KWAME NKRUMAH UNIVERSITY OF SCIENCE AND TECHNOLOGY
KUMASI, GHANA.

ETHNO-BOTANY, MORPHOLOGICAL AND MOLECULAR CHARACTERIZATION OF THE *PONA* YAMS IN THE *DIOSCOREA ROTUNDATA-CAYENENSIS* (POIR) COMPLEX IN GHANA.

A THESIS SUBMITTED TO THE DEPARTMENT OF CROP AND SOIL SCIENCES, COLLEGE OF AGRICULTURE AND NATURAL RESOURCES, KWAME NKRUMAH UNIVERSITY OF SCIENCE AND TECHNOLOGY, KUMASI, GHANA IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY IN PLANT BREEDING AND GENETICS

EMMANUEL OTOO

APRIL 2008

DECLARATION

I hereby declare that, except for references cited in relation to other peoples' work for which due acknowledgement has been made, this submission is my own work towards the PhD degree and that, to the best of my knowledge, it neither contains material previously published by another person nor material which has been accepted for the award of any other degree elsewhere.

.....

EMMANUEL OTOO

.....

DR. RICHARD AKROMAH (PRINCIPAL SUPERVISOR)

.....

DR. M. KOLESNISKOVA-ALLEN (CO-SUPERVISOR)

ACKNOWLEDGEMENTS

I extend my heartfelt gratitude to S.S. Alhassan, Sulley Bin Al-Hassan, Imoro Alhassan, Alhassan A. Mumuni, Dahamani Nantogmah, Samuel Tandoh and Kobina Yankson, Agricultural Extension Officers who facilitated the collection of the *Pona* complex accessions.

Technical assistance by Messrs Yaw Danso and Thomas Appiah-Danquah, Assistant Research Scientists (ARS) of the Root and Tuber Division of the Crops Research Institute cannot go without mention. The team is similarly indebted to Miss Habibah Aggrey and Mr. Obeng Asamoah, Technical Assistant and Chief Technical Officer respectively, of Root and Tuber Division for their assistance in data collection.

Misses Ruth Thompson and Priscilla Boateng, ARS of the Biotechnology Division of the Crops Research Institute, deserve special mention for the assistance during the DNA extraction stage in Ghana.

I am also heavily indebted to the laboratory assistants of the IITA Biotechnology Unit in 2007 who assisted in diverse ways to ensure the completion of the laboratory work in time.

The financial assistance of IFAD/WECARD/YAM Project is highly appreciated and so is the partial sponsorship provided by my co-supervisor, Dr. Maria Kolesnikova-Allen in facilitating my laboratory work at IITA.

Finally to my family for their forbearances during countless times that I was not around.

ABSTRACT

Ethno-botanical, morphological and molecular characterization approaches were used to determine the genetic relationship among 91 accessions assembled as *Pona* in the *Dioscorea rotundata-cayenensis* complex in Ghana. Ethno-botanical classification categorized the *Pona* complex into four groups: true *pona, kulunku, laribako-la* and *laribako-nya.* Ordination analysis using morphological data showed a wide dispersion among the accessions. Clustering using one hundred and eleven (111) characters revealed 5 morphotypes in the *Pona* complex and positively identified the check variety *Dente.* They are true *Pona, Laribako, Muchumudu, Kulunku, Fuseini,* and *Numbo.* Subsequent study of DNA variation among accessions using 13 polymorphic Simple Sequence Repeats (SSRs) revealed 3 morph-types and two sets of hybrids as well as the check, *Dente.* Combination of ethno-botany, morphological and molecular analysis besides identifying the check *Dente* and 'CRI-Kukrupa' also revealed 6 morphotypes of the *Pona* complex as authentic *Pona, Laribako, Muchumudu Kulunku, Fuseini and Numbo.*

TABLE OF CONTENTS

PAGE

LIST OF TABLES

PAGE

LIST OF FIGURES **PAGE**

LIST OF APPENDICES **PAGE**

LIST OF ABBREVIATIONS

AFLP – Amplified Fragment Length Polymorphism

CA – Correspondence Analysis

DNA – Deoxyribonucleic Acid

FAO – Food and Agriculture Organization

IITA – International Institute of Tropical Agriculture

NMDS – Non-metric Multi-Dimensionality

PCA – Principal Component Analysis

PCoA – Principal Coordinates Analysis

PCR – Polymerase Chain Reaction

RAPD – Random amplified polymorphic DNA

RFLP – Restricted Fragment Length Polymorphism

SSR – Simple Sequence Repeats

STMS – Sequence-Tagged Microsatellites

STR – Short Tandem Repeats

CHAPTER 1

1.0 INTRODUCTION

Yam (*Dioscorea* spp.) is a tuber crop, a polyploid and multi-species. The cultivated species are generally vegetatively propagated and the wild species sexually propagated. Ninety-six per cent (96%) of the world yam production occurs in the yam belt of West and Central Africa. Of this figure Nigeria (67.0%), Cote d'Ivoire (12.0%), Ghana (6.5%) and Benin (4.5%) account for 90% of the production (FAO, 2007). Yam is a staple food crop of over 300 million people in tropics and subtropics. It is the food of choice at festive occasions, and the only crop that is celebrated.

Yam is grown for their starchy tubers (both cultivated and wild). It contains aromatic substances that act as appetizers. It is also known to have medicinal properties and can improve fecundity. Yam's good potassium-sodium balance ensures protection against osteoporosis and heart disease. Yam products have a lower glycemic index than potato products; protecting against obesity and diabetes. The cooked yam has about 2% protein which is equal to that of potatoes but two times that of cassava.

Yam however is a crop that has suffered a lot from institutional neglect. Yam producing countries of the world are in the tropics and have poor economies that support little or inadequate agricultural research (IITA, 1995). Yams thus belong to a group of crops labeled "orphaned crops", which have not received research attention for a long period of time and very little improvement, has been made to the crop (Otoo, 2007). In Ghana, until May 2005, there had not been any formal release of any yam variety. Farmers grow what they perceive to be "good", selecting against "bad" crop, thus leading to the depletion of the genetic base of the crop. There are as many yam varieties as there are ethnic groups which care to name them (Otoo, 2001; Asemota *et al.*, 1996; Dansi *et al.*, 1999). Again, breeding and selection of yam cultivars with novel

or improved characteristics currently suffers from the fact that traditional cultivars have not been adequately characterized (Asiedu *et al.*, 1998). This makes reference to varieties ambiguous, unreliable and impossible to determine the true genetic variation in yams. Some elite cultivars of the *D. rotundata-cayenensis* complex, such as *Pona* and *Laribako* have sweet, floury and fragrant tuber flesh and are the most preferred cultivars by consumers in Ghana, yet there is no easy means of identifying such yams morphologically. This seriously affects the reliable identification of cultivars for germplasm management and improvement. Folklore from the Guinea Savannah agroecology surrounding the name *Laribako* however suggests that it is a "feminine" or "smaller" type of *Pona*.

Knowledge of genetic variation and relationships between accessions or genotypes is important to understand the genetic variability available and its potential use in breeding programs, estimate any possible loss of genetic diversity, offer evidence of the evolutionary forces shaping the genotypic diversities, and to choose the genotypes to be given priority for conservation (Thormann *et al.*, 1994). Systems of classification and identification based on morphological characters (Dansi *et al.*, 1998, 1999, 2000), soluble tuber protein profiles (Ikediobi and Igboanusi, 1983) or isozyme patterns (Dansi *et al.*, 2000; Mignouna *et al.*, 2003a) have been used to characterize yam germplasm.

Characterization of genetic resource collections has thus been greatly facilitated by the availability of a number of molecular marker systems (Beyene *et al.*, 2005). Morphological traits were among the first markers used in germplasm management, but they have a number of limitations, including low polymorphism, late expression, and vulnerability to environmental influences (Smith and Smith, 1992). Yam genotypes classified in the same cultivar group based on morphology were often genetically different, emphasizing the need for molecular fingerprinting in yam germplasm

characterization (Mignouna *et al.*, 1998). On the other hand, DNA markers, though expensive to detect, do not have such limitations. They can be used to detect variation and have proven to be effective tools for distinguishing between closely related genotypes. Different types of molecular markers have been used to assess the genetic diversity in crop species, but no single technique is universally ideal (Mignouna *et al.*, 2003a). Therefore, the choice of the technique depends on the objective of the study, financial constraints, skills and facilities available. However, recent work done by Mignouna *et al.* (2003b) indicated that amplified fragment length polymorphisms (AFLPs) showed the highest efficiency in detecting polymorphism and revealed genetic relationships that most closely reflected morphological classification and could be done at a reasonable cost. AFLP is a multilocus marker technique developed by Vos *et al.* (1995). AFLP markers are genomic fragments detected after selective PCR amplification which provides a number of appealing features in the fingerprinting of genomes of different complexity (Vos *et al.*, 1995). This technique has been used to identify markers for disease resistance loci (Becker *et al.*, 1995) to fingerprint DNAs (Vos *et al.*, 1995) and to assess relationships between molecular polymorphism and hybrid performance in maize (Ajmone-Marsan *et al.*, 1998) and germplasm characterization in yams (Mignouna *et al.*, 2003c). It has high multiplex ratio, offering it a unique advantage when genome coverage is a major issue due to the presence of linkage disequilibrium, such as in inbred lines and breeding materials (Pejic *et al.*, 1998). It can also be automated, making it more preferable to RFLP.

Generally, AFLPs (Vos *et al.*, 1995) and microsatellites, or simple sequence repeats (SSRs) (Tautz, 1989) are the most frequently used molecular markers in the analysis of genetic resources, because they can be automated and so have great potential in large-scale genetic diversity studies. Simple Sequence Repeats (SSRs), also known as microsatellites or short tandem repeats (STRs), have proven to be a most

important resource for polymorphic markers and have been used very successfully in both mammalian and plant genome mapping studies (Pejic *et al.*, 1998).

SSR refers to the observed interspersion of eukaryotic genomes with tandem repeats of short DNA sequences. SSR polymorphisms have been extensively used as genetic markers in mammals (Tautz, 1989); they occur frequently also in plant genome, showing an extensive variation in different individuals and accessions (Akkaya *et al.*, 1992; Pejic *et al.*, 1998). SSR and AFLP profiling technologies are good candidates for replacing RFLP markers in genetic similarity estimates and variety description, and that they have comparable accuracy in grouping inbred lines selected by pedigree (Pejic *et al.*, 1998). They are generally much simpler to apply and more sensitive to morphological and biochemical methods or the RFLP-based fingerprinting techniques; yet they provide results correlated with those from RFLP analyses (Pejic *et al.*, 1998). SSR loci are co-dominant markers and are more informative than RAPDs and RFLPs (Antonio *et al.*, 2004). The major strength of microsatellites therefore lies in the expected high polymorphism, co-dominant inheritance, high abundance and an even distribution across the genome (Mignouna *et al.*, 2003c).

The chromosomal locations of SSR markers are frequently known, thus providing additional information in genetic diversity studies. Powell *et al.*, (1996) examined the utility of AFLP, RAPD, and SSR markers for soybean germplasm analysis by evaluating information content (expected heterozygosity), number of loci simultaneously analyzed per experiment (multiplex ratio) and effectiveness in assessing relationships between accessions. In their study SSR markers had the highest expected heterozygosity, while AFLP markers had the highest effective multiplex ratio.

The *Pona* yams of Ghana are a class of yam that belongs to the *Dioscorea rotundata-cayenensis* complex. The authentic *Pona* has unique rheological properties including aroma and taste. These yams are the choicest on both local and foreign

markets. The problem facing consumers and researchers is that the authentic *Pona* is not easily discernible by morphological characters. The authentic *Pona* is only known after cooking because of its peculiar aroma and taste. They are given the premium value wherever they are sold. Market women and farmers have different opinions about the true *Pona*, hence they tend to deceive buyers with any *D. rotundata* variety as the *Pona*, in order to sell at higher prices. The problem is that both sellers and consumers are not very certain of the true nature of the *Pona* yam.

If Ghana is to maintain her position as the leading exporter of the crop, it must ensure purity of its varieties that will instill confidence in the markets. Furthermore, the informal nature of the yam trade and exchange of planting materials among farmers, have led to duplication of planting materials whose ethnic or local names have changed in different localities. Thus, the same material may be called differently in another region.

The general objective of this study therefore was to use ethno-botany, morphological and molecular techniques (SSRs) to determine the true *Pona*. The specific objectives of the study were to:

(i) Investigate genetic diversity and relationships among 91 supposed *Pona* yam accessions using ethno-botany, morphological, and SSR markers.

(ii) Assess the correlation between genetic distance estimates of *Pona* complex based on morphological traits and molecular markers.

(iii) Identify the authentic *Pona* using morphological, molecular, rheological data and farmers' opinion.

(iv) Identify distinct genotypes and eliminate obvious duplicates from the germplasm, and

(vi) Select core collection of *Pona* accessions for conservation and future work.

CHAPTER 2

2.0. REVIEW OF LITERATURE.

2.1. Yam Diversity

Yams (Dioscorea *spp.*) constitute an economically important staple food for millions of people in the tropics and subtropics (Mignouna *et al.*, 2003b). Its production has therefore been important to the survival and welfare of many generations of people in the tropics and continues to be very important for ensuring sustainable food security and income generation.

Botanically, yam belongs to the genus *Dioscorea* in the family Dioscoreaceae *and Order Dioscoreales.* The family Dioscoreaceae now contain four distinct genera, *Dioscorea, Stenomeris, Tacca* (previously in *Taccaceae*), and *Trichopus* (Caddick *et al.*, 2002). And the dioecious Dioscoreaceae genera, *Borderea, Epipetrum, Nanarepenta, Rajania, Tamus,* and *Testudinaria,* are nested within *Dioscorea* in phylogenetic analyses (Caddick *et al.*, 2002). *Dioscorea is an economically important, annual or* perennial tuber-bearing, dioecious, climbing, tropical genus of monocots that looks like a dicots. The occurrence of a vestigial element of a second cotyledon in a number of species, the shape and venation of the leaves, and the nature of the inflorescence provide strong evidence that Dioscorea spp. or their ancestors are closely related to the dicotyledons. Again, yams exhibit distinct changes in shoot apical meristem (SAM) structure and phyllotaxy during phase transition from juvenile to adult, also a feature of dicots.

In West Africa, there exists a species complex, *D. rotundata* Poir. and *D. cayenensis* Lam. Phenotypically *D. rotundata* (white-fleshed tubers and 6-8 months growth period), and *D. cayenensis* (yellow-fleshed tubers and 8-12 months growth period) can easily be distinguished from each other, however, there exist many intermediate forms. Domestication is on-going (Vernier, 2003). There is an intensive

vegetative propagation of the collected plants from forest areas (which could be wild or inter-specific hybrids). Over time, tubers with new genetic characteristics emerge resulting in potential polyclonality of a given variety even at field level. Since there has been institutional neglect, and the genetics and genomics of this crop is least understood due to several biological constraints and research negligence. Several grey areas exist with respect to identification of the varieties.

Dioscorea rotundata-cayenensis complexes are indigenous to West Africa and have been cultivated from time immemorial in Ghana with little improvement to the crop from the formal sector (Otoo, 2007). Of about 600 known yam species, only about six are important as staples in the tropics, namely: *D. rotundata, D. cayenensis, D. alata, D. dumetorium, D. bulbifera* and *D.* esculenta (Coursey, 1969). Together they account for 90% of all yams grown in the tropics (Hahn *et al.*, 1987). All the edible yams species mentioned above are grown in Ghana (Akwaag *et al.*, 1996). Even though farmers have been selecting genotypes they perceive to be "good" and selected against "bad" genotypes (Otoo, 2001b), there is still a great deal of diversity within species and such diversity could easily be seen on individual farmer fields or farming village. For instance, in a study conducted by Akwaag *et al.*, 1996 in Brong Ahafo Region of Ghana, 26 yam cultivars were documented as being grown in just four villages.

Yam is also an open-pollinated crop and flowers profusely especially in the wild form. The fruits are allowed to fall freely and may germinate to produce new yam varieties which may be harvested later by farmers. This will add to the diversity within species existing within a farming setup.

2.2 Genetic Diversity Studies

Genetic diversity studies and analysis involve three basic steps: (i). description of variation within and between populations, regions, or area; (ii). assessment of

relationships individuals, populations, regions, area; and (iii). expression of relationships between results obtained from different sets of characters.

Describing the diversity may be done within a population or between populations and may also extend to larger units such as areas and regions (Hoogendijk and Williams, 2001). Calculating the distances (geometric or genetic) among all pairs of subjects of study can be used to assess the relationships between individuals, populations, regions or areas. Using any classification and /or ordination methods, the relationships among individuals, regions or areas can then be established. The results of genetic diversity studies using different approaches can be used to propose scientifically based recommendations for strategies and future actions that best promote the conservation and use of unique crop genetic resources (Hoogendijk and Williams, 2001).

2.2.1 Describing Diversity

There are four methods for measuring genetic diversity, namely farmers' perceptions and folk classification, morphological characterization, biochemical characterization and molecular characterization (Hoogendijk and Williams, 2001).

2.2.1.1 Farmers' perceptions and Ethno-botany Classification

An ethno-botany classification is a vernacular naming system, and can be contrasted with scientific taxonomy. Names are often derived from particular characteristics of the material, such as distinctive morphological or agronomic traits, place of origin or special uses (Hoogendijk and Williams, 2001). These names may also form part of informal but structured classification schemes or 'ethno-botany'/ 'folk taxonomy' reflecting or providing some insight into what may be relationships among varieties.

Folk taxonomies are therefore generated from social knowledge and are used in everyday language. They are distinguished from scientific taxonomies that claim to be disembedded from social relations and thus objective and universal. Anthropologists have observed that taxonomies are generally embedded in local cultural and social systems, and serve various social functions (Mekbib, 2007).

Folk taxonomy is not only the historical root of modern biological classification but it is also crucially important to modern research scientists, who often rely on traditional knowledge when investigating native species (Rosenberger, 2006). This is especially true as they examine biodiversity in complex tropical environments, where local people are apt to recognize a vast number of organisms (Rosenberger, 2006).

It is important to explore and document farmers' perceptions of diversity because it provides an entry point into the body of detailed knowledge held by farmers, who probably know more than anyone else, about the material to be studied (Hoogendijk and Williams, 2001). Other benefits identified by Hoogendijk and Williams (2001) to be derived by using folk taxonomy are:

1. Documentation of the farmer's perception of a local variety because that is the only unit of diversity that the farmer recognize, manage and conserve;
2. Providing clues for genetic affinities and distance;
3. Revealing with respect to uses, cultural value, nutritional qualities, and culinary importance of the different varieties; and
4. It also provides valuable indications for selecting an appropriate, more in-depth characterization method.

However, the following weaknesses can however been cited for folk taxonomy:

a) The use of different names for genetically similar material or the use of same name for genetically different material, and

b) Repeatability of research results due to inconsistencies in data collection and analysis.

2.2.1.2. Morphological characterization

The use of molecular techniques for analysis of genetic diversity and the structure of germplasm, the first conceptual step in marker-assisted breeding, has been fruitful for many species (Naylor *et al.*, 2007). However, before the advent of molecular markers morphological descriptors were employed for germplasm management. Morphological descriptions can provide unique identification of cultivated varieties. The application of morphological descriptor lists is the simplest of the formal, standardized, repeatable methods of measuring crop genetic diversity (Hoogendijk and Williams, 2001). Morphological traits were among the earliest markers used in germplasm management, but they have a number of limitations, including low polymorphism, low heritability, late expression, and vulnerability to environmental influences (Smith and Smith, 1992).

The main advantages of conducting morphological characterization are that published descriptor lists are readily available for most major crop species, it can be carried out *in situ* (on-farm), it is relatively inexpensive, and it is relatively easy to carry out (Hoogendijk and Williams, 2001).

Morphological characterization is therefore a highly recommended first step that should be made before more in-depth biochemical or molecular studies are attempted. Principal Components Analysis (PCA) of the characterization results can identify a few key or 'minimum' descriptors that effectively account for the majority of the diversity observed, saving time and effort for future characterization efforts. This

approach has been used successfully for characterizing sapote (*Pouteria sapota*) in Guatemala and Cuba (Hoogendijk and Williams, 2001). 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 (Hoogendijk and Williams, 2001).

Smith and Smith (1992) identified the following weaknesses as some of the limiting factors hindering the use of morphological characterization for diversity studies:

a) Difficulty in taking environmental influences into account in the case of quantitative characters;

b) Unavailability of descriptor lists for many neglected and underutilized crops; and

c) For most morphological traits, the genetic control is unknown, although it is known that multiple genotypes can give phenotypes of similar outward appearance.

2.2.1.3. Molecular Characterization

Molecular genetic markers have developed into powerful tools for analyzing genetic relationships and genetic diversity (Tatineni *et al.*, 1996). They can detect variation directly at the DNA level. Molecular marker data therefore can assist in taxonomic evaluation, particularly the accurate identification of germplasm (REF????). These methods are therefore being adopted rapidly by crop improvement researchers globally as an effective and appropriate tool for basic and applied studies addressing biological components in agricultural production systems (Hash and Bramel-Cox, 2000). There are several groups of techniques currently available. The most frequently

employed include RFLP, RAPD, AFLP, STMS (sequence-tagged microsatellites), and sequencing. Each technique has its own particular advantages and drawbacks in terms of their applicability for different research objectives. When choosing an appropriate technique, there are aspects that should be taken into account, such as the degree of comparability between experiments, the cost and availability of reagents and equipment, the availability of crop-specific protocols and technical expertise (Hoogendijk and Williams, 2001).

Depending on the research questions, another consideration may be whether it is of importance to detect co-dominance. Some techniques, such as RAPD and AFLP, do not detect co-dominance and therefore cannot measure allelic frequency. While AFLP and RAPD are both suitable for some diversity studies, it should be noted that the results obtained with RAPD are not always comparable between laboratories and sometimes even between experiments. RFLP is an excellent non-random technique, but rather expensive. Microsatellites, also known as SSRs, require specific primers that can be costly, and are not yet available for many species, particularly neglected and underutilized crops (Scotti *et al.*, 1999). Gene sequencing is a wonderfully detailed characterization of the DNA molecule, base pair by base pair, but the excessive investment of time and resources required to conduct such work remains far beyond the capacity of most national programmes. Generally, the molecular markers can be grouped into two: protein and DNA markers.

2.2.1.3.1 Protein markers

Protein markers were among the first group of molecular markers exploited for genetic diversity assessment and genetic linkage map development (Harsh and Bramel-Cox, 2000). They are the basis for a newly emerging research area called proteomics. They also provide some of the most cost-effective tools for data point generation;

especially when iso-electric focussing equipment is used to precisely distinguish between very similar versions of proteins. Harsh and Bramel-Cox (2000) listed the following major limitations of these markers as:

a) Much of the genome (including much of the polymorphic portions of it that are less subject to evolutionary restrictions) does not code for genes,

b) Different biochemical procedures are required to visualize allelic differences for enzymes having different functions, and

c) Many several post-transcriptional steps removed from underlying DNA sequence polymorphism and thus can mask variation present at that level. For example, differences in tri-nucleotide sequences coding for the same amino acid, intron sequences that are post-transcriptionally removed from the mRNA, and post-translational modification can all contribute to reduced polymorphism expression at the protein level compared to that at the DNA level.

2.2.1.3.2. Biochemical characterization

This technique involves separation of functionally equivalent enzyme molecules according to their differing electrostatic charges, sizes, and molecular conformations, followed by their staining (Gillet, 1999). Inheritance analysis of the resulting banding patterns enables inference of their mode of inheritance and, consequently, allows them to be used as genetic markers (REF???). Biochemical characterization, most frequently, involves conducting gel electrophoresis on easily extracted proteins, such as isozymes, seed storage proteins, flavonoids, and others (Hoogendijk and Williams, 2001).

Isoenzymes, the electrophoretically separable variants of one enzyme system (Bergmann *et al.*, 1989), are coded by genes at one or often several loci. Variants that are coded by alleles at the same locus are called allozymes (REF???). Multilocus

analysis considers the results for various loci belonging to one or, more commonly, a large number of enzyme systems. In fact, isoenzymes are still widely used as genetic markers because they are inexpensive compared to DNA markers, the laboratory protocols are well-established in numerous tree species, they are products of structural genes whose roles in metabolism are known in most cases, and, most importantly, their typical levels of variation makes them suitable markers for a number of purposes.

The usefulness of a marker completely depends on its characteristics. Enzyme molecules are direct products of genes, and thus of DNA, and play essential roles in the primary and secondary metabolism of organisms (Gillet, 1999). Enzyme molecules are composed of chains of amino acids as determined by the DNA sequences of the coding genes. Differences in the total electrostatic charges of their amino acid sequences indicate the existence of differences in the DNA sequences (Gillet, 1999). Allozymes almost always differ due to single nucleotide substitutions at the locus that cause the substitution of single amino acids of different charges; as a rule, isoenzymes coded by different loci differ in size. Size differences result from insertions or deletions of nucleotides that lead to a longer or shorter amino acid sequence.

Even though biochemical analyses are sometimes more expensive than morphological characterization, such studies are relatively simple to conduct, relatively inexpensive with regard to extraction, reagents and laboratory equipment required in comparison to molecular methods, and the results obtained have excellent comparability and repeatability (Hoogendijk and Williams, 2001).

One of the major advantages of biochemical characterization methods is that they are capable of detecting different alleles. Co-dominant markers such as isozymes, enable the researcher to determine allelic frequencies and thereby directly measure genetic diversity. Allelic frequency is extremely important information for population

genetics studies, for example, to determine the effective population size (Cervus, 2007). In case appropriate protocols are not available for the species investigated, existing protocols for related or similar species may be adapted.

Weaknesses of biochemical characterization include the fact that limited detection systems are available and they detect relatively few polymorphic loci and therefore are not very useful, particularly peanut (*Arachis hypogaea*) (Hoogendijk and Williams, 2001).

2.2.1.3.3. DNA markers

DNA markers are highly reliable selection tools as they as stable, not influenced by environmental conditions and relatively easy to score in an experienced laboratory (Peleman and van der Voort, 2003). DNA-based fingerprinting technologies have proven useful in genetic similarity studies (Pejic *et al.*, 1998). The DNA markers differ in their technique for polymorphism detection. They may be hybridisation-based (e.g., RFLP), or polymerase chain reaction (PCR)-based (e.g., RAPD and AFLP); they may detect single locus, oligo-locus, or multiple locus differences; and the markers detected may be inherited in a presence/ absence, or co-dominant manner (Mignouna *et. al.*, 2004).

Most points on molecular marker-based genetic linkage map are anonymous DNA polymorphisms (e.g., restriction fragment length polymorphism (RFLP), random amplified polymorphic DNA (RAPD), amplified fragment length polymorphism (AFLP), and microsatellite markers) and do not correspond to any known function, with the exception of few molecular markers including coding DNA (Hash and Bramel-Cox, 2000).

The general limitation of DNA markers therefore is that the anonymous DNA markers generated by a great variety of techniques, results in great diversity in

reliability (repeatability and robustness) of the system. Coupled with its difficulty, expense, and the nature of the polymorphism they detect makes the use of DNA markers unpopular. However, recent studies by Pejic *et al.*, (1998), has shown that apart from RAPD, the other DNA markers provide consistent information for germplasm identification and pedigree validation.

The DNA markers which have gained global acceptance and usage are AFLP, RFLP, RAPD and SSR. Apart from morphological traits (Dansi *et al.*, 1998, 1999), isozymic techniques (Dansi *et al.*, 2000; Mignouna and Dansi, 2003) and flow cytometry (Dansi *et al.*, 2000), molecular techniques provide opportunities to obtain high amplification of genetic traits for the development of genetic maps, variety identification and for the analysis of important morphological and agronomic traits (Dansi *et al.*, 2000; Tostain *et al.*, 2002; Tostain *et al.*, 2003). Molecular markers showing a high level of polymorphism on plant materials include microsatellites (Sonnante *et al.*, 1994; Akkaya *et al.*, 1995), RAPDs (Williams *et al.*, 1990; Williams *et al.*, 1993; Dansi *et al.*, 2000) and AFLP (Vos *et al.*, 1995; Tostain *et al.*, 2002; Tostain *et al.*, 2003; Kiambi *et al.*, 2005). RAPD markers have been shown to be useful in assessing intra-specific or inter-specific genetic variability in many crop plant species (Liu and Fumier, 1993; Haley *et al.*, 1994; Katsiotis *et al.*, 2003; Ravi *et al.*, 2003).

2.2.1.3.3.1 Amplified fragment length polymorphism (AFLP)

AFLP is a PCR-based genetic fingerprinting technique. It uses restriction enzymes to cut genomic DNA, followed by ligation of complementary double stranded adaptors to the ends of the restriction fragments. A subset of the restriction fragments are then amplified using 2 primers complementary to the adaptor and restriction site fragments.

The fragments are visualized on denaturing polyacrylamide gels through either autoradiographic or fluorescence methodologies.

2.2.1.3.3.2. Random amplified polymorphic DNA (RAPD)

The technique of random amplification of polymorphic DNA (RAPD), which is simply PCR amplification of genomic DNA by a single short oligonucleotide primer, produces complex patterns of anonymous polymorphic DNA fragments (Clark and Lanigan, 1993). The information provided by these banding patterns has proved to be of great utility (Clark and Lanigan, 1993).

RAPD by the PCR is a means of rapidly detecting polymorphisms for genetic mapping and strain identification (Welsh and McClelland, 1990; Williams *et al.*, 1990). It is a very useful technique to evaluate taxonomic identity and kinship (Hadrys *et al.*, 1992). They are technically simpler than other DNA-based markers (e.g. RFLPs), facilitate studies of large numbers of loci, and are expected to provide a far more random sample of genomic than do allozymes (Aagaard *et al.*, 1998). This method applies the PCR with a single short oligonucleotide primer to randomly amplified short fragments of genomic DNA, which are size-fractionated by agarose gel electrophoresis.

RAPDs survey numerous loci in the genome making the method particularly attractive for analysis of genetic distance and phylogeny reconstruction. Polymorphism is detected as band presence versus absence and may be caused either by failure to prime a site in some individuals because of nucleotide sequence differences or by insertions or deletions in the fragment between two conserved primer sites (Clark and Lanigan, 1993). The RAPD method therefore is only useful in genetic analysis if variation in banding patterns represents allelic segregation at independent loci. True

allelic segregation can also be confused with intermittent PCR artefacts (Riedy *et al.*, 1992). They are also quick (Welsh and McClelland, 1990) and well adapted for the efficient non-radioactive DNA fingerprinting of genotypes (dos Santos *et al.*, 1994). Its usage is however limited by the fact that its amplification and data scoring are not reproducible (Demeke *et al.*, 1997; Karp *et al.*, 1997).

2.2.1.3.3.3 Restriction fragment length polymorphism (RFLP)

The discrimination power of restriction fragment length polymorphisms (RLFPs) has been extensively studied in maize, as has their use in establishing relationships with yield and heterosis (Melchinger, 1990). There are however several drawbacks to RLFPs. Amongst them are that large quantities of DNA are required, is costly, the technique is difficult to automate, it requires sizeable laboratories and specialized equipment (Pejic *et al.*, 1998), and the use of radioactive probes makes it hazardous.

2.2.1.3.3.4. Simple Sequence Repeats (SSR)

The microsatellites (SSRs, or Simple Sequence Repeats) have proven to be very useful for the purpose of unveiling genetic diversity in many plants (Scotti *et al.*, 1999).

The variability at microsatellite loci is due to the differences in the number of repeat units. It is therefore easily detectable as variation in length of a DNA fragment obtained through PCR amplification from the whole genome (Scotti *et al.*, 1999). SSRs have been applied to several problems in which the fingerprinting of genetic diversity and differentiation is needed (Scotti *et al.*, 2000).

SSRs are codominant in character, making it possible to distinguish between plants that are homozygous and heterozygous at a gene locus (Powell *et. al.*, 1996). The microsatellites have advantages in among-population differentiation studies as well (Gillet, 1999). Yams generally display a low level of among-population divergence.

This is confirmed by analyses carried out by Dansi *et al.*, (1998), leading to classification of *Laribako* and *Pona* as one variety.

Data from SSR markers can be used either as single loci (if they are unlinked) or as haplotypes-if they are linked, as is the case with plastidial markers, (Scotti *et al.*, 2000). Microsatellites also convey an extra amount of information, compared to other classes of markers, thanks to the underlying mutational model (Stepwise Mutation Model), and that they often carry high numbers of alleles at very low frequencies or "private" alleles, that is, alleles present in only one or few populations (Scotti *et. al.*, 2000). This greatly contributes to the assessment of the genetic relationships among populations and among individual crops (Gillet, 1999). For instance in a survey on natural populations of Norway spruce (*Picea abies* Karst), only two loci with no allele had a frequency higher than 0.10 over a sample of more than 1300 individuals (Scotti *et al.*, 2000).

Solutions have been found for most of the often cited drawbacks in the interpretation of SSR amplification patterns: (A). Stutter bands can appear along with the band corresponding to the expected fragment. This will cause (i) uncertainty in the estimation of the allele size and (ii) the possibility of mistaking a heterozygote for a homozygote, if the two bands are so close on the gel that the ladders produced by the two alleles overlap. The former problem can be easily overcome by running an internal standard (an allele of known size), plus a molecular size marker, along with the test samples. The standard will help to identify the "true" band among those amplified in the PCR. Generally speaking, the introduction of internal size standards is a very efficient strategy in microsatellite gel runs, because they can help in all cases in which the scoring of fragment sizes is uncertain (Scotti *et al.*, 2000).

(B). The presence of *null* alleles. These alleles are not amplified and therefore are not scored at all on the gels. They can therefore lead to underestimates of heterozygosity. This problem has no direct troubleshooting, since there is no direct way to turn these (partially) dominant loci into codominant, and can have consistent effects (Scotti *et al.*, 2000).

In this study, SSR markers were used because of the following attributes: (a). In terms of equipment and costs, only basic molecular biology equipment is needed- a PCR thermal cycler, an electrophoresis cell for vertical acrylamide gels and the equipment for film/gel development.

(b). In terms of PCR efficiency, SSR PCR is a rather robust one and in general does not require highly purified DNA.

(c). In terms of technical skills and background necessary to perform SSRs, basic knowledge in molecular biology is all that is required to set up an experiment involving the use of SSRs. The only techniques involved are PCR and gel run.

(d). PCR reaction can be scaled down, saving money on the reagents and allowing to get a strong signal on the gel even with very small amounts of template genomic DNA -down to 1.5 ng (Scotti *et al.*, 2003).

2.2.2. Quantifying genetic diversity: measuring intra-population genetic diversity

There are two bases for measuring intra-population genetic diversity- on the number of variants and frequency of variants.

2.2.2.1 Intra-population genetic diversity on the basis of number of variants

On the basis of number of variants: one can assess polymorphism or rate of polymorphism, proportion of polymorphic loci, richness of allelic variants (A) and average number of alleles per locus.

2.2.2.1.1 Polymorphism or rate of polymorphism (Pj)

This measure provides criteria to demonstrate that a gene is showing variation. A gene is said to be polymorphic if the frequency of one of its alleles is less than or equal to 0.95 or 0.99 (Cornell University, 2003)

$$Pj = q \leq 0.95 \text{ or } Pj = q \leq 0.99.$$

where,

Pj = rate of polymorphism; q = allele frequency

Its calculation is through direct observation of whether the definition is fulfilled and can be used with codominant markers and, very restrictively, with dominant markers (Cornell University, 2003). This is because the estimate based on dominant markers would be biased below the real number .

A polymorphic gene is usually one for which the most common allele has a frequency of less than 0.95. Rare alleles are defined as those with frequencies of less than 0.005 (Cornell University, 2003). The limit of allele frequency, which is set at 0.95 (or 0.99) is arbitrary, its objective being to help identify those genes in which allelic variation is common (Cavalli–Sforza and Bodmer, 1981).

2.2.2.1.2 Proportion of polymorphic loci

This measure expresses the percentage of variable loci in a population. Its calculation is based on directly counting polymorphic and total loci and can be used with codominant markers and, very restrictively, with dominant markers (Cornell

University, 2003). This is the number of polymorphic loci divided by the total number of loci (polymorphic and monomorphic), that is:

$$P = n_{pj}/n_{total}$$

where,

P = proportion of polymorphic loci; n_{pj} = number of polymorphic loci and n_{total} = total number of loci

2.2.2.1.3 Richness of allelic variants (A)

This measure refers to the number of variants in a sample. The measure of diversity is (A - 1) variants because, within a monomorphic population, the degree of diversity is zero (A - 1 = 0). For a given gene in a sample, this measure tells how many allelic variants can be found. It is however sensitive to sample size and even though the distribution of alleles does not matter, the maximum number of alleles does (Cornell University, 2003). However, it can only be applied to co-dominant markers such as SSRs (Cornell University, 2003).

2.2.2.1.4 Average number of alleles per locus

Another measure of intra-population genetic diversity is the average number of alleles per locus. It is the sum of all detected alleles in all loci, divided by the total number of loci.

$$\bar{n} = (1/K)\sum_{i=1}^{K} n_i$$

where,

K = the number of loci; ni = the number of alleles detected per locus.

This measure provides complementary information to that of polymorphism, and requires only counting the number of alleles per locus and then calculating the average.

It is best applied with codominant markers such as SSRs, because dominant markers do not permit the detection of all alleles (Cornell University, 2003).

2.2.2.2. Intra-population genetic diversity on the basis of the frequency of variants

On the basis of the frequency of variants the intra-population genetic diversity can be measured by assessing the effective number of alleles (Ae), and average expected heterozygosity (He; Nei's genetic diversity)

2.2.2.2.1 Effective number of alleles (Ae)

This measure tells about the number of alleles that would be expected in a locus in each population. It is defined as the number of alleles that can be present in a population.

$$A_e = 1/(1 - h) = 1/\Sigma p_i^2$$

where,

pi = frequency of the ith allele in a locus; h = 1 – Σpi and 2 = heterozygosity in a locus

Effective number of alleles (Ae) is calculated by inverting the measure of homozygosity in a locus. It can also be used with codominant markers such as SSRs. Its calculation is however sampling size dependent.

This measure of diversity may also be informative for establishing collecting strategies. For example, we estimate it in a given sample. We then verify it in a different sample or the entire collection. If the figure obtained the second time is less than the first estimated number, this could mean that our collecting strategy needs revising.

Another measure of intra-population genetic diversity is the Average expected heterozygosity (He) (Nei's genetic diversity [D]). This is the probability that, at a single locus, any two alleles, chosen at random from the population, are different to each other

There are three calculations that are possible:

- A locus with two alleles: $h_j = 1 - p^2 - q^2$
- A locus j with i alleles: $h_j = 1 - \Sigma pi^2$
- Average for several loci: $H = \Sigma_j^L h_j/L$

Where,

hj = heterozygosity per locus; p and q = allele frequencies; H = average heterozygosity for several loci and L = total number of loci

The average expected heterozygosity is calculated by subtracting from 1 the expected frequencies of homozygotes in a locus. The operation is repeated for all loci and the average then performed. It can be applied to all markers, both codominants and dominants. The estimated value may be affected by those alleles present at higher frequencies. It ranges from 0 to 1, and it is maximized when there are many alleles at equal frequencies. A minimum of 30 loci in 20 individuals per population should be analyzed to reduce the risk of statistical bias (Cornell University, 2003).

The average H_E over all loci is an estimate of the extent of genetic variability in the population.

2.2.2.2.2. Calculating intra-population distance, using microsatellites

Intra-population distance is the average of the sum of squares of the differences in the number of repeats between alleles.

$$S_{W} = \frac{2}{2n(2n-1)} \Sigma_{i<j} (a_i - a_j)^2$$

The average intra-population distance may be calculated for all analyzed loci (ds).

$$S_W = (1/d_s)\sum_j S_{Wj}$$

Where,
aij = size of the allele of the i^{th} copy (i = 1, 2, …, 2n) in the j^{th} population (j = 1, 2, …, ds); n = number of individuals in the sample

There are two considerations:

The calculation of distance between two alleles is a transformation of the number of repeats. One difficulty in using SSRs to estimate genetic distances is their high rate of mutation.

2.2.3. Expression of relationships between results obtained from different sets of characters (classification or clustering)

There are two main methods for expressing relationships between results obtained from different sets of characters: clustering and ordination.

2.2.3.1 Clustering Method

Cluster analysis, also called segmentation analysis or taxonomy analysis is a group of multivariate techniques used to group objects (subjects, respondents, products, etc.) based on the characteristics they possess. Each object within the cluster will be similar to every other object, and different from objects in other clusters. In other words, homogeneity is maximized within clusters and heterogeneity is maximized between them. Cluster analysis is an exploratory data analysis tool for solving classification problems. Each cluster thus describes, in terms of the data collected, the class to which its members belong; and this description may be abstracted through use from the particular to the general class or type. It is therefore a multivariate procedure for

detecting natural groupings in data. This is the process of grouping (or clustering) objects in categories or classes based on their common attributes or relationships. Cluster analysis is therefore very useful because it allows one to visualize similarities among taxa by the levels at which they are grouped together (Crawford, 1990).

To measure distance among clusters a number of methods is available and varies according to the way in which 'closest' is defined at each stage of merging groups. The following possibilities are available.

• Single Link (Nearest neighbour) defines the similarity between two clusters as the maximum similarity between any two samples in those clusters.

• Complete Link ('farthest neighbour') defines the similarity between two clusters as the minimum similarity between any two samples in those clusters (Aldenderfer *et. al.*, 1984).

• Average Link (UPGMA- unweighted pair-group method using the arithmetic average') defines the similarity between a cluster and two merging clusters as the average of the similarities with each of the original clusters. It therefore replaces two merging clusters by their mean, unweighted by cluster size (Kumar *et. al.*, 1994).

• Group Average - an average is taken over all the samples in the two merging clusters. Thus, the original clusters are replaced by their mean, weighted by cluster size (Gower, 1967)

• Median Sorting can be thought of in terms of clusters being represented by points in a multidimensional space; when two clusters join, the new cluster is represented by the midpoint of the original cluster points (Gower, 1967).

Other methods available for measuring distances among data points are:

• Unweighted pair-group method using the centroid (UPGMC). It is based on the distance between the mean values for each group (Kumar *et.al.*, 1994).

• Weighted pair-group method using the centroid (WPGMC). It takes the Observable Taxonomic Units (OTUs) median value in the groups (Gower, 1967).

• Ward. It works with the sum of the squared distances for pairs of OTUs. It is also known as the method of minimal variance because, while taking the squared values, it becomes a very sensitive method (different OTUs will look more dissimilar and similar OTUs will look even closer). It may be used with Euclidian distances and molecular data when a high number of DNA bands are available.

Most data sets may be thought of as a matrix of n units (rows) by p variables (columns). The first step in cluster analysis is the establishment of the similarity or distance matrix (Aldenderfer and Blashfield, 1984). Here a matrix of n rows by n columns in which the element a_{ij} describes the association between the i^{th} and the jth unit. This matrix is a table in which both the rows and columns are the units of analysis and the cell entries are a measure of similarity or distance for any pair of cases. The more matches there are between two groups, the more similar they are. Statistical clustering methods use similarity measures to partition objects (Anderberg 1973, Jain and Dubes 1988) whereas conceptual clustering methods cluster objects according to the concepts objects carry (Michalski and Stepp, 1983, Fisher 1987). Therefore there are three ways that the groupings can be made: hierarchical, non-hierarchical or overlapping (Garcia *et al.*, 1995).

2.2.3.1.1 Hierarchical Clustering

Hierarchical clustering is a way to investigate grouping in one's data, simultaneously over a variety of scales, by creating a cluster tree. The tree is not a single set of clusters, but rather a multilevel hierarchy, where clusters at one level are joined to clusters at the next higher level. This allows one to decide what level or scale of clustering is most appropriate in an application (Hastie *et. al.*, 2009).

There are two ways of classifying hierarchical clustering – agglomerative and divisive. The agglomerative hierarchical methods operate on a matrix of similarities for a set of units. The divisive hierarchical method is a clustering procedure whereby all objects begin in a single cluster, then the subsequent steps, each cluster is divided into two clusters containing objects that are the most dissimilar. This is the opposite of agglomerative methods. It is often found that using this method provides a more acceptable classification into a small number of major groups – the groupings obtained are also more robust to any aberrant similarities between individual pairs of units.

In agglomeratie clustering, the procedure begins with each object in a separate cluster. In each subsequent step, the two object clusters that are the most similar are combined into a new cluster. This is repeated until all objects are combined into a single cluster. This method is however criticized for concentrating initial attention on individual units and small groups since important groupings could be influenced by the chance merging of two units at an early stage of clustering.

2.2.3.1.2. Non-hierarchical clustering

These methods generally operate on units by variates matrix and seek to partition the units into a specified number of groups to optimize some criterion. The most common criterion used is maximizing the between groups sum of squares, which is equivalent to minimizing the within-groups sum of squares. This method assigns each individual to a unique group by comparing it with the initial classes so that its positioning is the most appropriate.

Clustering starts with some initial classification and using an iterative procedure, units are moved between groups until no move gives an improvement in the criterion. This is taken as the optimum grouping.

There are three main objectives for clustering data; to find natural groupings as in taxonomy, to simplify data (data reduction) and to understand the data better than producing a distribution as a whole (relationship identification).

For a data set made up of *m* objects, there are *m (m-1)/2* pairs in the data set. The result of this computation is commonly known as a distance or dissimilarity matrix.

2.2.3.1.3. Overlapping clustering

In this method, unlike the non-hierarchical clustering where individuals are assigned to a unique grouping, individuals may belong to more than one group.

Associations may also be expressed as similarities or distances. Similarities are usually constrained to lie in the range [0, 1] with 'self-similarities' taking the value 1. Distances or dissimilarities are complementary to similarities –they are usually non-negative with self-distance zero. Similarities can easily be converted to distances and vice-versa.

Three types of data commonly encountered in practice are binary, qualitative (with more than two states) and quantitative (Kaufman *et. al.,* 1990).

Binary Data: Data is in a form of presence (1) and absence (0). Similarity coefficient is calculated by either simple matching or Jaccard's method. In the simple matching method,

Similarity coefficient = similar/ (similar + dissimilar);

and Jaccard's suggests that if a scoring in neither accession do not contribute to inter-site similarity then,

Similarity coefficient = only positive similar/ (similar + dissimilar).

Qualitative Data: Qualitative data is extremely varied in nature. It includes virtually any information that can be captured that is not numerical in nature. These refer to characters or qualities, and are either binary or categorical: binary, taking only two values: present (1) or absent (0); categorical, taking a value among many possibilities, and are either ordinal or nominal: Ordinal: categories that have an order and Nominal: categories that are unrelated (Aldenderfer *et. al.*, 1984).

For characters such as colour and shape there are often more than two states of the character, and coded values are used for estimation of similarity indices. When zero pairs are available, data is adjusted as binary data.

Quantitative Data: Distances rather than similarities are considered when dealing with quantitative data. The common measure is the Euclidean distance. The general formula for the (squared) Euclidean distance between units *I* and *j*, based on *p* quantitative variables x*k* (k = 1…p) is

$$d_{ij} = (1/p)\Sigma(x_{ik} - x_{jk})^2 r_k^2$$ (Summation from k =1 to k =p).

In the formula, the distance between units for the kth variable is divided by rk to allow for differing scales of measurement. When included, the value of r is usually set to either the standard deviation or the range of the variable, ensuring that distances do not exceed 1.

Euclidean distance is sensitive to unusually large differences between units for a single variable. A more robust alternative is the City-block or Manhattan distance measure:

$$d_{ij} = (1/p)\Sigma(x_{ik} - x_{jk})r_k$$

(Summation from k =1 to k =p).

2.2.3.2. Ordination

Ordination is the arrangement or 'ordering' of sample units along coordinate systems. The purpose of ordination, as well as classification methods, is to interpret patterns in the composition of samples. This is a multivariate method that complements clustering, and is usually considered to be an approach that is closer to biological reality. With ordination methods, we want to represent the relationships of samples in a simple way by reducing the real situation to a 'low dimensional space' (Gauch,1982).

In principle, ordination is both an exploratory and hypothesis-testing tool. Numerical taxonomic techniques have been successfully used by many workers to classify variation patterns at both intra and interspecific levels (Sneath and Sokal, 1973; Chheda and Fatokun, 1982; Ariyo, 1991).

Principal component analysis (PCA) is a descriptive technique which reveals the pattern of character variation among individual accession. It further reduces multivariate data into units or component thereby accounting for a meaningful amount of variation in a population. It is however not suitable for molecular data (IPGRI and Cornell University, 2003)

2.3. Analysis of Molecular Marker Data

For molecular marker data, there are three methods for displaying relationships: principal coordinates analysis (PCoA), nonmetric multi-dimensional scaling (NMDS) and Correspondence analysis (CA).

Principal Coordinates Analysis (PCoA) also known as Multidimensional Scaling (MDS), is a more general projection method than PCA. This is because PCoA can use any distance matrix. However, because PCoA uses a distance matrix all information about the original variables is lost and the analysis relates only to the cases. Principal coordinates analysis therefore attempts to represent distances between

samples and may accommodate matrices from different dissimilarity measures. It maximizes the linear correlation between sample distances. When used with Euclidean distances, the results are identical to PCA. If the distance matrix uses Euclidean distances a PCoA and a PCA analysis would produce identical projections of the cases, even though with some possible reflections or translations.

The differences between PCoA and PCA can be are summarized as:

1. PCA searches for patterns in the variables, whereas PCoA searches for similarities between cases.
2. PCA reduces variable dimensionality by an eigen analysis of a correlation or covariance matrix, whereas PCoA analyses a distance matrix. Many different distance matrices can be used as long as they metric, in particular they must obey the 'triangle' rule.
3. The result of a PCoA is a set of coordinates on a number of derived axes such that similar cases are close together. It is not possible to associate these axes with any variables. PCA however generates axes which can be associated with variables.

Nonmetric multidimensional scaling (NMDS) on the other hand works by maximizing the rank order correlation and attempting to find the best shape to accommodate the data. This technique uncovers the basic configuration from the dissimilarity sample matrix. With NMDS, only the pattern of points is relevant, not the origin, and the representation may be rotated.

Correspondence analysis (CA) repeats the averages of sample scores and finds spots where all samples falling in the same spot are as similar as possible and, simultaneously, samples at different spots are as different as possible.

2.4 Concept of Core Collection

Core collections were first defined as a limited set of accessions which represents, with a minimum of repetitiveness, the genetic diversity of a crop species and its wild relatives (Frankel, 1984). This was later modified as "a limited set of accessions derived from an existing germplasm collection, chosen to represent the genetic spectrum in the whole collection, and including as much as possible of its genetic diversity" (Brown, 1994). It is also defined as "a germplasm collection optimally representing specific genetic diversity", implying that size, type and origin of a core collection depends on the requirements of the compiler (van Hintum, 1990).

A core collection can be based on a single existing collection or on several existing collections, but it can also be a newly created entity. It can represent the diversity in a complete genus, including wild species, or the diversity in a small part of the genepool. It can contain as much diversity as possible, but can also give higher priority to a certain type of material, reducing the total amount of diversity captured (van Hintum, 1990).

Numerous examples of the establishment of core collections that have been published including a core collection representing the complete US germplasm collection of peanut (Holbrook *et al.*, 1993), cultivated *Brassica oleracea* in European collections (Boukema and Hintum, 1994), lentil accessions from Chile, Greece and Turkey (Erskine and Muehlbauer, 1991), the entire genepool of *Hordeum* (Hintum *et. al.*, 1990), local maize populations with a good combining ability (Radovic and Jelovac, 1994) and *Pisum sativum* germplasm with disease resistance (Matthews and Ambrose, 1994).

Numerous approaches have been adopted in an effort to select a representative sample of a collection- from random selection (random core), to sequential set selection

and equidistant core (selection of accessions with accession numbers ending with a five or zero).

The general procedure for creating a core collection can be divided in four steps:

1. Define the material that should be represented, i.e. the domain of the core collection. This can be a crop and its wild relatives, but can also be a small part of the diversity of a crop such as material with specific traits or from a specific origin region.
2. Divide the domain in types, which should be genetically as distinct as possible.
3. Choose the number of entries in the core, and allocate them over the types based on the relative importance of, and diversity in the type.
4. Select the entries from each type that are to be included in the core. This selection can be made randomly, or if additional data are available, on the basis of these data.

The objective should be to best represent the diversity in the group. Also practical considerations can play a role in this choice, such as the availability of the seeds and the reliability and quantity of the data on the accession, since it is desirable to have readily available, authentic and well documented material in the core.

A major problem of the management of *ex situ* plant genetic resources (PGR) is the low accessibility of the collections (van Hintum, 1990). These core collections were invented to increase the accessibility of germplasm collections; low accessibility implies low utilization and difficult management. A core collection therefore offers a good starting point for searches for new traits, and can be used for in-depth evaluation increasing the knowledge of the entire collection (van Hintum, 1990).

All in all the concept of core collection is relevant in germplasm management as it was first proposed.

CHAPTER 3

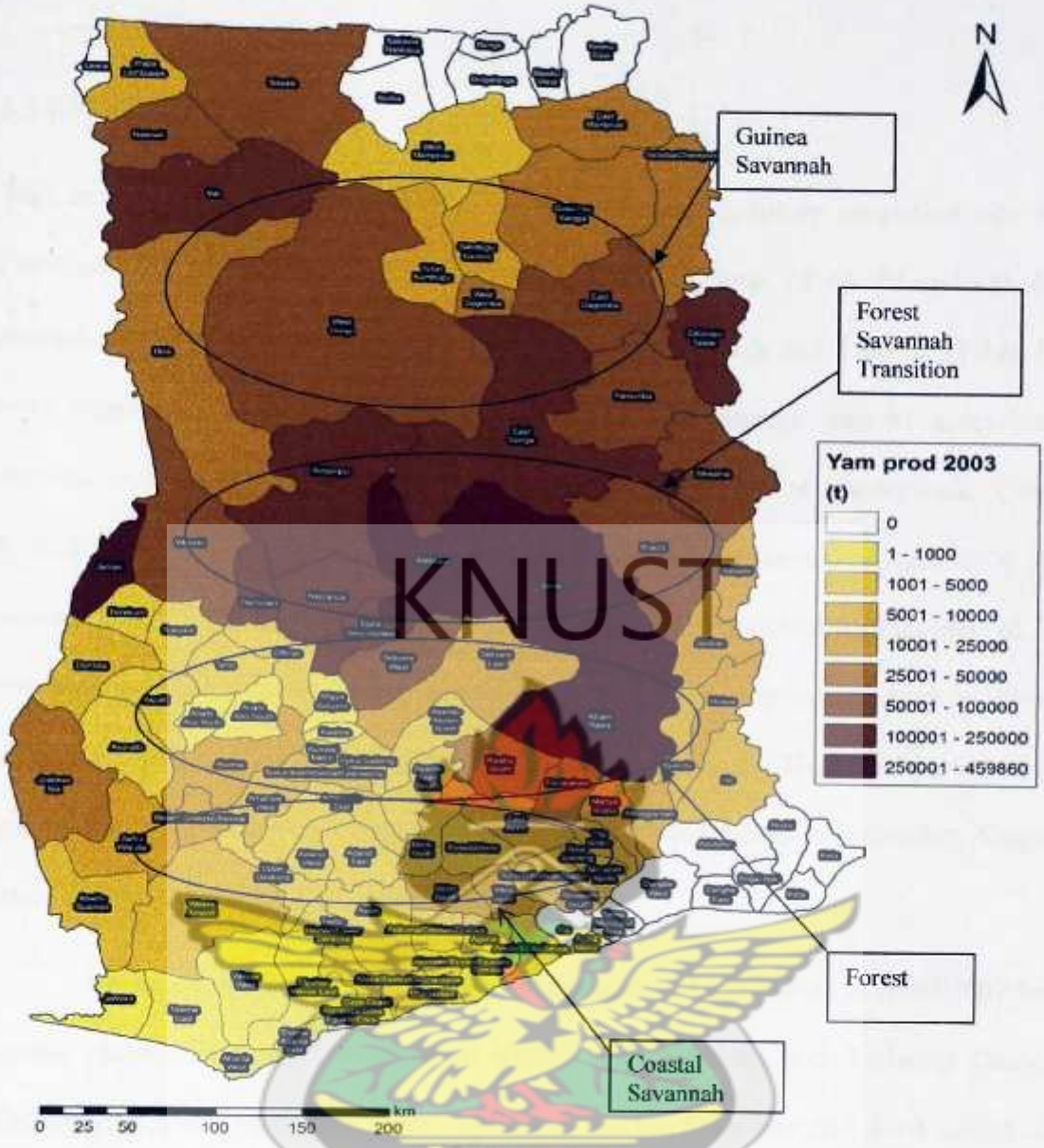
3.0 RESEARCH DESIGN AND METHODOLOGY

3.1 Study Areas

The accessions were collected from four yam growing agroecological zones in Ghana: Coastal Savannah (Bodwease, Mankrong and Bontrase), Forest (Fumesua and Ejisu), Forest-Savannah Transition Zone (Ejura, Wenchi, Techiman, and Kintampo), and Guinea Savannah (East Gonja, Tolon/Kubungu, Nanumba North and Nanumba South) (Fig. 1). These study areas were selected because the crop has co-existed with the people in these locations from time immemorial and rich in farmers' *Pona* varieties and *Pona* look-alikes. Hence they would be suitable for characterization and classification of the *Pona* collection.

3.2 Germplasm Collection

The general approach to the *Pona* diversity studies was to combine ethno-botany (indigenous or folk) taxonomy, morphological and molecular characterization methods. Ethno-botany classification was employed at the following stages of evaluation: at time of collection of accessions, vegetative and harvesting stages of the accessions. A total of 66 *Pona-Laribako* tubers or seedyams including *Laribako*-la and *Laribako*-nya (supposed male and female respectively) and 1 *Muchumudu* accession were collected from the study areas.



Source: FAO, 2005.

Fig. 1.Yam production map of Ghana showing where Pona accessions were collected.

3.3 Evaluation

This activity was conducted at CSIR-Crops Research Institute on-station site at Fumesua, Kumasi, Ghana and some farmer fields. A total of 66 *Pona-Laribako* accessions including 3 released *Pona* varieties, 1 *Muchumudu* and 1 *Dente* (Table 1) were planted in April 2006 and reclassified by expert farmers' into 91 accessions consisting of 43 *Laribako*, 40 true *Pona*, 5 *Kulunku* and 1 each of *Muchumudu*, 'CRI-Kukrupa', and *Dente* (Table 2). Tuber setts (300-400g) were planted in April 2006 on mounds at a spacing of 1.5m x 1.5m. Yam mounds was singly staked with bamboo of 2-3m high to enhance good canopy expression. Weeding was manually done as when necessary. No chemical amendments were applied to the soils. The aerial parts of the plant were assessed through May – August, the harvested tubers in December. Single harvesting was conducted in December 2006.

The study was repeated in 2007 by planting the 91 accessions in April 2007 for further characterization in Augmented RCBD with 3 blocks and 3 checks Dente, CRIKukrupa and MankrongPona as replicated checks. Each row had 8-10 clones of each accession, 5 plants were selected at random to record morphological traits. In 2007, the field was laid to reflect the order given by the dendogram of 2006 results, so that similar morphotypes could be grown close to each other and compared more accurately. These morphological descriptions were repeated over two years to verify the accuracy of the rated morphological traits.

Table 1: Original List of *Pona-Laribako* accessions collected and used for the study.

SNO	PEG NO	CODE FOR ACCESSION	ACCESSION NAME	DISTRICT	VILLAGE	COORDINATES	ELEVATION/M
01	101	EOYT/06/224-233	*Laribako*	TOLON/KUBUNGU	TINGOLI	09°22.514N;001°00.580W	162
02	102	EOYT/06/164-173	*Laribako*	TOLON/KUBUNGU	TINGOLI	09°22.514N;001°00.580W	162
03	103	EOZT/06/275-284	*Pona*	KINTAMPO	KINTAMPO	-	-
04	104	EOYT/06/154-163	*Laribako*	TOLON/KUBUNGU	TINGOLI	09°22.514N;001°00.580W	162
05	105	EOZKP/06/120-126	*Laribako*	KINTAMPO	KINTAMPO	-	-
06	106	EOZW/06/047-053	*Pona*	WINNEBA	BODWEASE	05°35'N 0°35'W	-
07	107	CRI KUKRUPA	KUKRUPA	WENCHI	WENCHI	07°44'N 2°7'W	-
08	108	EOYW/06/088-097	KULUKU	NANUMBA SOUTH	BINDA	08o56.083N;000o00.053E	196
09	109	EOYK/06/134-138	*Laribako*	NANUMBA NORTH	KALANDE	08°35.985N;000°29.084W	174
10	110	EOKULUKU/06/065-087	KULUKU	NANUMBA SOUTH	BINDA	08°56.083N;000°00.053E	196
11	111	EOZT/06/155-164	*Pona*	TOLON/KUBUNGU	TINGOLI	09°22.514N;001°00.580W	162
12	112	EOYB/06/004-013	*Laribako*	NANUMBA NORTH	BIMBILLA	08°53.669N;000°04.793E	186
13	113	EOYT/06/174-183	*Laribako*	TOLON/KUBUNGU	TINGOLI	09°22.514N;001°00.580W	162
14	114	EOZT/06/275-284	*Pona*	TECHIMAN	TECHIMAN	07°34.842N;001°56.846W	392
15	115	EOZT/06/295-304	*Pona*	TECHIMAN	TECHIMAN	07°34.842N;001°56.846W	392
16	116	EOZT/06/215-224	*Pona*	TOLON/KUBUNGU	TINGOLI	09°22.514N;001°00.580W	162
17	117	EOYT/06/274-283	*Laribako*	TECHIMAN	TECHIMAN	07°34.842N;001°56.846W	392
18	118	EOYT/06/254-263	*Laribako*	TOLON/KUBUNGU	KPALISOGU	09°22.513N;001°00560W	162
19	119	EOZT/06/195-204	*Pona*	TOLON/KUBUNGU	TINGOLI	09°22.514N;001°00.580W	162
20	120	EOZW/06/043-046	LARBRAKO-LAA (*Pona*)	NANUMBA SOUTH	BINDA	08°56.083N;000°00.053E	196

Table 1 continued.

21	121	EOZK/06/145-154	*Pona*	NANUMBA NORTH	KALANDE	08°35.985N;000°29.084W	174
22	122	EOZT/06/175-184	*Pona*	TOLON/KUBUNGU	TINGOLI	09°22.514N;001°00.580W	162
23	123	EOZT/06/225-234	*Pona*	TOLON/KUBUNGU	TINGOLI	09°22.514N;001°00.580W	162
24	124	EOZK/06/113-119	*Pona*	NANUMBA NORTH	KALANDE	08°35.985N;000°29.084W	174
25	125	EOYT/06/184-193	*Laribako*	TOLON/KUBUNGU	TINGOLI	09°22.514N;001°00.580W	162
26	126	EOZB/06/030-039	*Pona*	NANUMBA NORTH	DAMON NA YILI	08°56.085N;000°00.053E	202
27	127	EOZK/06/127-131	*Pona*	NANUMBA NORTH	KALANDE	08°35.985N;000°29.084W	174
28	128	EOZT/06/205-214	*Pona*	TOLON/KUBUNGU	TINGOLI	09°22.514N;001°00.580W	162
29	129	EOZS/06/103-112	*Pona*	NANUMBA NORTH	KALANDE	08°35.988N;000°29.085W	174
30	130	EOYK/06/124-133	*Laribako*	NANUMBA NORTH	KALANDE	08°35.985N;000°29.084W	174
31	131	EOYS/06/109-113	*Laribako*	NANUMBA NORTH	KALANDE	08°35.988N;000°29.085W	174
32	132	ASA	*Dente*	WINNEBA	BODWEASE	05°35'N 0°35'W	-
33	133	EOYE/06/001-004	*Laribako*	EAST GONJA	SALAGA	08°33.060N;000°31.198W	158
34	134	EOZT/06/185-194	*Pona*	TOLON/KUBUNGU	TINGOLI	09°22.514N;001°00.580W	162
35	135	EOZB/06/002-011	*Pona*	NANUMBA NORTH	BIMBILLA	08°53.669N;000°04.793E	186
36	136	EOZ/06/035-044	*Pona*	NANUMBA SOUTH	NAKPAYILI	08°56.083N;000°00.53E	163
37	137	EOYT/06/204-213	*Laribako*	TOLON/KUBUNGU	TINGOLI	09°22.514N;001°00.580W	162
38	138	EOYT/06/284-293	*Laribako*	TECHIMAN	TECHIMAN	07°34.842N;001°56.846W	392
39	139	EOZT/06/265-274	*Pona*	TOLON/KUBUNGU	KPALISOGU	09°23.930N;001°00.744W	156
40	140	EOZ/06/013-119	*Pona*	EAST GONJA	SALAGA	08°33.060N;000°31.198W	158
41	141	EOYW/06/064--	KULUKU	NANUMBA SOUTH	NAKPAYILI	08°56.083N;000°00.53E	163
42	142	EOYS/06/154-161	*Laribako*	NANUMBA NORTH	BIMBILLA	08°53.776N;000°02.536W	186
43	143	EOZT/06/235-244	*Pona*	TOLON/KUBUNGU	TINGOLI	09°22.514N;001°00.580W	162

Table 1 continued.

44	144	EOZKP/06/134-139	*Pona*	NANUMBA NORTH	BIMBILLA	08°53.776N;000°02.536W	186
45	145	EOZT/06/245-254	*Pona*	TOLON/KUBUNGU	TINGOLI	09°22.514N;001°00.580W	162
46	146	EOYW/06/098-102	LARBRAKO-LAA	NANUMBA SOUTH	NASAMBA	08°45.974N;000°03.329E	196
47	147	EOYT/06/234-243	*Laribako*	TOLON/KUBUNGU	TINGOLI	09°22.514N;001°00.580W	162
48	148	EOYW/06/103-108	LARBRAKO-NYA	NANUMBA SOUTH	NASAMBA	08°45.974N;000°03.329E	196
49	149	EOYB/06/004-011	*Laribako*	NANUMBA NORTH	PUSUGA	08°55.539N,000°.872W	209
50	150	EOYK/06/144-153	*Laribako*	NANUMBA NORTH	KALANDE	08°35.985N;000°29.084W	174
51	151	CRI *Pona*	*Pona*	EJISU	FUMESUA	06°41'N, 1°28'W	209
52	152	EOZS/06/155-160	*Pona*	NANUMBA NORTH	KALANDE	08°35.985N;000°29.084W	174
53	153	EOYW/06/055-063	KULUKU	NANUMBA SOUTH	NAKPAYILI	08°56.083N;000°00.53E	163
54	154	EOZT/06/285-294	*Pona*	TECHIMAN	TECHIMAN	07°34.842N;001°56.846W	392
55	155	EOZT/06/165-174	*Pona*	TOLON/KUBUNGU	TINGOLI	09°22.514N;001°00.580W	162
56	156	EOYT/06/264-273	*Laribako*	TOLON/KUBUNGU	KPALISOGU	09°23.930N;001°00.744W	156
57	157	EOZK/06/140-144	*Pona*	NANUMBA NORTH	KALANDE	08°35.985N;000°29.084W	174
58	158	EOYB/06/045-054	*Laribako*	NANUMBA NORTH	DAMONNA YILI	08°56.085N;000°00.053E	202
59	159	EOYT/06/194-203	*Laribako*	TOLON/KUBUNGU	TINGOLI	09°22.514N;001°00.580W	162
60	160	EOYT/06/244-253	*Laribako*	TOLON/KUBUNGU	TINGOLI	09°22.514N;001°00.580W	162
61	161	EOY/06/014-033	*Laribako*	NANUMBA NORTH	KALANDE	08°35.985N;000°29.084W	-
62	162	EOYK/06/144-123	*Laribako*	TOLON/KUBUNGU	TINGOLI	09°22.514N;001°00.580W	-
63	163	EOYT/06/214-223	*Laribako*	TOLON/KUBUNGU	TINGOLI	09°22.514N;001°00.580W	-
64	164	EOZ/06/255-264	*Pona*	TOLON/KUBUNGU	KPALISOGU	09°23.930N;001°00.744W	156
65	165	MANKRONG *Pona*	HYBRID	-	-	-	-
66	166	EOZB/06/020-029	*Pona*	TOLON/KUBUNGU	KPALISOGU	09°22.513N;001°00560W	-

Table 2. Reclassification of accessions based on expert farmer knowledge.

SNO	PEG NO.	CODE FOR ACCESSION	ACCESSION NAME	DISTRICT	VILLAGE	COORDINATES	ELEVATION/M
01	101	EOYT/06/224-233	*Laribako*	TOLON/KUBUNGU	TINGOLI	09°22.514N;001°00.580W	162
02	102	EOYT/06/164-173	*Laribako*	TOLON/KUBUNGU	TINGOLI	09°22.514N;001°00.580W	162
03	103	EOZT/06/275-284	*Pona*	KINTAMPO	KINTAMPO		
04	104	EOYT/06/154-163	*Laribako*	TOLON/KUBUNGU	TINGOLI	09°22.514N;001°00.580W	162
05	105	EOZKP/06/120-126	*Laribako*	KINTAMPO	KINTAMPO		
06	106	EOZW/06/047-053	*Pona*	WINNEBA	BODWEASE	05°35'N 0°35'W	
07	106*	EOZW/06/054-059	*Laribako*	WINNEBA	BODWEASE	05°35'N 0°35'W	
08	107	CRI KUKRUPA	KUKRUPA	WENCHI	WENCHI	07°44'N 2°7'W	
09	107*	CRI KUKRUPA	KUKRUPA	WENCHI	WENCHI	07°44'N 2°7'W	
10	108	EOYW/06/088-097	KULUKU	NANUMBA SOUTH	BINDA	08o56.083N;000o00.053E	196
11	108*	EOYW/06/093-097	*Pona*	NANUMBA SOUTH	BINDA	08o56.083N;000o00.053E	196
12	109	EOYK/06/134-138	*Laribako*	NANUMBA NORTH	KALANDE	08°35.985N;000°29.084W	174
13	110	EOKULUKU/06/065-087	KULUKU	NANUMBA SOUTH	BINDA	08°56.083N;000°00.053E	196
14	111	EOZT/06/155-164	*Pona*	TOLON/KUBUNGU	TINGOLI	09°22.514N;001°00.580W	162
15	112	EOYB/06/004-013	*Laribako*	NANUMBA NORTH	BIMBILLA	08°53.669N;000°04.793E	186
16	113	EOYT/06/174-183	*Laribako*	TOLON/KUBUNGU	TINGOLI	09°22.514N;001°00.580W	162
17	114	EOZT/06/275-284	*Pona*	TECHIMAN	TECHIMAN	07°34.842N;001°56.846W	392
18	115	EOZT/06/295-301	*Pona*	TECHIMAN	TECHIMAN	07°34.842N;001°56.846W	392
19	115*	EOZT/06/302-304	*Laribako*	TECHIMAN	TECHIMAN	07°34.842N;001°56.846W	392
20	116	EOZT/06/215-224	*Pona*	TOLON/KUBUNGU	TINGOLI	09°22.514N;001°00.580W	162
21	117	EOYT/06/274-278	*Laribako*	TECHIMAN	TECHIMAN	07°34.842N;001°56.846W	392

Table 2 continued.

22	117*	EOYT/06/279-283	*Laribako*	TECHIMAN	TECHIMAN	07°34.842N;001°56.846W	392
23	118	EOYT/06/254-263	*Laribako*	TOLON/KUBUNGU	KPALISOGU	09°22.513N;001°00560W	162
24	119	EOZT/06/195-204	*Pona*	TOLON/KUBUNGU	TINGOLI	09°22.514N;001°00.580W	162
25	120	EOZW/06/043-046	LARRBAKO-LAA (*Pona*)	NANUMBA SOUTH	BINDA	08°56.083N;000°00.053E	196
26	121	EOZK/06/145-154	*Pona*	NANUMBA NORTH	KALANDE	08°35.985N;000°29.084W	174
27	122	EOZT/06/175-184	*Pona*	TOLON/KUBUNGU	TINGOLI	09°22.514N;001°00.580W	162
28	123	EOZT/06/225-234	*Pona*	TOLON/KUBUNGU	TINGOLI	09°22.514N;001°00.580W	162
29	124	EOZK/06/113-119	*Pona*	NANUMBA NORTH	KALANDE	08°35.985N;000°29.084W	174
30	125	EOYT/06/184-193	*Laribako*	TOLON/KUBUNGU	TINGOLI	09°22.514N;001°00.580W	162
31	126	EOZB/06/030-034	*Pona*	NANUMBA NORTH	DAMON NA YILI	08°56.085N;000°00.053E	202
32	126*	EOZB/06/035-039	*Laribako*	NANUMBA NORTH	DAMON NA YILI	08°56.085N;000°00.053E	202
33	127	EOZK/06/127-131	*Pona*	NANUMBA NORTH	KALANDE	08°35.985N;000°29.084W	174
34	128	EOZT/06/205,206,210-214	*Pona*	TOLON/KUBUNGU	TINGOLI	09°22.514N;001°00.580W	162
35	128*	EOZT/06/207-209	*Laribako*	TOLON/KUBUNGU	TINGOLI	09°22.514N;001°00.580W	162
36	129	EOZS/06/103-112	*Pona*	NANUMBA NORTH	KALANDE	08°35.988N;000°29.085W	174
37	130	EOYK/06/124-133	*Laribako*	NANUMBA NORTH	KALANDE	08°35.985N;000°29.084W	174
38	131	EOYS/06/109-111	*Laribako*	NANUMBA NORTH	KALANDE	08°35.988N;000°29.085W	174
39	131*	EOYS/06/112-113	*Pona*	NANUMBA NORTH	KALANDE	08°35.988N;000°29.085W	174
40	132	ASA	*Dente*	WINNEBA	BODWEASE	05°35'N 0°35'W	
41	132*	ASAOUT	*Muchumudu*	WINNEBA	BODWEASE	05°35'N 0°35'W	
42	133	EOYE/06/001-004	*Laribako*	EAST GONJA	SALAGA	08°33.060N;000°31.198W	158

Table 2 continued.

43	134	EOZT/06/185-193	*Pona*	TOLON/KUBUNGU	TINGOLI	09°22.514N;001°00.580W	162
44	134*	EOZT/06/194	*Laribako*	TOLON/KUBUNGU	TINGOLI	09°22.514N;001°00.580W	162
45	135	EOZB/06/002-011	*Pona*	NANUMBA NORTH	BIMBILLA	08°53.669N;000°04.793E	186
46	136	EOZ/06/035-044	*Pona*	NANUMBA SOUTH	NAKPAYILI	08°56.083N;000°00.53E	163
47	137	EOYT/06/204-211,213	*Laribako*	TOLON/KUBUNGU	TINGOLI	09°22.514N;001°00.580W	162
48	137*	EOYT/06/212	*Pona*	TOLON/KUBUNGU	TINGOLI	09°22.514N;001°00.580W	162
49	138	EOYT/06/284-288	*Laribako*	TECHIMAN	TECHIMAN	07°34.842N;001°56.846W	392
50	138*	EOYT/06/289-293	*Laribako*	TECHIMAN	TECHIMAN	07°34.842N;001°56.846W	392
51	139	EOZT/06/265,271-273	*Pona*	TOLON/KUBUNGU	KPALISOGU	09°23.930N;001°00.744W	156
52	139*	EOZT/06/266,269,270,274	*Pona*	TOLON/KUBUNGU	KPALISOGU	09°23.930N;001°00.744W	156
53	140	EOZ/06/013-119	*Pona*	EAST GONJA	SALAGA	08°33.060N;000°31.198W	158
54	140*	EOZ/06/014,019, 020	*Laribako*	EAST GONJA	SALAGA	08°33.060N;000°31.198W	158
55	141	EOYW/06/064-068	KULUKU	NANUMBA SOUTH	NAKPAYILI	08°56.083N;000°00.53E	163
56	141*	EOYW/06/069-073	KULUKU	NANUMBA SOUTH	NAKPAYILI	08°56.083N;000°00.53E	
57	142	EOYS/06/154,157,158,160,161	*Laribako*	NANUMBA NORTH	BIMBILLA	08°53.776N;000°02.536W	186
58	142*	EOYS/06/155,156,159	*Laribako*	NANUMBA NORTH	BIMBILLA	08°53.776N;000°02.536W	186
59	143	EOZT/06/235-244	*Pona*	TOLON/KUBUNGU	TINGOLI	09°22.514N;001°00.580W	162
60	143*	EOZT/06/240-242	*Laribako*	TOLON/KUBUNGU	TINGOLI	09°22.514N;001°00.580W	162
61	144	EOZKP/06/134-140	*Pona*	NANUMBA NORTH	BIMBILLA	08°53.776N;000°02.536W	186
62	144*	EOZP/06/141-144	*Laribako*	NANUMBA NORTH	BIMBILLA	08°53.776N;000°02.536W	186
63	145	EOZT/06/245-254	*Pona*	TOLON/KUBUNGU	TINGOLI	09°22.514N;001°00.580W	162
64	146	EOYW/06/098-106	LARBRAKO-LAA	NANUMBA SOUTH	NASAMBA	08°45.974N;000°03.329E	196

Table 2 continued.

65	146*	EOYW/06/103-104	*Pona*	NANUMBA SOUTH	NASAMBA	08°45.974N;000°03.329E	196
66	147	EOYT/06/234-243	*Laribako*	TOLON/KUBUNGU	TINGOLI	09°22.514N;001°00.580W	162
67	147*	EOYT/06/234-243	*Laribako*	TOLON/KUBUNGU	TINGOLI	09°22.514N;001°00.580W	162
68	148	EOYW/06/103-108	LARBRAKO-NYA	NANUMBA SOUTH	NASAMBA	08°45.974N;000°03.329E	196
69	148*	EOYW/06/109-114	*Pona*	NANUMBA SOUTH	NASAMBA	08°45.974N;000°03.329E	196
70	149	EOYB/06/004-008	*Laribako*	NANUMBA NORTH	PUSUGA	08°55.539N,000°.872W	209
71	149*	EOYB/06/009-011	*Laribako*	NANUMBA NORTH	PUSUGA	08°55.539N,000°.872W	209
72	150	EOYK/06/144-153	*Laribako*	NANUMBA NORTH	KALANDE	08°35.985N;000°29.084W	174
73	151	CRI*Pona*	*Pona*	EJISU	FUMESUA	06°41'N, 1°28'W	209
74	152	EOZS/06/155-160	*Pona*	NANUMBA NORTH	KALANDE	08°35.985N;000°29.084W	174
75	153	EOYW/06/055-063	KULUKU	NANUMBA SOUTH	NAKPAYILI	08°56.083N;000°00.53E	163
76	154	EOZT/06/285-294	*Pona*	TECHIMAN	TECHIMAN	07°34.842N;001°56.846W	392
77	155	EOZT/06/165-170	*Pona*	TOLON/KUBUNGU	TINGOLI	09°22.514N;001°00.580W	162
78	155*	EOZT/06/171-174	*Laribako*	TOLON/KUBUNGU	TINGOLI	09°22.514N;001°00.580W	162
79	156	EOYT/06/265-273	*Laribako*	TOLON/KUBUNGU	KPALISOGU	09°23.930N;001°00.744W	156
80	156*	EOYT/06/264	*Pona*	TOLON/KUBUNGU	KPALISOGU	09°23.930N;001°00.744W	156
81	157	EOZK/06/140-144	*Pona*	NANUMBA NORTH	KALANDE	08°35.985N;000°29.084W	174
82	158	EOYB/06/045-054	*Laribako*	NANUMBA NORTH	DAMON NA YILI	08°56.085N;000°00.053E	202
83	159	EOYT/06/194-203	*Laribako*	TOLON/KUBUNGU	TINGOLI	09°22.514N;001°00.580W	162
84	160	EOYT/06/244-253	*Laribako*	TOLON/KUBUNGU	TINGOLI	09°22.514N;001°00.580W	162
85	161	EOY/06/014-033	*Laribako*	NANUMBA NORTH	KALANDE	08°35.985N;000°29.084W	
86	162	EOYK/06/147-123	*Laribako*	TOLON/KUBUNGU	TINGOLI	09°22.514N;001°00.580W	

Table 2 continued.

87	162*	EOYK/06/144-146	*Pona*	TOLON/KUBUNGU	TINGOLI	09°22.514N;001°00.580W	
88	163	EOYT/06/214-223	*Laribako*	TOLON/KUBUNGU	TINGOLI	09°22.514N;001°00.580W	
89	164	EOZ/06/255-264	*Pona*	TOLON/KUBUNGU	KPALISOGU	09°23.930N;001°00.744W	156
90	165	MANKRONG *Pona*	*Pona*	-	-	-	
91	166	EOZB/06/020-029	*Pona*	TOLON/KUBUNGU	KPALISOGU	09°22.513N;001°00560W	

* It is the off-type of the grouping as given at time of collection.

3.3.1 Ethno-botany or Folk Taxonomy

Various research methods were used in order to have a clear picture of *Pona* classification. Focus group interviews (287 farmers), community-based Participatory Rural Appraisals (PRA) in 9 districts and direct participatory on-farm assessment (20 farmers). Key informants interviews involving 50 farmers and 30 agricultural extension agents and yam sellers were done. Semi-structured interviews with 200 farmers were carried out to quantify the indigenous taxonomy.

3.3.2. Morphological Characterization

Morphological characterization was conducted by combining *in-situ* characterization in farmers' fields at Kwamang Pepease –Forest zone and Bodwease (Coastal Savannah) and *ex-situ* characterization under controlled the controlled conditions at Crops Research Institute Experimental Station at Fumesua in 2006 and 2007. The *ex-situ* characterization was carried out to mop up some accessions that the researcher missed at collection time because farmers had planted already.

The assessment of the accessions at Crops Research Institute was conducted with active participation of 7 expert *Pona* farmers including, the Ashanti Regional Best Yam Farmer (2006) and farmers with not less than 15 years of *Pona* cultivation experience.

Fig. 2: In-situ characterization at Kwamang –Pepease, Ashanti Region.

The descriptors used for morphological characterization (Appendix 1) consisted mainly of descriptors selected by Hamon (1987) as the most pertinent for identification and description of cultivated yams in West Africa and those recommended by Martin and Rhodes (1978) and IPGRI/IITA (1997) as cited by Dansi *et al.* (1999) and modified to suit the study. Additional farmers' descriptors such as tuber striping or streaking, presence of specific Fig., presence of colour bands (Dansi *et al.*, 1999) were also incorporated. A total of 111 characters (Appendix 1) were scored on binary basis for differentiating the accessions.

Two farmer field days were organized each year at vegetative, flowering and harvesting stages to incorporate farmer knowledge of true *Pona* characteristics. Harvested yam tubers were cooked and its rheological properties (aroma, texture and taste) assessed. After assessment for year one, the expert yam farmers confirm and declassified some of

the accessions leading to reclassification of the collected accessions to 91 accessions (Table 2).

3.3.2.1. Data analysis

All data were converted to binary form using ALS Binary and clustering according to the Darwin 5.1.531 neighbour-joining approach.

A measure of diversity which takes into account proportional abundance and evenness of categories, the Shannon diversity index (H) (Shannon and Weaver, 1949), which assesses the diversity of each trait in the *Pona* yam collection, was assessed using the formula:

$$H = -\Sigma pi \ln pi$$

where:

H is the diversity index; ln is the natural logarithm; i is an index number for each trait present in a sample; pi is the number of individuals with the trait (ni) divided by the total number of individuals (N) present in the entire sample.

The means and standard errors of 26 quantitative traits were calculated for each trait. Principal Component analysis (PCA) was used to assess the range of genotypic variability in the population. In PCA analysis, the eigen values were correlation coefficient between original variable, (morphological or molecular data) with respective principal component (PC). Variation percentage expressed by each PC and accumulated variability percentage were also generated. A two-dimensional scatter diagram of the PCs with

variability >10% was then plotted to observe the dispersion of the accessions on the two PCs.

Cluster analysis was used to estimate pair-wise genetic similarity values among accessions using Jaccard's similarity coefficient. The similarity matrix was then used to construct a phenogram.

3.3.3 Molecular Analysis

3.3.3.1. DNA Extraction

Some accessions were lost (failure to germinate) before the molecular analysis, hence total genomic DNA was extracted from young freshly harvested leaves of 72 yam accessions from the experimental fields of the Crops Research Institute, Fumesua, Ghana using Qiagen DNeasy Plant Mini Protocol (Qiagen, 2006).

Hundred milligram (100mg) fresh weight of sample material was grinded in a mortar and pestle. This was followed by the addition of 400µl and 4µl of Buffer AP1 and RNase respectively. It was then vortexed and incubated in a preheated water bath at 65°C for 10 min during which the tube was inverted 3 times. Buffer AP2 was added to the tube, thoroughly mixed and incubated on ice for 5 minutes. The lysate was centrifuged for 5 min at 14,000 rpm. The lysate was pipetted into QIAshredder Mini spin column in a 2 ml collection tube and centrifuged at 14,000 rpm for 2 min. The flow-through fraction was transferred into a new tube, and 1.5 volumes of Buffer AP3/E added and mixed by pipetting. 650 µl of the mixture was transferred into a DNeasy Mini spin column in a 2ml collection tube and centrifuged at 8,000 rpm for 1 min and the flow-through discarded. Another 500 µl Buffer AW was added and centrifuged for 2 min at 14,000 rpm. The spin column was transferred to a new 2 ml microcentrifuge tube and 100 µl of Buffer AE added

for elution. The sample was then incubated at room temperature for 5 min, and centrifuged for 1 min at 8000 rpm.

3.3.3.2. DNA Quality Testing and Quantification

DNA extracts of 72 accessions were transferred to IITA in Nigeria for the molecular analysis. DNA quantity and quality was determined at the Biotechnology laboratory of the International Institute of Tropical Agriculture (IITA) using spectrophotometer (Beckman Coulter DU530) and the absorbance read at 260nm and 280nm (A_{260} and A_{280} respectively) levels. Each DNA sample was diluted ten times (2 µl DNA+ 18 µl Nuclease-free milli-Q water). The milli-Q water was used as a reference sample to set the spectrophotometer at 260 nm wavelength (blanking). A 10 µl volume of the diluted DNA sample was loaded to the curvette of the spectrophotometer for estimation of the concentration.

The quality of DNA was assessed using the absorbance ratio at 260 and 280 wavelengths (A_{260}/A_{280}). A DNeasy purified DNA has an A_{260}/A_{280} ratio of 1.7-1.9. An A_{260}/A_{280} ratio equal to 1.8 is generally considered to be pure (Qiagen, 2006). On the other hand an A_{260}/A_{280} value lower than 1.8 indicates the presence of impurities. The DNA concentration of the samples was determined using the double-stranded DNA standard of 1 A_{260} = 50 µg/µl of DNA. The working concentration of 2.5µg/µl DNA was prepared based on the estimated concentration obtained from the spectrophotometer reading. The volume of sample needed to be diluted for the working sample concentration was calculated based on the formula

$$V_1=C_2V_2/C_1$$

where C_1= initial concentration of the sample; C_2= required concentration of the sample; V_2 = required volume of the sample; and V_1 = volume of the initial concentration needed to be diluted to the required volume.

For samples with very weak concentration which required V_1 of greater or equal to 100 µl no further dilution was done. For samples requiring V_1 of <100 µl, the volume was taken and topped up to a final volume of 100µl.

3.3.3.3. Molecular Markers and Polymerase Chain Reactions

Amplifications were carried out in an automated thermal cycler (Peltier Thermal Cycler 200). The PCR conditions described by Kawchuk *et al.* (1996) were used with some modifications (Mignouna *et al.*, 2003). Twenty-one (21) SSR primer pairs were used for this study (Table 3). Amplification reactions were carried out in 20 µl reaction volumes each containing 5µl of master mix DNA, buffer (ammonium sulphate), deoxynucleoside triphosphates (dNTP), 1.7 mM $MgCl_2$, 0.4 mM each of forward and reverse SSR primer, and two units of Taq DNA polymerase (Promega) (Table 4.).

Reactions were conducted in a Thermal Cycler of Promega programmed for the following procedures- initial denaturing at 94°C for 4 min for one cycle, followed by 35 cycles of 94°C for 30 sec, 53.1°C for 1 min (annealing temperature depending on marker) and 72°C for 1 min. After the 35 cycles, that is the primer extension stage, the samples were held at 72°C for 7 min followed by 60°C for 30 min and then stored at 4°C. PCR optimization was performed for all markers and best performing conditions identified. The optimization PCR reactions were carried out in 5 µl.

3.3.3.4 DNA Fragment Analysis

Capillary electrophoresis was performed using a semi-automated system of ABI 3100 Genetic Analyzer in a 36 cm capillary array using POP 4 (Performance Optimized Polymer) matrix to separate amplified PCR products. Genotyping was done using Liz + formamide at a ratio of 1: 94 per reaction as a size standard. Four negative controls (W1-W4) and four already genotyped yam genotypes were deliberately added to the 72 accessions to test the extent to which the GeneMapper™ procedures could classify the accessions.

Table 3: Set of Yam Microsatellite markers (SSR) used in fingerprinting

SNo	Marker name	Primer sequence (5'-3')[a]	Dye*
1	*YM-13*	*TTCCCTAATTGTTCCTCTTGTTG (F)* *GTCCTCGTTTTCCCTCTGTGT (R)*	
2	*YM-26*	*AATTCGTGACATCGGTTTCTCC (F)* *ACTCCCTGCCCACTCTGCT (R)*	*PET*
3	*Dpr3D06*	*ATAGGAAGGCAATCAGG (F)* *ACCCATCGTCTTACCC (R)*	*VIC*
4	*Da1F08*	*AATGCTTCGTAATCCAAC (F)* *CTATAAGGAATTGGTGCC (R)*	*PET*
5	*YM-15*	*TACGGCCTCACTCCAAACACTA (F)* *AAAATGGCCACGTCTAATCCTA (R)*	*VIC*
6	*Dab2E09*	*AACATATAAAGAGAGATCA (F)* *ATAACCCTTAACTCCA (R)*	*PET*
7	*Da1D08*	*GATGCTATGAACACAACTAA (F)* *TTTGACAGTGAGAATGGA (R)*	*6-FAM*
8	*Da1A01*	*TATAATCGGCCAGAGG (F)* *TGTTGGAAGCATAGAGAA (R)*	*VIC*
9	*Dpr3F04*	*AGACTCTTGCTCATGT (F)* *GCCTTGTTACTTTATTC (R)*	*PET*
10	*Dab2D06*	*TGTAAGATGCCCACATT (F)* *TCTCAGGCTTCAGGG (R)*	*VIC*
11	*Dab2C05*	*CCCATGCTTGTAGTTGT (F)* *TGCTCACCTCTTTACTTG (R)*	*PET*
12	*Da1C12*	*GCCTTTGTGCGTATCT (F)* *AATCGGCTACACTCATCT (R)*	*NED*
13	*Dpr3F10*	*TCAAAGGAATGTTGGG (F)* *ACGCACATAGGGATTG (R)*	*VIC*
14	Dpr3B12	CATCAATCTTTCTCTGCTT (F) CCATCACACAATCCATC (R)	NED
15	Dpr3F12	TCCCCATAGAAACAAAGT (F) TCAAGCAAGAGAAGGTG (R)	NED
16	Dab2D08	ACAAGAGAACCGACATAGT (F) GATTTGCTTTGAGTCCTT (R)	6-FAM
17	YM-5	AATGAAGAAACGGGTGAGGAAGT (F) CAGCCCAGTAGTTAGCCCATCT (R)	6-FAM
18	YM-28	GGAGTGCGGGGAGAGGAG (F) CGGCGTGAGCTATTGGTGTGT (R)	VIC
19	YM-1	TTGTCAGCGAAATAAGCAGAGA (F) CAACAGACGCAGCCCAACT (R)	6-FAM
20	YM-19	CCACCCTCTACCTCAAGT (F) GAGGCTTCTCCCACTAAGT (R)	NED
21	Dab2E07	TTGAACCTTGACTTTGGT (F) GAGTTCCTGTCCTTGGT (R)	PET

NB: Italicized primers were polymorphic and vice versa.

[a]F, forward primer; R, reverse primer.

Table 4: Optimized grid for yam PCR: Volume of reagents for a total volume of 20µl per reaction.

Serial Number	Marker	T°C	WATER	10X BUFFER (NH_4)	$MgCl_2$ 25mM	dNTP 2mM	F 10µM	R 10µM	TAQ 5µ/µ	DNA 2.5ng/µl
1	*YM-13*	*47.4*	*12.7*	*2*	*2.2*	*1*	*0.5*	*0.5*	*0.1*	*1*
2	*YM26*	*55.3*	*12.7*	*2*	*2.2*	*1*	*0.5*	*0.5*	*0.1*	*1*
3	*Dpr3D06*	*42*	*12*	*2*	*1.2*	*2*	*0.75*	*0.75*	*0.3*	*1*
4	*Da1F08*	*47.4*	*12.7*	*2*	*2.2*	*1*	*0.5*	*0.5*	*0.1*	*1*
5	*YM-15*	*53*	*12.1*	*2*	*1.2*	*1.5*	*0.5*	*0.5*	*0.2*	*2*
6	*Dab2E09*	*42*	*12.7*	*2*	*2.2*	*1*	*0.5*	*0.5*	*0.1*	*1*
7	*Da1D08*	*47.6*	*12.7*	*2*	*2.2*	*1*	*0.5*	*0.5*	*0.1*	*1*
8	*Da1A01*	*47.1*	*12.7*	*2*	*2.2*	*1*	*0.5*	*0.5*	*0.1*	*1*
9	*Dpr3F04*	*41.9*	*11.6*	*2*	*1.7*	*1*	*0.75*	*0.75*	*0.2*	*2*
10	*Dab2D06*	*48*	*12.1*	*2*	*1.2*	*1.5*	*0.5*	*0.5*	*0.2*	*2*
11	*Dab2C05*	*49*	*12.7*	*2*	*2.2*	*1*	*0.5*	*0.5*	*0.1*	*1*
12	*Da1C12*	*50*	*12.7*	*2*	*2.2*	*1*	*0.5*	*0.5*	*0.1*	*1*
13	*Dpr3F10*	*45.9*	*12.7*	*2*	*2.2*	*1*	*0.5*	*0.5*	*0.1*	*1*
14	*Dpr3B12*	*47.4*	*9.5*	*2*	*2.2*	*2*	*1*	*1*	*0.3*	*2*
15	*Dpr3F12*	*47.7*	*12.7*	*2*	*2.2*	*1*	*0.5*	*0.5*	*0.1*	*1*
16	*Dab2D08*	*47.4*	*12.7*	*2*	*2.2*	*1*	*0.5*	*0.5*	*0.1*	*1*
17	*YM-5*	*56.5*	*12.7*	*2*	*2.2*	*1*	*0.5*	*0.5*	*0.1*	*1*
18	*YM-28*	*60*	*9.5*	*2*	*2.2*	*2*	*1*	*1*	*0.3*	*2*
19	YM-1	54.3	12.1	2	1.2	1.5	0.5	0.5	0.2	2
20	YM-19	53.1	12.1	2	1.2	1.5	0.5	0.5	0.2	2
21	Dab2E07	48.8	9.5	2	2.2	2	1	1	0.3	2

NB: Italicized primers were polymorphic.

3.3.3.4.2. Size matching, binning, allele size calling and verification

The GeneMapper™ version 3.7 software (Applied Biosystems Inc., Foster City, CA 94404, USA) was used to generate genotype plots. Size matching/calling was based on Local Southern Method algorithm with reference to a defined standard range, GS75-500(-250) Liz base pairs. To reduce bin sizes and increase inter-bin distances to enhance efficient automated binning, Ghosh *et al.* (1997) external adjustments was used to verify and augment the internal GS75-500(-250) Liz size standards. For each marker, alleles for the data set was sorted according to size and "tolerance level" of 0.4 base pairs selected as the minimum allowable distance between adjacent bins in base pairs. When the difference between two sequentially sized alleles is greater than the set tolerance level a new bin is created. This procedure was conducted for each marker until all alleles were binned with the smallest and largest sized alleles for any marker representing the start of the first bin and the end of the last bin. After grouping the alleles, the mean and ranges were calculated for all bins. The bin labels, which represent the mean sizes rounded up to the nearest whole number was assigned to each group. This data was then submitted manually to the GeneMapperTM software to adjust the bins. Figure 3 is a representative SSR profile obtained for three yam accessions with marker Dab2E07, as analyzed using genotyping GeneMapperTM v.3.7 (Applied Biosystems, USA).

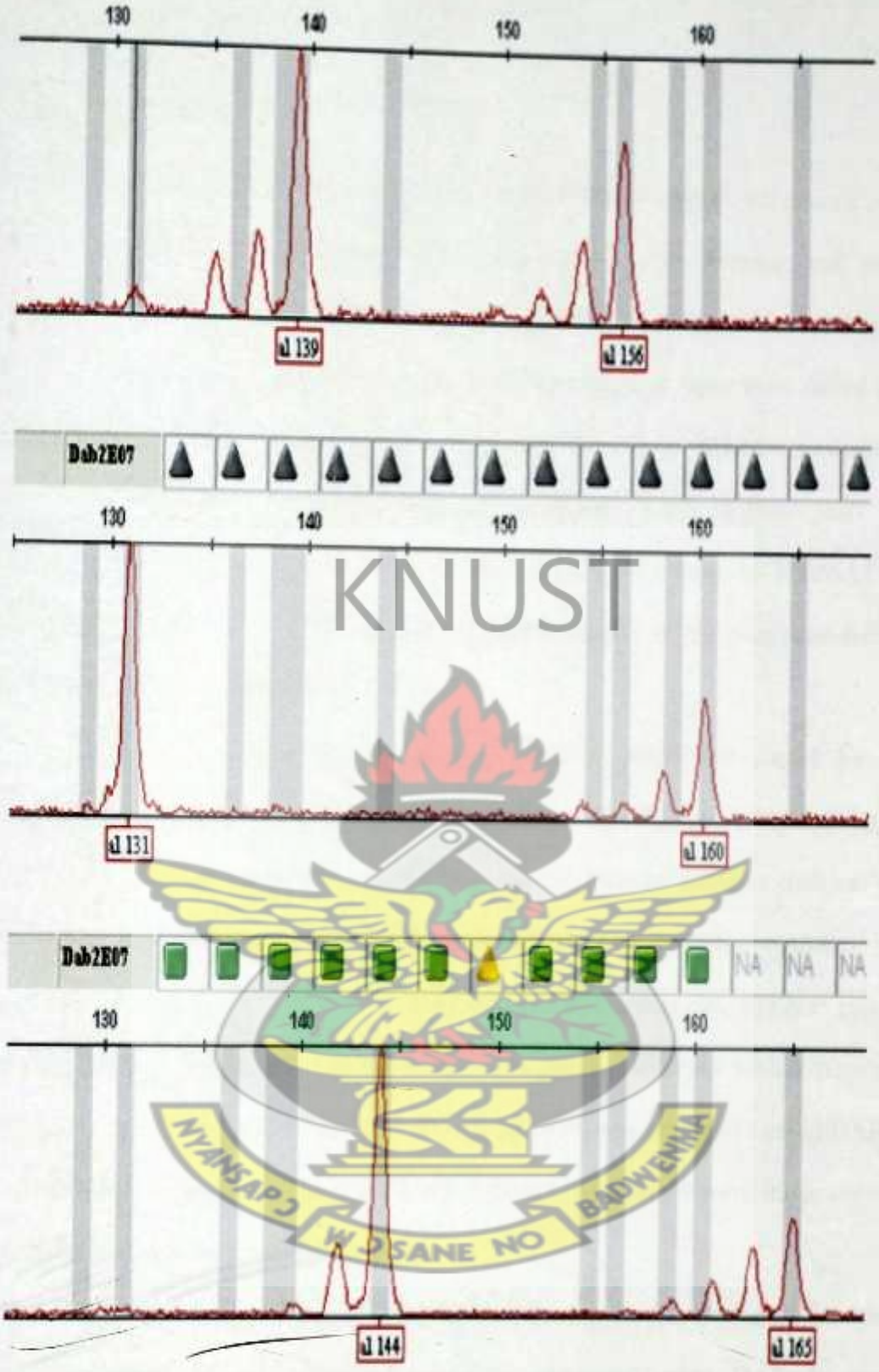


Fig. 3: Sample of SSR profiles obtained for three yam accessions with marker Dab2E07 and analysed using genotyping GeneMapperTM v. 3.7 (Applied Biosystems, USA)

3.3.3.4.2 Data Analysis

The polymorphic 13 SSR markers (Table 3) and 72 *Pona* complex accessions plus 4 checks from IITA yam collections were subjected to gene diversity and genetic differentiation analysis with 4.0% missing data over all loci and accessions. Observed allelic data were binned into discrete units and SSR fragment sizes were called using GeneMapperTM software v.3.7 (Figure 4). For the following statistical analysis, the fragment sizes generated by GeneMapperTM software v.3.7 in base pairs were converted to binary data using the software "ALS Binary" developed by ICRISAT. The use of binary format was determined by polyploid nature of the crop and different ploidy levels of studied accessions.

The presence (1) or absence (0) of individual allele was scored for each genotype across all SSR markers used for the study. Missing data accounted for less than 5% (i.e. marker × genotype) of the entire data set. Pairwise distance matrices were computed using the Jaccard similarity coefficient. The resulting matrices were subjected to unweighted Neighbour-Joining method (Saitou and Nei, 1987) to generate structured tree (dendogram). The structure of the genetic diversity within population was further analysed by factor analysis (PCA). Analysis was performed using DARwin 5.0.153 software (developed by CIRAD) and SAS v 9.1. Neighbour-joining approach was employed in classifying the accessions.

CERVUS 3.0.3 software was used in conducting allelic frequency analysis. Factorial analysis of correspondence was also conducted using DARwin version 5.0.153 to determine the genetic relationships among individuals. To determine the similarity among the individuals, a neighbour-joining tree (Saitou and Nei, 1987) was constructed based on the shared allele distance (D_{AS}) between individuals using the DARwin version 5.0.153 computer software. Cophenetic correlation values were calculated to evaluate the robustness of the resulting tree topologies.

3.4 Core Collection Determination

The core collection was defined as *Pona* complex core collection and its size 10% of the whole accession. Based on the ethno-botany, morphological and molecular data this domain was further categorized based on the six classes identified: authentic *Pona*, *Laribako*, *Pona* hybrids (*Numbo*), *Laribako* hybrids (*Fuseini*), *muchumudu* and *kulunku*. The core collection size was assigned as 10% each for the stated categories.

CHAPTER 4

4.0 RESULTS AND DISCUSSIONS

4.1 Folk Taxonomy and Classification Based on Indigenous Expertise

Most named yam landraces are morphologically distinct (Tamuri *et al.*, 2008). Farmers over the years have also identified the nutritional, medicinal and other values of each folk species (Mekbib, 2007). The *Dioscorea rotundata-cayenensis* complex is indigenous to West Africa. Yam farmers in this region including Ghana have therefore grown the crop from time immemorial and therefore are in a unique position to contribute positively to plant breeding in general and yam breeding in particular (Otoo, 2001). The benefits that are implicit from folk names or taxonomy are the reflection of associated values of each folk species (Mekbib, 2007).

Generally, there were two schools of thought among the farmers from whom the yams were collected. Some farmers (60%; N= 287) claimed *Pona* was a big group comprising *Pona*, *Laribako*-nya and *Laribako*-la, others (40%; N= 287) claimed *Laribako* was rather the main group with *Pona*, laribriko-nya (female *Laribako*) and *Laribako*-la (male *Laribako*) as sub-groups; hence both accessions of *Laribako* and *Pona* represent the *Pona* complex (Table 5). In both instances, therefore, farmers claimed that *Laribako* and *Pona* are closely linked, agreeing with Dansi *et al.*, (1999) grouping of *Pona* and *Laribako* as cultivars of *Pona* group. Another group of yams *Kulunku* (Gonja language meaning fast-fast), which farmers claim belong to *Pona* group but matures faster (3 months) than *Pona* (4-6 months) and produces a characteristic one tuber even though several sprouts will emerge from a seedyam planted, was also included in the *Pona* complex. Farmers' classification of the *Pona* complex can therefore be schematically represented as in Fig.4.

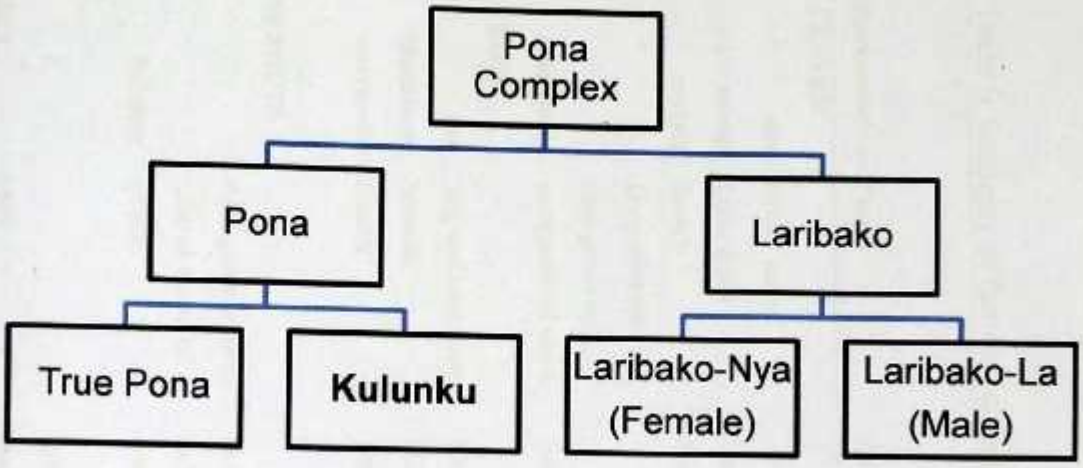


Fig. 4: Schematic representation of folk taxonomic classification of the *Pona* complex.

A total of 27 characters were listed by farmers whom the yams were collected from and the expert farmers as basis for classification of *Pona* complex (Table 5). These characteristics ranged from leaves through vines and fresh tuber to seedyam differences. In each agro-ecology, farmers generally agreed in most of the characteristics listed with a few exceptions, such as shade of leaf colour, and colour of tuber flesh. This was understandable due to differences in individual perception of colour. There were complete agreement with regard to the external characteristics such as tuber shape, presence of cracks on tuber, and characteristics of cooked tuber.

From the results of indigenous taxonomy and knowledge it can be concluded that the term *Pona* is a complex and refers to a distinct group of *D. rotundata* variety with great morphological diversity. This concept is similar to the assertion regarding complexity of Guinea yam in general (Dansi *et al.*, 2000). Different authors consider Guinea yam to be represented by one species, two species, or even a species complex (Martin and Rhodes, 1978; Onyilagha and Lowe, 1985; Hamon and Toure, 1990; Hamon *et al.*, 1992; Asemota *et al.*, 1996).

Table 5: Summary of farmers' criteria for classifying *Pona* and *Laribako*

Parameter	*Pona*	*Laribako*	Per Cent Farmers by Agroecology (%) CS (N=40)	F (N=60)	FST (N=60)	GS (N=120)	Expert Yam Farmers (N=7)
LEAVES							
size	Broad leaves	small and narrow leaves	100	100	100	100	100
colour	Light green	Dark green	80	50	70	70	100
surface	Rough	Smooth and shiny leaves	60	70	90	80	90
emergence	Long after sprout emergence and elongation of vines	Emerge with sprouts	100	100	100	100	100
VINES							
size	Big and long vines	Slender and long vines	100	100	100	100	100
thorniness	Sporadic	Nil	100	100	100	100	100
internodes	Shorter	longer	60	60	80	80	100
SPROUTS							
Number	*Pona* gives as many tuber as the no of sprouts	No matter the no of sprouts, larbrako will always give just a tuber	60	80	70	80	100
SEX	Male	Both sexes: female-*Laribako* nya and male-larebrako la	100	100	100	100	100

NB: CS = COASTAL SAVANNAH; F = FOREST; FST = FOREST SAVANNAH TRANSITION; GS = GUINEA SAVANNAH; N = POPULATION SIZE

Table 5 continued.

FRESH TUBER SKIN							
size	Thick	Thin	100	100	100	100	100
colour	Dark	Light	100	100	100	100	100
texture	Rough	Smooth	100	100	100	100	100
Number and size	Few but big tubers	Several small tubers	100	100	100	100	100
Flesh colour	White –yellow	White	80				
Maturity	Late maturing (7-8 months)	Early maturing (6-7 months)	100	100	100	100	100
Surface	Rough and deep grooves	Smooth and shallow grooves	100	100	100	100	100
Shape	Big head and tapers to the tail	Wide head with narrow crown	100	100	100	100	100
Hardiness	Hard	Soft	60	100	100	100	100
Crown							
Size	Big	Small	100	100	100	100	100
Position	Inserted in the head of tuber	Attached to the head of tuber	70	90	80	80	100

NB: CS = COASTAL SAVANNAH; F = FOREST; FST = FOREST SAVANNAH TRANSITION; GS = GUINEA SAVANNAH; N = POPULATION SIZE

Table 5 continued.

Cooked tuber							
Texture	Grainy	Very smooth	100	100	100	100	100
Hardiness	Hard	Soft	100	100	100	100	100
Taste	Less tasty and starchy than larbrako	Tastier and more starchy	100	100	100	100	100
Colour	Off-white to yellow	White	100	100	100	100	100
Aroma	Sweet scented	More sweet scented	100	100	100	100	100
Seedyam							
Colour	Red-headed	White-headed	70	90	80	100	100
Special Feature	*Kulunku* type of *Pona* *Kulunku* means fast fast						

The results also lend credence to the folklore from the Guinea Savannah agroecology surrounding the name *Laribako* which suggests that it is a "feminine" or "smaller" type of *Pona*. There are two schools of thoughts explaining the name *Laribako*. One school of thought explained that a yam farmer in the Guinea Savannah who had a daughter "Lariba" whom he reserved the most beautiful and small tubers for. He will often call his daughter and ask her in Gonja language "ko" meaning take; hence *Laribako*. The second school explained that the yams were as beautiful as the farmer's daughter, Lariba, hence *Laribako* meaning "*Laribako*'s own".

4.2 Morphological characterization

Morphological analysis based on aerial and underground organs has shown significant diversity within the cultivated yams of the *D. cayenensis/D. rotundata* complex (Dansi *et al.*, 1999). Generally, the *Pona* complex is in *D. rotundata* group and as such their vines are characteristically wingless and stems climbed clockwise. Results of the morphological analysis also confirmed the great morphological diversity among the *Pona* accessions as suggested by indigenous taxonomy (Table 6).

Some of the accessions that farmers classified as *Laribako* were morphologically identical to *Pona* and therefore clustered together, and vice versa.

Table 6: Range of morphological variability in *Pona* complex accessions in Ghana.

Parameter	Range of variability	
	Laribako	**True *Pona***
Leaves	Small and saggitate (Fig. 5a) light green shiny smooth surface already formed at emergence of sprout	Large and cordate (Fig. 5b) deep green rough and dull develops long after emergence of sprout
Vines	Spineless	vary from long slender to long big, with or without spines (Figs. 6a and 6b)
Sprout	sole sprouting	sole sprouting (true *Pona*) to multiple sprouting (*Kulunku*)
Tuber Shape	Cylindrical (Fig. 7)	oval-oblong (Fig. 7).
Tuber flesh colour	White	white with yellowish tint (Fig. 8)
Seedyam crown colour	White (Fig 9)	Purple (Fig 9)
Cooked tuber	very soft and very mealy	soft and mealy

A. Characteristic narrow sagitate long leaves of *Laribako*

B. Characteristic broad cordate leaves of *Pona*

Fig. 5: Range of leaf shape of the *Pona* complex

A. Smooth and spineless vines of *Laribako*

B. Rough and spiny vines of *Pona*

Fig. 6: Range of vine characteristics of *Pona* complex

Fig. 7: Varying tuber shape of *Pona* complex: characteristic oval-oblong shape (big-head and tapering tail) of *Pona* (below), and cylindrical shape of *Laribako* (above).

Fig. 8: Varying tuber flesh characteristics of *Pona* complex

Fig. 9: Varying head colouration of Pona complex seedyam: purple of Pona and white of Laribako.

4.3 Diversity Assessment

The diversity of the collection was also assessed using Shannon's diversity index. This index provides a measurement of both allelic richness and evenness. The Shannon's diversity values (H) can range from 0 to 1; with a value near 0 indicating that the trait was same for all accessions. Conversely, a value near 1 indicates that the trait was very diverse among the accessions. Hence, the higher the value of Shannon's diversity index, the greater the genetic diversity within the population. There was a wide polymorphism range in quantitative the descriptors recorded (Table 7).

Results from Shannon's diversity index analysis show that there were wide genetic diversity between number of internodes to first branching, stem diameter, spine length,

stem length, internode length, female flower length, male flower length, internode number, tuber length and tuber width, suggesting that the accession can be classified based on the aforementioned parameters.

Again, wide polymorphism was recorded for several of the qualitative morphological characteristics which reflected in their Shannon's Diversity Index values (Table 8). Wide range of polymorphism was observed in the following qualitative characteristics such as stem length, , stem diameter, internode number internode length, branching, presence of waxiness/ pruinose, first leaf emergence, number of leaves, maturity (tubers) after emergence in months, tuber size, corm position on tuber, tuber shape, tendency of tuber to branch, hardness of tubers, skin colour at the head of the tuber, total number of tubers harvested, total weight of harvested tubers, ease of peeling, cooking time to softness, texture of cooked tuber and overall assessment of the cooked tuber.

Days to emergence for instance, had wide range of polymorphism (6 classes; SDI= 1.00). *Kulunku* accessions emerged first (mean =7 DAP) followed by *Laribako* (mean =10 DAP.), *Pona* (Mean =12 DAP.), Mankrong*Pona* (Mean = 13.7DAP), *Muchumudu* (Mean =14 DAP) and *Dente* (Mean =16 DAP) in decreasing order of emergence.

Again, accessions with more internodes had larger stems than the accessions with fewer internodes. The *Muchumudu* accessions, for instance, had larger stems with few internodes as compared with other accessions and also had a characteristic striation across the cross-section.

All the *Pona* and *Laribako* accessions, all the accessions that exhibited their sexuality were males. All *Pona* and *Laribako* accessions including *Laribako*-nya (which) farmers claimed were female did not show their sexuality. This could be explained that this

group of accessions may be the female *Pona-Laribako* accessions which failed to exhibit their sexuality because of unfavourable environmental conditions.

Table 7: Phenotypic diversity and Shannon diversity index of quantitative traits of *Pona* Complex in Ghana

Sno	Quantitative traits	Mean	SE	No of Classes	Shannon's DI
1	Number of sprouts/seedyam	1.1	0.6	6	0.82
2	Number of internodes to first branching	1.9	0.4	3	1.00
3	Number of surviving plants at 4WAP	6.3	2.6	8	0.67
4	Number of leaves (30 days after emergence)	53.5	4.8	6	0.36
5	Number of veins per leaf	6.9	0.2	10	0.50
6	Stem diameter at 15 cm from the base(cm)	4.8	0.5	3	1.00
7	Terminal leaflet length (cm)	11.7	0.1	6	0.51
8	Terminal leaflet width (cm)	8.6	0.2	6	0.64
9	Number of branches	17.4	4.7	15	0.57
10	Spine length (cm)	1.1	0.3	3	1.00
11	Stem length (cm)	110.1	5.8	3	1.00
12	Internode length (cm)	10.7	1.3	5	1.00
13	Female flower length (mm)	5.1	0.1	3	1.00
14	Female flower diameter (cm)	0.5	0.1	3	0.39
15	Male flower length (cm)	4.5	1.5	4	1.00
16	Male flower diameter (mm)	0.7	0.1	3	0.61
17	Internode number	8.7	3.1	8	1.00
18	Days to flowering from emergence	72.6	1.7	6	0.61
19	Number of inflorescence per plant	1.5	0.7	3	1.00
20	Average length of inflorescence (cm)	1.3	0.5	3	0.44
21	Number of tubers per hill	1.3	0.48	3	0.64
22	Tuber length (cm)	40.8	0.86	4	1.00
23	Tuber width (cm)	27.5	0.56	4	1.00

Table 8: Phenotypic diversity and Shannon diversity index of qualitative traits of *Pona* Complex in Ghana

Descriptor No.	Descriptor	Number of Classes	Shannon's Diversity Index
7.1.2	Stem length [cm]. Assessed at 20 days after emergence. Mean of 10 plants	3	1.00
7.1.3	Internode number [number]	9	1.00
7.1.4	Stem colour	3	0.69
7.1.5	Absence/presence of waxiness	2	0.69
7.1.8	Absence/presence of hairs	1	0.54
7.1.9	Absence/presence of spines	2	0.03
7.1.10	Absence/presence of coloured spot at spine base	2	0.03
7.1.11	Absence/presence of barky patches	2	0.01
7.1.12	Plant type	1	0.03
7.1.13	Vigour	3	1.00
7.1.14	Twining habit	1	0.00
7.1.15	Twining direction	1	0.00
7.1.16	Stem height	3	1.00
7.1.17	Number of stems per plant	3	1.00
7.1.18	Stem colour	3	0.61
7.1.19	Number of internodes to first branching	3	1.10
7.1.20	Branching (2.4.4)	12	0.42
7.1.21	Stem diameter [cm] (recorded at 15 cm from the base of the plant	6	1.00
7.1.22	Stem cross-section shape at base	1	0.70
7.1.23	Internode length [cm] (Recorded at 1m height. Average of 5 plants).	4	1.00
7.1.24	Absence/presence of waxiness	2	1.00
7.1.34	Spines on stem base (2.4.6, 2.4.7)	3	1.00
7.1.35	Spines on stem above base	3	0.03
7.1.36	Spine position	4	0.02
7.1.37	Spine shape	4	1.00
7.1.38	Spine length	4	0.01
7.1.39	Absence/presence of coalescent spines	4	0.02
7.1.40	Colour of spot at spine base	4	0.02
Mature leaves			
7.2.1	First leaf emergence	3	1.00
7.2.2	Number of leaves	7	0.00

Table 8 continued.

7.2.3	Leaf colour	4	0.01
7.2.4	Leaf margin colour	1	0.00
7.2.5	Vein colour	1	0.02
7.2.6	Petiole colour	3	0.01
7.2.7	Petiole wing colour	3	0.01
7.2.9	Position of leaves	1	0.00
7.2.10	Leaf density	4	0.02
7.2.11	Internode number to fully expanded leaf [number]	3	0.00
7.2.12	Leaf type (2.5.3)	1	0.70
7.2.12.1	Leaf margin	1	0.70
7.2.12.2	Leaf lobation	2	1.00
7.2.14	Leatheriness	1	0.70
7.2.15	Leaf colour	4	1.00
7.2.16	Leaf vein colour (upper surface)	2	1.00
7.2.17	Leaf vein colour (lower surface)	2	0.04
7.2.18	Leaf margin colour	1	0.70
7.2.19	Hairiness of upper surface (2.5.6)	1	0.00
7.2.20	Hairiness of lower surface (2.5.7)	2	0.00
7.2.21	Waxiness of upper/lower surface (2.5.8, 2.5.9)	2	0.00
7.2.22	Leaf shape (2.5.5)	2	0.69
7.2.23	Leaf apex shape	1	0.70
7.2.24	Undulation of leaf	2	0.87
7.2.25	Distance between lobes	2	0.54
7.2.26	Upward folding of leaf along main vein	1	0.42
7.2.27	Downward arching of leaf along main vein	1	0.56
7.2.28	Upward folding of leaf lobes to form a cup	2	1.00
7.2.29	Downward arching of leaf lobes	1	0.68
7.2.30	Leaf measurement [cm] (2.5.4) Observed on 20 adult leaves	3	1.00
7.2.31	Position of the widest part of the leaf	3	0.74
7.2.32	Tip length (2.6.5)	3	0.65
7.2.33	Tip colour	2	0.12
7.2.34	Petiole length	3	0.11
7.2.35	Petiole length in correlation to leaf blade	3	0.14
7.2.36	Hairiness of petiole	1	0.00

Table 8 continued.

7.2.37	Petiole colour	3	0.09
7.2.38	Petiole wing colour	2	0.03
7.2.39	Spininess of petiole (2.5.13)	1	0.00
7.2.40	Absence/presence of stipules (2.6.1)	1	0.70
7.3.2	Days to flowering after emergence [d]	2	0.61
7.3.3	Sex (2.7.2)	2	0.63
7.3.4	Inflorescence position	1	0.63
7.3.5	Number of inflorescences per plant (2.7.4)	3	1.00
7.3.6	Inflorescence smell	1	0.63
7.3.7	Inflorescence type	3	0.68
7.3.8	Average length of inflorescence (2.7.3)	2	1.00
7.3.9	Number of inflorescences per internode	1	1.00
7.3.10	Number of female flowers per inflorescence (2.7.5)	1	1.00
7.3.11	Flower colour (2.7.6)	2	0.03
7.3.12	Female flower length (2.7.7)	2	1.00
7.3.13	Female flower diameter [mm]	2	0.00
7.3.14	Male flower diameter (2.7.8)	3	0.69
7.3.15	Male flower length [cm]	8	1.00
7.4.1	Fruit formation (2.8.1)	2	0.18
Adult Plants			
7.6.2	Maturity (tubers) after emergence in months; ≤6=0; >6=1	3	0.61
7.6.4	Number of tubers per hill; 0-1=1; 2-5=1	2	0.64
7.6.5	Relationship of tubers; 1-2=0;3=1	3	0.26
7.6.6	Absence/presence of corms; 1=0; 2=1	1	0.39
7.6.7	Corm size (In relation to tuber size); 3-5 =0, 7 =1	3	1.00
7.6.8	Corm ability to be separated from tuber; NO =0, YES = 1	2	0.26
7.6.9	Corm type; 1 =0, 2 =1	2	1.00
7.6.10	Absence/presence of rhizome	1	0.00
7.6.11	Spininess of roots; 3=0; 7=1	2	1.00
7.6.12	Absence/presence of anchor roots	2	0.50

Table 8 continued.

7.6.13	Sprouting at harvest; no=0; yes= 1	1	0.00
7.6.14	Tuber shape; 2-3=0; 4=1	4	1.00
7.6.15	Tendency of tuber to branch; 3 =0, 5-7 = 1	3	0.32
7.6.16	Place where tuber branches; 1=0, 2=1	3	0.55
7.6.19	Roots on the tuber surface; 3 =0, 7 =1	3	0.26
7.6.19.1	Spiny roots on the tuber surface; 3=0, 7=1	1	0.00
7.6.20	Place of roots on the tuber; 1 =0, 2 =1	3	1.00
7.6.21	Prickly appearance of the tuber; NO = 0, YES =1	1	1.00
7.6.22	Wrinkles on tuber surface; 3 =0, 7=1	2	1.00
7.6.23	Absence/presence of blisters on tuber surface	3	3.00
7.6.24	Absence/presence of cracks on the tuber surface	2	1.00
7.6.25	Tuber skin thickness; <1mm=0, ≥1 mm	2	0.60
7.6.26	Tuber skin colour (beneath the bark); 1=0, 2=1	3	0.00
7.6.27	Sprouting [%] <80=0; ≥80=1	2	1.00
7.6.27.1	Number of months	3	1.00
7.6.28	Hardness of tuber; (When cut with a knife) 1=0, 2=1	2	1.00
7.6.29	Skin colour at head of the tuber; white=0; purple=1	2	1.00
7.6.30	Flesh colour at central transverse cross-section; 1-0; 2-3 =1	3	0.69
7.6.31	Flesh colour of lower part of tuber; 1=0;2-3=1	3	0.69
7.6.32	Uniformity of flesh colour in cross-section; (From cortex to centre), no=0; yes= 1	2	0.69
7.6.33	Texture of flesh; 1=0;2-3=1	3	0.69
7.6.34	Time for flesh oxidation after cutting; 1=0;2-3=1	3	0.69
7.6.35	Flesh oxidation colour; 1=0;2-3=1	1	0.69
7.6.36	Amount of gum released by cut tuber; 3=0, 5-7= 1	3	0.48
7.6.37	Ability of cut tuber to irritate human skin (When tuber is rubbed on the arm). 3=0, 7=1	2	1.00
8.1.1	Total number of harvested tubers; ≤10 =0, >10 =1	2	1.00
8.1.2	Total weight of harvested tubers [kg]; <25=0, ≥25 =1	2	1.00
8.3.1	Ease of peeling; 1=0, 2=1	2	0.89
8.3.2	Preferred cooking method; 1=0, 2=1	1	0.33
8.3.2.1	Poundability of boiled tuber; 1=0, 2=1	1	0.44
8.3.3	Cooking time to softness [min]; <30 =0, ≥30 =1	2	1.00
8.3.4	Discolouration of cooking water; 1=0;2-3=1	2	0.69

Table 8 Continued.

8.3.5	Appearance of tuber after cooking; 3-5 =0, 7=1	2	1.00
8.3.6	Colour of tuber after cooking; 1= 0, 2=1	2	0.69
8.3.7	Attractiveness of cooked tuber; 3-5 =0, 7=1	2	0.69
8.3.8	Erosion of tuber upon cooking; no=0; yes= 1	2	0.69
8.3.9	Texture of cooked tuber; 1 =0, 2-3 =1	3	0.00
8.3.10	Stickiness of cooked tuber; 1= 0, 2=1	2	0.69
8.3.11	Flavour of cooked tuber; 0=0, 1-2 =1	2	0.69
8.3.14	Absence/presence of moisture on cooked tuber; no=0; yes= 1	2	0.69
8.3.15	Overall assessment of cooked tuber; 3-5 =0, 7=1	2	0.69

The information obtained from scoring the 111 morphological traits was translated into qualitative data with different character states being present or absence (Table 2).

In multivariate analysis, the diversity of a dataset can be represented using factorial analysis and tree methods. The factorial analyses aim mainly to give an overall representation of diversity and not in the individual effects. On the other had, tree methods tend to represent individual relations faithfully and may be less accurate for the global structure. Thus, two different ways of viewing the data and must be considered complementary rather than concurrent (Perrier *et al.*, 2003). Hence, ordination method was combined with clustering procedure.

Factorial analysis using the principal component analysis showed that out of the 82 Principal Component Axes only 30 accounted for all the variation and only 8 had Eigen values greater than 1.00 accounting for 92.44%. The scree test, however, revealed first 3 PCAs were the most important accounting for 81.35% of variation (Table 10). Hence assessment was based on the first three principal component axes.

Table 9: Eigen values of morphological data and per cent variation accounted for by the first three principal component axes of the ordination of accessions.

PC	Eigenvalue	Difference	% of variance	Accumulated variability (%)
1	36.21	11.34	42.12	42.12
2	24.88	16.01	28.93	71.04
3	8.87	5.74	10.31	81.35

In yams, general appearances of the plant and the tuber are of great importance in identification of cultivars (Dansi *et al.*, 1999). Tuber characteristics- tuber shape, presence of wrinkles on tuber, presence of cracks on tuber surface, tuber skin colour, tuber flesh colour, tuber flesh texture, skin thickness, tuber hardness and maturity period accounted for the variation of PCA1. Spine characteristics - presence of spines, spine length,position and size accounted for variation observed in PCA2. Leaf characteristics - leaf colour, leaf vein colour (upper and lower surface) and other parameters such as stem colour sex, sprouts, seedyam head colour, number of tuber per hill including other tuber external morphological characteristics such as forking and forking positions accounted for variation in PCA3.

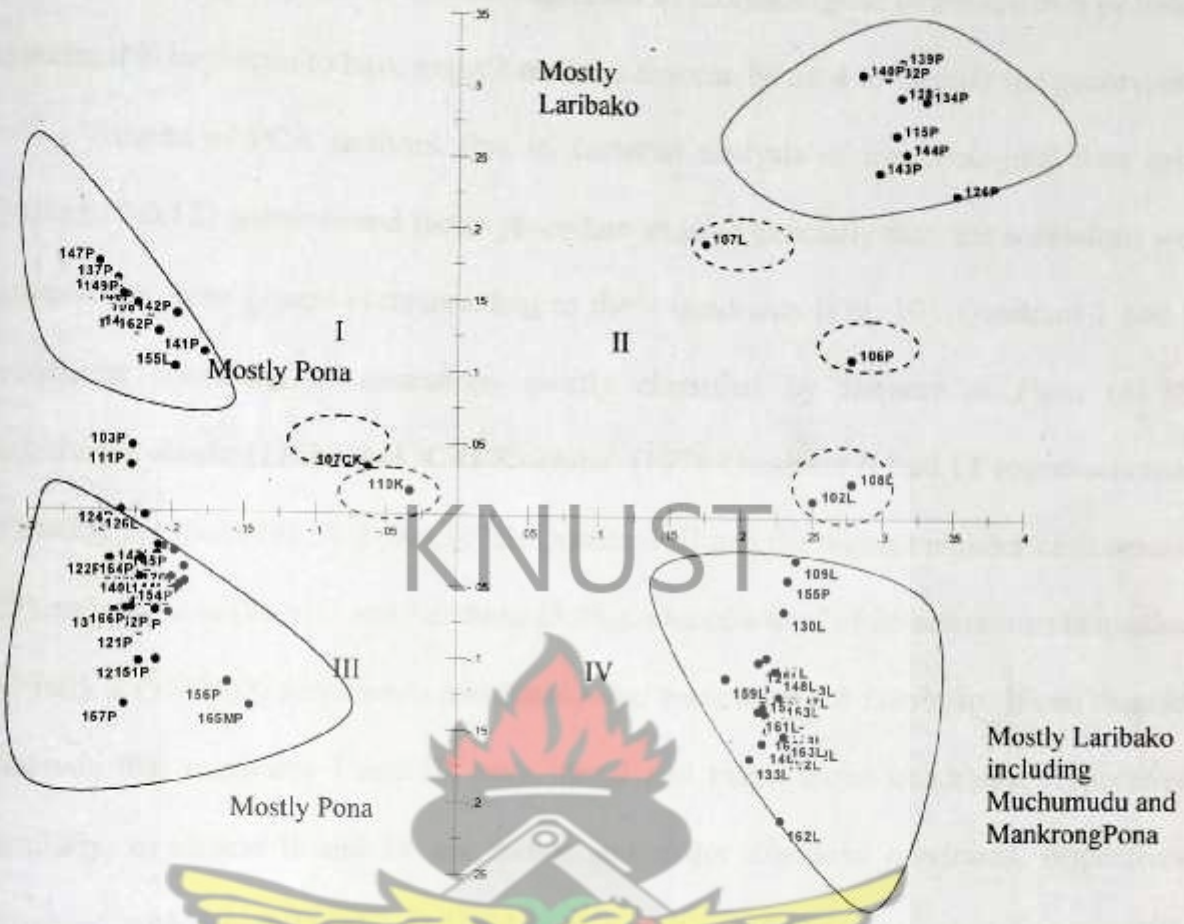


Fig. 10: 2D representation of genetic relationship among 91 accessions of yam revealed by morphological data

4.4 Phenetic Analysis Based on Morphological Data

Germplasm collection of vegetatively propagated species, such as *D. rotundata-cayenensis* often contains accessions which, although morphologically similar have different genetic origins and vice versa (Lebot *et al.*, 1998).. Identical cultivars may have different names in different collections and areas due to the numerous vernacular names. Hence putative duplicates and cultivar misclassification were identified and removed. The majority of cultivars were most likely clones of a common source. Since agronomically desirable traits

such as yield, high dry matter are not expressed as morphological characteristics or linked to them, it is important to have genetic markers that can be used to identify the genotypes.

Results of PCA analysis, that is, factorial analysis of morphological data using DARwin 5.0.153 software and factor procedure showed generally that, the accessions were grouped into four groups corresponding to the 4 quadrants (Fig. 10). Quadrant 1 had 18 accessions consisting of accessions mostly classified by farmers as *Pona* (88.8%) including *Kulunku* (110K) and 'CRI-Kukrupa' (107). Quadrant II had 13 accessions made of mainly *Laribako* (92.3%) plus 107D. Quadrant III had the highest number of accessions (29) mainly *Pona* (96.6%) and *Laribako* (3.4%). Out of a total of 26 accessions in quadrant IV, only 2 (7.7 %) (*Muchumudu* and Mankrong*Pona*) were not *Laribako*. It can therefore conclude that quadrants I and III were minor and major *Pona* quadrants, respectively. Similarly, quadrants II and IV are minor and major *Laribako* quadrants, respectively. However, within both Quadrants I and II, the position of some accessions suggested that there could be a sub-group each for the two quadrants. For instance, accessions 107 and 110K in Quadrant I were so close to the origin compared to the others. Similarly, accessions 102L and 107L in Quadrant II were so close to the positive abscissa, whilst the rest of the accessions in the quadrant were all further away. Not all the agroecologies were represented in the four quadrants suggesting that agroecology had effect on the spread of the *Pona* complex accessions. Unlike in Benin where *Kponan* group had two cultivars with distinct distributions: *Kponan* in the north and *Laribako* in the south of the country, in Ghana, *Pona* are widely distributed from the Coastal Savannah, Forest through Forest-Savannah Transition to Guinea Savannah in the north, with *Laribako* however, limited to the north.

Further clustering of the morphological data with the Tree analysis concept using the DARwin 5.0.153 software and tree construction procedure with the neighbour-joining approach gave 4 main clusters *Labariko* (39 accessions), *Pona* (46 accessions) with two 110K (*Kulunku*-very early variety) and 107 *Dente* and 'CRI-Kukrupa' –late maturing variety, as outliers (Fig.11). *Muchumudu* clustered with other *Laribako* accessions. Similarly, CRI*Pona* and Mankrong*Pona* clustered with other *Pona* accessions.

Assessment of the morphological analysis fit for factorial and tree data, fit criterion for tree had edge length sum of 5.0316; Mean error of -0.011; Mean absolute error of 0.037; Maximum absolute error of 0.2561; Mean square error: of 0.0026 and Cophenetic r: 0.9794. A cophenetic value of 0.9 shows that the phenogram truly represents the genetic structure of the population and no errors was generated by the methodology used.

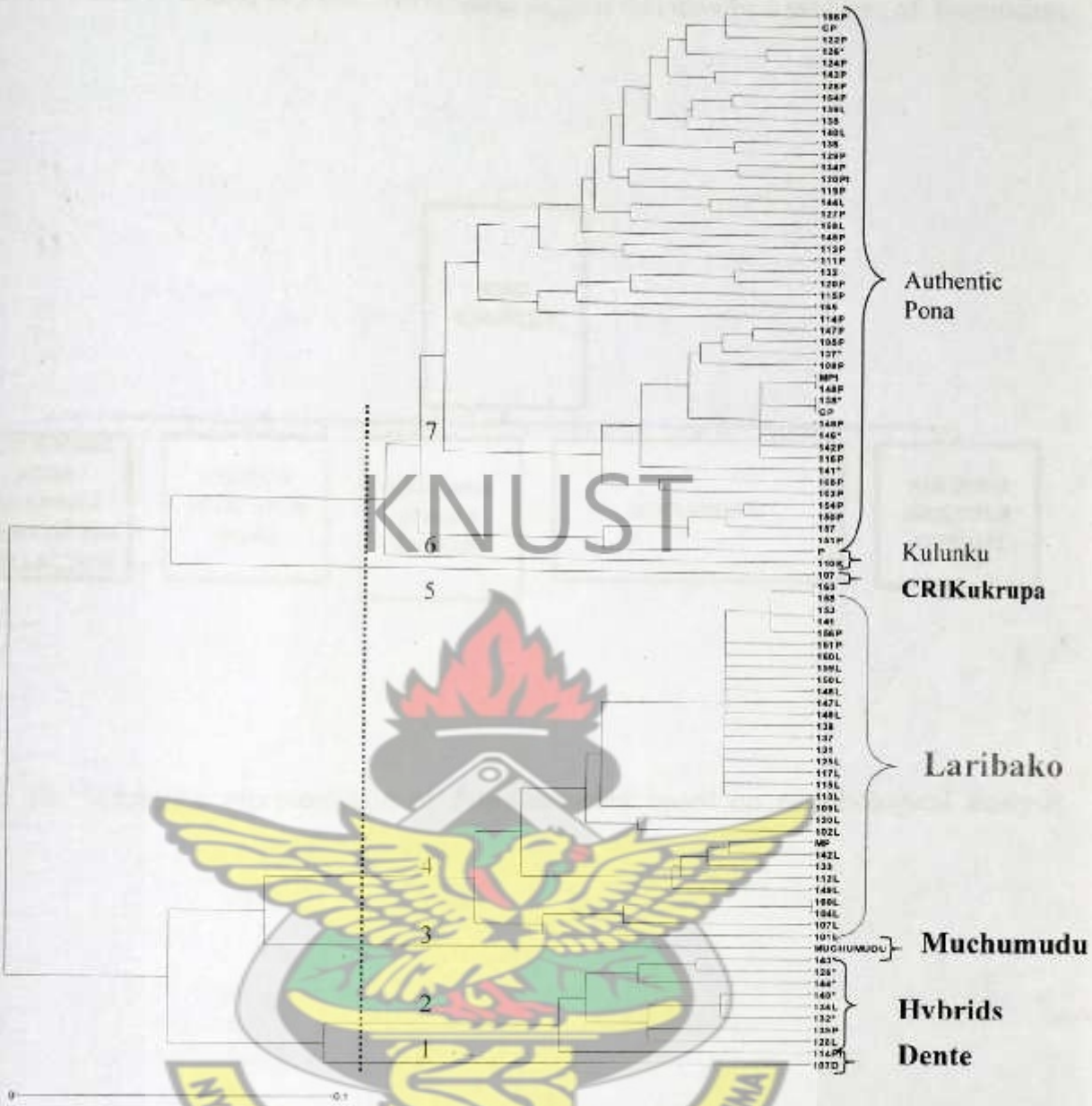


Fig. 11: Genetic diversity tree of 91 pona complex accessions based on Jaccard's index genetic similarity among morphotypes. Similarity matrix is based on presence or absence of 111 morphological traits.

Based on the morphological characterization, *Pona* complex can be schematically represented as (Fig.12). The difference between the folk taxonomic classification and

morphological characterization is that *Kulunku* which seems to be a *Laribako* was classified as a new kind of *Pona*. This results suggest this may be a problem of linguistics.

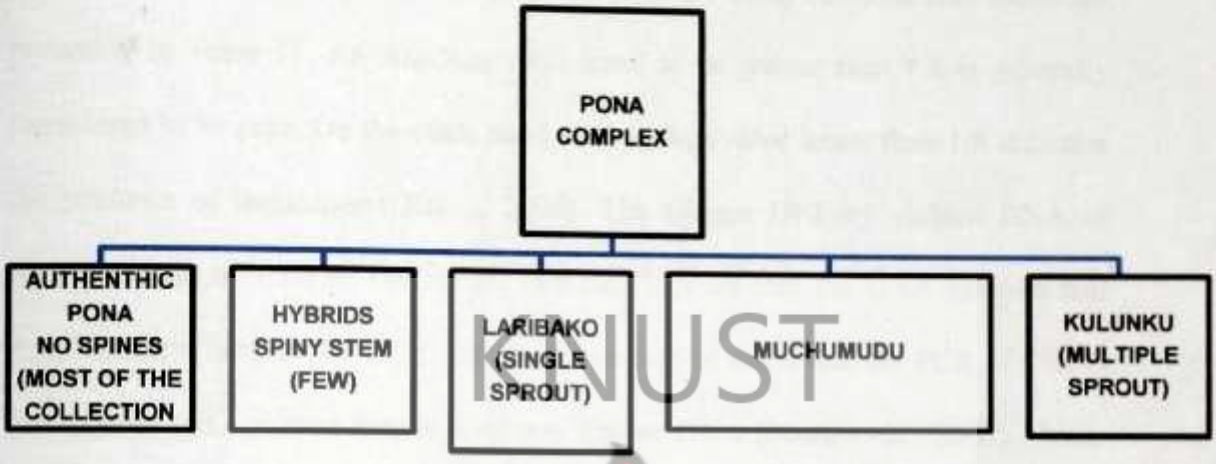


Fig. 12: Schematic representation of *Pona* complex based on morphological analysis.

4.5 Molecular Characterization

4.5.1 DNA Quantification and Quality Testing.

The quantities and quality of DNA extracted from the *Pona* complex accessions are presented in Table 11. An A_{250}/A_{280} ratio equal to or greater than 1.8 is generally considered to be pure. On the other hand an A_{260}/A_{280} value lesser than 1.8 indicates the presence of impurities (Qiagen, 2006). The Qiagen DNEasy yielded DNA of A_{260}/A_{280} ratio of 1.7-1.9. The results obtained indicate that the DNA samples had A_{260}/A_{280} ratio ranged of 1.02 -1.4. It must however be noted that the PCR of SSR is very robust and can even handle relatively impure DNA (Scotti *et al.*, 2003). Hence the DNA samples were pure enough for the analysis.

4.5.2 Allele Frequency Analysis

An electrophenogram of some accessions studied with microsatellite markers is pasted at Figure 3. The number of peaks corresponds to the number of alleles at each locus. A total of 27 loci were detected from the 13 markers used in this study with an average of 4.26 alleles per locus ranging from 2 to 15 alleles per locus for Dab2E09 and Dpr3D06, respectively (Table 11). The mean allelic richness ranged from 2 alleles per locus to 12 with an average of 5.23 per locus, indicating there were many allelic variants per locus.

The allele frequency analysis calculates two common measures of variation for each locus: expected heterozygosity and polymorphic information content (PIC) for each locus. Expected heterozygosity is calculated using an unbiased formula from allele frequencies assuming Hardy-Weinberg equilibrium (Nei 1987). This is a useful measure of informativeness of a locus: loci with expected heterozygosity of 0.5 or less are in general, not very useful for large-scale parentage analysis.

Table 10: Range of sizes and number of alleles detected using the 13 SSRs primers *Pona* Complex accessions.

SSR Marker Name	Min size detected (bp)	Max size detected (bp)	Number of alleles detected	Allele sizes identified
Da1F08	166	179	5	166, 170, 172, 175, 179
Dab2C05	178	193	*3*	*178, 190, 193*
Dab2D06	165	186	4	165, 171, 176, 186
Dab2E09	117	197	*3*	*117, 120, 197*
Dpr3D06	125	170	12	**125, 127, 131, 133, 137, 139, 143, 145, 148, 150, 160, 170**
Dpr3F04	81	131	11	81, 88, 95, 97, 99, 103, 118, 121, 124, 127, 131
Da1A01	212	225	*3*	*212, 214, 225*
YM13	175	227	4	175, 212, 214, 225
YM15	170	293	8	170, 172, 186,197, 211, 223, 228, 293
YM26	102	174	5	102, 107, 172, 175, 179
Da1D08	223	337	7	223, 229, 300, 304, 308, 321, 337
Da1C12	140	160	*3*	*140, 158, 160*
Dpr3F10	102	173	**12**	**102, 107, 111, 127, 129, 133, 136, 142, 149, 155, 168, 173**

Genetic diversity indicated by expected heterozygosity (H_E) ranged from 0.514 for Dab2E09 to 1.00 for Dpr3D06, Da1C12, YM15 and YM26 with a mean of 0.6279. Generally the expected heterozygosity of the loci was greater than 0.5, except in DA1A01 indicating that a good parentage analysis can be obtained from the molecular analysis. This is the probability that, at a single locus, any two alleles, chosen at random from the population are different from each other. The average H_E over all loci therefore estimates the extent of genetic variability in the population. At most loci the observed heterozygosity (Ho) was higher than the expected heterozygosity (H_E). The mean H_E value of 0.63 meant that there was some degree of genetic variation among the population. Mean proportion of individuals typed was 0.20; Mean expected heterozygosity was 0.63 and Mean polymorphic information content (PIC) was 0.46 (Table 12). Mean polymorphic information content (PIC) values for all markers used was 0.53 and ranged from 0.00 to 0.89. Polymorphic information content (PIC) is a measure of informativeness related to expected heterozygosity and likewise is calculated from allele frequencies (Botstein *et al.* 1980; Hearne *et al.* 1992). It is commonly used in linkage mapping.

Deviations from Hardy-Weinberg equilibrium at many or all loci are an indicator of population substructure. The population might be divided into a series of closely related or inbred family groups or there may be hybridisation between genetically differentiated species, subspecies or races (i.e. a mix of two or more populations).

Table 11: Summary statistics of allele frequency analysis of SSRs of *Pona* Complex in Ghana.

Marker	Count	Heterozygotes	Homozygotes	HObs	HExp	PIC	H-W[a]	Chi-Square[b]	Df[c]	P-value	NF[d]
YM13	68	66	1	0.971	0.628	0.545	***	18.1147	1		-0.2494
YM26	124	124	0	1.0000	0.5693	0.469	***	46.8500	1	<0.0000001	-0.2944
Dpr3D06	148	145	0	1.0000	0.8560	0.8309	***	18.1112	1	2E-07	-0.881
YM15	56	56	0	1.0000	0.8859	0.8454	***	20.221	1	<0.0000001	-0.0799
Dab2E09	76	76	0	1.0000	0.5953	0.5018	***	29.1068	1	<0.0000001	-0.2873
Da1A01	76	76	0	0	0	0	ND	-	-	-	ND
Da1D08	48	48	1	1.0000	0.7167	0.6116	**	10.5612	1	<0.0000001	-0.122
Dpr3F04	48	48	0	1.0000	1.0000	0.375	ND	-	-	-	ND
Dab2D06	116	104	6	1	0.5354	0.4088	***	17.1685	1	3E-05	-0.3201
Da1C12	120	120	0	1	1	0.375	ND	-	-	-	ND
Dpr3F10	157	155	1	1	0.8174	0.7804	ND	-	-	-	ND
Dab2C05	36	30	3	0.8333	0.5619	0.4482	ND	-	-	-	-0.2127
Da1F08	74	74	0	1.0000	0.5709	0.4670	***	26.3008	1	3E-07	-0.2938

Count : Number of occurrences of the allele in the genotype file; H_O: Observed heterozygosity; H_E: Expected heterozygosity; **PIC**: Polymorphic information content. **HW[a]**: Significance of deviation from Hardy-Weinberg equilibrium. Key: NS = not significant, * = significant at the 5% level, ** = significant at the 1% level, *** = significant at the 0.1% level, ND = not done. These significance levels include a Bonferroni correction.

Chi-Square[b] the chi-square value and the number of degrees of freedom to calculate the significance of any deviation from Hardy-Weinberg equilibrium.

Df[c]: The number of degrees of freedom is equal to ½ $n(n - 1)$, where n is the number of allelic classes remaining after rare alleles have been combined. Yates' correction for continuity (subtracting 0.5 from the absolute value of the difference between observed and expected frequencies) is applied when there is only one degree of freedom.

NF[d]: Null allele frequency- Estimates the null allele frequency; the frequency of the allele taking account of any null allele present. They are meaningless if the if the estimated null allele frequency is negative.

Deviations from Hardy-Weinberg equilibrium (HWE) were assessed using a chi-square goodness-of-fit test, which compares observed genotype frequencies with expected genotype frequencies that are calculated from allele frequencies assuming Hardy-Weinberg equilibrium (Nei, 1987).

From this study, deviation from HWE was highly significant ($p<0.001$) for primers YM13, YM26, Dpr3D06, YM15, Dab2E09, Da1D08, Dab2D06 and Da1F08. The relatively high number of loci that significantly deviated from the H-W equilibrium confirms that the *Pona* complex population had substructures and that the population is made up of closely related types such as *Pona*, larbako, *Kulunku* and *Muchumudu* which can easily be mistaken for *Pona*. It could also be attributed the occurrence of hybridization between species or sub-species or isolate breaking effect (mixing of two or more previously isolated populations).

A null allele is any allele that cannot be detected by the assay used to genotype individuals at a particular locus. Null alleles are a common cause of apparent deviations from Hardy-Weinberg equilibrium at microsatellite loci (Pemberton *et al.* 1995), particularly where only a single locus shows a deviation. With microsatellite loci, a null allele most often occurs because of mutations in one or both primer binding sites, sufficient to prevent effective amplification of the microsatellite allele. This problem is particularly common when the microsatellite locus is cloned in one species and typed in a different species using the same microsatellite primers (Cervus, 2007).

The Null allele frequency also estimates frequency of the allele taking account of any null allele present. They are meaningless if the estimated null allele frequency is negative. From our study the null allele had no effect on the analysis since all the

estimates were negative, except for a few which had low population numbers and, as such, could not be determined.

A gene is said to be polymorphic if the frequency of one of its alleles is less than or equal to 0.95 or 0.99. Results obtained from allelic frequency analysis showed that all the 13 primers were polymorphic (Appendix V). No rare allele (alleles with allelic frequencies of less than 0.005) was obtained; this could be attributed to the genetic relatedness of the accessions being studied. The proportion of polymorphic loci (the number of polymorphic loci divided by the total number of loci) was 0.71. All the loci except YM13_212, Dpr3D06_127, Da1D08_337 and Dpr3F10_107 were heterozygotes.

4.5. 3. Ordination Analysis

Principal Coordinates analysis of the molecular data showed that the first three coordinates were important (Table 12). PCoA axes 1, 2 and 3 accounted for 40.51% of observed variation. The genetic distances generated using PCO software (Appendix V) was used in generating the PCoA plots.

Table 12: Principal Coordinates Analysis of Molecular data.

Principal Coordinates	Percentage of variation explained	
	Individual	Cumulative
Axis 1	18.05%	18.05%
Axis 2	11.80%	29.85%
Axis 3	10.66%	40.51%

The plots of PCoA1 versus PCoA2 using PCO software showed a wide dispersion of accessions along the four quadrants (Fig. 13). *Pona* and *Laribako* accessions could be found in all four quadrants suggesting that some of *Pona* accessions clustered with the

Laribako accessions and vice versa. Quadrants **I** had 19 accessions of mostly *Laribako* with 161P as the most distinct member of this group. Quadrant II had a few (7 accessions) with 2 IITA checks (TDr1929 and TDr2689) grouped with 3 *Pona* and 9 *Laribako* accessions respectively. TDr1929 was the most distinct accession. Mankrong *Pona*, a hybrid from IITA released as a new variety in Ghana (circled) was on the horizontal line separating Quadrants I and II. Quandrant III had the most (26) accessions which was a mixture of *Pona*, *Laribako*, *Muchumudu* (115M) among others. All the 15 accessions grouped in Quadrant IV were *Pona*.

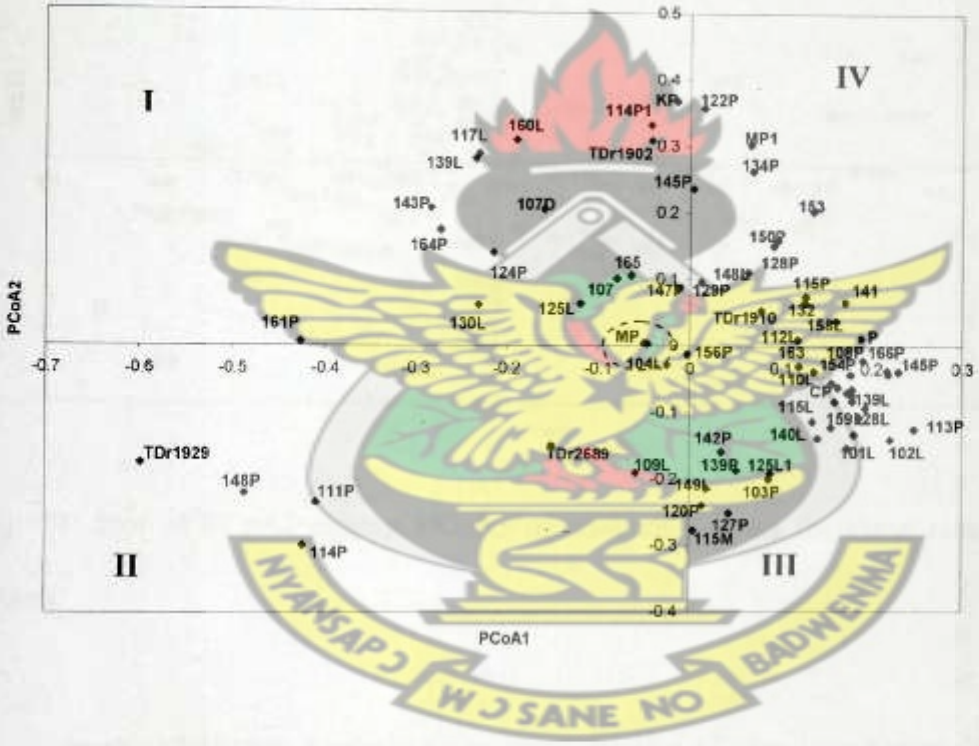


Fig. 13: PCoA1 versus PCoA 2 of SSR allelic data for *Pona* complex in Ghana.

A similar trend was obtained when PCoA 2 was plotted against PCoA3 (Fig. 14) except that 4 accessions occupied the midpoint between Quadrants I and II. Again TDr1929 was the most distinct accession.

Tree analysis of the molecular data using the DARwin 5.0.153 software and tree construction procedure with the neighbour-joining approach showed large number of inter- and intra-specific polymorphisms that enabled reliable discrimination between the samples (Figure 15). Again, some of the *Laribako* accessions clustered with *Pona* and vice versa.

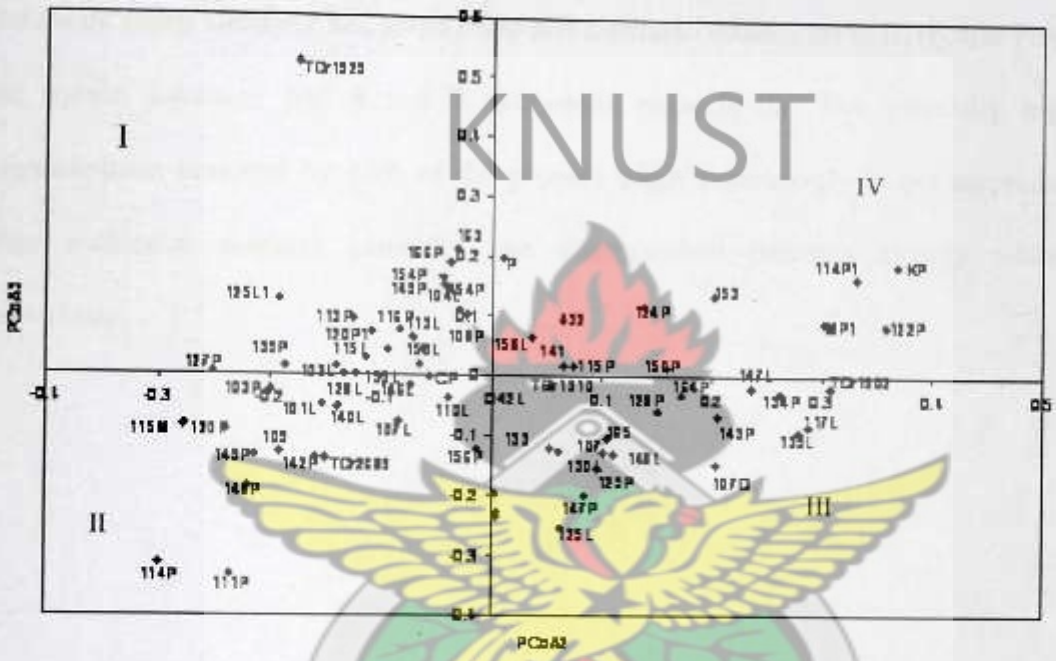


Fig. 14: Plot of PCoA2 against PCoA3 of SSR allelic data for *Pona* complex in Ghana.

In the Molecular Analysis Fit criterion for tree of edge length sum: 2570.40; Mean error: -1.28; Mean absolute error: 7.77; Maximum absolute error: 50.60; Mean square error: 107.40 and cophenetic r: 0.94. The phenogram truly represents the genetic structure of the population and no errors were generated by the methodology as indicated by the cophenetic value of 0.9. Hence the groupings of the accessions are a true reflection of the relationships.

Four main groups and two small ones were identified from the allelic data: authenthic *Pona*, *Laribako*, *Muchumudu*, and *Dente*, and the minor groups were Hybrid *Pona* and Hybrid *Laribako* (Fig. 15). All the IITA checks clustered with the authenthic *Pona*. The *Muchumudu* group had 5 *Laribako* accessions in it. The *Laribako* group also had some *Pona* accessions in it and vice versa.

The *Dente* group similarly had some *Pona* and *Laribako* accessions in it. Hybrid *Pona* and hybrid *Laribako* had 8 and 7 accessions respectively. The generally high polymorphism revealed by each of the primers taken separatively is not surprising since molecular markers generally can distinguished between closely related individuals.

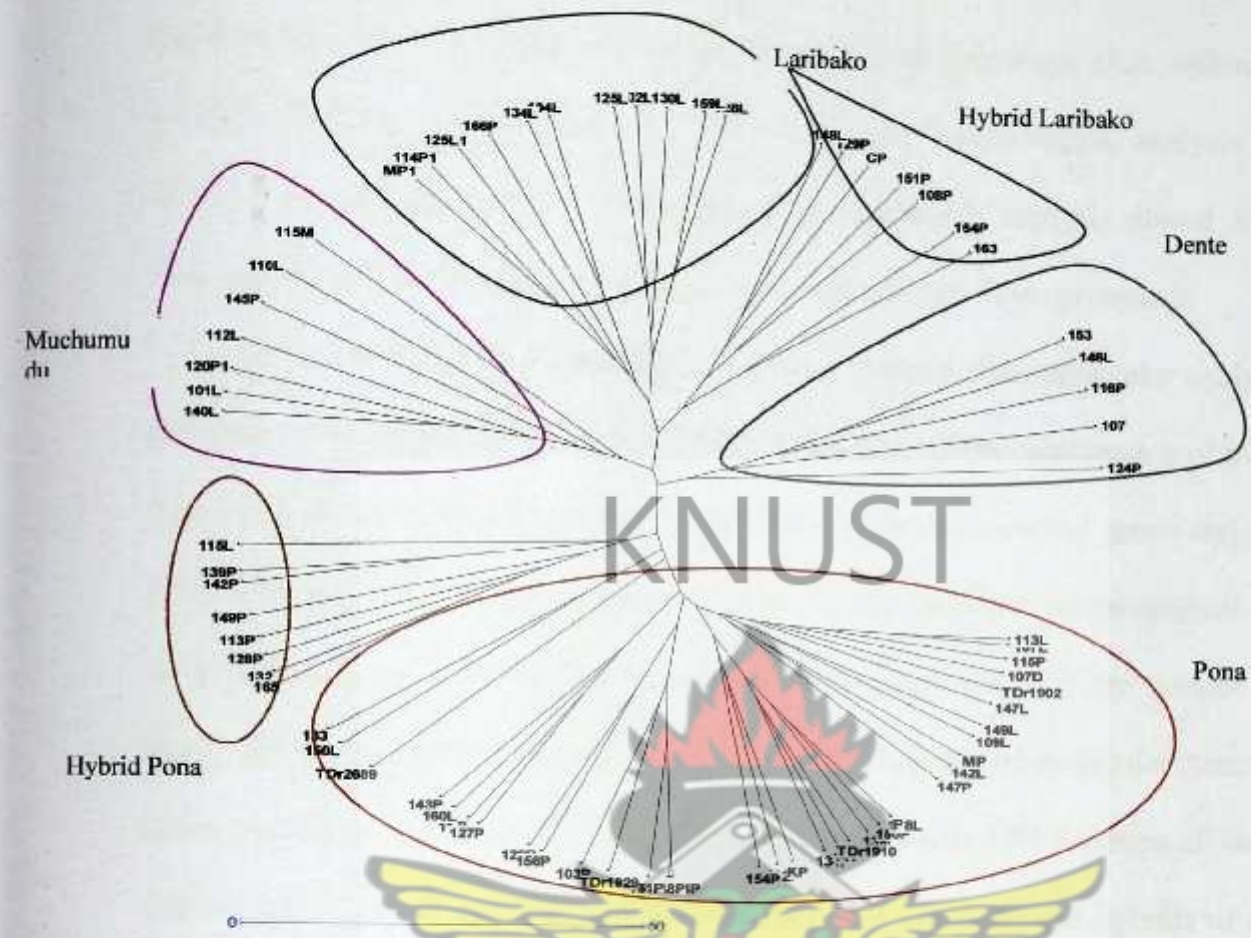


Fig.15: Genetic diversity tree of 72 accessions plus 4 IITA checks based on SSR data using unweighted neighbour-joining analysis

4. 6. Comparison of morphological and molecular characterization

Morphological and molecular analysis identified 4 main groupings each within the *Pona* complex and two minor groupings. However, whilst morphological analysis identified *Pona*, *Laribako*, *Dente*, and *Kulunku* groupings; molecular analysis placed *Kulunku* in *Pona* group and placed more accessions in *Muchumudu* and *Dente* groupings.

In instances where morphological analysis differed from molecular analysis in the grouping of accessions, there are two options available: (1) the combination of information from different types of markers-that is, those that refer to functional genes and those that show polymorphism in the genomic regions or (2) combine morphological data and molecular data, to give the best approximation to knowledge on the genetic variation (Schneider *et al.*, 2000). Combination of the morphological and molecular data, however, led to formation of 6 clusters: *Dente*, *Muchumudu*, *Kulunku*, CRI Kukrupa all stood alone and the rest of the accessions grouped into *Pona* and *Laribako* and hybrids of *Pona* and *Laribako* (Fig. 16). The resultant 6 groupings truly represent the true genetic groupings of the *Pona* complex since all available relationships have been captured by the data generated.

A characterization only based on morphological or agronomic traits is known to mask important genetic information (Zannou *et al.*, 2006). The importance of employing novel and improved technologies such as molecular approaches in combination to morphological analysis cannot therefore be over-emphasized.

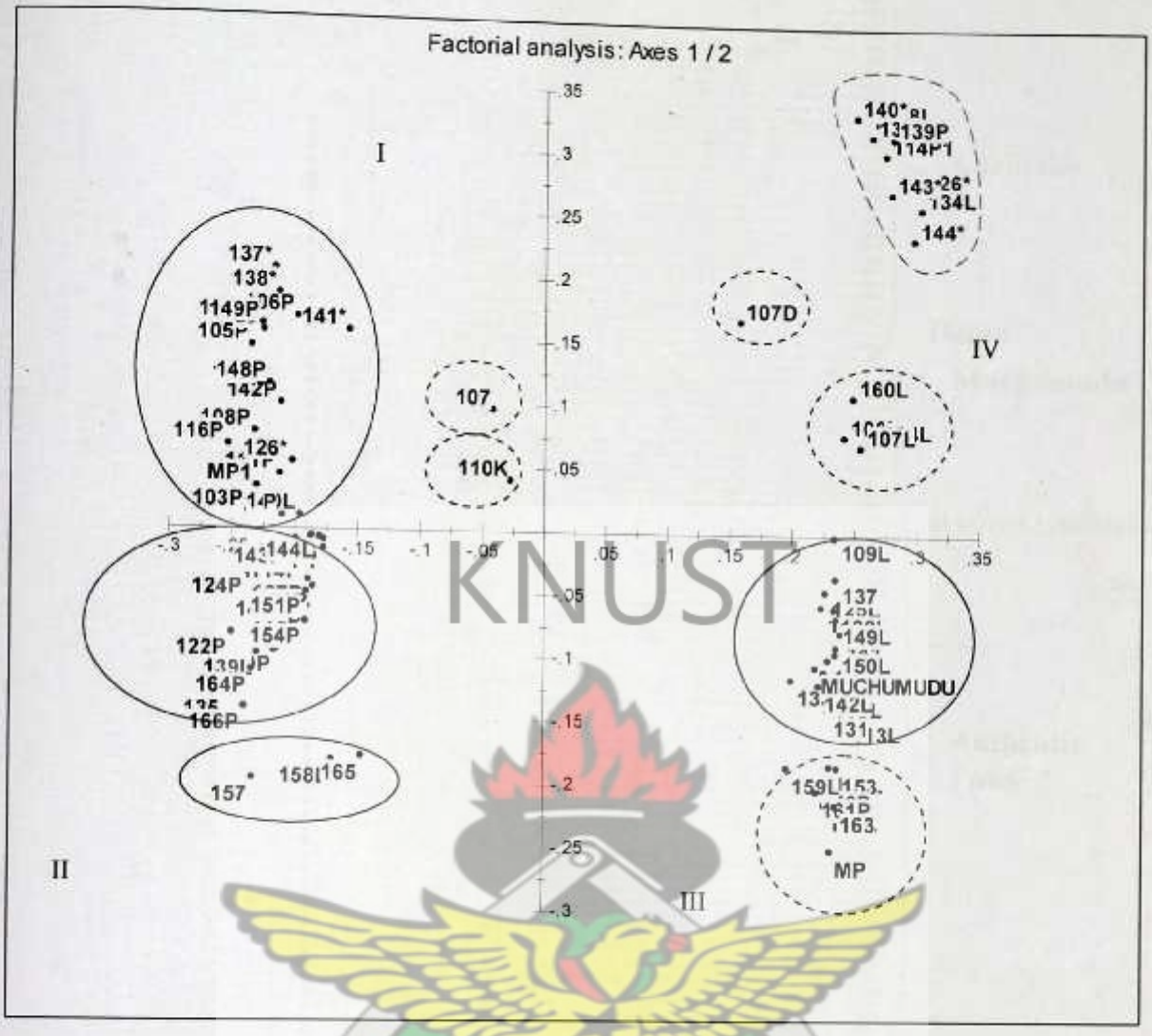


Fig. 16: Configuration of the combined morphological and molecular similarity data using DARwin, v. 5.0.153.

Morphological analysis of the accessions of the *Pona* complex germplasm resulted in 4 main morphotypes (cultivars groups): *Pona*, *Laribako*, *Dente*, and *Kulunku* groups (Fig. 17). The identity of these groupings is presented in Table 14. A group as defined by Dansi *et al.*, (1999) is "*a very particular cultivar or set of morphologically similar cultivars such that the intra-group variability is lower than the inter-group variability*". In terms of maturity, all the accessions were early maturing except *Muchumudu* and *Dente* which was not part of the *Pona* complex but a check).

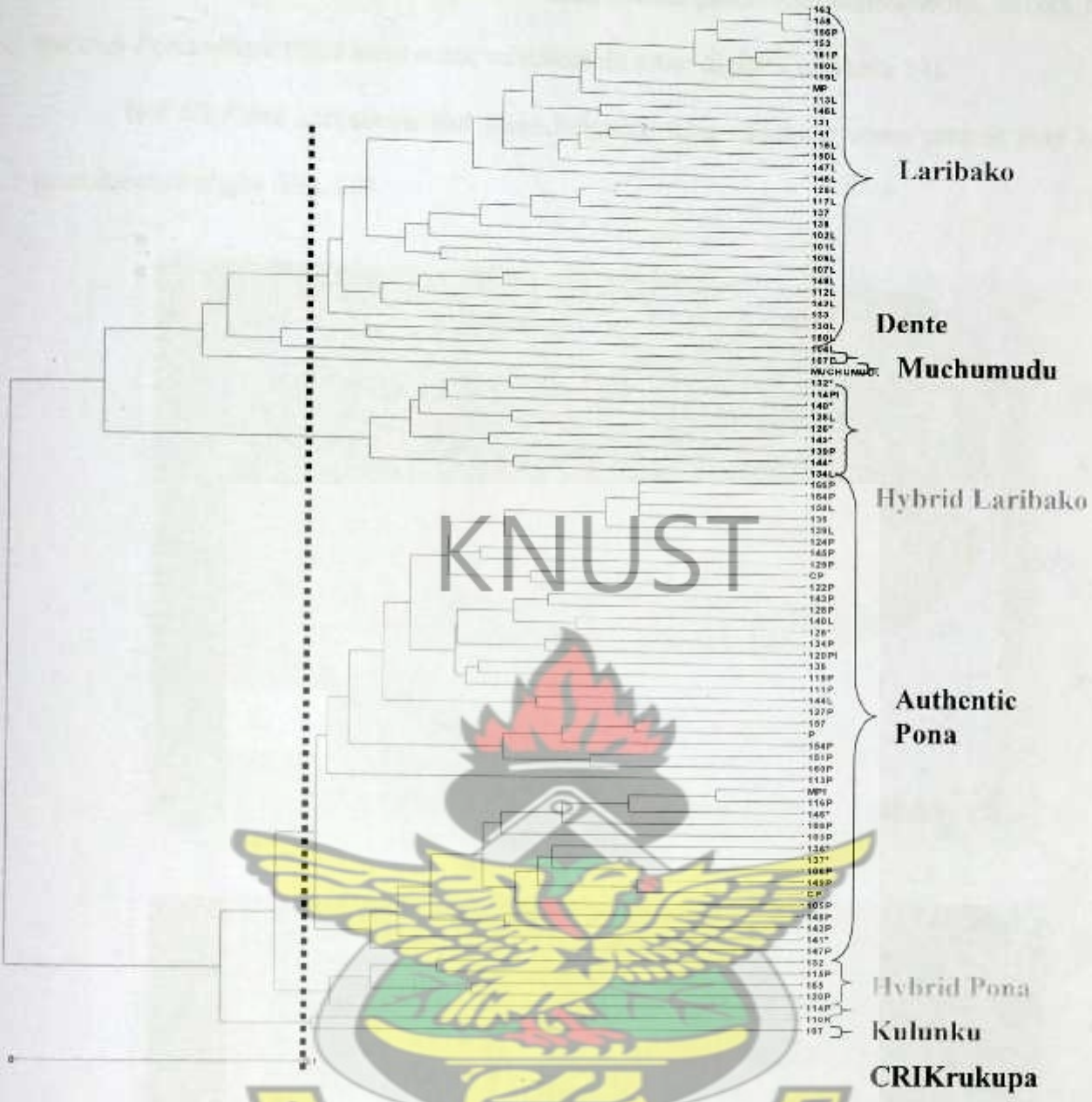


Fig. 17: Genetic diversity tree of pona complex based on combined morphological and molecular data.

For all the groups, the aerial and underground parts were homogenous, except in the true *Pona* where there were some variations in tuber digitation (Table 14).

Not all *Pona* accessions had forked/digitate tuber. Forking when present may be prominent or slight (Fig. 18).

Fig. 18: Sample of pona tubers showing digitations; digitate (A) to no digitation (B).

In terms of tuber shape, the true *Pona* and its 'look-alikes' differed principally in terms of crown position on tuber, shape of tail and size of head in relation to tail.

Fig. 19: Close up view of head region of *Pona* and its 'look-alikes' showing position of crowns on tuber.

The true *Pona* differ from its 'look-alikes' in having its crown well inserted into the head of the tuber, as oppose to the crown standing on the head of the tuber (Fig. 19).

The tail region also differed significantly among *Pona* and its 'look-alikes' (Fig. 20).

Fig. 20: Varying shape of tail of tuber *Pona* and its 'look-alikes'.

Muchumudu tubers have evenly shaped big tail (Fig. 20). *Numbo* tubers have characteristic conical-shaped tail. Kulunku tubers have digitate tail. The tail region of Fuseini looks like human foot. The tubers of the true *Pona*, on the other hand, are sagitate.

Fig. 21 shows the 'size of head in relation to tail' variation between *Pona* and its 'look-alikes'. The true *Pona* has characteristics big head and tapers to the tail region. *Muchumudu* on the other hand has big head which continues to the tail. *Fuseini* and *Laribako* all have medium head which continues to the tail. In contrast *Numbo* has small head compared to broad tail. Kunluku also has small head and broad.

Flowering was also profuse when they occurred. The *Muchumudu* group had similar vegetative characteristics as the true *Pona* but differed in terms of tuber quality in having more watery flesh of freshly harvested tubers than the true *Pona*.

True *Pona*

Fuseini

Numbo

Muchumudu

Laribako

Kulunku

Fig. 21: The varying 'size of head in relation to tail' and tail shape of *Pona* and its 'look-alikes'.

They also had characteristic striations in the cross-section of the cut surface and were similar to "Ahimon" in Benin (Dansi *et al.*, 1999). Its flowering was complex: either male or female or monoecious. An intraclonal variation in sex determination have been observed in "Ahimon" where two individuals derived from the same original tuber had flowered differently as male and female (Dansi *et al.*, 1999).

After losing some water content after harvesting, *Muchumudu* tasted and smelled like *Pona*. *Muchumudu* being late maturing usually comes to the market when the true *Pona* has been exhausted, due to its similarity with *Pona*, it is marketed as *Pona*. This accounts for its usage as *Pona* by yam sellers after storing it for a while after harvesting. It is worth noting that *Dente* was not part of accessions farmers claimed to be *Pona*. *Dente* together with 'CRI-Kukrupa' (late maturing variety) were deliberately included to assess the effectiveness of the morphological characteristics in discriminating among the accessions; the analysis identified these unique accessions. It is only in the cooked tuber state that the unique taste and flavour (a culinary characteristic of *Pona*) was undeniably observed. This explains why most farmers in the market, where one cannot determine the culinary characteristics, pass off most yam varieties as *Pona*.

Externally, the tubers of the true *Pona* can be identified by looking at the insertion of crown in the tuber, tail shape, presence of grooves on the tuber surface and the size of head in relation to tail (Fig. 22).

Fig. 22. External tuber characteristics of authentic *Pona* (A) Inserted crown; (B) relatively big head compare to tail (C); and (D) characteristic grooves on tuber surface.

4.7 Combined Morphological and Molecular Characterization

The combination of morphological and molecular data (Appendix III) and their subsequent ordination and clustering is presented in Fig. 16 and Fig. 17 respectively. The ordination analysis of the combined morphological and molecular data (Fig. 16) showed that there were subgroups within each quadrant. For instance in Quadrant I, there was a main Pona group consisting of 16 accessions and 2 stand alone groups of 'CRI-Kukrupa' (107) and Kulunku (110K). Similarly, 2 groups were observed in Quadrant II, with a major group of about 13 Pona accessions and a smaller group of 3 accessions of Pona. Quadarnt III had one main Laribako and minor group of Pona. A distinct Dente (107D) group, a 4-member Laribako group and a major Laribako group were found in Quadrant IV.

Further analysis of the molecular data with the Tree analysis of the molecular data using the DARwin 5.0.153 software and tree construction procedure with the neighbour-joining approach showed large number of inter- and intra-specific polymorphisms that reliably discriminate between the samples (Figure 17). Again, some of the *Laribako* accessions clustered with *Pona* and vice versa indicating the close relationship between the two varieties.

The local checks *Dente* (107D), and *CRI-Kukrupa* (107) were separated from the rest and so was *Muchumudu* and *Kulunku*. A small group of 4 accessions including Mankrong*Pona* were named Hybrid *Laribako* due to their closeness to the main *Laribako* group. Similarly, a minor group consisting of 4 accessions were also named Hybrid *Pona* due to their closeness to the main authentic *Pona*. It has been suggested that hybridization between congeneric diploid taxa (often species) followed by selection for one or more recombinant types can lead to the formation of new species (Grant, 1981). This might have accounted for the emergence of "hybrids" of pona and Laribako. The structure and reproductive biology of these groups are presented in Table 13. The characteristics of the groupings are also presented in Table 14.

Table 13: Structure of the cultivar groups.

Cultivar Group	Intragroup variability: aerial part	Tuber	Cultivar names	Ea	LL Cm	LW cm	TS cm	sex	Flowering	Fruit setting
1. True *Pona*	Homogenous	heterogenous	*Pona*	E	13	9.2	100	M	Profuse	None
2. Hybrid *Pona*	Heterogenous	heterogenous	H*Pona*	E	12	7.8	90	M	Profuse	None
3. *Laribako*	Homogenous	heterogenous	*Laribako*	E	13	9.2	60	M/F?	Profuse	None
4. Hybrid *Laribako*	Heterogenous	heterogenous	H*Laribako*	E	11	75	70	M/F?	Profuse	None
5. Kunlunku	Homogenous	homogenous	*Kulunku*	E	12	6.1	60	-	-	-
6. *Dente*	Homogenous	heterogenous	*Dente*	L	11	6.4	70	F	Profuse	Yes
7. 'CRI-Kukrupa'	Homogenous	homogenous	'CRI-Kukrupa'	L	10.9	5.6	62	M	Profuse	None

NB: Ea = Earliness; E = Early; L = Late; LL= Leaf Length; LW = Leaf Width (average of 20 observations); TS= Tuber Size (average of 3 tubers);

Table 14: Characteristics of the 6 groupings of yams identified within *Pona* complex.

Cultivar Groups	Major morphological characteristics
Pona	Stem smooth and unarmed. Leaf green and cordiform. Tuber long and oval-oblong; roots unarmed; vines emerge long before leaf emergence; Mealy tuber; aromatic; thick tuber skin relative to *Laribako*; White-yellowish tuber flesh colour. Crown of seed yam is red in colour.
Laribako	Stem smooth and unarmed. Leaf green and cordiform. Tuber long and cylindrical; roots unarmed; vines emerge with leaves; thin tuber skin; Very mealy cooked tuber and aromatic. Crown of seedyam is white in colour.
Muchumudu	Stem smooth but armed. Leaf light green with smooth border. Leaf cordiform; Late maturing; Tuber yellowish and finely striated. Tuber big and unbranched; roots armed; vines emerge long before leaf emergence
Kunlunku	Numerous vines (12-14) yet single tuber per mound. Stem smooth and unarmed. Leaf green and cordiform. Tuber long and cylindrical; roots unarmed; vines emerge with leaves. Thin tuber skin; Cooked tuber mealy, aromatic.
Dente	Stem rough and armed. Brown thorn; Leaf light green sagittate leaf. Tuber short with big head. Thick tuber skin; mealy cooked tuber; grainy texture and non-aromatic
'CRI-Kukrupa'	Stem rough and armed. Brown thorn; Leaf light green sagittate leaf. Tuber short with big head. Thick tuber skin; mealy cooked tuber; grainy texture and aromatic. Multiple tubering

4.9. Core Collection Determination

Based on the morphological and molecular data, a core *Pona* collection was made consisting of 5 accessions each for Laribako (102l, 101L, 108L, 107L and 149L) and *Pona* (128P, 134P, 114P and 111P). Accessions 124L, 115P, 110K and *Muchumudu* were also selected to represent *Laribako* hybrid, *Pona* hybrid, *Kulunku* and *Muchumudu* respectively. These materials would be conserved under slow growth in vitro (tissue culture) systems to ensure that they do not become extinct under field genebank conditions.

CHAPTER 5

5.0 CONCLUSIONS AND RECOMMENDATIONS

From this study the following conclusions can be drawn:

1. There is only one true Pona.
2. *Pona* complex is made up of six cultivars- *Pona, Laribako, Kulunku,* 'Fuseini', 'Numbo' and *Muchumudu* due to their close similarity with respect to their culinary attributes.
3. Morphological and folk taxonomy because they rely mostly on phenotypic and functional variation, classified these four cultivars as *Pona* at different times of the year.
4. Molecular characterization was able to differentiate between closely related individuals.
5. The genetic structure of this yam group was clearly defined when morphological data was combined with molecular analysis. This study confirms the need to combine both morphological and molecular data for such analysis.
6. Germplasm collection of vegetatively propagated species such as *D. rotundata-cayenensis* often contains accessions which although morphologically similar have different genetic origins and vice versa. Identical cultivars may have different names in different collections and areas due to the numerous vernacular names hence putative duplicates and cultivar misclassification were identified and removed.
7. The true *Pona* has the following characteristics: smooth and unarmed stem; green and cordiform leaf; unarmed roots; at sprouting, vines emerge long before leaf emergence; long and cylindrical tuber; thick tuber skin relative to *Laribako*; cooked tuber is mealy and aromatic; tuber flesh is white with yellowish tint; the

crown of the seedyam of authentic *Pona* is purple; and the crown of the tuber is also well inserted in the tuber.

8. A total of 13 DNA markers facilitated the identification of the authentic *Pona*. They are Da1F08, Dab2C05, Dab2D06, Dab2E09, Dpr3D06, Dpr3F04, Da1A01, YM13, YM15, YM26, Da1D08, Da1C12, and Dpr3F10.
9. A core collection of *Pona* collection was determined consisting of 5 accessions each for Laribako (102l, 101L, 108L, 107L and 149L) and *Pona* (128P, 134P, 114P and 111P). Accessions 124L, 115P, 110K and *Muchumudu* for were selected to represent *Laribako* hybrid, *Pona* hybrid, *Kulunku* and *Muchumudu* respectively.

RECOMMENDATIONS

1. The true Pona having been identified must be multiplied and disseminated to farmers.
2. If Ghana is to maintain its place as the leading exporter of yams, it must ensure the purity of the varieties. In this regard, this findings will be made available to the Ghana Standards Board to facilitate the yam export trade with respect to varietal identification and ultimately to ensure purity of our varieties.
3. A regulatory body equivalent to Ghana Grains Development Board must be set up to regulate the trade in planting materials of root and tuber crops in general and yams in particular.
4. Some selected farmers and traders must also be trained in maintenance of varietal purity and certified to produce such tubers.
5. A factsheet and poster of *Pona* and its look-alikes will be developed and made available to all yam stakeholders.

REFERENCES

Aagaard, J.E., Krutovskii, K.V. and Strauss, S.H. 1998. RAPDs and allozymes exhibit similar levels of diversity and differentiation among populations and races of Douglas-fir. *Heredity* **81**: 69-78.

Ajmone-Marsan, P., Egidy, G., Monfredini, G. ,Di Silvestro, S. and Motto, M. 1998. RAPD markers in maize genetic analysis. *Maydica*. **38**: 259-264.

Akkaya, M.S., Shoemaker, R.C., Specht, J.E., Bhagwat, A.A. and Gregan, P.B. 1995. Integration of simple sequence repeats DNA markers into a soybean linkage map. *Crop Science* **35**: 1439-1445.

Akwaag, A.A., Berchie, J.N., Echavez, M.L., Dapaah H.K., Kebede, T., Njue, S.W., Paraoan, I. and Labrada H.R. 2000. Sustainable Farming Practices –Towards reducing the expansion of the Forest Savanna Transition Zone of the Sekyere-West and Ejura-Sekyedumase districts of the Ashanti Region of Ghana. International Centre for development oriented Research in Agriculture and. Crops Research Institute- Working Document Series 86, Ghana- 2000.

Antonio, A. F., Garcia, I., Luciana, L. B., Antônia, M., Barbosa, M., Isaias, O. G., Cláudio, L. S. and Anete P. de Souza. 2004. Comparison of RAPD, RFLP, AFLP and SSR markers for diversity studies in tropical maize inbred lines. Genet. Mol. Biol. vol.27 no.4 São Paulo.

Aldenderfer, M. S. and Blashfield, R. K. 1984. Cluster analysis. Thousand Oaks, CA: Sage Publications, Quantitative Applications in the Social Sciences Series No. 44.

Anderberg, M. R. 1973. Cluster Analysis for Applications, Academic Press.

Ariyo, O. J. 1991. Numerical analysis of variation among accessions of okro (*Abelmoschus esculentus*). *Ann. Bot.* **67**:527-531.

Asemota, H.N., Ramser, J., López-Peralta, C., Weising, K., and Kahl, G. 1996. Genetic variation and cultivar identifi cation of Jamaican yam germplasm by random amplified polymorphic DNA analysis. *Euphytica* **92**:341–351.

Asiedu, R., Ng, S.Y.C., Bai, K.V., Ekanayake, I.J. and Wanyera N.M.W. 1998. Genetic Improvement. In: *Food Yams: Advances in research.* Orkwor, G.C., Asiedu, R. and Ekanayake, I.J eds. 1998. pp. 63-104.

Becker, J., Vos, P., Kuiper, M., Salamini, F. and Heun, M. 1995. Combined mapping of AFLP and RFLP markers in barley. *Mol. Gen. Gent.* **249**: 65-73.

Bergmann F, Gregorius H-R, Scholz F (1989) Isoenzymes, indicators of environmental impacts on plants or environmentally stable gene markers? In: Scholz F, Gregorius H-R, Rudin D (eds.). *Genetic Effects of Air Pollutants in Forest Tree Populations.* Springer-Verlag Heidelberg, New York, Tokyo, pp. 17-28.

Beyene, Y., Botha A and Myburg A. A. 2005. A comparative study of molecular and morphological methods of describing genetic relationships in traditional Ethiopian highland maize. *African Journal of Biotechnology* Vol. **4** (7), pp. 586-595.

Botstein, D., White, R.L., Skolnick, M and Davis, R.W. (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *American Journal of Human Genetics* **32**, 314-331

Boukema, I.W. and Van Hintum, Th.J.L. 1994: *Brassica oleracea*, a case of an integral approach to genetic resources conservation. *In: F. Balfourier And M.R. Perretant (Eds) Evaluation and Exploitation of Genetic Resources, Pre-Breeding. Proceedings of the Genetic Resources Section.* Meeting of Eucarpia, 123-129.

Brown, A.H.D. 1994: The core collection at the crossroads. *In: T. Hodgkin, A.H.D. Brown, TH.J.L. Van Hintum AND E.A.V. Morales (EDS.) Core Collections of Plant Genetic Resources.* John Wiley and Sons, UK.

Caddick, L. R., Wilkin, P., Rudall, P. J., Hedderson, T. A. J., Chase, M. W. 2002. Yams reclassified: a recircumscription of Dioscoreaceae and Dioscoreales. *Taxon,* Volume **51**, Number 1, pp. 103-114(12).

Cavalli–Sforza, L.L. and W.F. Bodmer. 1981. Genética de las Poblaciones Humanas. Ed. Omega, Barcelona.

Cervus 2007. Cervus 3.0.3 by Tristan Marshall at Field Genetics Ltd. Cervus is © Copyright Tristan Marshall 1998-2007.

Chheda, H.R. and Fatokun, C.A. 1982. Numerical analysis of variation patterns in okra (*Abelmoschus esculentus*). *Bot. Gazette* **143**: 253-261.

Clark, A.G. and Lanigan, C.M.S. 1993. Prospects for estimating nucleotide divergence with RAPDs. *Mol. Biol. Evol.* **10**: 1096–1111.

Coursey, D.G. 1969. The future of the yam in world agriculture: an appraisal of some problems. Tropical Root and Tuber crops Newsletter, 2: 4-8.

Crawford, D.J. 1990. Generation and Analysis of Data from enzyme electrophoresis. *Plant Molecular Systematics, Macromolecular Approaches.* A Wiley-InterScience Publication. John Wiley and Sons. New York.

Dansi, A., Mignouna, H.D., Zoundjihekpon, J., Sangare, A., Asiedu, R and Quinn F.M. 1998. Varietal identification key of the cultivated yams (*Dioscorea cayenensis/ Dioscorea rotundata complex*) of Benin Republic. *Plant Genetic Resources Newsletter*. No. **116**: 18-25.

Dansi, A., Mignouna, H.D., Zoundjihekpon, J., Sangare, A., Asiedu, R and Quinn F.M. 1999. Morphological diversity, cultivar groups and possible descent in the

cultivated yams (Dioscorea cayenensis/D. rotunadata) complex in Benin Republic. *Genetic Resources and Crop Evolution* **46:** 371-388.

Dansi, A., Mignouna, H.D., Zoundjihekpon, J., Sangare, A., Asiedu, R., Ahoussou. N. 2000. Using isozyme polymorphism to assess genetic variation within cultivated yams (Dioscorea cayenensis / Dioscorea rotundata complex) of the Republic of Benin. *Genetic Resources and Crop Evolution* **47**: 371-383.

DARwin, 2006. Published by Perrier, X. and Jacquemoud-Collet, J.P. (2006). DARwin software. http://darwin.cirad.fr/darwin

Demeke, T., Sasikumar, B., Huci, P. and Chibbar, R.N. 1997. Random amplified polymorphic DNA (RAPD) in cereal improvement. *Maydica* **42**: 133-142.

Dos Santos, J.B., Nienhuis, J., Skroch, P., Tivang, J and Slocum, M. K. 1994. Comparison of RAPD and RFLP genetic markers in determining genetic similarity among Brassica oleracea L. genotypes. *Theoretical and Applied Genetic* **87**:909-915.

Erskine, W. and Muehlbauer, F.J. 1991: Allozyme and morphological variability, outcrossing rate and core collection formation in lentil germplasm. *Theoretical and Applied Genetics* **83**, 119-125.

FAO. 2005. Production Yearbook for 2004. FAO, Rome.

Fisher, D. H. 1987. Knowledge Acquisition via Incremental Conceptual Clustering, Machine Learning, 2(2), pp.139-172.

Frankel, O.H. 1984: Genetic Perspectives of Germplasm Conservation. *In: W. Arber, K. Llimensee, W.J. Peacock and P. Starlinger (Eds.), Genetic Manipulation: Impact on Man and Society,* Cambridge University Press, Cambridge, 161-170.

Garcia, J.A., Duque, M.C., Tohme, J., Xu, S and Levy, M. 1995. SAS for Classification Analysis; Agrobiotecnology Course, October 1995. Working document. Centro Internacional de Agricultura Tropical (CIAT), Cali, Colombia.

Gauch, H.G., Jr. 1982. Multivariate Analysis and Community Structure. Cambridge University Press, UK.

Ghosh, S., Karanjawala, Z. E., Hauser, E. R., Ally, D., Knapp, J. I., Rayman, J.B., Musick, A., Tannenbaum, J., Te, C., Shapiro, S., Eldridge, W., Musick, T., Martin, C., Smith, J. R., Carpten, J. D., Brownstein, M. J., Powell, J. I., Whiten, R., Chines, P., Nylund, S. J., Magnuson, V. L., Boehnke, M and Collins, F. S. 1997. Methods for precise sizing, automated binning of alleles, and reduction of error rates in large-scale genotyping using fluorescently labeled dinucleotide markers. *Genome Res.* 7:165-178.

Gillet, E.M. (ed.). 1999. DNA Markers - Concepts and characteristics. Which DNA Marker for Which Purpose? Final Compendium of the Research Project Development, optimisation and validation of molecular tools for assessment of biodiversity in forest trees in the European Union DGXII Biotechnology FW IV Research Programme Molecular Tools for Biodiversity. Gillet, E.M. (ed.). 1999. URI. http://webdoc.sub.gwdg.de/ebook/y/1999/whichmarker/index.htm.

Gower, J.C. 1967. A comparison of some methods of Cluster Analysis. *Biometrics*. Vol. 23, No. 4, pp. 623-637.

Grant, V. 1981. "Plant Speciation," 2nd ed. New York, Columbia University Press.

Hadrys, H., M. Balick, and B. Schierwater. 1992. Applications of random amplified polymorphic DNA (RAPD) in molecular ecology. *Mol. Ecol.* **1**:55-63.

Hahn, S.K., Osiru D.S.O., Akoroda, M.O., Otoo, J.A. 1987. Yam production and its future prospects. *Outlook on Agriculture*. **16(3)**: 105-110.

Haley, S., Miklas, P.N., Afanador, L. and Kelly, J.D. 1994. Random Amplified Polymorphic DNA (RAPD) marker variability between and within gene pools

of common bean. *Journal of the American Society for Horticultural Science* **119(1):** 122-125.

Hamon P. 1987. Structure, origine genetique des ignames cultivees du complexe Dioscorea cayenesis-rotundata et domestication des ignames en Afrique de l'Ouest. These de Doctorat es-Sciences, Universite Paris XI, Centre d'Orsay,

Hamon P. and Toure B. 1990. The classification of the cultivated yams (Dioscorea cayenensis–rotundata complex) of West Africa. *Euphytica* **47**: 179–187.

Hamon, P., Brizard, J-P., Zoundjihekpon, J., Duperay, C. and Borgel, A. 1992. Etude des index d'ADN de huit espèces d'ignames (Dioscorea sp.) par cytometrie en flux. *Can. J. Bot.* **70**: 996–1000.

Hash, C.T and Bramel-Cox, P.J. 2000. Survey of marker application. In: *Haussmann, H.H. Geiger, D.E. Hess, C.T. Hash, and P. Bramel-Cox (eds.).2000. Application of molecular markers in plant breeding.* Training manual seminar held at IITA, Ibadan, Nigeria. 16-17, August 1999.

Hastie, T., Tibshirani, R. and Friedman, J. 2009. Hierachical Clustering. The Elements of Statistical Learning (2ed). New York: York: Springer.pp. 520-528.

Hearne, C.M, Ghosh, S and Todd, J.A. 1992 Microsatellites for linkage analysis of genetic traits. *Trends in Ecology and Evolution* **8**: 288-94.

Hintum, Th.Van., Von, R.. Bothmer, G., Fischbeck, G. and H. Knüpffer. 1990: The Establishment of the Barley Core Collection. *Barley Newsletter* **34**, 41-42.

Hintum, Th.Van. 1990. The Concept and Application of core Collection in Genebanks.

Holbrook, C.C., Anderson, W.F. and Pittman. R.N. 1993: Selection of a core collection from the U.S. germplasm collection of peanut. *Crop Science* **33**, 859-861.

Hoogendijk, M. and D. Williams. 2001. Characterizing genetic diversity of home garden crop species: some examples from the Americas. *In. Proceedings of the*

Second International Home Gardens Workshop, 17–19 July 2001, Witzenhausen, Federal Republic of Germany. Home gardens and *in situ* conservation of plant genetic resources in farming systems. Eds. J.W. Watson and P.B. Eyzaguirre. Pp. 34-40.

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

International Institute of Tropical Agriculture. 1995. *Yams Research at IITA: 1971-1993.* Crop Improvement Division, IITA, Ibadan, Nigeria. 38pp.

Cornell University, 2003. Genetic Diversity Analysis- Learning Module-Molecular Markers 2.

IPGRI/IITA. 1997. Descriptors for Yam (Dioscorea spp.). International Institute of Tropical Agriculture, Ibadan Nigeria/International plant genetic resources institute, Rome, Italy.

Jain, A. K. and Dubes, R. C. 1988. Algorithms for Clustering Data. Prentice Hall.

Karp, A., Edwards, K., Bruford, M., Vosmann, B., Morgante, M., Seberg, O., Kremer, A., Boursot, P., Arctander, P., Tautz, D. and Hewitt, G. 1997. Newer molecular technologies for biodiversity evaluation: opportunities and challenges. *Nature Biotechnol.* **15**:625-628.

Katsiotis, A., Hagidimitriou, M., Drossou, A., Pontikis, C. and Loukas, M. 2003. Genetic relationships among species and cultivars of Pistacia using RAPDs and AFLPs. *Euphytica* **132**: 279-286.

Kaufman, L. and Rousseeuw P. J. 1990. *Finding groups in data: An introduction to cluster analysis* NY: John Wiley & Sons.

Kawchuk, L. M., Lynch, D. R., Thomas, J., Penner, B., Sillito, D., and Kulcsar, F. 1996. Characterization of Solanum tuberosum simple sequence repeats and

application to potato cultivar identification. *American Potato Journal* **73**:325-335.

Kiambi, D.K., Newbury, H.J., Ford-Llord, B.V. and Dawson, I. 2005. Contrasting genetic diversity among Oryza longistaminata (A. Chev et Roehr) proportions from different geographic origins using AFLP. *African Journal of Biotechnology* **4(4)**: 308-317.

Kumar, S., Tamura, K and Nei, M. 1994. MEGA: molecular evolutionary genetics analysis software for microcomputers. *Bioinformatics*. Oxford University Press.

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

Liu, Z. and Furnier, G.R. 1993. Comparison of allozyme, RFLP, and RAPD markers for revealing genetic variation within and between trembling aspen and big tooth aspen. *Theoretical and Applied Genetics* **87**: 97-105.

Martin, F.W. and A.M. Rhodes. 1978. The relationship of *Dioscorea cayenensis and D. rotundata. Trop. Agric.* (Trinidad) **55**: 193-206.

Matthews, P. and M.J. Ambrose. 1994: Development and use of a „core" collection for the John Innes Pisum collection. *In: F. Balfourier And M.R. Perretant (EDS) Evaluation and Exploitation of Genetic Resources, Pre-breeding. Proceedings of the Genetic Resources Section*. Meeting of Eucarpia, 99-107.

Mekbib, F. 2007. Infra-specific folk taxonomy in sorghum (Sorghum bicolor (L.) Moench) in Ethiopia: folk nomenclature, classification, and criteria. *Journal of Ethnobiology and Ethnomedicine*, **3**:38

Melchinger, A.E. 1990. Use of molecular markers in breeding for oligogenic disease resistance. *Plant Breeding* **104**:1-19.

Michalski, R. S. and Stepp, R. E. 1983. Automated Construction of Classifications: Conceptual Clustering Versus Numerical Taxonomy, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **5(4)**, pp. 396-410.

Mignouna, H. D., Abang M. M. and R. Asiedu. 2003a. Harnessing modern biotechnology for tropical tuber crop improvement: Yam (Dioscorea spp.) molecular breeding. *African Journal of Biotechnology* Vol. **2**, No. **12**. Pp. 478-485.

Mignouna, H.D, Ellis, N. T.H., Knox, M. R., Asiedu, R. and Quat N.N. 1998. Analysis of genetic diversity in Guinea yams (Dioscorea spp) Using AFLP Fingerprinting. *Trop. Agric.* (Trinidad) Vol. 75. No.2.

Mignouna, H. D., Abang M.M. and Fagbemi S.A. 2003b. A comparative assessment of molecular marker assays (AFLP, RAPD and SSR) for white yam (Dioscorea rotundata) germplasm characterization. *Ann. Appl. Biol.* 142: 269-276

Mignouna, H. D, Abang, M.M and Fagbemi, S.A. 2003c. A comparative assessment of molecular marker assays (AFLP, RAPD and SSR) for white yam (Dioscorea rotundata) germplasm characterization. *Ann. Appl. Biol.* **142**: 269-276.

Naylor, R.L., Falcon, W.P., Goodman, R.M., Jahn, M.M., Sengooba, T., Tefera, H and Nelson, R.J. 2007. Biotechnology in the developing world: a case for increased investments in orphan crops. New approaches to plant breeding of orphan crops in Africa. Location: Bern, Switzerland Date: 19 - 21 September 2007.

Nei M and Li WH. 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proc. Natl. Acad. Sci.* USA **76**: 5269-5273.

Nei, M. 1987. Molecular Evolutionary Genetics. Columbia University Press: New York.

Onyilagha J.C. and Lowe J. 1985. Studies on the relationship of Dioscorea cayenensis and Dioscorea rotundata cultivars. *Euphytica* **35**: 733–739.

Otoo E. 2001. Yam production in Ghana: A food security enhancer or an environmental degrader. In. M. Akoroda, Ed.. *Proceedings of the 8th Symposium of the International Society for Tropical Root Crops-Africa Branch Triennial Workshop*, Ibadan, Nigeria. November 12-16, 2001. pp. 387-391.

Otoo E. 2001. Yam Production in Ghana: To Stake or Not To Stake. *In M. Akoroda, ed. Proc. of the 8th Symposium of the International Society for Tropical Root Crops –Africa Branch (ISTRC-AB). Triennial Workshop*, Ibadan, Nigeria. November 12-16, 2001. pp. 387-391.

Otoo E. 2007. Farmer Participatory Breeding- A Waste or a Must: The Case of Yam Breeding in Ghana. *In. M. Akoroda, ed Proc. 13th Symposium of International Society of Tropical Root Crops (ISTRC)*, Arusha, Tanzania. 9-15th November 2003.

Pejic I, Ajmon-Marson P, Morgante M, Kozumplick V, Castiglioni P , Tarmino G and Motto M. 1998. Comparative analysis of genetic similarity among maize inbred lines detected by RFLPs, RAPDs, and AFLPs. *Theor. Appl Genet*. **97**: 1248-1255.

Peleman, J.D. and van der Voort, J.R. 2003. The challenges in Marker Assisted Breeding. *Excopia Leafy Vegetables* (eds). Th.J.L. van Hiutum, A. Lebeda, D. Pink and J.W. Schut. Pp. 125-130.

Pemberton, J.M., Slate, J. Baneroft, D.R and Barrett, J.A. 1995 Non-amplifying alleles at microsatellite loci: a caution for parentage and population studies. *Molecular Ecology* **4**: 249-252.

Perrier, X., Flori, A. and Bonnot, F. 2003. Data analysis methods. *In: Hamon, P., Seguin, M., Perrier, X. , Glaszmann, J. C. Ed., Genetic diversity of cultivated tropical plants. Enfield,* Science Publishers. Montpellier.pp 43 - 76.).

Powell, W., Morgante, M., Andre, C., Hanafey, M., Vogel, I., Ting, Y. S. and Rafalski A. 1996. The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) makers for germplasm analysis. *Mol. Breed.* **2**: 225-238.

Qiagen. 2006. DNeasy Plant Handbook. 07/2006. www.qiagen.com.

Radovic, G. And D. Jelovac 1994: The possible approach in maize „core collection" development. In: F. Balfourier And M.R. Perretant (Eds) Evaluation and Exploitation of Genetic Resources, Pre-breeding. Proceedings of the Genetic Resources Section Meeting of Eucarpia. Pp 109-115.

Ravi, M., Geethanjali, S., Sameeyafarheen, F. and Maheswaran, M. 2003. Molecular marker based genetic diversity analysis in rice (Oryza sativa L.) using RAPD and SSR markers. *Euphytica* **133**: 243-252.

Riedy, M. F., W. J. Hamilton, and C. F. Aquadro. 1992. Excess of non-parental bands in offspring from known primate pedigrees assayed using RAPD PCR. *Nucleic Acids Res.* 20:918.

Rosenberger, A. 2006. http://www.visionlearning.com/library/module_viewer.php?mid=70. Visited on 7th July 2006.

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

SAS. 1989. SAS/SAT User Guide version 6.4 th edn. SAS Institute Inc., Cary, North Carolina

Schneider, S., D. Roessli and L. Excoffier. 2000. Arlequin. A Software for population Genetics Data Analysis. Version 2.000. Genetics and Biometry Laboratory, Dept of Anthropology, University of Geneva, Switzerland.

Scotti, I., Paglia, G., Magni, F and Morgante M. 1999. Microsatellite markers as a tool for the detection of intra- and interpopulational genetic structure. In: *Which DNA Marker for Which Purpose? Final Compendium of the Research Project Development, optimisation and validation of molecular tools for assessment of biodiversity in forest trees in the European Union DGXII Biotechnology FW IV Research Programme Molecular Tools for Biodiversity.* Gillet, E.M. (ed.). 1999. http://webdoc.sub.gwdg.de/ebook/y/1999/whichmarker/index.htm

Shannon, C. E., and Weaver, W. 1949. The Mathematics Theory of Communication. Univ. Illinois Press, Urbana

Smith, J.S.C and Smith, O.S. 1992. Fingerprinting crop varieties. *Adv. Agron* **47**: 85-140.

Sneath, H.A. and Sokal, R.R. 1973. Numerical taxonomy. The principles and practice of numerical classification. Freeman WH and Co. San Francisco. P. 173.

Sonnante, G., Stockton, T., Nodari, R.O., Becerra Velàsquez, V.I. and Gepts, P. (1994). Evolution of genetic diversity during the domestication of common-bean (Phaseolus vulgaris L.). *Theoretical and Applied Genetics* 89: 629-635.

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

Tautz, D. 1989. Hypervariability of simple sequences as a general source for polymorphic DNA markers. *Nucleic Acids Res.* **17**: 6463- 6471.

Thormann, C.E., Ferreira, M. E., Camargo L.E.A., Tivanga, J.G., Osborn, T. C. 1994. Comparison of RFLP and RAPD markers for estimating genetic relationships within and among cruciferous species. *Theor. Appl. Gene.***88**:973-980.

Tostain, S., Agbangla, C. and Daïnou, O. 2002. Diversité AFPL des ignames sauvages D. abyssinica et D. praehensilis. *Annales des Sciences Agronomiques du Bénin* **3(2)**: 1-20.

Tostain, S., Okry, F.K., Baco, N.M., Mongbo, R.L., Agbangla, C. and Daïnou, O. 2003. La domestication des ignames Dioscorea abyssinica dans les sous préfectures de Sinendé et de Bantè au Bénin (Afrique de l'Ouest). *Annales des Sciences Agronomiques du Bénin* **4(1):** 33-54.

Vernier, P., Orkwor, G.C., Dossou, A.R. 2003. Studies on yam domestication and farmers' practices in Benin and Nigeria. *Outlook Agri.* 32: 35-41

Vos P, Hogers R, Bleeker M, Reijans M, VanDe Lee T, Hornes M, Frijters A, Pot J, Poleman I, Kuiper M, Zabeau M. 1995. AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Res.* **23**: 4407- 4414.

Ward, J. H. 1963 Hierarchical grouping to optimize an objective function. *J Amer Stat Ass* 58:236–244.

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

Williams, J.G., Hanafey, M.K., Rafalski, J.A. and Tingey S.V. 1993. Genetic analysis using random amplified polymorphic DNA markers. *Methods in Enzymology* **218:** 704-740.

Williams, J.G.K., Kubelik, A.R., Livak, K.J., Rafalsky, J.A. and Tingey, S.V. 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Research* **18**: 6531-6535.

Zannou, A. 2006. Socio-economic, agronomic and molecular analysis of yam and cowpea diversity in the Guinea-Sudan transition zone of Benin. PhD thesis, Wageningen University.

APPENDICES

Appendix 1: List of descriptors used for the study

SNO	IPGRI NO	IPGRI DESCRIPTOR	CODE
1	7.1.1	Days to emergence [d];<14=EARLY=0; ≥14=LATE=1	A
2	7.1.17	No of sprouts/seedyam min;<5=0, ≥5=1	B
3	7.1.18	Stem colour; green= 0, non-green=1	C
4	7.1.19	Number of internodes to first branching; <3 =0, ≥3=1	D
5	7.1.2	Stem length [cm]; >120 =0, ≤120=1	E
6	7.1.24	Absence/presence of waxiness; absence =0, present =1	F
7	7.1.34	Spines on stem base; Few =0, Many =1	G
8	7.1.35	Spines on stem above base; Few =0, Many =1	H
9	7.1.36	Spine position; wing =0, else =1	I
10	7.1.37	Spine shape; straight =0, else= 1	J
11	7.1.38	Spine length; short =0, long =1	K
12	7.1.39	Absence/presence of coalescent spines; absence =0, present =1	L
13	7.1.40	Colour of spot at spine base; Red =0, else =1	M
14	7.1.5	Absence or presence of waxiness; absence =0, present =1	N
15	7.1.7	Wing colour; green =0, non-green =1	O
16	7.2.1	First leaf emergence (after sprouting in days); 1 (early) =0, >1(late) =1	P
17	7.2.10	Leaf density; low = 0, high = 1	Q
18	7.2.12	Leaf type ; simple = 0, compound =1	R
19	7.2.12.1	Leaf margin; entire = 0, serrate =1	S
20	7.2.12.2	Leaf lobation; shallowly lobed = 0, deeply lobed = 1	T
21	7.2.14	Leatheriness; No =0, Yes = 1	U
22	7.2.15	Leaf colour, yellowish = 0, non-yellowish = 1	V
23	7.2.16	Leaf vein colour (upper surface) ; yellowish = 0, non-yellowish = 1	W
24	7.2.17	Leaf vein colour (lower surface); yellowish = 0, non-yellowish = 1	X

Appendix 1, Continued

25	7.2.18	Leaf margin colour; green = 0, non-green = 1	Y
26	7.2.2	Number of leaves (30 days after emergence); <60 = 0,≥60=1	Z
27	7.2.22	Leaf shape; ovate = 0, non-ovate = 1	AA
28	7.2.24	Undulation of leaf; few = 0, many = 1	AB
29	7.2.25	Distance between lobes; none = 0, some = 1	AC
30	7.2.27	Downward arching of leaf along main vein; no = 0, yes = 1	AG
31	7.2.28	Upward folding of leaf lobes to form a cup; no = 0, yes = 1	AH
32	7.2.29	Downward arching of leaf lobes; no = 0, yes = 1	AI
33	7.2.3	Leaf colour; light green = 0, other =1	AJ
34	7.2.30	Position of the widest part of the leaf; third upper = 0, other = 1	AK
35	7.2.32	Tip length; <2mm = 0, ≥2mm = 1	AL
36	7.2.33	Tip colour; light green = 0, other = 1	AM
37	7.2.34	Petiole length; ≤5cm = 0, >5cm =1	AN
38	7.2.35	Petiole length in correlation to leaf blade; <2mm = 0, ≥2mm = 1	AO
39	7.2.37	Mature Petiole colour; all green with purplish base =0, other = 1	AP
40	7.2.38	Petiole wing colour; green = 0, non-green = 1	AQ
41	7.2.40	Absence/presence of stipules; Absence = 0, present = 1	AR
42	7.2.5	Vein colour; yellowish =0, Non-yellowish=1	AS
43	7.2.6	Young Petiole colour; all green with purplish base =0, other = 1	AT
44	7.2.7	Petiole wing colour; green = 0, non-green = 1	AU
45	7.3.1	Flowering: none =0, flowered=1	AV
46	7.3.10	Number of female flowers per inflorescence; ≤10 = 0, >10 = 1	AW
47	7.3.11	Flower colour; yellowish =0; light green =1)	AX
48	7.3.12	Female flower length; ≤2.5cm = 0, >2.5cm =1	AY
49	7.3.13	Female flower diameter [mm]; ≤2.0mm = 0, >2.0mm =1	AZ
50	7.3.14	Male flower diameter; ≤2.0mm = 0, >2.0mm =1	BA

Appendix 1. Continued

51	7.3.15	Male flower length [cm]<3=0;≥3=1	BB
52	7.3.2	Days to flowering after emergence [d]; <50 =0, ≥50	BC
53	7.3.3	Sex; female=0;male=1	BD
54	7.3.4	Inflorescence position; 0=up, 1=down	BE
55	7.3.6	Inflorescence smell; absence = 0, present = 1	BF
56	7.3.8	Average length of inflorescence; ≤5cm = 0, >5cm =1	BG
57	7.3.9	Number of inflorescences per internode; <4=0;≥4=1	BH
58	7.4.1	Fruit formation; No = 0, Yes = 1	BI
59	7.6.10	Absence/presence of rhizome; absence = 0, present =1	BJ
60	7.6.11	Spininess of roots; sparse = 0, dense = 1	BK
61	7.6.12	Absence/presence of anchor roots; absence = 0, presence = 1	BL
62	7.6.13	Sprouting at harvest; no=0; yes= 1	BM
63	7.6.14	Tuber shape; round = 0, else = 1	BN
64	7.6.15	Tendency of tuber to branch; non-branched = 0, branched =1	BO
65	7.6.16	Place where tuber branches; upper third =0, other=1	BP
66	7.6.17	Tuber length; ≤20cm = 0, >20cm = 1	BQ
67	7.6.18	Tuber width [cm], <8 =0, ≥8 =1	BR
68	7.6.19	Roots on the tuber surface; few = 0, many =1	BS
69	7.6.19.1	Spiny roots on the tuber surface; few =0, many=1	BT
70	7.6.2:	Maturity (tubers) after emergence in months; ≤6=0; >6=1	BU
71	7.6.20	Place of roots on the tuber; lower =0, other =1	BV
72	7.6.21	Prickly appearance of the tuber; NO = 0, YES =1	BW
73	7.6.22	Wrinkles on tuber surface; few =0, many=1	BX
74	7.6.23	Absence/presence of blisters on tuber surface; absence = 0, present = 1	BY
75	7.6.24	Absence/presence of cracks on the tuber surface; absence = 0, present = 1	BZ
76	7.6.25	Tuber skin thickness; <1mm=0, ≥1 mm	CA

Appendix 1. Continued

77	7.6.26	Tuber skin colour (beneath the bark); light maroon =0, other =1	CB
78	7.6.27	Sprouting [%] <80=0; ≥80=1	CC
79	7.6.27.1	Number of months before sprouting; <2 months = 0, ≥2 months	CD
80	7.6.28	Hardness of tuber; (When cut with a knife) hard = 0, easy =1	CE
81	7.6.29	Skin colour at head of the tuber; white=0; purple=1	CF
82	7.6.30	Flesh colour at central transverse cross-section; white = 0; non-white =1	CG
83	7.6.31	Flesh colour of lower part of tuber; white = 0; non-white =1	CH
84	7.6.32	Uniformity of flesh colour in cross-section; (From cortex to centre), no=0; yes= 1	CI
85	7.6.33	Texture of flesh; smooth =0, non-smooth =1	CJ
86	7.6.34	Time for flesh oxidation after cutting; <1min = 0, ≥=1	CK
87	7.6.35	Flesh oxidation colour; grey = 0, non-grey =1	CL
88	7.6.36	Amount of gum released by cut tuber; low = 0, high= 1	CM
89	7.6.37	Ability of cut tuber to irritate human skin (When tuber is rubbed on the arm). Low = 0, high =1	CN
90	7.6.4	Number of tubers per hill; ≤1 = 0; >1=1	CO
91	7.6.5	Relationship of tubers; fused = 0; non-fused = 1	CP
92	7.6.6	Absence/presence of corms; absence = 0; present =1	CQ
93	7.6.7	Corm size (In relation to tuber size); small = 0, big =1	CR
94	7.6.8	Corm ability to be separated from tuber; NO =0, YES = 1	CS
95	7.6.9	Corm type; irregular = 0, regular = 1	CT
96	8.1.1	Total number of harvested tubers; ≤10 =0, >10 =1	CU
97	8.1.2	Total weight of harvested tubers [kg]; <25=0, ≥25 =1	CV
98	8.3.1	Ease of peeling; difficult = 0, easy = 1	CW
99	8.3.10	Stickiness of cooked tuber; non-stick = 0, sticky=1	CX
100	8.3.11	Flavour of cooked tuber; non-flavoured =0, flavoured =1	CY

Appendix 1. Continued

101	8.3.14	Absence/presence of moisture on cooked tuber; absence = 0; present= 1	CZ
102	8.3.15	Overall assessment of cooked tuber; low = 0, high = 1	DA
103	8.3.2	Preferred cooking method; boiled = 0, other=1	DB
104	8.3.2.1	Poundability of boiled tuber; poor = 0, good = 1	DC
105	8.3.3	Cooking time to softness [min]; <20 =0, ≥20 =1	DD
106	8.3.4	Discolouration of cooking water; low = 0; high =1	DE
107	8.3.5	Appearance of tuber after cooking; poor = 0,good =1	DF
108	8.3.6	Colour of tuber after cooking; white, non-coloured = 0, highly coloured=1	DG
109	8.3.7	Attractiveness of cooked tuber; non-attractive = 0, attractive =1	DH
110	8.3.8	Erosion of tuber upon cooking; no=0; yes= 1	DI
111	8.3.9	Texture of cooked tuber; smooth = 0, grainy =1	DJ

Appendix II: Similarity matrix of morphological data

Appendix III: Similarity matrix of molecular data

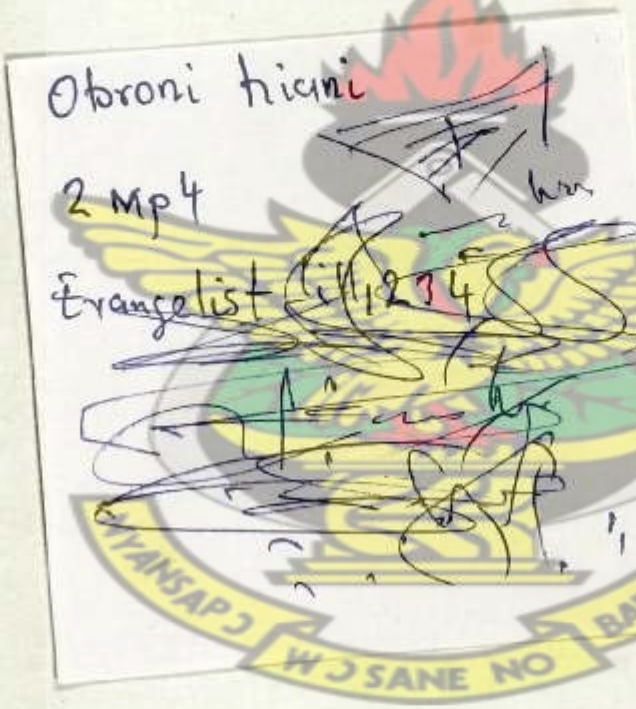
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39
2	84.9																																						
3	125.6	114.1																																					
4	97.8	83.0	116.5																																				
5	110.0	89.9	84.8	101.1																																			
6	108.6	77.6	83.7	112.9	51.1																																		
7	117.2	108.3	129.2	107.6	127.6	127.8																																	
8	103.0	80.6	111.1	97.7	89.7	83.1	121.1																																
9	118.0	105.5	81.2	112.6	73.1	76.7	132.3	101.4																															
10	91.0	98.5	145.2	124.6	123.6	118.8	136.6	125.6	139.3																														
11	126.1	109.4	72.9	104.6	58.8	77.2	138.5	106.3	69.9	140.7																													
12	71.8	93.0	122.7	100.5	115.0	103.4	117.2	106.1	128.7	79.9	127.7																												
13	111.9	82.5	78.3	108.3	50.3	28.2	132.4	78.4	70.9	122.0	70.0	107.0																											
14	127.5	124.4	105.5	125.6	104.9	97.2	124.1	105.3	123.8	145.9	119.3	113.4	101.2																										
15	107.9	66.6	104.5	70.7	109.4	103.2	118.7	96.6	109.3	119.3	113.9	103.2	99.4	122.1																									
16	126.1	109.4	72.9	104.6	58.8	77.2	138.5	106.3	69.9	140.7	0.0	127.7	70.0	119.3	113.9																								
17	100.3	109.3	102.6	112.2	82.8	81.6	123.0	87.3	79.9	134.3	93.5	114.1	88.0	92.9	119.7	93.5																							
18	85.1	91.5	120.5	100.6	103.9	112.0	129.6	112.8	116.3	106.7	108.3	103.9	108.0	144.1	103.5	108.3	126.8																						
19	102.7	94.2	83.4	106.1	68.1	53.4	127.9	75.5	78.4	130.2	84.5	105.0	45.5	97.6	109.3	84.5	75.3	117.5																					
20	119.8	107.4	136.4	106.7	132.5	130.2	103.6	115.8	151.1	119.8	148.6	95.7	133.1	89.9	104.5	146.6	129.0	136.9	132.8																				
21	123.1	118.2	73.9	110.5	91.3	92.3	123.2	103.3	93.8	130.4	98.0	102.6	87.6	109.1	108.3	98.0	101.4	124.6	83.2	126.6																			
22	136.0	120.7	52.2	104.1	86.6	92.5	138.5	117.9	86.5	148.6	50.9	127.7	86.6	107.8	101.8	50.9	106.5	119.7	98.7	137.4	83.8																		
23	55.1	87.1	135.6	92.2	106.4	108.0	124.7	105.0	129.0	88.4	123.1	71.9	110.7	122.5	109.8	123.1	112.7	88.3	107.4	112.2	126.9	133.2																	
24	123.7	116.3	61.7	116.3	85.5	89.9	132.9	119.5	99.8	139.0	75.9	117.9	92.5	94.9	113.4	75.9	100.5	123.4	104.0	127.5	86.9	56.3	120.2															
25	128.3	119.7	74.6	118.9	96.3	91.1	128.1	96.8	104.0	146.9	103.1	119.1	90.8	84.5	113.9	103.1	93.2	141.0	81.9	119.6	91.3	89.7	127.7	91.5															
26	111.2	125.3	127.9	106.2	134.8	139.8	108.7	126.0	133.6	136.6	131.6	115.9	136.7	120.1	106.4	131.6	132.4	103.7	136.6	114.0	123.9	121.3	114.5	131.9	132.3													
27	110.5	82.1	96.4	86.0	116.8	113.2	110.3	117.2	113.6	128.3	105.2	113.3	118.3	130.1	68.2	105.2	121.6	107.8	127.1	114.2	112.5	92.0	118.1	94.3	117.9	110.5													
28	85.8	55.5	109.4	92.1	98.9	95.5	99.9	100.4	93.0	113.2	109.3	108.4	101.5	135.7	89.5	109.3	109.1	91.9	105.0	120.5	119.1	120.5	101.4	113.5	124.6	125.8	78.2											
29	130.3	126.2	85.6	108.3	94.5	103.2	130.0	112.2	99.1	143.3	99.4	112.0	98.0	104.8	116.9	99.4	110.1	126.6	94.6	124.3	46.8	81.8	125.8	84.1	100.7	126.9	119.5	125.7											
30	94.6	63.0	130.2	93.8	99.5	100.5	114.3	102.2	122.8	83.5	121.7	82.4	104.3	127.0	91.5	121.7	117.0	105.2	113.7	87.7	119.8	131.9	85.1	118.7	132.1	135.7	102.6	82.9	123.5										
31	121.8	122.2	112.1	128.9	107.9	96.0	106.1	104.1	115.2	144.9	124.1	135.2	101.9	73.7	129.3	124.1	84.2	145.9	93.3	115.8	110.8	123.7	125.0	109.6	96.0	118.9	137.4	129.1	118.6	129.2									
32	102.1	77.4	112.5	111.4	91.4	75.3	133.5	79.2	106.1	113.8	106.0	100.0	49.2	114.3	101.3	106.0	99.4	107.3	92.3	121.8	119.1	117.6	104.4	107.2	118.1	144.4	112.0	95.7	127.4	89.5	114.5								
33	100.9	69.6	117.7	92.8	102.7	96.1	104.9	109.6	95.7	103.8	103.9	98.1	104.2	138.4	87.9	103.9	107.9	98.4	114.0	113.8	105.3	115.7	110.3	117.6	135.8	124.2	75.6	65.6	115.5	72.5	135.7	104.1							
34	123.5	121.5	87.5	121.9	99.5	93.5	103.5	105.9	105.3	151.1	134.0	126.5	97.8	69.4	119.6	114.0	88.5	140.4	88.9	113.0	90.3	102.0	133.6	99.1	79.1	106.9	120.9	122.0	100.7	137.2	61.2	120.0	129.0						
35	119.7	106.9	114.5	125.7	92.4	73.6	120.7	101.7	107.3	139.8	102.7	117.0	82.0	110.8	128.4	107.7	98.4	126.5	76.8	142.8	108.3	119.9	119.2	116.9	109.2	152.3	133.4	119.2	117.8	124.6	112.3	90.8	120.1	109.2					
36	93.7	68.3	112.5	54.7	103.4	100.1	108.4	84.4	120.9	118.8	114.2	84.9	96.0	112.4	70.6	114.2	118.5	101.8	90.6	96.8	100.3	113.8	88.7	118.4	109.7	111.8	96.1	90.6	109.6	86.6	119.5	100.6	101.4	117.0	114.0				
37	112.0	101.4	74.2	105.6	78.0	70.7	119.6	94.7	93.3	133.0	103.2	104.9	75.9	82.9	98.0	103.2	86.4	126.5	65.5	118.3	72.6	89.8	110.7	82.8	64.4	129.8	108.3	109.8	82.9	115.1	88.4	103.2	117.5	73.6	90.6	98.4			
38	121.8	103.8	48.2	106.6	70.6	73.6	125.3	99.8	74.7	137.8	78.6	116.0	67.3	97.7	92.3	78.6	97.7	110.2	81.8	130.4	57.0	59.8	125.3	83.5	87.6	119.2	97.0	104.6	69.8	120.8	107.4	105.2	107.2	86.0	107.3	105.2	66.1		
39	130.1	123.9	107.1	125.5	92.7	96.6	120.6	98.1	101.7	150.4	102.3	133.6	92.1	93.4	135.7	102.1	96.1	130.4	90.1	126.0	102.4	114.1	132.8	118.7	104.8	126.8	139.1	125.1	112.4	136.3	102.0	122.5	127.1	92.5	113.0	123.4	107.0	99.4	
40	54.9	95.6	136.3	108.0	116.5	115.0	116.8	94.6	126.9	160.7	132.5	83.8	118.2	118.8	116.5	132.5	94.9	95.8	114.9	112.8	128.7	142.0	70.3	130.0	123.8	106.2	119.1	102.6	136.2	104.8	118.2	112.1	109.4	121.3	125.6	104.4	122.7	127.5	121.3
41	117.7	109.9	76.3	110.9	88.8	77.9	130.6	98.4	87.5	141.9	104.3	107.8	77.7	86.1	103.6	104.3	94.9	132.8	83.8	124.5	75.6	91.0	120.4	92.8	59.9	131.5	119.0	115.3	81.1	126.5	96.6	108.3	127.4	78.8	96.4	96.8	48.9	73.8	110.0
42	97.4	107.6	86.4	116.4	79.8	74.6	123.5	99.9	73.2	132.2	88.7	107.1	79.0	97.0	125.2	88.7	68.1	120.9	88.1	132.3	85.7	99.6	106.8	85.4	91.4	137.6	120.3	93.7	95.6	115.5	96.2	95.2	109.1	92.7	90.1	111.0	80.9	84.2	95.3
43	124.7	116.8	108.6	126.8	89.5	87.4	110.5	101.5	103.5	144.4	102.3	128.1	93.9	85.0	137.1	102.3	88.5	131.2	94.0	119.3	105.8	114.3	128.2	109.9	104.0	128.7	129.5	114.3	114.1	129.3	94.7	115.3	116.8	83.3	102.8	125.6	99.9	101.0	48.5
44	135.9	129.7	75.8	119.6	99.7	106.9	135.1	118.1	107.8	145.8	93.7	117.0	102.8	110.4	120.6	93.7	116.3	130.8	101.7	128.7	97.7	78.6	136.8	81.6	95.3	131.8	113.4	129.9	85.2	126.0	124.7	131.1	118.2	107.2	123.2	116.3	91.6	77.1	117.7
45	136.0	120.7	52.2	104.1	86.6	92.5	138.5	117.9	86.5	149.6	50.9	127.7	86.6	107.8	101.8	50.9	106.5	119.7	98.7	137.4	83.8	0.0	133.2	56.3	89.7	121.3	92.0	120.5	81.8	131.9	123.7	117.6	115.7	102.0	119.9	113.8	89.8	59.8	114.1
46	82.3	111.3	128.1	116.4	136.5	129.2	133.4	126.7	138.3	115.0	147.1	83.7	129.2	123.4	108.6	147.1	128.9	118.8	121.0	117.2	127.9	138.0	89.9	128.1	119.0	117.7	115.3	114.2	135.2	119.4	130.1	123.0	130.5	126.0	136.5	103.1	110.3	126.9	150.9
47	91.7	74.3	133.7	101.6	116.2	104.3	93.8	89.7	129.5	114.0	129.9	97.7	108.1	103.2	100.5	129.8	106.1	118.3	107.6	84.9	131.3	139.4	94.4	129.6	122.7	107.3	111.3	93.2	139.3	86.3	90.8	93.7	102.4	104.1	119.0	85.6	118.9	127.7	113.5
48	117.7	96.4	89.0	120.1	71.4	61.8	123.6	92.1	68.1	132.3	90.1	120.4	55.2	115.2	110.6	90.1	92.0	118.8	62.1	144.1	101.2	103.5	121.3	106.1	94.1	146.8	125.3	101.2	110.1	115.1	107.3	97.4	115.4	106.8	77.8	110.7	83.3	83.9	107.7
49	103.1	97.2	96.9	108.2	66.1	64.4	125.9	87.6	85.9	133.0	90.3	110.9	69.7	104.0	118.0	90.3	75.6	118.8	64.0	137.5	85.7	103.7	108.9	96.0	99.6	138.4	119.7	103.3	95.3	115.6	102.0	98.9	105.8	93.8	81.7	105.4	76.7	83.7	96.1
50	97.4	64.1	97.1	103.6	79.6	61.5	121.6	66.2	91.3	114.9	97.9	101.3	67.6	114.8	93.1	97.8	95.4	109.4	73.3	122.4	110.9	110.3	105.3	108.1	104.5	138.2	104.7	77.6	119.4	90.2	110.1	56.0	96.2	106.1	95.9	90.1	87.0	95.8	114.5
51	126.4	109.7	72.9	106.6	69.3	77.6	140.1	106.6	70.4	140.9	7.9	129.5	70.5	119.2	113.8	7.9	93.8	108.6	84.9	145.5	98.0	50.9	123.4	75.9	103.1	131.6	105.1	109.5	96.3	122.0	125.8	106.3	104.2	113.9	108.0	116.0	103.2	78.5	102.4
52	113.7	100.4	74.6	116.6	65.2	69.1	128.4	96.1	59.4	135.3	76.0	124.3	62.6	119.2	113.9	76.0	86.6	107.5	69.4	147.3	87.9	91.5	124.4	93.6	105.5	137.2	117.9	95.0	97.8	118.3	110.8	102.2	103.7	103.5	103.5	115.2	87.6	87.1	96.1
53	127.1	119.6	125.1	123.5	123.3	115.5	114.6	99.0	138.7	142.4	138.9	115.8	118.9	82.2	117.0	138.9	114.1	139.7	117.0	93.8	119.1	129.3	125.4	118.7	103.9	115.8	126.4	132.1	127.8	123.8	83.3	91.1	138.8	84.3	116.2	114.2	101.3	116.5	115.4
54	118.6	102.2	97.7	131.1	84.2	71.5	126.3	105.7	100.5	128.6	92.6	116.5	76.9	111.1	121.8	92.6	93.6	129.8	88.3	140.9	113.6	105.4	117.5	88.3	104.2	155.0	120.5	115.9	121.8	110.6	108.0	92.9	118.0	117.4	76.0	122.7	95.9	98.4	120.2
55	115.9	102.7	73.1	102.1	81.0	71.3	129.0	85.1	93.2	137.0	97.4	103.4	65.8	81.2	91.1	97.4	92.5	124.1	59.1	121.3	58.5	83.0	114.8	86.2	70.1	123.4	112.0	118.6	73.5	120.4	94.0	103.7	120.7	79.9	91.2	88.4	43.0	55.9	101.8
56	103.7	96.8	112.1	99.9	94.4	92.1	128.7	65.8	91.9	129.4	112.4	111.6	87.7	122.7	105.4	112.4	103.1	119.9	75.1	129.4	109.6	123.4	106.9	125.6	102.0	139.8	124.4	96.7	113.2	106.8	117.0	88.4	117.5	117.1	107.3	90.3	92.4	108.1	107.4
57	119.5	97.9	129.3	103.0	130.1	122.2	86.4	106.7	144.2	135.7	144.5	113.4	125.4	101.4	94.7	144.5	129.5	134.4	123.7	78.0	125.4	135.2	118.1	133.4	111.6	90.1	105.9	112.7	133.7	113.6	97.0	122.4	120.4	87.7	134.3	91.5	109.2	123.0	121.2
58	132.8	114.2	62.6	110.6	88.6	86.7	138.6	110.7	80.2	142.6	82.8	124.7	81.9	107.7	93.4	82.8	101.0	130.8	87.4	134.7	93.4	65.3	131.5	83.5	84.5	132.4	101.6	118.6	101.7	126.2	112.6	113.5	122.1	99.8	115.5	114.3	68.8	74.1	119.6
59	110.1	113.0	74.7	108.0	81.7	84.9	123.2	106.7	95.9	136.6	92.9	104.2	88.9	75.4	110.0	92.9	65.1	121.0	86.9	117.1	65.5	77.7	111.5	54.4	89.6	121.4	104.1	111.1	71.2	115.5	96.2	104.0	114.3	84.0	101.6	103.6	65.4	59.0	105.2
60	122.0	95.5	63.9	108.2	71.7	61.8	131.9	90.9	84.9	131.8	90.7	109.1	55.3	89.9	82.7	90.7	97.9	118.1	71.5	124.9	69.9	75.0	120.9	77.6	78.3	126.5	105.3	112.3	82.6	114.2	99.4	97.4	114.5	86.5	99.7	97.9	52.3	41.0	107.7
61	92.9	66.2	131.5	81.2	101.1	102.1	106.1	91.5	124.2	97.4	123.5	82.8	106.0	125.9	93.8	123.5	117.0	107.2	101.9	89.8	108.6	133.5	83.5	129.2	123.5	129.0	104.7	85.5	113.2	50.3	129.3	102.9	75.2	129.7	113.9	89.8	105.9	122.3	127.9
62	129.2	112.6	71.6	110.6	75.6	81.7	132.4	109.6	96.7	130.4	90.4	104.8	76.4	98.4	102.9	90.4	107.7	121.5	89.1	116.4	56.5	74.7	130.7	81.6	94.1	128.6	114.1	120.4	65.6	109.0	116.5	111.5	107.7	92.8	113.1	106.0	79.6	59.5	113.1
63	118.9	100.7	78.0	100.6	54.8	64.2	132.2	97.3	55.2	134.4	41.0	121.0	56.7	111.0	105.5	41.0	83.3	100.2	73.3	140.7	88.1	65.4	119.8	86.3	107.9	124.4	109.0	101.3	92.0	113.9	117.0	97.7	94.6	106.1	98.8	106.8	94.7	67.0	92.6
64	115.5	98.0	83.2	90.1	95.8	97.8	128.7	90.8	90.6	136.9	99.5	116.8	93.4	117.1	72.1	99.5	102.5	116.1	98.5	125.5	95.6	85.5	127.5	99.7	103.9	118.5	88.8	101.9	103.9	115.9	121.1	94.5	106.5	101.8	124.1	99.3	87.0	76.9	119.8
65	122.3	100.9	80.8	88.3	88.8	99.3	124.1	94.3	104.3	134.8	92.1	116.0	93.8	113.8	89.5	92.1	121.4	108.8	105.1	116.7	91.2	76.8	118.3	83.6	105.0	118.1	93.5	101.1	89.5	110.2	128.9	95.5	112.3	108.2	126.4	92.9	92.2	69.9	116.0
66	136.5	129.7	95.0	129.7	105.6	103.9	108.2	113.1	114.5	157.9	114.8	130.1	106.0	80.5	126.5	114.8	105.5	140.6	105.4	120.2	91.0	102.9	142.1	101.9	92.1	100.1	125.0	130.9	102.1	144.6	74.9	128.4	132.8	48.2	121.0	124.6	93.6	86.3	90.4
67	97.6	91.9	122.9	82.1	125.4	120.4	116.2	111.6	122.6	123.3	134.3	99.0	123.6	117.1	74.9	134.3	111.9	127.1	115.5	97.1	120.0	124.3	99.1	119.8	109.8	120.6	88.7	95.7	126.6	96.3	122.4	109.3	106.2	120.7	130.1	84.7	96.8	118.6	138.0
68	110.9	75.8	100.2	92.5	78.3	77.5	123.8	74.8	104.2	121.9	103.0	98.5	80.8	107.7	85.4	103.0	109.1	109.8	92.6	116.0	105.3	102.9	108.1	99.2	104.0	129.5	101.5	96.0	109.3	92.0	110.6	74.0	106.6	110.7	109.6	78.8	87.0	88.1	122.7
69	126.4	119.2	88.3	109.3	101.0	94.2	130.1	90.3	86.1	149.7	135.0	120.6	89.9	87.5	99.2	105.0	77.7	136.0	78.4	125.3	84.9	91.8	129.4	104.6	59.3	120.8	118.7	122.9	94.7	134.6	93.3	119.3	126.9	81.0	109.2	108.1	66.2	82.1	106.9
70	101.4	66.1	96.2	87.5	105.5	97.9	113.3	97.5	106.1	118.4	121.0	100.4	97.8	117.6	53.4	121.0	119.7	108.9	101.9	102.8	108.0	109.8	108.7	109.3	97.1	118.1	80.3	73.5	116.5	91.4	121.2	98.1	98.8	109.4	125.3	73.4	87.3	92.1	134.6
71	107.0	107.2	66.5	101.2	75.4	74.0	113.9	95.3	76.9	139.9	84.7	104.6	73.2	97.8	112.9	84.7	81.2	117.6	58.1	132.9	66.4	84.7	118.7	86.6	82.0	127.0	113.5	102.7	79.9	124.7	92.9	105.5	111.0	80.3	92.7	94.8	65.7	64.5	90.6
72	122.0	95.4	63.6	107.7	71.5	61.7	131.8	91.0	84.9	131.7	90.1	109.0	55.0	89.7	82.7	90.1	97.9	118.0	71.3	124.7	69.9	74.3	120.5	77.0	78.1	126.5	105.2	112.3	82.1	114.2	99.4	97.3	114.5	86.3	99.6	97.7	52.2	40.8	107.6
73	121.8	94.9	64.1	110.7	69.9	54.7	131.0	95.4	89.9	130.8	75.2	104.8	54.7	99.2	101.3	75.2	98.2	121.4	71.0	131.6	90.1	75.2	121.1	77.3	73.5	138.4	105.9	112.4	101.3	114.4	99.6	93.1	114.6	96.1	94.9	97.6	73.5	70.8	107.4
74	110.9	111.0	59.6	100.4	80.1	79.4	124.7	95.7	87.5	142.8	93.3	108.9	73.9	87.5	101.0	93.3	86.2	124.3	58.8	125.4	59.4	78.1	121.4	84.8	70.4	123.0	112.9	114.1	73.0	127.5	97.3	109.3	121.7	68.1	97.6	94.2	50.4	57.1	99.6
75	136.0	120.7	52.2	104.1	86.6	92.5	138.5	117.9	86.5	149.6	50.9	127.7	86.6	107.8	101.8	50.9	106.5	119.7	98.7	137.4	83.8	0.0	133.2	56.3	89.7	121.3	92.0	120.5	81.8	131.9	123.7	117.6	115.7	102.0	119.9	113.8	89.8	59.8	114.1
76	126.9	122.7	98.5	119.5	88.3	95.1	137.6	112.5	102.6	139.5	75.1	117.6	93.6	122.9	136.4	75.1	105.5	116.6	90.0	139.7	87.9	93.7	123.9	97.0	106.5	146.0	128.4	122.6	87.2	120.3	127.9	106.5	110.7	119.4	96.1	119.2	109.1	102.6	108.5

	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75
	128.1																																			
	110.0	75.6																																		
	115.5	109.6	87.7																																	
	141.4	94.9	103.2	119.9																																
	142.0	91.0	99.6	114.3	78.6																															
	99.2	107.4	115.5	146.2	140.6	138.0																														
	86.5	121.7	110.2	106.3	143.4	139.4	114.0																													
	126.2	81.4	82.0	109.3	114.4	103.5	131.0	121.3																												
	109.5	85.4	71.1	88.5	88.8	103.7	131.0	113.5	85.4																											
	113.4	93.2	90.5	106.7	123.1	110.3	119.0	95.2	79.7	88.9																										
	132.8	104.3	88.7	102.8	93.8	50.8	147.1	130.0	90.4	90.7	98.1																									
	124.7	94.3	76.7	98.0	88.7	91.5	139.5	125.2	71.5	80.2	85.5	76.5																								
	121.6	108.3	117.1	109.1	132.5	129.3	125.0	88.6	131.0	122.4	107.3	138.9	134.6																							
	124.2	105.4	92.3	112.9	115.0	105.4	137.2	117.7	87.7	91.8	94.4	93.1	95.1	127.4																						
	121.7	47.8	85.0	105.3	82.5	83.0	113.8	117.6	82.6	75.3	94.2	97.4	87.2	103.7	97.3																					
	118.4	86.6	92.6	110.4	123.6	123.4	121.4	137.7	87.4	85.1	88.6	112.6	98.8	113.9	113.8	92.6																				
	113.3	115.4	132.0	115.1	138.4	135.2	118.9	55.5	137.0	128.8	114.5	144.4	140.3	77.0	141.4	111.1	120.9																			
	141.9	72.5	99.7	120.9	94.3	65.3	125.8	135.6	82.6	104.8	100.6	82.8	94.2	121.4	97.4	73.3	104.0	128.1																		
	116.9	73.6	84.2	97.6	79.8	77.7	115.1	118.7	102.9	80.9	104.6	92.9	90.0	105.1	97.7	64.6	112.5	121.7	94.2																	
	127.4	62.5	94.1	109.4	87.9	75.0	120.7	121.3	73.5	84.2	87.2	90.6	78.4	109.4	89.5	38.3	100.3	116.5	83.1	72.8																
	103.2	116.1	113.9	121.9	117.8	133.5	118.0	88.6	116.6	104.2	92.5	123.7	116.8	123.9	121.4	109.4	94.4	103.7	127.5	115.5	115.8															
	135.1	81.2	99.4	114.1	69.3	74.7	131.7	132.5	94.4	95.4	102.0	90.4	90.7	121.8	108.2	70.8	116.1	128.3	84.9	78.1	81.2	110.5														
	125.5	95.9	78.7	93.7	97.1	65.4	141.3	122.5	79.7	80.5	86.5	41.8	63.3	132.1	95.1	87.4	104.4	138.1	87.6	82.3	78.9	115.4	79.6													
	126.4	92.5	108.9	121.2	108.9	85.5	122.7	122.9	100.7	106.3	78.6	99.5	96.4	111.3	118.1	85.9	85.6	118.2	81.4	98.7	77.1	117.7	87.5	90.6												
	131.5	98.0	106.0	116.1	100.0	76.8	126.8	123.0	109.6	109.8	84.6	92.1	98.4	107.2	122.1	90.6	97.4	114.6	95.2	85.7	83.3	112.3	83.0	47.8	68.5											
	129.2	92.5	101.3	86.8	107.8	132.9	138.9	108.4	119.8	104.6	120.8	114.8	109.5	89.3	125.9	90.3	129.4	90.8	113.0	87.0	95.7	137.7	101.8	106.5	115.1	109.1										
	112.1	101.1	109.7	132.7	132.8	124.3	90.7	98.2	123.9	122.7	106.7	134.3	132.2	109.6	129.4	104.8	95.8	97.8	105.4	106.2	111.6	94.4	125.9	127.9	91.3	110.5	133.8									
	120.2	92.4	106.5	118.5	112.7	102.9	122.5	105.2	94.8	97.2	58.7	105.1	98.9	105.5	101.8	87.5	85.1	113.1	100.0	94.4	78.9	94.2	92.5	92.8	77.5	74.7	114.3	107.0								
	122.2	60.9	92.0	110.0	102.8	91.6	119.4	124.4	88.3	96.9	106.7	104.9	99.9	109.7	114.6	61.6	98.9	116.7	67.0	89.2	71.4	125.4	93.8	96.4	88.2	108.6	97.1	101.1	106.4							
	115.9	54.8	109.3	133.1	120.5	109.8	93.8	98.9	97.4	114.7	80.6	121.0	108.3	112.9	118.7	90.8	92.9	92.9	90.7	105.9	82.7	93.6	101.2	113.4	78.8	88.0	120.0	72.1	78.1	98.1						
	119.0	64.0	60.0	90.2	88.5	84.7	121.1	119.2	85.1	70.5	89.4	87.3	69.9	117.2	102.5	65.4	94.9	123.6	91.3	66.4	76.8	114.1	84.2	73.8	94.0	92.1	85.5	115.6	89.4	83.2	102.0					
	127.4	52.2	93.9	109.2	87.8	74.3	120.5	121.2	73.4	84.1	87.1	90.0	78.4	109.4	89.5	38.3	100.2	116.4	62.9	72.5	1.6	115.7	60.9	79.6	77.0	82.6	95.8	111.5	78.8	71.3	82.7	78.7				
	127.4	75.3	92.1	105.5	92.4	75.2	127.8	117.8	77.6	84.5	82.4	78.1	83.5	117.2	74.4	68.5	103.3	123.8	67.9	87.2	58.5	115.9	79.2	79.4	95.8	97.3	102.0	122.1	74.4	92.2	95.8	70.6	58.3			
	122.3	48.2	74.9	102.9	82.6	78.1	113.6	122.6	85.5	76.3	94.0	93.2	81.4	109.4	106.5	41.9	96.3	116.4	76.4	63.4	58.1	117.1	66.1	63.8	79.1	86.1	87.9	107.7	96.4	68.4	93.0	50.8	57.8	76.8		
	142.0	91.0	99.6	114.3	78.6	0.0	138.0	139.4	103.5	103.7	110.3	50.9	91.5	129.3	105.4	83.0	123.4	135.2	65.3	77.7	75.0	133.5	74.7	65.4	85.5	76.8	102.9	124.3	102.9	91.8	109.6	84.7	74.3	75.2	78.1	
	133.3	107.8	90.1	106.8	85.9	93.7	146.6	133.3	109.4	91.4	112.5	75.4	98.1	132.2	108.1	104.9	116.6	148.3	114.2	98.8	112.3	109.6	95.5	85.6	126.8	113.5	118.4	143.4	116.3	120.2	135.2	87.9	111.8	95.4	101.0	93.7

Column	Values (top to bottom)
33.00	0.37, 0.88, 0.30, 0.56, 0.56, 0.37, 0.39, 0.70, 0.67, 0.32, 0.68, 0.50, 0.34, 0.53, 0.38, 0.52, 0.37, 0.51, 0.70, 0.35, 0.67, 0.55, 0.41, 0.61, 0.46, 0.35, 0.67, 0.19, 0.65, 0.27, 0.55, 0.36, 0.53, 0.40, 0.53, 0.41, 0.57, 0.39, 0.54, 0.34, 0.38, 0.58, 0.38, 0.60, 0.50, 0.53, 0.28, 0.38, 0.55, 0.49, 0.54, 0.54, 0.58, 0.37, 0.55, 0.25, 0.37, 0.32
34.00	0.69, 0.32, 0.54, 0.48, 0.29, 0.39, 0.62, 0.55, 0.27, 0.58, 0.27, 0.28, 0.46, 0.33, 0.60, 0.53, 0.28, 0.68, 0.22, 0.62, 0.47, 0.39, 0.53, 0.41, 0.10, 0.61, 0.32, 0.64, 0.31, 0.55, 0.30, 0.45, 0.48, 0.51, 0.54, 0.51, 0.29, 0.46, 0.32, 0.61, 0.59, 0.36, 0.40, 0.62, 0.55, 0.32, 0.37, 0.55, 0.52, 0.58, 0.53, 0.58, 0.38, 0.55, 0.30, 0.40, 0.29
35.00	0.73, 0.53, 0.54, 0.59, 0.67, 0.26, 0.58, 0.60, 0.34, 0.73, 0.69, 0.40, 0.66, 0.51, 0.57, 0.11, 0.20, 0.59, 0.25, 0.45, 0.58, 0.59, 0.66, 0.62, 0.75, 0.64, 0.33, 0.58, 0.59, 0.62, 0.60, 0.61, 0.49, 0.54, 0.51, 0.58, 0.48, 0.64, 0.66, 0.57, 0.63, 0.65, 0.53, 0.56, 0.70, 0.71, 0.56, 0.55, 0.57, 0.67, 0.61, 0.54, 0.68, 0.70, 0.72, 0.73
36.00	0.56, 0.47, 0.34, 0.39, 0.69, 0.52, 0.27, 0.72, 0.24, 0.25, 0.52, 0.38, 0.49, 0.34, 0.30, 0.71, 0.33, 0.72, 0.51, 0.39, 0.54, 0.41, 0.30, 0.67, 0.28, 0.71, 0.24, 0.52, 0.30, 0.51, 0.42, 0.54, 0.46, 0.55, 0.26, 0.52, 0.34, 0.31, 0.51, 0.36, 0.39, 0.61, 0.50, 0.29, 0.33, 0.52, 0.61, 0.53, 0.51, 0.57, 0.36, 0.52, 0.29, 0.36, 0.26
37.00	0.28, 0.63, 0.62, 0.50, 0.38, 0.55, 0.54, 0.56, 0.51, 0.37, 0.61, 0.34, 0.54, 0.56, 0.57, 0.56, 0.52, 0.34, 0.60, 0.47, 0.58, 0.53, 0.45, 0.51, 0.48, 0.53, 0.77, 0.60, 0.31, 0.56, 0.33, 0.62, 0.38, 0.63, 0.27, 0.57, 0.56, 0.35, 0.59, 0.52, 0.46, 0.35, 0.61, 0.62, 0.37, 0.29, 0.71, 0.53, 0.40, 0.60, 0.37, 0.55, 0.53, 0.57
38.00	0.57, 0.59, 0.51, 0.25, 0.48, 0.50, 0.50, 0.49, 0.37, 0.58, 0.19, 0.57, 0.51, 0.53, 0.54, 0.55, 0.21, 0.57, 0.33, 0.58, 0.48, 0.47, 0.53, 0.48, 0.47, 0.16, 0.53, 0.16, 0.52, 0.23, 0.61, 0.36, 0.57, 0.16, 0.53, 0.51, 0.18, 0.53, 0.56, 0.38, 0.21, 0.44, 0.51, 0.19, 0.20, 0.16, 0.15, 0.27, 0.56, 0.19, 0.52, 0.48, 0.48
39.00	0.38, 0.59, 0.58, 0.28, 0.01, 0.51, 0.38, 0.58, 0.18, 0.55, 0.11, 0.38, 0.57, 0.30, 0.59, 0.53, 0.26, 0.61, 0.28, 0.29, 0.53, 0.35, 0.61, 0.32, 0.62, 0.17, 0.53, 0.32, 0.59, 0.27, 0.68, 0.12, 0.56, 0.38, 0.38, 0.60, 0.39, 0.29, 0.68, 0.59, 0.33, 0.43, 0.61, 0.56, 0.60, 0.58, 0.64, 0.18, 0.61, 0.50, 0.42, 0.32
40.00	0.68, 0.47, 0.25, 0.73, 0.39, 0.30, 0.54, 0.43, 0.55, 0.36, 0.39, 0.71, 0.33, 0.67, 0.59, 0.45, 0.47, 0.52, 0.52, 0.70, 0.34, 0.71, 0.40, 0.63, 0.42, 0.59, 0.45, 0.58, 0.51, 0.54, 0.42, 0.60, 0.42, 0.42, 0.63, 0.44, 0.51, 0.68, 0.61, 0.48, 0.48, 0.63, 0.62, 0.61, 0.61, 0.55, 0.45, 0.62, 0.40, 0.34, 0.39
41.00	0.51, 0.64, 0.29, 0.72, 0.60, 0.52, 0.58, 0.51, 0.59, 0.69, 0.31, 0.63, 0.34, 0.47, 0.58, 0.58, 0.62, 0.65, 0.26, 0.64, 0.31, 0.68, 0.55, 0.62, 0.54, 0.57, 0.52, 0.65, 0.52, 0.55, 0.48, 0.69, 0.63, 0.57, 0.59, 0.67, 0.62, 0.56, 0.73, 0.77, 0.57, 0.55, 0.58, 0.56, 0.61, 0.64, 0.57, 0.66, 0.68, 0.69
42.00	0.61, 0.59, 0.57, 0.51, 0.35, 0.60, 0.28, 0.58, 0.56, 0.60, 0.53, 0.57, 0.35, 0.54, 0.29, 0.61, 0.51, 0.50, 0.59, 0.57, 0.54, 0.35, 0.50, 0.35, 0.60, 0.38, 0.67, 0.30, 0.63, 0.36, 0.57, 0.50, 0.36, 0.55, 0.60, 0.46, 0.33, 0.52, 0.59, 0.34, 0.32, 0.34, 0.31, 0.25, 0.01, 0.34, 0.68, 0.43, 0.58
43.00	0.60, 0.39, 0.27, 0.51, 0.34, 0.46, 0.37, 0.27, 0.55, 0.30, 0.65, 0.50, 0.60, 0.57, 0.42, 0.22, 0.64, 0.29, 0.61, 0.23, 0.54, 0.34, 0.48, 0.37, 0.52, 0.46, 0.59, 0.33, 0.53, 0.31, 0.28, 0.52, 0.31, 0.26, 0.60, 0.51, 0.31, 0.38, 0.50, 0.61, 0.54, 0.54, 0.58, 0.33, 0.53, 0.26, 0.38, 0.28
44.00	0.76, 0.70, 0.66, 0.65, 0.65, 0.64, 0.67, 0.79, 0.70, 0.29, 0.61, 0.65, 0.60, 0.64, 0.67, 0.31, 0.68, 0.21, 0.69, 0.28, 0.49, 0.27, 0.53, 0.61, 0.70, 0.52, 0.65, 0.56, 0.61, 0.49, 0.72, 0.71, 0.51, 0.69, 0.66, 0.49, 0.51, 0.69, 0.78, 0.53, 0.60, 0.62, 0.62, 0.57, 0.61, 0.53, 0.68, 0.68, 0.70
45.00	0.32, 0.56, 0.35, 0.58, 0.34, 0.17, 0.74, 0.31, 0.70, 0.55, 0.44, 0.60, 0.32, 0.25, 0.78, 0.32, 0.70, 0.25, 0.49, 0.27, 0.53, 0.41, 0.68, 0.18, 0.58, 0.38, 0.61, 0.32, 0.36, 0.49, 0.34, 0.32, 0.58, 0.47, 0.17, 0.26, 0.49, 0.43, 0.46, 0.48, 0.61, 0.22, 0.40, 0.11, 0.26, 0.11
46.00	0.45, 0.37, 0.66, 0.59, 0.32, 0.69, 0.22, 0.60, 0.45, 0.36, 0.53, 0.43, 0.25, 0.60, 0.25, 0.67, 0.50, 0.54, 0.37, 0.50, 0.33, 0.51, 0.42, 0.51, 0.33, 0.40, 0.29, 0.29, 0.55, 0.35, 0.43, 0.61, 0.53, 0.33, 0.37, 0.54, 0.53, 0.54, 0.54, 0.57, 0.40, 0.54, 0.34, 0.38, 0.29
47.00	0.56, 0.34, 0.52, 0.59, 0.52, 0.45, 0.48, 0.25, 0.53, 0.62, 0.65, 0.46, 0.46, 0.47, 0.53, 0.54, 0.23, 0.59, 0.24, 0.61, 0.25, 0.60, 0.23, 0.58, 0.24, 0.51, 0.56, 0.34, 0.59, 0.53, 0.49, 0.36, 0.53, 0.56, 0.36, 0.35, 0.93, 0.97, 0.40, 0.65, 0.35, 0.60, 0.56, 0.59
48.00	0.59, 0.17, 0.37, 0.58, 0.34, 0.59, 0.64, 0.16, 0.64, 0.21, 0.30, 0.54, 0.34, 0.59, 0.31, 0.08, 0.26, 0.56, 0.32, 0.58, 0.29, 0.55, 0.21, 0.55, 0.32, 0.38, 0.62, 0.41, 0.30, 0.68, 0.65, 0.39, 0.45, 0.64, 0.61, 0.64, 0.62, 0.67, 0.37, 0.64, 0.35, 0.47, 0.38
49.00	0.58, 0.56, 0.63, 0.52, 0.56, 0.29, 0.55, 0.36, 0.51, 0.50, 0.51, 0.40, 0.59, 0.49, 0.25, 0.56, 0.24, 0.62, 0.50, 0.64, 0.59, 0.55, 0.21, 0.51, 0.51, 0.26, 0.54, 0.59, 0.59, 0.44, 0.25, 0.45, 0.55, 0.24, 0.26, 0.24, 0.33, 0.59, 0.24, 0.55, 0.53, 0.53
50.00	0.35, 0.61, 0.31, 0.56, 0.55, 0.25, 0.67, 0.29, 0.28, 0.57, 0.29, 0.63, 0.36, 0.60, 0.18, 0.59, 0.75, 0.58, 0.38, 0.57, 0.19, 0.56, 0.35, 0.44, 0.52, 0.42, 0.34, 0.71, 0.63, 0.30, 0.46, 0.64, 0.65, 0.61, 0.67, 0.26, 0.64, 0.35, 0.46, 0.35
51.00	0.37, 0.37, 0.67, 0.61, 0.45, 0.61, 0.34, 0.26, 0.61, 0.33, 0.67, 0.31, 0.53, 0.28, 0.53, 0.42, 0.52, 0.35, 0.63, 0.35, 0.53, 0.35, 0.34, 0.47, 0.32, 0.36, 0.59, 0.48, 0.18, 0.33, 0.50, 0.45, 0.47, 0.49, 0.52, 0.25, 0.50, 0.12, 0.38, 0.15
52.00	0.70, 0.34, 0.52, 0.58, 0.62, 0.66, 0.67, 0.25, 0.63, 0.34, 0.64, 0.61, 0.62, 0.52, 0.66, 0.62, 0.65, 0.64, 0.67, 0.50, 0.65, 0.68, 0.57, 0.71, 0.65, 0.49, 0.56, 0.72, 0.75, 0.57, 0.55, 0.56, 0.56, 0.61, 0.64, 0.57, 0.70, 0.70, 0.70
53.00	0.59, 0.48, 0.40, 0.52, 0.42, 0.24, 0.63, 0.29, 0.69, 0.35, 0.57, 0.38, 0.51, 0.38, 0.55, 0.37, 0.32, 0.32, 0.53, 0.30, 0.38, 0.57, 0.34, 0.43, 0.64, 0.54, 0.35, 0.40, 0.55, 0.54, 0.57, 0.55, 0.59, 0.37, 0.56, 0.31, 0.41, 0.32
54.00	0.49, 0.63, 0.58, 0.64, 0.60, 0.26, 0.65, 0.34, 0.69, 0.61, 0.62, 0.53, 0.59, 0.50, 0.59, 0.62, 0.59, 0.50, 0.65, 0.73, 0.57, 0.69, 0.67, 0.64, 0.69, 0.72, 0.70, 0.60, 0.57, 0.59, 0.59, 0.63, 0.64, 0.60, 0.69, 0.70, 0.70
55.00	0.51, 0.34, 0.59, 0.51, 0.40, 0.52, 0.47, 0.54, 0.25, 0.56, 0.17, 0.59, 0.24, 0.58, 0.26, 0.54, 0.15, 0.46, 0.48, 0.16, 0.51, 0.69, 0.44, 0.19, 0.44, 0.54, 0.18, 0.22, 0.23, 0.20, 0.28, 0.56, 0.18, 0.55, 0.53, 0.54
56.00	0.61, 0.32, 0.43, 0.61, 0.39, 0.61, 0.38, 0.59, 0.28, 0.58, 0.31, 0.53, 0.33, 0.56, 0.34, 0.53, 0.38, 0.36, 0.60, 0.43, 0.32, 0.55, 0.59, 0.38, 0.47, 0.50, 0.58, 0.56, 0.63, 0.31, 0.60, 0.43, 0.45, 0.44
57.00	0.65, 0.55, 0.55, 0.58, 0.57, 0.60, 0.37, 0.60, 0.35, 0.65, 0.38, 0.68, 0.28, 0.63, 0.35, 0.57, 0.59, 0.35, 0.58, 0.63, 0.53, 0.37, 0.55, 0.61, 0.39, 0.33, 0.30, 0.28, 0.64, 0.38, 0.62, 0.50, 0.61
58.00	0.38, 0.62, 0.48, 0.64, 0.38, 0.57, 0.34, 0.52, 0.15, 0.50, 0.22, 0.61, 0.27, 0.59, 0.45, 0.48, 0.67, 0.48, 0.33, 0.65, 0.58, 0.33, 0.44, 0.59, 0.55, 0.69, 0.67, 0.63, 0.20, 0.59, 0.35, 0.42, 0.55
59.00	0.62, 0.30, 0.64, 0.26, 0.55, 0.27, 0.47, 0.39, 0.51, 0.39, 0.53, 0.30, 0.49, 0.30, 0.29, 0.53, 0.30, 0.35, 0.57, 0.53, 0.38, 0.32, 0.55, 0.40, 0.53, 0.53, 0.38, 0.33, 0.35, 0.27, 0.37, 0.27
60.00	0.68, 0.23, 0.61, 0.53, 0.60, 0.45, 0.61, 0.48, 0.62, 0.45, 0.60, 0.41, 0.56, 0.59, 0.50, 0.63, 0.58, 0.43, 0.52, 0.66, 0.66, 0.53, 0.51, 0.59, 0.52, 0.57, 0.64, 0.53, 0.53, 0.65, 0.73
61.00	0.64, 0.39, 0.56, 0.33, 0.55, 0.39, 0.49, 0.41, 0.52, 0.37, 0.49, 0.62, 0.55, 0.57, 0.34, 0.36, 0.58, 0.54, 0.28, 0.36, 0.66, 0.49, 0.52, 0.54, 0.57, 0.37, 0.56, 0.27, 0.56, 0.29
62.00	0.64, 0.52, 0.58, 0.42, 0.66, 0.50, 0.60, 0.52, 0.61, 0.46, 0.67, 0.65, 0.49, 0.66, 0.61, 0.45, 0.49, 0.65, 0.72, 0.51, 0.48, 0.50, 0.50, 0.56, 0.61, 0.51, 0.66, 0.66, 0.70
63.00	0.51, 0.31, 0.52, 0.39, 0.51, 0.41, 0.59, 0.37, 0.51, 0.30, 0.28, 0.50, 0.30, 0.29, 0.31, 0.31, 0.26, 0.26, 0.52, 0.47, 0.50, 0.50, 0.56, 0.32, 0.52, 0.25, 0.32, 0.25
64.00	0.59, 0.29, 0.60, 0.27, 0.61, 0.33, 0.62, 0.21, 0.54, 0.53, 0.20, 0.57, 0.57, 0.41, 0.22, 0.43, 0.53, 0.21, 0.59, 0.18, 0.18, 0.28, 0.57, 0.21, 0.51, 0.49, 0.51
65.00	0.56, 0.31, 0.56, 0.25, 0.59, 0.18, 0.54, 0.39, 0.39, 0.56, 0.37, 0.25, 0.66, 0.58, 0.39, 0.38, 0.59, 0.55, 0.58, 0.59, 0.69, 0.18, 0.59, 0.28, 0.41, 0.31

Appendix IV: Similarity matrix of combined morphological and molecular data

	1.00	2.00	3.00	4.00	5.00	6.00	7.00	8.00	9.00	10.00	11.00	12.00	13.00	14.00	15.00	16.00	17.00	18.00	19.00	20.00	21.00	22.00	23.00	24.00	25.00	26.00	27.00	28.00	29.00	30.00	31.00
2.00	0.25																														
3.00	0.59	0.55																													
4.00	0.36	0.30	0.59																												
5.00	0.58	0.35	0.21	0.56																											
6.00	0.63	0.48	0.25	0.56	0.28																										
7.00	0.18	0.38	0.50	0.52	0.56	0.59																									
8.00	0.53	0.48	0.50	0.53	0.44	0.39	0.52																								
9.00	0.46	0.39	0.58	0.33	0.52	0.49	0.43	0.49																							
10.00	0.32	0.27	0.57	0.35	0.57	0.52	0.35	0.55	0.47																						
11.00	0.57	0.35	0.21	0.53	0.19	0.18	0.56	0.43	0.47	0.60																					
12.00	0.36	0.31	0.56	0.36	0.57	0.53	0.31	0.55	0.49	0.27	0.59																				
13.00	0.45	0.40	0.40	0.44	0.39	0.37	0.46	0.35	0.43	0.47	0.37	0.41																			
14.00	0.56	0.55	0.34	0.61	0.31	0.25	0.59	0.48	0.56	0.57	0.35	0.52	0.37																		
15.00	0.40	0.39	0.62	0.38	0.64	0.53	0.40	0.43	0.46	0.43	0.62	0.40	0.47	0.59																	
16.00	0.33	0.35	0.57	0.31	0.66	0.64	0.41	0.63	0.41	0.39	0.52	0.57	0.38	0.58	0.28																
17.00	0.62	0.59	0.42	0.61	0.39	0.34	0.60	0.41	0.55	0.59	0.35	0.54	0.39	0.28	0.54	0.52															
18.00	0.58	0.59	0.37	0.59	0.40	0.37	0.62	0.45	0.60	0.61	0.38	0.56	0.40	0.37	0.54	0.55	0.37														
19.00	0.47	0.37	0.65	0.38	0.61	0.53	0.43	0.52	0.38	0.45	0.61	0.48	0.49	0.61	0.51	0.49	0.68	0.64													
20.00	0.60	0.57	0.41	0.63	0.38	0.37	0.60	0.45	0.61	0.55	0.44	0.50	0.43	0.27	0.51	0.61	0.33	0.34	0.69												
21.00	0.37	0.31	0.54	0.39	0.60	0.56	0.44	0.52	0.46	0.43	0.50	0.36	0.40	0.55	0.30	0.24	0.57	0.56	0.47	0.56											
22.00	0.61	0.60	0.23	0.55	0.28	0.35	0.62	0.54	0.53	0.64	0.15	0.62	0.37	0.38	0.64	0.56	0.44	0.45	0.66	0.47	0.60										
23.00	0.23	0.26	0.58	0.34	0.57	0.52	0.35	0.49	0.49	0.29	0.57	0.24	0.39	0.48	0.33	0.24	0.48	0.50	0.46	0.52	0.26	0.62									
24.00	0.55	0.54	0.29	0.56	0.33	0.36	0.57	0.46	0.57	0.58	0.33	0.52	0.35	0.25	0.58	0.50	0.32	0.37	0.65	0.33	0.48	0.34	0.49								
25.00	0.60	0.55	0.36	0.60	0.36	0.32	0.55	0.46	0.59	0.59	0.36	0.49	0.38	0.28	0.55	0.54	0.34	0.39	0.59	0.29	0.50	0.40	0.51	0.25							
26.00	0.62	0.61	0.38	0.60	0.44	0.44	0.61	0.43	0.61	0.67	0.40	0.61	0.42	0.35	0.50	0.57	0.36	0.32	0.69	0.30	0.52	0.39	0.54	0.37	0.35						
27.00	0.59	0.60	0.29	0.61	0.38	0.45	0.61	0.55	0.59	0.63	0.38	0.56	0.40	0.27	0.63	0.57	0.38	0.35	0.72	0.35	0.56	0.33	0.53	0.24	0.33	0.30					
28.00	0.52	0.49	0.27	0.57	0.32	0.28	0.57	0.48	0.65	0.57	0.34	0.53	0.33	0.30	0.58	0.54	0.36	0.31	0.62	0.35	0.51	0.39	0.46	0.30	0.33	0.33	0.30				
29.00	0.60	0.59	0.24	0.62	0.29	0.34	0.62	0.49	0.57	0.64	0.31	0.58	0.33	0.25	0.62	0.58	0.34	0.33	0.65	0.35	0.46	0.32	0.54	0.25	0.29	0.30	0.27	0.25			
30.00	0.28	0.18	0.58	0.36	0.57	0.50	0.38	0.52	0.47	0.22	0.57	0.24	0.37	0.50	0.34	0.27	0.53	0.53	0.46	0.48	0.24	0.62	0.14	0.47	0.51	0.57	0.53	0.46	0.54		
31.00	0.55	0.51	0.32	0.60	0.31	0.28	0.66	0.44	0.57	0.55	0.35	0.53	0.33	0.21	0.58	0.57	0.30	0.40	0.59	0.31	0.51	0.39	0.48	0.25	0.24	0.33	0.32	0.22	0.23	0.46	
32.00	0.40	0.34	0.66	0.35	0.62	0.58	0.36	0.58	0.45	0.37	0.63	0.50	0.48	0.63	0.53	0.48	0.62	0.63	0.32	0.63	0.53	0.68	0.39	0.62	0.63	0.74	0.70	0.63	0.69	0.39	0.62
33.00	0.58	0.55	0.36	0.58	0.41	0.37	0.58	0.52	0.56	0.52	0.37	0.50	0.33	0.37	0.58	0.55	0.33	0.40	0.68	0.38	0.55	0.40	0.52	0.33	0.30	0.36	0.31	0.27	0.36	0.47	0.31
34.00	0.59	0.52	0.28	0.59	0.33	0.31	0.57	0.48	0.59	0.56	0.37	0.54	0.36	0.29	0.56	0.54	0.29	0.38	0.62	0.38	0.47	0.37	0.51	0.29	0.25	0.32	0.21	0.22	0.26	0.48	0.18
35.00	0.48	0.41	0.69	0.44	0.56	0.58	0.44	0.53	0.44	0.47	0.63	0.46	0.53	0.61	0.54	0.65	0.67	0.60	0.24	0.69	0.50	0.67	0.46	0.65	0.60	0.71	0.72	0.63	0.65	0.49	0.59
36.00	0.58	0.56	0.31	0.58	0.37	0.37	0.61	0.49	0.51	0.60	0.37	0.53	0.33	0.33	0.58	0.53	0.35	0.30	0.70	0.34	0.51	0.37	0.60	0.29	0.31	0.32	0.29	0.24	0.26	0.62	0.33
37.00	0.38	0.32	0.58	0.38	0.62	0.58	0.39	0.52	0.52	0.39	0.61	0.34	0.42	0.56	0.40	0.36	0.55	0.63	0.52	0.60	0.37	0.61	0.32	0.55	0.54	0.60	0.54	0.58	0.58	0.52	0.55
38.00	0.38	0.32	0.50	0.38	0.58	0.57	0.38	0.55	0.47	0.37	0.57	0.50	0.38	0.54	0.33	0.19	0.57	0.54	0.51	0.55	0.16	0.58	0.25	0.47	0.48	0.55	0.52	0.50	0.48	0.25	0.53
39.00	0.55	0.51	0.26	0.57	0.23	0.22	0.56	0.41	0.49	0.60	0.20	0.57	0.32	0.32	0.62	0.58	0.39	0.39	0.53	0.43	0.53	0.21	0.58	0.36	0.34	0.36	0.40	0.29	0.23	0.56	0.24
40.00	0.59	0.57	0.44	0.65	0.43	0.42	0.62	0.41	0.65	0.61	0.46	0.59	0.44	0.31	0.56	0.60	0.35	0.34	0.66	0.27	0.37	0.47	0.49	0.35	0.38	0.28	0.40	0.35	0.37	0.51	0.30
41.00	0.51	0.44	0.54	0.43	0.58	0.58	0.43	0.59	0.46	0.49	0.67	0.40	0.51	0.59	0.56	0.56	0.68	0.69	0.22	0.65	0.52	0.61	0.50	0.67	0.58	0.70	0.69	0.62	0.64	0.50	0.59
42.00	0.40	0.40	0.57	0.47	0.60	0.59	0.45	0.51	0.57	0.42	0.61	0.37	0.44	0.55	0.33	0.36	0.53	0.53	0.53	0.49	0.30	0.62	0.30	0.52	0.54	0.51	0.59	0.51	0.55	0.32	0.51
43.00	0.55	0.55	0.30	0.60	0.34	0.56	0.54	0.49	0.56	0.59	0.32	0.51	0.34	0.29	0.61	0.52	0.32	0.35	0.62	0.29	0.52	0.38	0.49	0.23	0.19	0.33	0.25	0.21	0.23	0.49	0.22
44.00	0.55	0.48	0.66	0.42	0.65	0.65	0.45	0.64	0.45	0.52	0.65	0.49	0.52	0.66	0.57	0.51	0.74	0.71	0.33	0.69	0.40	0.67	0.50	0.67	0.67	0.70	0.73	0.68	0.67	0.48	0.68
45.00	0.63	0.60	0.25	0.62	0.32	0.36	0.66	0.52	0.60	0.65	0.25	0.60	0.37	0.26	0.56	0.61	0.33	0.38	0.71	0.38	0.53	0.22	0.57	0.26	0.29	0.27	0.19	0.29	0.23	0.57	0.31
46.00	0.53	0.48	0.37	0.60	0.38	0.33	0.58	0.43	0.60	0.58	0.44	0.51	0.34	0.24	0.53	0.55	0.39	0.36	0.63	0.26	0.48	0.43	0.45	0.24	0.24	0.33	0.28	0.24	0.26	0.44	0.22
47.00	0.30	0.21	0.63	0.35	0.60	0.57	0.37	0.49	0.47	0.21	0.60	0.36	0.41	0.53	0.34	0.29	0.53	0.55	0.46	0.53	0.38	0.65	0.17	0.51	0.51	0.57	0.56	0.45	0.57	0.17	0.46
48.00	0.57	0.50	0.26	0.57	0.23	0.21	0.58	0.44	0.47	0.58	0.22	0.53	0.33	0.36	0.61	0.60	0.40	0.40	0.51	0.42	0.58	0.27	0.58	0.38	0.33	0.43	0.43	0.33	0.30	0.54	0.28
49.00	0.32	0.29	0.57	0.41	0.58	0.54	0.41	0.54	0.47	0.28	0.56	0.31	0.35	0.54	0.36	0.24	0.53	0.56	0.49	0.55	0.21	0.63	0.19	0.49	0.48	0.55	0.55	0.48	0.53	0.19	0.50
50.00	0.49	0.45	0.27	0.55	0.22	0.17	0.55	0.43	0.53	0.51	0.24	0.52	0.35	0.31	0.60	0.59	0.37	0.40	0.55	0.38	0.59	0.25	0.50	0.34	0.35	0.42	0.35	0.33	0.34	0.50	0.27
51.00	0.62	0.59	0.29	0.58	0.38	0.40	0.64	0.54	0.57	0.66	0.31	0.60	0.36	0.32	0.59	0.49	0.38	0.39	0.68	0.41	0.53	0.25	0.27	0.24	0.30	0.30	0.21	0.33	0.24	0.55	0.32
52.00	0.51	0.44	0.68	0.43	0.62	0.60	0.46	0.60	0.35	0.53	0.63	0.54	0.51	0.65	0.56	0.52	0.68	0.66	0.29	0.71	0.52	0.67	0.52	0.69	0.66	0.73	0.74	0.61	0.65	0.52	0.65
53.00	0.57	0.55	0.36	0.61	0.37	0.32	0.61	0.46	0.51	0.60	0.41	0.52	0.37	0.28	0.56	0.56	0.36	0.40	0.63	0.37	0.51	0.40	0.48	0.29	0.28	0.33	0.34	0.28	0.29	0.50	0.17
54.00	0.51	0.42	0.66	0.40	0.62	0.56	0.46	0.61	0.46	0.45	0.63	0.48	0.55	0.53	0.56	0.56	0.68	0.71	0.28	0.65	0.54	0.65	0.50	0.62	0.61	0.71	0.63	0.64	0.67	0.46	0.66
55.00	0.36	0.29	0.54	0.37	0.59	0.53	0.46	0.52	0.46	0.36	0.59	0.31	0.39	0.53	0.36	0.26	0.59	0.59	0.42	0.57	0.15	0.60	0.26	0.46	0.46	0.57	0.60	0.52	0.53	0.23	0.49
56.00	0.51	0.46	0.39	0.54	0.30	0.31	0.55	0.41	0.53	0.53	0.32	0.56	0.33	0.41	0.57	0.59	0.43	0.45	0.54	0.47	0.54	0.34	0.51	0.41	0.39	0.46	0.46	0.37	0.32	0.51	0.33
57.00	0.39	0.41	0.61	0.40	0.61	0.60	0.44	0.48	0.56	0.44	0.64	0.36	0.50	0.57	0.29	0.38	0.57	0.55	0.58	0.51	0.36	0.65	0.12	0.54	0.56	0.51	0.61	0.58	0.60	0.34	0.55
58.00	0.61	0.58	0.32	0.56	0.34	0.35	0.60	0.52	0.54	0.64	0.26	0.60	0.35	0.42	0.61	0.59	0.48	0.46	0.62	0.48	0.70	0.19	0.51	0.41	0.39	0.45	0.38	0.41	0.37	0.51	0.41
59.00	0.55	0.52	0.25	0.60	0.33	0.29	0.57	0.45	0.59	0.57	0.32	0.51	0.36	0.24	0.58	0.54	0.33	0.30	0.63	0.34	0.49	0.33	0.49	0.22	0.27	0.34	0.26	0.22	0.22	0.48	0.20
60.00	0.50	0.38	0.64	0.37	0.60	0.58	0.42	0.52	0.44	0.47	0.61	0.46	0.47	0.59	0.46	0.50	0.65	0.66	0.24	0.66	0.45	0.65	0.46	0.62	0.60	0.70	0.72	0.64	0.64	0.46	0.62
61.00	0.50	0.48	0.35	0.56	0.35	0.33	0.60	0.45	0.59	0.49	0.33	0.49	0.31	0.22	0.59	0.55	0.21	0.36	0.62	0.34	0.52	0.40	0.44	0.29	0.33	0.36	0.31	0.25	0.28	0.42	0.23
62.00	0.54	0.44	0.68	0.35	0.64	0.61	0.48	0.61	0.46	0.46	0.63	0.48	0.61	0.65	0.51	0.49	0.71	0.66	0.31	0.66	0.47	0.63	0.52	0.67	0.66	0.70	0.72	0.66	0.65	0.46	0.66
63.00	0.57	0.55	0.24	0.55	0.35	0.36	0.59	0.45	0.53	0.59	0.22	0.54	0.20	0.32	0.56	0.46	0.25	0.31	0.65	0.37	0.52	0.39	0.52	0.26	0.30	0.33	0.27	0.26	0.22	0.51	0.33
64.00	0.41	0.36	0.55	0.37	0.60	0.60	0.42	0.50	0.51	0.41	0.58	0.33	0.44	0.52	0.32	0.24	0.59	0.60	0.55	0.60	0.23	0.69	0.31	0.63	0.61	0.67	0.59	0.58	0.58	0.31	0.58
65.00	0.55	0.54	0.21	0.53	0.32	0.28	0.56	0.48	0.53	0.56	0.21	0.56	0.30	0.34	0.61	0.58	0.45	0.36	0.61	0.41	0.56	0.12	0.56	0.33	0.38	0.37	0.34	0.33	0.25	0.56	0.35
66.00	0.37	0.31	0.54	0.37	0.60	0.56	0.40	0.56	0.46	0.47	0.58	0.53	0.40	0.53	0.34	0.27	0.59	0.57	0.49	0.54	0.16	0.60	0.29	0.48	0.52	0.55	0.56	0.48	0.52	0.24	0.47
67.00	0.56	0.57	0.34	0.63	0.29	0.32	0.61	0.52	0.55	0.50	0.33	0.57	0.44	0.34	0.62	0.67	0.40	0.49	0.64	0.39	0.62	0.33	0.57	0.38	0.36	0.43	0.34	0.38	0.38	0.57	0.29
68.00	0.26	0.25	0.55	0.36	0.56	0.55	0.38	0.55	0.47	0.32	0.54	0.30	0.40	0.52	0.30	0.25	0.56	0.55	0.46	0.55	0.28	0.61	0.23	0.49	0.51	0.59	0.53	0.48	0.54	0.20	0.53
69.00	0.62	0.57	0.31	0.59	0.33	0.31	0.64	0.52	0.53	0.62	0.21	0.63	0.41	0.39	0.59	0.61	0.47	0.49	0.63	0.49	0.60	0.20	0.61	0.43	0.41	0.44	0.42	0.42	0.41	0.60	0.36
70.00	0.36	0.38	0.58	0.41	0.60	0.55	0.47	0.49	0.40	0.35	0.62	0.35	0.43	0.55	0.22	0.34	0.56	0.53	0.50	0.50	0.29	0.64	0.29	0.56	0.54	0.52	0.61	0.54	0.59	0.29	0.54
71.00	0.66	0.51	0.22	0.59	0.17	0.21	0.58	0.46	0.49	0.56	0.25	0.56	0.34	0.32	0.63	0.60	0.41	0.41	0.57	0.43	0.53	0.20	0.56	0.36	0.36	0.40	0.37	0.28	0.28	0.56	0.29
72.00	0.31	0.24	0.63	0.35	0.57	0.63	0.42	0.54	0.47	0.36	0.58	0.32	0.36	0.52	0.30	0.35	0.56	0.55	0.45	0.50	0.19	0.59	0.22	0.49	0.46	0.56	0.53	0.49	0.51	0.16	0.49
73.00	0.59	0.52	0.32	0.60	0.34	0.30	0.63	0.46	0.61	0.60	0.34	0.64	0.38	0.38	0.66	0.64	0.40	0.37	0.63	0.36	0.52	0.43	0.52	0.36	0.34	0.43	0.39	0.31	0.33	0.51	0.31
74.00	0.60	0.56	0.30	0.60	0.34	0.36	0.62	0.46	0.60	0.63	0.38	0.56	0.38	0.38	0.59	0.56	0.41	0.35	0.64	0.39	0.51	0.45	0.55	0.30	0.26	0.40	0.36	0.28	0.28	0.53	0.30
75.00	0.40	0.37	0.53	0.38	0.60	0.60	0.43	0.60	0.45	0.41	0.54	0.34	0.43	0.58	0.30	0.13	0.59	0.60	0.54	0.61	0.23	0.54	0.31	0.48	0.52	0.57	0.55	0.56	0.54	0.29	0.58
76.00	0.60	0.55	0.30	0.62	0.34	0.38	0.60	0.50	0.63	0.61	0.34	0.56	0.41	0.38	0.60	0.55	0.41	0.35	0.67	0.42	0.53	0.41	0.55	0.24	0.31	0.42	0.39	0.37	0.28	0.53	0.34
77.00	0.58	0.58	0.28	0.54	0.28	0.33	0.59	0.47	0.55	0.61	0.19	0.59	0.41	0.41	0.61	0.55	0.41	0.40	0.63	0.46	0.55	0.23	0.59	0.31	0.38	0.42	0.40	0.42	0.29	0.59	0.39
78.00	0.57	0.51	0.59	0.50	0.64	0.64	0.58	0.62	0.56	0.59	0.59	0.50	0.53	0.62	0.49	0.39	0.60	0.59	0.49	0.54	0.39	0.69	0.46	0.55	0.56	0.51	0.61	0.61	0.54	0.46	0.60
79.00	0.42	0.39	0.53	0.35	0.62	0.61	0.44	0.57	0.48	0.44	0.55	0.34	0.43	0.57	0.32	0.17	0.56	0.56	0.53	0.59	0.22	0.56	0.31	0.49	0.49	0.54	0.55	0.57	0.51	0.31	0.56
80.00	0.54	0.53	0.27	0.54	0.38	0.39	0.59	0.53	0.57	0.58	0.29	0.52	0.34	0.35	0.53	0.44	0.37	0.40	0.69	0.42	0.43	0.24	0.49	0.25	0.31	0.33	0.27	0.32	0.25	0.48	0.35
81.00	0.64	0.66	0.29	0.64	0.38	0.39	0.66	0.52	0.64	0.66	0.29	0.62	0.43	0.39	0.59	0.52	0.37	0.37	0.74	0.44	0.51	0.38	0.56	0.29	0.34	0.38	0.37	0.30	0.26	0.57	0.38
82.00	0.61	0.40	0.56	0.37	0.62	0.62	0.45	0.60	0.48	0.43	0.67	0.36	0.66	0.59	0.31	0.15	0.58	0.60	0.55	0.61	0.12	0.57	0.30	0.51	0.51	0.56	0.57	0.60	0.53	0.30	0.58
83.00	0.37	0.39	0.49	0.34	0.56	0.57	0.43	0.57	0.48	0.41	0.50	0.39	0.40	0.54	0.34	0.20	0.55	0.57	0.52	0.60	0.33	0.52	0.30	0.47	0.51	0.53	0.53	0.52	0.47	0.28	0.53
84.00	0.36	0.35	0.56	0.39	0.61	0.62	0.45	0.58	0.47	0.42	0.56	0.37	0.43	0.57	0.33	0.15	0.56	0.58	0.54	0.60	0.15	0.56	0.27	0.50	0.52	0.54	0.53	0.54	0.51	0.27	0.58
85.00	0.38	0.37	0.59	0.39	0.60	0.60	0.45	0.59	0.47	0.40	0.54	0.35	0.43	0.57	0.30	0.15	0.56	0.59	0.54	0.60	0.15	0.54	0.27	0.50	0.52	0.54	0.55	0.56	0.51	0.27	0.56
86.00	0.43	0.43	0.59	0.42	0.65	0.65	0.49	0.52	0.52	0.47	0.60	0.43	0.49	0.62	0.22	0.24	0.58	0.57	0.59	0.55	0.24	0.61	0.36	0.54	0.58	0.49	0.59	0.60	0.57	0.36	0.61
87.00	0.59	0.59	0.23	0.54	0.25	0.32	0.59	0.51	0.51	0.62	0.16	0.59	0.37	0.36	0.62	0.55	0.41	0.44	0.64	0.44	0.55	0.06	0.56	0.34	0.37	0.38	0.32	0.39	0.20	0.59	0.39
88.00	0.41	0.42	0.55	0.37	0.62	0.62	0.45	0.59	0.48	0.43	0.57	0.35	0.45	0.59	0.31	0.15	0.58	0.60	0.55	0.61	0.13	0.57	0.30	0.51	0.51	0.56	0.57	0.56	0.59	0.30	0.59
89.00	0.61	0.60	0.28	0.60	0.36	0.39	0.64	0.55	0.57	0.63	0.30	0.59	0.35	0.31	0.60	0.53	0.37	0.40	0.68	0.41	0.51	0.23	0.57	0.28	0.31	0.29	0.22	0.31	0.20	0.66	0.31
90.00	0.59	0.58	0.33	0.53	0.46	0.49	0.64	0.40	0.61	0.64	0.41	0.59	0.36	0.41	0.47	0.49	0.37	0.37	0.68	0.35	0.45	0.38	0.53	0.34	0.39	0.24	0.35	0.40	0.32	0.54	0.40
91.00	0.60	0.58	0.23	0.61	0.30	0.37	0.65	0.51	0.57	0.66	0.39	0.60	0.38	0.28	0.59	0.50	0.32	0.36	0.70	0.34	0.43	0.26	0.54	0.25	0.28	0.28	0.19	0.28	0.17	0.55	0.30



Appendix V: Detailed Results of Allelic frequency analysis of *Pona* Complex in Ghana using SSR allelic data

Marker	Allele	Count	Het.ero zygotes	Homo zygotes	Freq	HObs	HExp	PIC	H-W[a]	Chi-Square[b]	Df[c]	P value	Na[d]
YM13						0.971	0.628	0.545	***	18.1147	1		-0.2494
	175	2	2	0	0.0294								
	212	31	29	1	0.4559								
	220	30	30	0	0.4412								
	227	3	3	0	0.0441								
	250	2	2	0	0.0294								
YM26						1.0000	0.5693	0.469	***	46.8500	1	<0.0000001	-0.2944
	102	4	4	0	0.03252								
	107	1	1	0	0.00813								
	127	1	1	0	0.00813								
	133	56	56	0	0.45528								
	135	2	2	0	0.01626								
	141	54	54	0	0.43902								
	162	1	1	0	0.00813								
	174	5	5	0	0.04065								
Dpr3D06						1.0000	0.8560	0.8309	***	18.1112	1	0.0000002	-0.881
	125	18	18	0	0.1217								
	127	1	1	0	0.0068								
	131	5	5	0	0.0338								
	133	6	6	0	0.0406								
	137	7	7	0	0.0473								
	143	4	4	0	0.0270								
	145	2	2	0	0.0135								
	148	13	10	0	0.0879								
	150	11	11	0	0.0744								
	160	3	3	0	0.0203								

Appendix V: continued

	166	4	4	0	0.0270						
	170	36	36	0	0.2434						
	175	35	35	0	0.2366						
	179	1	1	0	0.0068						
	197	2	2	0	0.0135						
YM15						1.0000	0.8859	0.8454	***	20.221	1	<0.0000001	-0.0799
	117	2	2	0	0.0357						
	179	2	2	0	0.0357						
	186	4	4	0	0.0714						
	197	6	6	0	0.1071						
	211	10	10	0	0.1786						
	223	4	4	0	0.0714						
	228	10	10	0	0.1786						
	230	4	4	0	0.0714						
	240	2	2	0	0.0357						
	293	12	12	0	0.2143						
Dab2E09						1.0000	0.5953	0.5018	***	29.1068	1	<0.0000001	-0.2873
	117	38	38	0	0.5000						
	120	38	38	0	0.5000						
Da1A01						0	0	0	ND	-	-	-	ND
	212	38	38	0	0.5						
	214	19	19	0	0.2						
	225	19	19	0	0.2						

Appendix V: continued

Da1D08					1.0000	0.7167	0.6116	**	10.5612	1	<0.0000001	-0.122
	223	6	6	0	0.1250							
	229	3	3	0	0.0625							
	250	1	1	0	0.0208							
	260	1	1	0	0.0208							
	300	12	12	0	0.2500							
	304	7	7	0	0.1458							
	308	2	2	0	0.0417							
	337	16	14	1	0.3333							
Dpr3F04					1.0000	1.0000	0.375	ND	-	-	-	ND
	81	2	2	0	0.0417							
	86	2	2	0	0.0417							
	88	2	2	0	0.0417							
	95	2	2	0	0.0417							
	97	4	4	0	0.0833							
	99	2	2	0	0.0417							
	103	2	2	0	0.0417							
	118	2	2	0	0.0417							
	121	8	8	0	0.1667							
	124	6	6	0	0.1250							
	127	8	8	0	0.1667							
	129	2	2	0	0.0417							
	131	6	6	0	0.1250							
Dab2D06					1	0.5354	0.4088	***	17.1685	1	0.0000346	-0.3201
	165	2	2	0	0.0345							
	171	2	2	0	0.0345							
	176	9	9	0	0.1552							

Appendix V: continued

	178	15	15	0	0.2586							
	186	9	9	0	0.1552							
	192	20	14	3	0.3448							
	200	1	1	0	0.0172							
Da1C12						1	1	0.375	ND	-	-	-	ND
	127	2	2	0	0.0333							
	140	2	2	0	0.0333							
	158	28	28	0	0.4665							
	160	28	28	0	0.4665							
Dpr3F10						1	0.8174	0.7804	ND	-	-	-	ND
	107	1	1	0	0.0064							
	111	4	4	0	0.0255							
	122	2	2	0	0.0127							
	127	25	23	1	0.1592							
	129	2	2	0	0.0127							
	133	16	16	0	0.1019							
	136	4	4	0	0.0255							
	142	5	5	0	0.0318							
	149	1	1	0	0.0064							
	153	2	2	0	0.0127							
	155	37	37	0	0.2357							
	163	1	1		0.0064							
	168	31	31	0	0.1975							
	170	1	1	0	0.0064							
	173	22	22	0	0.1401							
	192	1	1		0.0064							
	198	2	2	0	0.0127							

Appendix V: continued

Dab2C05					0.83333	0.5619	0.4482	ND	-	-	-	-0.2127
	178	15	15	0	0.4167							
	193	19	13	3	0.5586							
	200	1	1	0	0.0278							
Da1F08					1.0000	0.5709	0.4670	***	26.3008	1	0.0000003	-0.2938
	166	3	3	0	0.0405							
	170	35	35	0	0.473							
	172	1	1	0	0.0135							
	175	34	34	0	0.4595							
	179	1	1	0	0.0135							

APPENDIX VI: PRINCIPAL COORDINATES (METRIC –DIMENSIONAL SCALING) SCORES

	Principal Coordinates Scores		
Sample	**1**	**2**	**3**
1	0.17354	-0.1537	-0.0489
2	0.22102	-0.141	0.01514
3	0.08705	-0.2006	-0.0251
4	-0.0249	-0.0289	0.11045
5	-0.1585	0.20379	-0.1474
6	0.15963	-0.084	-0.0782
7	-0.0795	0.10069	-0.127
8	0.19089	-0.0222	0.10151
9	-0.0586	-0.1912	-0.1274
10	0.13649	-0.0385	-0.0374
11	-0.4096	-0.2351	-0.331
12	0.11868	0.00876	-0.0381
13	0.17311	-0.0707	0.06346
14	0.24697	-0.1254	0.09478
15	-0.0426	0.33125	0.16457
16	-0.4235	-0.2994	-0.3122
17	0.13484	-0.1136	0.02896
18	0.00432	-0.2778	-0.0825
19	0.12711	0.07398	0.01759
20	0.17908	-0.0832	0.07605
21	-0.2306	0.28777	-0.0829
22	-0.5973	-0.1753	0.5209
23	0.18676	-0.1077	0.07263
24	0.01402	-0.24	-0.0908
25	0.0154	0.35727	0.08448
26	-0.2143	0.13846	0.11554
27	0.08851	-0.1925	0.12656
28	-0.1191	0.06226	-0.2534
29	0.04324	-0.252	0.00508
30	0.18096	-0.1331	0.00097
31	-0.09228	0.15014	-0.0583
32	0.01256	0.09614	-0.1551
33	-0.2307	0.06067	-0.1257
34	0.1259	0.06432	0.09177
35	0.07841	0.05264	-0.1209
36	0.19363	-0.0938	0.04309
37	0.06921	0.26197	-0.0285
38	-0.2342	0.27875	-0.0941
39	0.0517	-0.1877	0.01555
40	0.14085	-0.139	-0.0525

APPENDIX VI: CONTINUED

41	0.17097	0.06569	0.01714
42	-0.0497	0.00378	-0.2343
43	0.03519	-0.1597	-0.1367
44	-0.2837	0.20508	-0.0666
45	-0.5973	-0.1753	0.5209
46	0.2297	-0.0386	0.14364
47	0.17741	-0.0732	-0.0167
48	0.00372	0.23589	-0.02
49	-0.0126	0.08427	-0.1985
50	0.06368	0.11043	-0.1311
51	-0.4865	-0.2212	-0.1847
52	0.01833	-0.2142	-0.1314
53	0.21849	-0.0427	0.15248
54	0.16279	-0.0623	0.04461
55	0.09708	0.15879	0.01285
56	0.17885	-0.0654	0.01807
57	0.13583	0.20247	0.13465
58	0.1774	-0.0436	0.16447
59	-0.0029	-0.0132	-0.1232
60	0.16098	0.03676	0.06377
61	0.15551	-0.1222	0.00221
62	-0.189	0.30778	-0.0189
63	-0.4269	0.0055	-0.2271
64	0.12018	-0.0308	0.21085
65	-0.273	0.17222	-0.0322
66	-0.0642	0.10507	-0.1031
67	0.21761	-0.0376	0.18897
68	0.15591	-0.0558	-0.0025
69	-0.0147	0.36758	0.18627
70	0.06669	0.3019	0.08906
71	-0.0465	0.00327	-0.197
72	0.18842	0.01106	0.197
73	-0.0419	0.30829	-0.0201
74	0.07891	0.0529	-0.018
75	-0.5973	-0.1753	0.5209
76	-0.1502	-0.1501	-0.1379