3.058 ± 1.227	1	6	5	1.506
WC	1.404 ± 0.846	0	3	3	0.716
FCCCS	1.231 + 0.614	1	3	2	0.377

Appendix 5. Standardized data of 28 morphological traits of 52 yam genotypes

Variety	std-DE	std-NS	std-SC	std-IL	std-NB	std-APW	std-WC	std-PL	std-LD	std-LL	std-LC	std-LVCUS	std-LVCLS	std-LMC
NR07/045	-2.323	-1.881	-0.838	1.544	1.865	-0.197	2.234	-1.421	-0.513	0.270	1.017	1.886	2.171	-1.014
NR07/040	-2.323	-1.881	1.234	1.544	1.865	-0.197	2.234	-1.421	-0.513	1.350	1.017	1.886	-0.064	-1.014
WR07/025	0.422	-0.035	1.234	1.544	0.554	-0.197	0.745	0.690	-0.513	1.350	1.017	1.886	2.171	0.108
WR07/024	0.422	-0.035	-0.484	0.228	0.554	-0.197	0.745	0.690	-0.513	1.350	1.017	0.771	1.053	0.108
NR07/059	0.422	-1.881	-0.181	0.228	1.210	1.667	2.234	0.690	1.912	-0.810	1.017	-0.343	-2.300	-3.258
SR07/075	0.422	-1.881	-0.484	0.228	0.554	-0.197	-0.745	0.690	-0.513	1.350	1.017	0.771	-0.064	0.108
NR07/042	0.422	-1.881	0.830	1.544	0.554	-0.197	0.745	0.690	-0.513	1.350	-0.339	0.771	1.053	0.108
SR07/074	-2.323	1.810	1.841	-1.088	-1.411	1.667	0.745	-1.421	1.912	-0.810	-1.695	-0.343	-0.064	1.230
NR07/069	0.422	-0.035	0.931	1.544	-0.101	-0.197	-0.745	0.690	-0.513	-0.810	-0.339	-1.457	-0.064	0.108
NR07/043	0.422	-0.035	-0.484	1.544	-0.101	1.667	2.234	0.690	1.912	1.350	-0.339	-0.343	-0.064	0.108
ER07/030	0.422	-0.035	2.296	0.228	0.554	-0.197	-0.745	0.690	-0.513	1.350	2.372	1.886	1.053	0.108
ER07/032	0.422	-0.035	0.982	0.228	0.554	-0.197	-0.745	0.690	-0.513	1.350	1.017	0.771	1.053	0.108
NR07/041	0.422	-0.035	-0.484	1.544	-1.411	-0.197	0.745	0.690	-0.513	0.270	1.017	0.771	1.053	0.108
NR07/067	0.422	-0.035	1.841	0.228	-1.411	-0.197	-0.745	-1.421	-0.513	-0.810	-0.339	-0.343	-1.182	0.108
SR07/085	0.422	-0.035	-0.484	1.544	-1.411	1.667	2.234	0.690	1.912	1.350	1.017	-0.343	-1.182	0.108
ER07/038	0.422	-0.035	-0.484	0.228	-1.411	-0.197	0.745	-1.421	-0.513	0.270	-0.339	-0.343	1.053	1.230
ER07/033	0.422	-0.035	1.184	1.544	0.554	-0.197	-0.745	-1.421	-0.513	2.429	1.017	0.771	1.053	0.108
WR07/010	0.422	-0.035	1.184	1.544	-1.411	-0.197	0.745	0.690	-0.513	1.350	1.017	0.771	-0.064	0.108
NR07/052	-2.323	1.810	1.993	0.228	-1.411	1.667	0.745	-1.421	1.912	-0.810	-1.695	-0.343	-0.064	1.230
SR07/079	0.422	-0.035	1.841	1.544	0.554	-0.197	-0.745	0.690	-0.513	1.350	1.017	1.886	2.171	0.108
WR07/013	0.422	-1.881	-1.647	0.228	-1.411	1.667	0.745	0.690	1.912	1.350	1.017	0.771	-0.064	-3.258
SR07/073	0.422	-0.035	1.993	0.228	0.554	-0.197	-0.745	0.690	-0.513	1.350	1.017	0.771	1.053	0.108
WR07/016	0.422	-0.035	-1.394	-1.088	0.554	-0.197	-0.745	-1.421	-0.513	0.270	-0.339	0.771	1.053	0.108
WR07/020	0.422	-0.035	-1.394	-1.088	0.554	-0.197	-0.745	0.690	-0.513	-0.810	-0.339	-0.343	-1.182	0.108
ER07/034	0.422	-0.035	-1.192	-1.088	0.554	-0.197	-0.745	0.690	-0.513	0.270	-0.339	0.771	-0.064	0.108
ER07/029	0.422	-0.035	-1.192	-1.088	0.554	-0.197	-0.745	0.690	-0.513	-0.810	1.017	0.771	-0.064	0.108
WR07/008	0.422	-0.035	-1.091	0.228	0.554	-0.197	-0.745	-1.421	-0.513	0.270	1.017	0.771	1.053	0.108
ER07/031	0.422	-0.035	-0.989	0.228	0.554	-0.197	-0.745	0.690	-0.513	0.270	-0.339	-0.343	-0.064	0.108
WR07/015	0.422	-0.035	-0.888	-1.088	0.554	-1.130	-0.745	0.690	-0.513	-0.810	-0.339	-0.343	-1.182	0.108
WR07/014	0.422	-0.035	-0.787	0.228	0.554	-0.197	-0.745	-1.421	-0.513	0.270	-0.339	-0.343	1.053	0.108
SR07/080	0.422	-0.035	-0.787	-1.088	0.554	-1.130	-0.745	0.690	-0.513	-0.810	-0.339	0.771	1.053	0.108

WR07/028	0.422	-0.035	-0.737	-1.088	0.554	-0.197	-0.745	0.690	-0.513	0.270	1.017	0.771	-0.064	0.108
SR07/076	0.422	-0.035	-0.636	0.228	0.554	1.667	-0.745	0.690	-0.513	-0.810	-0.339	-0.343	-1.182	0.108
NR07/047	0.422	-1.881	-0.737	0.228	1.210	-1.130	-0.745	0.690	-0.513	0.270	-0.339	0.771	-1.182	0.108
SR07/084	0.422	1.810	-0.535	0.228	0.554	-1.130	-0.745	-1.421	-0.513	-0.810	-0.339	-1.457	-1.182	0.108
WR07/007	0.422	-0.035	-0.433	-1.088	0.554	-1.130	-0.745	0.690	-0.513	-0.810	-0.339	-1.457	-0.064	0.108
NR07/068	0.422	-0.035	-0.433	-1.088	-1.411	-1.130	0.745	0.690	-0.513	-0.810	-0.339	-0.343	-1.182	-3.258
WR07/004	0.422	-0.035	-0.282	0.228	-1.411	-1.130	-0.745	0.690	-0.513	-0.810	-0.339	-1.457	-1.182	0.108
SR07/081	0.422	-0.035	-0.080	-1.088	0.554	-1.130	-0.745	0.690	-0.513	-0.810	-0.339	-1.457	-0.064	0.108
ER07/036	0.422	-0.035	-0.282	0.228	-1.411	-0.197	-0.745	0.690	-0.513	-0.810	1.017	-1.457	-0.064	-2.136
WR07/022	0.422	-0.035	-0.181	-1.088	0.554	-1.130	-0.745	0.690	-0.513	0.270	1.017	-1.457	-0.064	0.108
NR07/054	0.422	-0.035	-0.181	-1.088	0.554	-0.197	-0.745	0.690	-0.513	-1.889	-1.695	-1.457	-1.182	0.108
NR07/060	-2.323	1.810	-0.029	-1.088	-1.411	1.667	0.745	-1.421	1.912	-0.810	-1.695	-0.343	-0.064	1.230
TDr95/00005	0.422	1.810	0.325	-1.088	-1.411	1.667	0.745	-1.421	1.912	-0.810	-1.695	-0.343	-0.064	1.230
SR07/082	0.422	-0.035	-0.585	0.228	0.554	-1.130	-0.745	-1.421	-0.513	-0.810	-0.339	-0.343	-0.064	0.108
ER07/037	0.422	-0.035	-0.535	-1.088	0.554	-1.130	-0.745	0.690	-0.513	-0.810	-0.339	-0.343	-1.182	0.108
NR07/071	-2.323	1.810	0.173	-1.088	-1.411	1.667	0.745	-1.421	1.912	-0.810	-1.695	-0.343	-0.064	1.230
SR07/072	-2.323	1.810	-0.029	-1.088	-1.411	1.667	0.745	-1.421	1.912	-0.810	-1.695	-0.343	-0.064	1.230
WR 07/001	0.422	-0.035	0.224	0.228	0.554	-1.130	-0.745	0.690	-0.513	-0.810	-0.339	-1.457	-0.064	0.108
ER 07/039	0.422	-0.035	0.224	-1.088	0.554	-1.130	0.745	0.690	-0.513	-0.810	-0.339	-0.343	-1.182	0.108
NR 07/057	0.422	-0.035	0.224	0.228	0.554	-0.197	-0.745	0.690	-0.513	-0.810	-0.339	-1.457	-1.182	0.108
TDr95/18544	-2.323	1.810	0.426	-1.088	-1.411	1.667	0.745	-1.421	1.912	-0.810	-1.695	-0.343	-0.064	1.230

Appendix 5. Continued

Variety	std-LS	std-LAS	std-DBL	std-LL1	std-LL2	std-LW1	std-LW2	std-TLM	std-TC	std-PLM	std-PC	std-PWC	std-TS	std-tfleshc
NR07/045	-0.270	0.092	2.352	0.617	-2.547	-1.504	1.001	-1.459	-0.807	0.959	0.698	-1.677	-1.659	1.253
NR07/040	-0.270	0.092	2.352	0.617	-2.547	-1.504	-0.530	-1.459	-0.807	0.959	-1.376	-1.677	-1.659	1.253
WR07/025	-0.270	0.092	-0.706	0.617	0.521	0.699	2.531	0.988	-0.807	0.479	1.735	-0.862	0.705	-0.376
WR07/024	-0.270	-1.103	-0.706	0.617	0.521	0.699	1.001	0.988	-0.807	0.479	1.735	0.768	0.705	-0.376
NR07/059	2.071	1.287	0.314	0.617	-1.524	-1.504	-0.530	2.211	0.392	1.438	-0.339	0.768	1.886	2.881
SR07/075	-0.270	0.092	1.333	-0.666	0.521	0.699	1.001	-0.235	-0.807	0.479	1.735	-1.677	-0.477	-0.376
NR07/042	-0.270	0.092	-0.706	0.617	0.521	0.699	1.001	-0.235	-0.807	0.479	0.698	-0.047	-0.477	-0.376
SR07/074	-0.270	1.287	1.333	0.617	-1.013	0.699	-0.530	-1.459	1.592	-1.438	-1.376	0.768	-1.659	-0.376
NR07/069	-0.270	0.092	-0.706	-0.666	0.521	0.699	-0.530	-0.235	-0.807	0.479	0.698	0.768	-0.477	-0.376
NR07/043	-0.270	0.092	0.314	0.617	0.010	-1.504	-0.530	2.211	0.392	1.438	0.698	0.768	1.886	2.881
ER07/030	-0.270	-1.103	-0.706	-0.666	0.521	0.699	1.001	0.988	-0.807	0.479	1.735	-0.047	0.705	-0.376
ER07/032	-0.270	-1.103	-0.706	-0.666	0.521	-1.504	1.001	0.988	-0.807	0.479	0.698	-1.677	-0.477	-0.376
NR07/041	-0.270	0.092	2.352	0.617	0.521	0.699	1.001	-0.235	-0.807	-0.959	0.698	0.768	-0.477	-0.376
NR07/067	2.071	1.287	0.314	0.617	0.521	-1.504	-0.530	-0.235	-0.807	-1.438	-0.339	0.768	-0.477	-0.376
SR07/085	4.412	1.287	0.314	3.183	0.010	0.699	-0.530	0.988	0.392	1.438	0.698	-0.047	1.886	2.881
ER07/038	-0.270	0.092	1.333	0.617	0.521	0.699	1.001	-0.235	-0.807	-0.959	-1.376	-0.047	0.705	-0.376
ER07/033	-0.270	0.092	-0.706	-0.666	0.521	0.699	1.001	0.988	0.392	0.479	1.735	-1.677	0.705	-0.376
WR07/010	-0.270	0.092	1.333	0.617	0.521	-1.504	-0.530	-0.235	-0.807	-1.438	-1.376	2.397	-0.477	-0.376
NR07/052	-0.270	1.287	1.333	0.617	1.033	0.699	-0.530	-1.459	3.991	-0.479	-1.376	0.768	-1.659	-0.376
SR07/079	-0.270	0.092	0.314	-0.666	0.521	0.699	1.001	2.211	0.392	0.479	1.735	-0.047	0.705	-0.376
WR07/013	4.412	2.483	0.314	0.617	0.010	0.699	-0.530	2.211	0.392	1.438	0.698	-0.047	1.886	2.881
SR07/073	-0.270	-1.103	-0.706	-0.666	0.521	-1.504	1.001	0.988	0.392	0.479	1.735	-1.677	0.705	-0.376
WR07/016	-0.270	0.092	-0.706	-0.666	0.521	-1.504	1.001	-0.235	0.392	0.479	0.698	-0.047	0.705	-0.376
WR07/020	-0.270	-1.103	-0.706	-0.666	0.521	-1.504	1.001	-0.235	0.392	0.479	-0.339	-0.047	0.705	-0.376
ER07/034	-0.270	0.092	-0.706	0.617	0.521	-1.504	-0.530	-0.235	0.392	0.479	0.698	-0.047	0.705	-0.376
ER07/029	-0.270	0.092	-0.706	-0.666	0.521	-1.504	1.001	-0.235	0.392	0.479	-0.339	-0.047	0.705	-0.376
WR07/008	-0.270	0.092	-0.706	-0.666	0.521	-1.504	1.001	-0.235	0.392	0.479	-0.339	2.397	0.705	-0.376
ER07/031	-0.270	-1.103	-0.706	-0.666	0.521	-1.504	1.001	-0.235	0.392	0.479	0.698	0.768	0.705	-0.376
WR07/015	-0.270	-1.103	-0.706	-0.666	0.521	-1.504	-0.530	-0.235	0.392	0.479	-0.339	-0.047	0.705	-0.376
WR07/014	-0.270	-1.103	-0.706	-0.666	0.521	0.699	1.001	-0.235	0.392	0.479	0.698	-0.047	0.705	-0.376
SR07/080	-0.270	-1.103	-0.706	-0.666	0.521	0.699	1.001	-0.235	0.392	0.959	-0.339	-0.047	0.705	-0.376

WR07/028	-0.270	-1.103	-0.706	-0.666	0.521	0.699	1.001	-0.235	0.392	0.959	-0.339	0.768	-0.477	-0.376
SR07/076	-0.270	-1.103	-0.706	-0.666	0.521	0.699	-0.530	-0.235	-0.807	0.479	0.698	0.768	0.705	2.881
NR07/047	-0.270	-1.103	-0.706	-0.666	0.521	0.699	1.001	-0.235	-0.807	0.479	-0.339	0.768	0.705	-0.376
SR07/084	-0.270	0.092	-0.706	-0.666	0.521	0.699	-2.060	-0.235	-0.807	-1.438	-0.339	-0.047	-0.477	-0.376
WR07/007	-0.270	0.092	-0.706	-0.666	0.521	-0.402	-2.060	-0.235	-0.807	-1.438	-0.339	-0.047	-0.477	-0.376
NR07/068	-0.270	0.092	0.314	0.617	0.521	0.699	-0.530	-0.235	0.392	-0.479	-1.376	-1.677	-0.477	-0.376
WR07/004	-0.270	-1.103	-0.706	-0.666	0.521	-0.402	-2.060	-0.235	-0.807	-1.438	-0.339	-0.047	-0.477	-0.376
SR07/081	-0.270	-1.103	-0.706	-0.666	0.521	-0.402	-2.060	-0.235	-0.807	-1.438	-0.339	-1.677	-0.477	-0.376
ER07/036	-0.270	2.483	-0.706	-0.666	-0.502	0.699	-0.530	-0.235	-0.807	0.479	-0.339	-1.677	-0.477	-0.376
WR07/022	-0.270	-1.103	-0.706	-0.666	-2.547	0.699	-0.530	-0.235	-0.807	-1.438	-0.339	-0.047	0.705	-0.376
NR07/054	-0.270	-1.103	-0.706	-0.666	-2.547	0.699	-0.530	-0.235	-0.807	-1.438	-0.339	0.768	-0.477	-0.376
NR07/060	-0.270	1.287	1.333	0.617	-1.013	0.699	-0.530	-1.459	1.592	-1.438	-1.376	0.768	-1.659	-0.376
TDr95/00005	-0.270	1.287	1.333	0.617	0.521	0.699	-0.530	-0.235	1.592	-1.438	-1.376	0.768	-0.477	-0.376
SR07/082	-0.270	-1.103	-0.706	-0.666	0.521	0.699	-0.530	-0.235	-0.807	0.479	0.698	-1.677	0.705	-0.376
ER07/037	-0.270	-1.103	-0.706	-0.666	0.521	0.699	-0.530	-0.235	0.392	0.479	-0.339	-0.047	0.705	-0.376
NR07/071	-0.270	1.287	1.333	0.617	-1.013	0.699	-0.530	-1.459	1.592	-1.438	-1.376	0.768	-1.659	-0.376
SR07/072	-0.270	1.287	1.333	0.617	-1.013	0.699	-0.530	-1.459	1.592	-1.438	-1.376	0.768	-1.659	-0.376
WR 07/001	-0.270	0.092	-0.706	-0.666	-2.547	0.699	-0.530	2.211	-0.807	-1.438	-0.339	-0.047	-0.477	-0.376
ER 07/039	-0.270	0.092	-0.706	-0.666	0.521	0.699	-0.530	-0.235	-0.807	0.479	-0.339	-0.047	-0.477	-0.376
NR 07/057	-0.270	0.092	-0.706	-0.666	0.521	0.699	-0.530	0.988	-0.807	0.479	-0.339	-0.047	-0.477	-0.376
TDr95/18544	-0.270	1.287	1.333	4.466	-1.013	0.699	-0.530	-1.459	1.592	1.438	-1.376	0.768	-1.659	-0.376

Appendix 6. Genstat 12.1 commands for calculation of standardized first 10 principal component scores for 28 morphological traits of 52 yam genotypes

```
  8  DELETE [REDEFINE=yes] variety,std1,std2,std3,std4,std5,std6,std7,std8,std9,\
  9  std10,std11,std12,std13,std_x14,std15,std16,std17,std18,std19,std20,std21,\
 10  std22,std23,std24,std25,std26,std27,std28
 11  UNITS [NVALUES=*]
 12  FACTOR [MODIFY=yes; NVALUES=52; LEVELS=52; LABELS=!t('ER 07/039','ER07/029',\
 13  'ER07/030','ER07/031','ER07/032','ER07/033','ER07/034','ER07/036',\
 14  'ER07/037','ER07/038','NR 07/057','NR07/040','NR07/041','NR07/042',\
 15  'NR07/043','NR07/045','NR07/047','NR07/052','NR07/054','NR07/059',\
 16  'NR07/060','NR07/067','NR07/068','NR07/069','NR07/071','SR07/072',\
 17  'SR07/073','SR07/074','SR07/075','SR07/076','SR07/079','SR07/080',\
 18  'SR07/081','SR07/082','SR07/084','SR07/085','TDr 95/18544','TDr95/00005',\
 19  'WR 07/001','WR07/004','WR07/007','WR07/008','WR07/010','WR07/013',\
 20  'WR07/014','WR07/015','WR07/016','WR07/020','WR07/022','WR07/024',\
 21  'WR07/025','WR07/028'); REFERENCE=1] variety
 22  READ variety; FREPRESENTATION=ordinal
221  %PostMessage 1129; 0; 39494688 "Sheet Update Completed"
 222  DELETE [REDEFINE=yes] _lrv
 223  PCP [PRINT=roots,loadings,scores,tests; NROOTS=10; METHOD=correlation] !p(std_APW,\
 224 std_DBL,std_DE,std_IL,std_LAS,std_LC,std_LD,std_LL,std_LMC,std_LL1,std_LL2,std_LW1,\
 225
std_LW2,std_LS,std_LVCLS,std_LVCUS,std_NB,std_NS,std_PC,std_PL,std_PLM,std_PWC,std_SC,\
 226  std_TC,std_TLM,std_TS,std_WC,std_tfleshc); SAVE=_pcpsave; LRV=_lrv
```

Principal components analysis

Latent roots

1	2	3	4	5	6
7.672	5.025	3.419	2.161	1.636	1.067
7	8	9	10		
0.992	0.869	0.711	0.701		

Percentage variation

1	2	3	4	5	6
27.40	17.95	12.21	7.72	5.84	3.81
7	8	9	10		
3.54	3.10	2.54	2.50		

Trace

28.00

Latent vectors (loadings)

	1	2	3	4	5
_pcpsave['data']					
std1	0.27365	-0.09268	-0.20542	0.21262	0.12731
std2	-0.25140	-0.14901	0.07128	0.30075	0.04279
std3	-0.05508	0.04778	0.25151	0.05951	0.43107
std4	0.13005	0.24820	0.14219	-0.03630	0.32634
std5	0.23854	-0.05474	0.06720	-0.24589	-0.29276
std6	-0.21699	0.27314	-0.03541	0.19127	-0.10223
std7	-0.13785	0.33273	0.00235	-0.18239	-0.00174
std8	0.21406	-0.02388	-0.21318	0.05177	0.10906
std9	-0.26394	0.22649	-0.11044	0.15722	-0.06829
std10	0.18197	0.25842	0.22648	0.10469	0.10530
std11	0.27116	0.16879	0.06144	-0.07138	0.11880
std12	0.12852	0.22587	0.32894	0.00689	-0.19788
std13	0.08975	0.09077	0.42493	0.09272	0.02496

stdx14	-0.13897	-0.16292	0.25929	0.26613	-0.04547
std15	-0.00123	0.28183	-0.29461	0.07821	0.10177
std16	-0.21143	0.22483	-0.07616	0.01863	0.18258
std17	-0.21628	0.21129	0.17199	-0.20465	0.06201
std18	-0.19240	0.26539	0.00511	0.02743	0.02288
std19	0.10996	-0.07546	0.02169	0.44554	-0.00841
std20	-0.07806	-0.04779	0.01426	0.17008	0.33403
std21	0.16624	0.10902	0.31335	0.15989	-0.22905
std22	0.21448	0.15569	-0.17825	0.20824	0.19665
std23	-0.23817	0.07490	0.04754	0.26423	-0.29529
std24	0.17129	0.23698	-0.02457	0.03795	-0.34527
std25	0.27389	0.10843	0.10542	0.17580	0.06545
std26	-0.13107	-0.00235	-0.08220	0.27762	-0.12768
std27	0.23246	0.11871	-0.22487	0.27228	-0.15482
std28	0.01852	0.32754	-0.26577	-0.08612	-0.08980

	6	7	8	9	10
_pcpsave['data']					
std1	0.09949	-0.15238	0.05709	0.04356	-0.07064
std2	-0.03883	0.11307	-0.12532	0.02420	-0.08629
std3	0.08945	0.28375	-0.43544	-0.08405	0.35663
std4	0.25806	0.07182	0.05548	-0.26377	-0.21417
std5	0.00734	0.40121	-0.04379	-0.05135	-0.02849
std6	0.00719	0.20403	-0.03637	0.04261	0.00195
std7	0.13477	0.02120	0.14994	-0.13388	0.08427
std8	0.08605	0.06490	0.12145	0.00718	0.65263
std9	-0.08613	0.18159	-0.06369	0.06392	0.03301
std10	0.07373	-0.01736	-0.04001	-0.03420	-0.06900
std11	0.04482	-0.15392	-0.08797	0.19282	0.08120
std12	0.02211	-0.16942	-0.02470	0.08931	0.21377
std13	-0.11598	-0.01540	0.08218	0.27984	-0.22331
stdx14	0.24770	0.25096	0.14751	-0.20908	-0.17209
std15	-0.07613	-0.16061	-0.04011	0.06273	-0.14780
std16	-0.19164	-0.23622	-0.13806	0.23901	-0.07142
std17	0.12934	-0.19802	0.10873	0.01295	0.01307
std18	0.03216	-0.11563	0.20373	-0.26503	0.10679
std19	0.00974	-0.42711	-0.15529	-0.48743	0.02846
std20	-0.47410	0.16902	0.62308	-0.03757	0.11843
std21	-0.06661	-0.08844	0.22521	0.24444	0.14477
std22	0.00420	0.25885	-0.13230	0.27378	0.00931
std23	-0.19250	0.03876	-0.27403	0.07153	0.15788
std24	-0.23622	0.06042	0.01416	-0.33609	0.18213
std25	-0.15065	0.21139	0.00121	-0.08885	-0.18100
std26	0.62043	0.01553	0.27168	0.25801	0.12593
std27	0.04038	0.03367	0.02879	0.08465	-0.19029
std28	0.08029	0.22576	0.02941	-0.14536	-0.16035

Principal component scores

	1	2	3	4	5
1	0.447	3.584	3.419	-5.421	-1.543
2	-0.493	3.368	2.537	-6.001	-0.377
3	3.226	2.085	3.073	1.252	1.018
4	2.691	0.884	1.089	1.271	-0.035
5	0.547	5.121	-5.191	-1.628	-1.157
6	2.596	0.596	1.417	-0.801	0.431
7	1.919	1.147	1.743	-0.119	0.918
8	-6.512	-0.044	1.417	0.256	0.410
9	0.604	-1.255	-0.398	0.476	1.638
10	0.424	4.281	-1.886	1.270	-0.828
11	3.607	0.663	2.295	1.358	1.181

12	2.996	-0.024	1.774	0.038	0.262
13	0.215	1.170	1.745	0.177	1.635
14	-1.502	-0.243	-0.795	0.097	1.913
15	-0.701	6.456	-3.506	1.744	1.041
16	-1.041	-0.068	1.256	0.443	0.706
17	2.749	1.148	2.548	1.224	1.146
18	-0.507	0.808	0.907	-0.055	1.813
19	-6.532	0.543	1.777	1.799	-0.152
20	3.040	1.605	2.683	1.869	1.449
21	1.076	6.533	-4.474	0.862	0.395
22	3.213	0.367	1.929	0.919	0.228
23	1.200	-0.617	0.896	0.677	-2.464
24	1.080	-1.782	-1.133	0.254	-2.469
25	1.039	-0.586	-0.451	0.485	-1.795
26	1.427	-0.922	-0.249	0.303	-2.195
27	1.117	-0.165	0.932	1.047	-2.123
28	1.703	-0.945	-0.083	0.855	-1.760
29	1.000	-2.180	-1.453	-0.139	-1.806
30	1.275	-0.887	0.992	1.010	-1.036
31	1.503	-1.545	0.388	0.720	-1.707
32	1.380	-1.024	0.420	0.718	-1.455
33	0.902	-0.095	-2.150	0.549	-0.792
34	2.384	-1.106	-0.317	-0.244	-1.053
35	-0.904	-3.000	-1.245	-0.287	1.349
36	0.022	-2.943	-1.529	-0.763	0.775
37	-0.633	-0.763	-1.903	-1.565	0.817
38	-0.132	-2.872	-1.820	-0.444	1.599
39	0.468	-3.191	-1.215	-1.216	0.917
40	0.285	-0.466	-1.590	-1.192	2.021
41	0.930	-2.206	-0.884	-1.367	0.800
42	-0.882	-3.211	-1.848	-1.420	0.207
43	-6.409	-0.133	0.947	0.145	-0.397
44	-4.972	-0.156	0.021	2.009	0.150
45	1.414	-1.765	-0.047	-0.387	-0.082
46	0.809	-2.268	-1.332	0.257	-0.917
47	-6.420	-0.123	0.997	0.157	-0.309
48	-6.409	-0.133	0.947	0.145	-0.397
49	0.512	-1.859	-1.184	-1.197	2.012
50	0.320	-1.698	-1.020	-0.586	0.162
51	0.613	-1.673	-1.454	0.063	0.960
52	-6.682	1.592	1.010	0.386	-1.105

	6	7	8	9	10
1	-0.236	0.043	0.582	0.063	-1.121
2	0.703	0.344	-0.894	-0.963	0.197
3	-0.594	0.344	0.638	0.206	0.479
4	0.359	0.411	1.512	0.046	0.228
5	0.631	1.089	-1.137	1.063	0.884
6	-1.216	-0.744	0.687	-0.526	0.222
7	0.196	0.008	1.104	-0.985	0.677
8	-0.160	0.869	-0.575	0.541	0.493
9	0.416	0.679	0.299	-0.978	0.114
10	1.801	1.984	-0.370	-0.832	-0.678
11	-0.054	0.906	-0.551	0.502	1.203
12	-0.115	0.274	-1.682	-0.094	0.332
13	1.215	-1.815	1.988	0.302	0.171
14	1.539	-1.707	-1.776	0.056	-0.836
15	-0.076	-0.022	0.959	-1.545	-0.637
16	0.608	-1.577	1.670	0.099	-1.432
17	-1.472	0.666	-1.083	-0.214	-1.377

18	4.024	-1.820	-0.268	0.273	0.938
19	-0.475	0.284	-1.530	-0.967	0.877
20	-0.268	0.942	-0.614	0.964	0.528
21	-1.894	-1.406	0.118	1.471	-0.516
22	-0.364	0.867	-2.416	-0.085	0.470
23	-0.396	-0.891	-0.670	0.770	-1.573
24	0.326	-0.449	-0.363	-0.096	0.413
25	0.077	-0.693	-0.677	-0.259	0.042
26	0.071	-1.089	-0.671	0.846	0.498
27	1.704	-1.100	-0.186	1.382	-1.142
28	1.001	-0.044	-0.195	-0.084	-0.134
29	0.467	-0.360	-0.894	-0.553	0.370
30	-0.842	0.219	0.704	-0.096	-1.531
31	-1.018	-0.289	0.942	0.294	0.715
32	-0.279	-0.322	0.819	0.352	1.347
33	0.647	1.669	0.987	-1.326	-0.352
34	0.591	-0.173	1.512	-0.512	0.842
35	0.025	0.118	-0.305	-0.712	-1.921
36	0.340	-0.223	-0.529	-0.049	-0.446
37	-1.916	-2.436	-0.191	-0.216	1.651
38	1.037	-0.574	-0.362	-0.906	-0.283
39	-0.410	0.135	-0.960	-0.785	-0.440
40	-2.366	-1.442	-0.476	0.800	0.176
41	0.126	1.305	0.905	1.797	-0.195
42	0.439	1.940	1.292	1.185	0.313
43	-0.327	0.339	0.239	0.698	-0.174
44	0.045	-0.278	-0.124	0.475	-0.411
45	-1.458	0.185	0.142	-1.309	-1.753
46	-0.546	0.112	0.324	-0.665	0.757
47	-0.309	0.396	0.151	0.681	-0.102
48	-0.327	0.339	0.239	0.698	-0.174
49	0.095	2.053	0.442	2.047	-0.206
50	-0.323	-0.010	0.347	-0.828	1.103
51	-0.197	0.749	0.029	-0.701	0.470
52	-0.842	0.196	0.866	-1.327	0.923

304 LRVSCREE [PRINT=scree; PLOT=*] _lrv

No	Root	%%	Cum	%	Scree Diagram (* represents 1%)
1	7.672	274	274	27	**************************
2	5.025	179	453	18	******************
3	3.419	122	576	12	************
4	2.161	77	653	8	********
5	1.636	58	711	6	******
6	1.067	38	749	4	****
7	0.992	35	785	4	****
8	0.869	31	816	3	***
9	0.711	25	841	3	***
10	0.701	25	866	3	***

Scale: 1 asterisk represents 1 unit.

Appendix 7. Genstat 12.1 commands for calculation of standardized first six factor loadings for 28 morphological traits of 52 yam genotypes

```
228  FCA [PRINT=loadings,scores,tests,communalities; METHOD=correlation; MAXCYCLE=50;
TOLERANCE=1e-006;\
 229   NDIMENSION=6]
!p(std_APW,std_DBL,std_DE,std_IL,std_LAS,std_LC,std_LD,std_LL,std_LMC,\
 230  std_LL1,std_LL2,std_LW1,std_LW2,std_LS,std_LVCLS,std_LVCUS,std_NB,std_NS,std_PC,\
 231  std_PL,std_PLM,std_PWC,std_SC,std_TC,std_TLM,std_TS,std_WC,std_tfleshc)
```

Factor analysis

Warning 1, code OP 8, statement 1 on line 231

Command: FCA [PRINT=loadings,scores,tests,communalities; METHOD=correlation; MAX
The convergence is not certain but a lower point could not be found.

Factor loadings

1	0.7081	0.0959	0.6355	-0.0895	0.0544
2	-0.6071	-0.4599	0.1090	-0.0871	0.4153
3	-0.1707	-0.0351	-0.2377	-0.3440	0.0036
4	0.2065	0.4953	-0.1399	-0.3618	-0.3238
5	0.7589	0.1253	-0.4187	0.4185	0.1088
6	-0.7714	0.4614	-0.0352	0.0539	0.1519
7	-0.5178	0.5270	-0.2606	0.0425	-0.3422
8	0.4938	0.1671	0.3904	0.0571	-0.0900
9	-0.8797	0.3536	0.0550	0.1795	0.1409
10	0.2932	0.6210	-0.2106	-0.4911	0.0313
11	0.5929	0.4898	-0.0479	-0.3006	-0.1985
12	0.2297	0.4974	-0.5539	-0.4343	0.1314
13	0.1835	0.1602	-0.4571	-0.5799	0.2321
14	-0.2516	-0.4091	-0.1940	-0.1737	0.4500
15	-0.2240	0.6135	0.4010	0.0706	-0.2413
16	-0.7061	0.2404	0.0803	-0.1019	-0.2164
17	-0.6407	0.1444	-0.4200	-0.2204	-0.3513
18	-0.6472	0.3306	-0.0796	-0.1323	-0.1660
19	0.2376	-0.0312	0.3676	-0.3925	0.3608
20	-0.2157	-0.1560	0.1515	-0.1395	0.0482
21	0.3677	0.3192	-0.3449	-0.4519	0.3484
22	0.3863	0.5824	0.4338	-0.0107	0.0818
23	-0.7118	0.0528	-0.1052	0.0524	0.5070
24	0.2988	0.6545	-0.1452	0.1123	0.1465
25	0.6008	0.4542	-0.0002	-0.2410	0.2048
26	-0.3410	-0.0226	0.1976	0.0232	0.1983
27	0.4650	0.5780	0.4900	0.0734	0.2476
28	-0.1542	0.7674	0.1247	0.3606	-0.2693

1	0.0116
2	-0.1819
3	-0.4793
4	-0.3424
5	-0.0552
6	-0.0603
7	0.0143
8	-0.0084
9	-0.0545
10	-0.1619
11	-0.0060

12	0.2173
13	-0.1628
14	-0.2334
15	0.1101
16	-0.0011
17	0.0745
18	0.0726
19	0.1730
20	-0.2189
21	0.1847
22	-0.3023
23	0.1731
24	0.2340
25	-0.2553
26	0.0809
27	0.1236
28	0.0013

Factor communalities

1	0.9256
2	0.8051
3	0.4350
4	0.6605
5	0.9570
6	0.8387
7	0.7329
8	0.4355
9	0.9569
10	0.7844
11	0.7235
12	0.8601
13	0.6849
14	0.5554
15	0.6627
16	0.6201
17	0.7854
18	0.5848
19	0.5068
20	0.1635
21	0.7158
22	0.7748
23	0.8103
24	0.6276
25	0.7325
26	0.2023
27	0.8723
28	0.8308

Factor score coefficients

	1	2	3	4	5
1	0.15716	0.05143	0.43775	-0.09278	0.08454
2	-0.05148	-0.09422	0.02868	-0.03449	0.24640
3	-0.00499	-0.00248	-0.02158	-0.04697	0.00074
4	0.01005	0.05824	-0.02113	-0.08220	-0.11025
5	0.29163	0.11629	-0.49930	0.75069	0.29243
6	-0.07902	0.11418	-0.01118	0.02577	0.10886
7	-0.03203	0.07876	-0.05003	0.01229	-0.14809
8	0.01446	0.01181	0.03546	0.00780	-0.01843
9	-0.33752	0.32771	0.06552	0.32144	0.37813

10	0.02247	0.11498	-0.05010	-0.17573	0.01678
11	0.03543	0.07070	-0.00889	-0.08384	-0.08296
12	0.02713	0.14196	-0.20309	-0.23955	0.10861
13	0.00962	0.02029	-0.07439	-0.14197	0.08516
14	-0.00935	-0.03673	-0.02237	-0.03014	0.11699
15	-0.01098	0.07260	0.06096	0.01614	-0.08272
16	-0.03071	0.02525	0.01084	-0.02069	-0.06585
17	-0.04933	0.02686	-0.10035	-0.07922	-0.18924
18	-0.02575	0.03178	-0.00983	-0.02458	-0.04622
19	0.00796	-0.00252	0.03822	-0.06139	0.08456
20	-0.00426	-0.00745	0.00929	-0.01286	0.00666
21	0.02138	0.04483	-0.06222	-0.12264	0.14170
22	0.02834	0.10323	0.09879	-0.00367	0.04198
23	-0.06200	0.01110	-0.02844	0.02130	0.30891
24	0.01326	0.07017	-0.02000	0.02326	0.04549
25	0.03711	0.06777	-0.00003	-0.06950	0.08847
26	-0.00706	-0.00113	0.01270	0.00225	0.02874
27	0.06019	0.18070	0.19679	0.04433	0.22416
28	-0.01506	0.18104	0.03780	0.16440	-0.18395

	6
1	0.03764
2	-0.22501
3	-0.20453
4	-0.24316
5	-0.30950
6	-0.09010
7	0.01292
8	-0.00358
9	-0.30513
10	-0.18107
11	-0.00522
12	0.37463
13	-0.12453
14	-0.12657
15	0.07867
16	-0.00069
17	0.08374
18	0.04216
19	0.08456
20	-0.06310
21	0.15667
22	-0.32363
23	0.22003
24	0.15151
25	-0.23006
26	0.02446
27	0.23348
28	0.00190

Factor analysis test statistics

Log-likelihood: -185.4
Goodness of fit statistic: 267.7
Degrees of freedom: 225
Probability: 0.0

Appendix 8. Genstat 12.1 commands for calculation of cluster analysis of first two principal components for 28 morphological traits of 52 yam genotypes

```
316  GETATTRIBUTE [ATTRIBUTE=rows] distance; _ps
 317  LRV [ROWS=_ps[1]; COLUMNS=2] _lrv
 318  PCO distance; LRV=_lrv
 319  DMST [TITLE='Minimum Spanning Tree'] _lrv[1]; SIMILARITY=distance
 320  DDENDROGRAM [ORDER=given; DSIMILARITY=yes] DATA=_ddmst; PERM=_perm;
WINDOW=1; LABELS=variety
```

Appendix 9. Genstat 12.1 commands and analysis of variance for 28 morphological traits meaned across 52 yam genotypes using regression analysis

```
  8  DELETE [REDEFINE=yes] var,RFI_log,DNA,PLOIDY
  9  UNITS [NVALUES=*]
 10  FACTOR [MODIFY=yes; NVALUES=52; LEVELS=52; LABELS=!t('ER 07/039','ER07/029',\
 11  'ER07/030','ER07/031','ER07/032','ER07/033','ER07/034','ER07/036',\
 12  'ER07/037','ER07/038','NR 07/057','NR07/040','NR07/041','NR07/042',\
 13  'NR07/043','NR07/045','NR07/047','NR07/052','NR07/054','NR07/059',\
 14  'NR07/060','NR07/067','NR07/068','NR07/069','NR07/071','SR07/072',\
 15  'SR07/073','SR07/074','SR07/075','SR07/076','SR07/079','SR07/080',\
 16  'SR07/081','SR07/082','SR07/084','SR07/085','TDr 95/18544','TDr95/00005',\
 17  'WR 07/001','WR07/004','WR07/007','WR07/008','WR07/010','WR07/013',\
 18  'WR07/014','WR07/015','WR07/016','WR07/020','WR07/022','WR07/024',\
 19  'WR07/025','WR07/028'); REFERENCE=1] var
 20  READ var; FREPRESENTATION=ordinal
42  %PostMessage 1129; 0; 42837024 "Sheet Update Completed"
 43  "Simple Linear Regression"
 44  MODEL PLOIDY
 45  TERMS DNA
 46  FIT [PRINT=model,summary,correlations,estimates,accumulated; CONSTANT=estimate; FPROB=yes;\
 47   TPROB=yes] DNA
```

Regression analysis

Response variate: PLOIDY
Fitted terms: Constant, DNA

Summary of analysis

Source	d.f.	s.s.	m.s.	v.r.	F pr.
Regression	1	13.332	13.3324	77.60	<.001
Residual	50	8.591	0.1718		
Total	51	21.923	0.4299		

Percentage variance accounted for 60.0
Standard error of observations is estimated to be 0.415.

Message: the following units have large standardized residuals.

Unit	Response	Residual
23	2.000	-2.96
30	2.000	-2.71

Estimates of parameters

Parameter	estimate	s.e.	t(50)	t pr.
Constant	-6.30	1.17	-5.40	<.001
DNA	5.521	0.627	8.81	<.001

Correlations between parameter estimates

Parameter	ref	correlations	
Constant	1	1.000	
DNA	2	-0.999	1.000
		1	2

Accumulated analysis of variance

Change	d.f.	s.s.	m.s.	v.r.	F pr.
+ DNA	1	13.3324	13.3324	77.60	<.001
Residual	50	8.5906	0.1718		
Total	51	21.9231	0.4299		

Appendix 10. Genstat 12.1 commands and ANOVA in DNA content among three species of yam (*D. alata*, *D. bulbifera* and *D. rotundata*) with 4*x* genotypes

```
8  DELETE [REDEFINE=yes] variety,Species,rep,PLOIDY,DNA
  9  UNITS [NVALUES=*]
  10  FACTOR [MODIFY=yes; NVALUES=135; LEVELS=45; LABELS=!t('ER-29','ER-31',\
  11  'ER-32','ER-33','ER-34','ER-36','ER-37','ER-38','ER-39','NR-41','NR-42',\
  12  'NR-43','NR-45','NR-47','NR-52','NR-54','NR-57','NR-59','NR-60','NR-67',\
  13  'NR-69','NR-71','NR-72','SR-73','SR-74','SR-75','SR-76','SR-79','SR-80',\
  14  'SR-81','SR-82','SR-85','TDr 95/00005','WR-01','WR-04','WR-07','WR-08',\
  15  'WR-10','WR-13','WR-15','WR-20','WR-22','WR-24','WR-25','WR-28')\
  16  ; REFERENCE=1] variety
  17  READ variety; FREPRESENTATION=ordinal
78  "General Analysis of Variance."
  79  BLOCK rep
  80  TREATMENTS Species
  81  COVARIATE "No Covariate"
  82  ANOVA [PRINT=aovtable,information,means,%cv; FACT=32; CONTRASTS=7; PCONTRASTS=7;
FPROB=yes;\
  83   PSE=diff,lsd; LSDLEVEL=5] DNA
```

Analysis of variance

Variate: DNA

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
rep stratum	2	0.502386	0.251193	51.60	
rep.*Units* stratum					
Species	2	0.054730	0.027365	5.62	0.005
Residual	130	0.632886	0.004868		
Total	134	1.190001			

Message: the following units have large residuals.

rep 3 *units* 12	-0.1936	approx. s.e.	0.0685
rep 3 *units* 36	-0.2106	approx. s.e.	0.0685

Tables of means

Variate: DNA

Grand mean 1.8546

Species	1	2	3
	1.8460	1.9050	1.9009
rep.	114	3	18

Standard errors of differences of means

Table	Species	
rep.	unequal	
d.f.	130	
s.e.d.	0.05697X	min.rep
	0.04081	max-min
	0.00924X	max.rep

(No comparisons in categories where s.e.d. marked with an X)

Least significant differences of means (5% level)

Table	Species	
rep.	unequal	
d.f.	130	
l.s.d.	0.11271X	min.rep
	0.08074	max-min
	0.01828X	max.rep

(No comparisons in categories where l.s.d. marked with an X)

Stratum standard errors and coefficients of variation

Variate: DNA

Stratum	d.f.	s.e.	cv%
rep	2	0.07471	4.0
rep.*Units*	130	0.06977	3.8

Appendix 11. Genstat 12.1 commands and ANOVA in DNA content of 45, 4*x* genotypes

```
 84  "General Analysis of Variance."
 85  BLOCK rep
 86  TREATMENTS variety
 87  COVARIATE "No Covariate"
 88  ANOVA [PRINT=aovtable,information,means,%cv; FACT=32; CONTRASTS=7; PCONTRASTS=7; FPROB=yes;\
 89   PSE=diff,lsd; LSDLEVEL=5] DNA
```

Analysis of variance

Variate: DNA

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
rep stratum	2	0.502386	0.251193	57.41	
rep.*Units* stratum					
variety	44	0.302567	0.006877	1.57	0.037
Residual	88	0.385048	0.004376		
Total	134	1.190001			

Tables of means

Variate: DNA

Grand mean 1.8546

variety	ER-29	ER-31	ER-32	ER-33	ER-34
	1.8300	1.8880	1.8900	1.8480	1.8380
variety	ER-36	ER-37	ER-38	ER-39	NR-41
	1.8400	1.9270	1.8000	1.8860	1.8200
variety	NR-42	NR-43	NR-45	NR-47	NR-52
	1.9180	1.8340	1.9050	1.8733	1.8557
variety	NR-54	NR-57	NR-59	NR-60	NR-67
	1.7720	1.8590	1.8320	1.9380	1.8260
variety	NR-69	NR-71	NR-72	SR-73	SR-74
	1.8270	1.8800	1.9230	1.8970	1.8720
variety	SR-75	SR-76	SR-79	SR-80	SR-81
	1.7980	1.8970	1.8717	1.7880	1.7810
variety	SR-82	SR-85	TDr 95/00005	WR-01	WR-04
	1.7900	1.8310	1.9370	1.8527	1.8730
variety	WR-07	WR-08	WR-10	WR-13	WR-15
	1.8740	1.8990	1.8200	1.8090	1.9280
variety	WR-20	WR-22	WR-24	WR-25	WR-28
	1.8420	1.7760	1.7790	1.9210	1.8120

Standard errors of differences of means

Table	variety
rep.	3
d.f.	88
s.e.d.	0.05401

Least significant differences of means (5% level)

Table	variety
rep.	3
d.f.	88
l.s.d.	0.10733

Stratum standard errors and coefficients of variation

Variate: DNA

Stratum	d.f.	s.e.	cv%
rep	2	0.07471	4.0

rep.*Units*

Appendix 12. Correlation matrix of the two cytological and three morphological trait means across 52 yam genotypes

```
  8  DELETE [REDEFINE=yes] variety,DNA,Ploidy,APW,NS,WC
  9  UNITS [NVALUES=*]
 10  FACTOR [MODIFY=yes; NVALUES=52; LEVELS=52; LABELS=!t('ER 07/029',\
 11  'ER 07/030','ER 07/031','ER 07/032','ER 07/033','ER 07/034','ER 07/036',\
 12  'ER 07/037','ER 07/038','ER 07/039','NR 07/040','NR 07/041','NR 07/042',\
 13  'NR 07/043','NR 07/045','NR 07/047','NR 07/052','NR 07/054','NR 07/057',\
 14  'NR 07/059','NR 07/060','NR 07/067','NR 07/068','NR 07/069','NR 07/071',\
 15  'SR 07/072','SR 07/073','SR 07/074','SR 07/075','SR 07/076','SR 07/079',\
 16  'SR 07/080','SR 07/081','SR 07/082','SR 07/084','SR 07/085','TDr 95/00005',\
 17  'TDr 95/18544','WR 07/001','WR 07/004','WR 07/007','WR 07/008','WR 07/010',\
 18  'WR 07/013','WR 07/014','WR 07/015','WR 07/016','WR 07/020','WR 07/022',\
 19  'WR 07/024','WR 07/025','WR 07/028'); REFERENCE=1] variety
 20  READ variety; FREPRESENTATION=ordinal

54  %PostMessage 1129; 0; 52601888 "Sheet Update Completed"
 55  FCORRELATION [PRINT=correlations,test; METHOD=twosided] DNA,PLOIDY,APW,NS,WC
Correlations

                DNA  1          -
             PLOIDY  2     0.7798          -
                APW  3    -0.3928    -0.2715          -
                 NS  4     0.2882     0.3468    -0.4727          -
                 WC  5    -0.3578    -0.2895     0.7143    -0.1755          -
                          1          2          3          4          5
Number of observations: 52
Two-sided test of correlations different from zero

                DNA  1          -
             PLOIDY  2    <0.001          -
                APW  3    0.0040     0.0515          -
                 NS  4    0.0383     0.0118    <0.001          -
                 WC  5    0.0092     0.0374    <0.001     0.2133          -
                          1          2          3          4          5
```

Appendix 13. Summary statistics of the two cytological and three morphological traits using canonical correlation analysis

```
203  CANCOR [PRINT=pcoeff,pscores,qcoeff,qscores,correlations]
!p(stdDNA,stdPLODY);!p(stdAPW,\
 204  stdNS,stdWC)
```

Canonical correlation analysis

Canonical correlations

	CA_Corrs	%Corrs	Cum%Corrs
1	0.4441	60.91	60.91
2	0.2850	39.09	100.00

Loadings for the P-set of variates

	1	2
stdDNA	0.0890	-0.2052
stdPLODY	0.0591	0.2157

Loadings for the Q-set of variates

	1	2
stdAPW	-0.01590	0.19713
stdNS	0.08153	0.13435
stdWC	-0.08452	-0.07874

Appendix 14. Summary statistics of the three morphological and two cytological traits using multiple regression analysis

```
187  "Multiple Linear Regression"
 188  MODEL APW
 189  TERMS [FACT=9] DNA,PLOIDY
 190  FIT [PRINT=model,summary,correlations,estimates,accumulated; CONSTANT=estimate;
FPROB=yes;\
 191   TPROB=yes; FACT=9] DNA,PLOIDY
```

Regression analysis

Response variate: APW
Fitted terms: Constant, DNA, PLOIDY

Summary of analysis

Source	d.f.	s.s.	m.s.	v.r.	F pr.
Regression	2	1.066	0.5328	4.58	0.015
Residual	49	5.704	0.1164		
Total	51	6.769	0.1327		

Percentage variance accounted for 12.3
Standard error of observations is estimated to be 0.341.

Message: the following units have large standardized residuals.

Unit	Response	Residual
8	0.000	-2.44
19	0.000	-2.66
47	0.000	-2.39

Message: the residuals do not appear to be random; for example, fitted values in the range 0.899 to 0.995 are consistently smaller than observed values and fitted values in the range 1.004 to 1.240 are consistently larger than observed values.

Message: the error variance does not appear to be constant; large responses are less variable than small responses.

Message: the following units have high leverage.

Unit	Response	Leverage
11	1.000	0.212
23	1.000	0.229
30	1.000	0.213
35	1.000	0.211
52	0.000	0.216

Estimates of parameters

Parameter	estimate	s.e.	t(49)	t pr.
Constant	4.03	1.21	3.34	0.002
DNA	-1.818	0.824	-2.21	0.032
PLOIDY	0.049	0.116	0.42	0.673

Correlations between parameter estimates

Parameter	ref	correlations		
Constant	1	1.000		
DNA	2	-0.970	1.000	
PLOIDY	3	0.607	-0.780	1.000
	1	2	3	

Accumulated analysis of variance

Change	d.f.	s.s.	m.s.	v.r.	F pr.
+ DNA	1	1.0446	1.0446	8.97	0.004
+ PLOIDY	1	0.0210	0.0210	0.18	0.673
Residual	49	5.7036	0.1164		
Total	51	6.7692	0.1327		

```
 192  "Multiple Linear Regression"
 193  MODEL NS
 194  TERMS [FACT=9] DNA,PLOIDY
 195  FIT [PRINT=model,summary,correlations,estimates,accumulated; CONSTANT=estimate;
FPROB=yes;\
 196   TPROB=yes; FACT=9] DNA,PLOIDY
```

Regression analysis

Response variate: NS
Fitted terms: Constant, DNA, PLOIDY

Summary of analysis

Source	d.f.	s.s.	m.s.	v.r.	F pr.
Regression	2	3.75	1.8756	3.38	0.042
Residual	49	27.23	0.5557		
Total	51	30.98	0.6075		

Percentage variance accounted for 8.5
Standard error of observations is estimated to be 0.745.

Message: the following units have large standardized residuals.

Unit	Response	Residual
15	4.000	3.36
52	5.000	3.98

Message: the following units have high leverage.

Unit	Response	Leverage
11	1.000	0.212
23	1.000	0.229
30	1.000	0.213
35	1.000	0.211
52	5.000	0.216

Estimates of parameters

Parameter	estimate	s.e.	t(49)	t pr.
Constant	-0.66	2.64	-0.25	0.805
DNA	0.38	1.80	0.21	0.833
PLOIDY	0.370	0.254	1.46	0.152

Correlations between parameter estimates

Parameter	ref	correlations		
Constant	1	1.000		
DNA	2	-0.970	1.000	
PLOIDY	3	0.607	-0.780	1.000
		1	2	3

Accumulated analysis of variance

Change	d.f.	s.s.	m.s.	v.r.	F pr.
+ DNA	1	2.5734	2.5734	4.63	0.036
+ PLOIDY	1	1.1779	1.1779	2.12	0.152
Residual	49	27.2295	0.5557		
Total	51	30.9808	0.6075		

```
 197  "Multiple Linear Regression"
 198  MODEL WC
 199  TERMS [FACT=9] DNA,PLOIDY
 200  FIT [PRINT=model,summary,correlations,estimates,accumulated; CONSTANT=estimate;
FPROB=yes;\
 201    TPROB=yes; FACT=9] DNA,PLOIDY
```

Regression analysis

Response variate: WC
Fitted terms: Constant, DNA, PLOIDY

Summary of analysis

Source	d.f.	s.s.	m.s.	v.r.	F pr.
Regression	2	4.68	2.3425	3.61	0.035
Residual	49	31.83	0.6497		
Total	51	36.52	0.7161		

Percentage variance accounted for 9.3
Standard error of observations is estimated to be 0.806.

Message: the following units have high leverage.

Unit	Response	Leverage
11	2.000	0.212
23	2.000	0.229
30	2.000	0.213
35	1.000	0.211
52	0.000	0.216

Estimates of parameters

Parameter	estimate	s.e.	t(49)	t pr.
Constant	7.26	2.85	2.54	0.014
DNA	-3.08	1.95	-1.58	0.120
PLOIDY	-0.035	0.275	-0.13	0.900

Correlations between parameter estimates

Parameter	ref	correlations		
Constant	1	1.000		
DNA	2	-0.970	1.000	
PLOIDY	3	0.607	-0.780	1.000
		1	2	3

Accumulated analysis of variance

Change	d.f.	s.s.	m.s.	v.r.	F pr.
+ DNA	1	4.6746	4.6746	7.20	0.010
+ PLOIDY	1	0.0103	0.0103	0.02	0.900
Residual	49	31.8343	0.6497		
Total	51	36.5192	0.7161		