Hierarchical Approaches to the Analysis of Genetic Diversity in Plants: A Systematic Overview

M E Osawaru
Department of Biological Sciences
University of Abuja
Gwagwalada,
Nigeria
E-mail: edwinosawaru@yahoo.com

M C Ogwu*
Department of Plant Biology and Biotechnology
Faculty of Life Sciences
University of Benin
Nigeria
E-mail: matthew.ogwu@gmail.com

R O Aiwansoba
Department of Plant Biology and Biotechnology
Faculty of Life Sciences
University of Benin
Nigeria
E-mail: raymond_pedro@yahoo.com

Paper accepted on 14 April 2015

Abstract

Hierarchical analysis highlights the nature of relationship between and among type samples as outlined by standard descriptors. It produces an output called dendrogram, which depicts the hierarchical structure of genetic interaction in clusters/groups. Genetic diversity is the variation of heritable characteristics in a population. It results from one or more of the following; evolution, mutation, migration, domestication, plant breeding and selection. Knowledge about genetic diversity and relationships among plants may be an invaluable aid in plant breeding and classification. The analyses of genetic diversity rely on pedigree, morphological, biochemical and most recently molecular (DNA-based) data through PCR or non-PCR techniques. For accurate and unbiased assessment of genetic diversity, adequate attention has to be devoted to: sampling strategies, choice of genetic distance as well as judicious combination of analytical

techniques and statistical tools. A critical distinction in ways of analyzing genetic diversity in plant lies in the data used for analysis. Based on the data type, the hierarchy may be expressed with; phenogram, cladogram or phylogram, which are forms of dendrogram. The resulting dendrogram is often used to predict if a population is genetically diverse or homogenous.

Keywords: *Genetic diversity, Hierarchical approach, Plant, Clustering, Descriptive Characters, Markers*

**For correspondences and reprints*

1. INTRODUCTION

Hierarchical analysis highlights the nature of relationship between any types of samples as outlined by standard descriptors and it usually produce a graphical output known as dendrogram that shows this hierarchical clustering structure (Holland, 2006). Study of genetic diversity is the process by which variation among individuals or groups of individuals or populations is analyzed by a specific method or a combination of methods (Mohammadi & Prasanna, 2003). These data often involve numerical measurement and in many cases, combinations of different types of variables. Diverse data sets have been used by researchers to analyze genetic diversity in crop plants, most important among such data sets are pedigree data (Bernardo, 1993; Messmer *et al*, 1993; Van Hintum and Haalman, 1994), passport data-morphological data (Smith and Smith, 1992; Bar-Hen *et al*., 1995), biochemical data obtained by analysis of isozymes (Hamrick & Godt, 1997) and storage proteins (Smith *et al*., 1987) and recently, DNA-based marker data that allow more reliable differentiation of genotypes. Since each of these data sets provide different types of information, the choice of analytical method(s) depends on the objective(s) of the experiment, the level of resolution required, the resources and technological infrastructure available, and the operational and time constraints, if any (Karp *et al*., 1997).

Genetic diversity is not randomly distributed over plant populations. As a result of processes in natural evolution, domestication, modern plant breeding, mutation, migration and selection, genetic diversity has a structure that can generally be summarized in a hierarchical model or a tree. Hierarchical methods are particularly useful in that they are not limited to a pre-determined number of clusters and can display similarity of samples across a wide range of scale (Holland, 2006). The selection of accessions for a core collection is based on the structure of the genetic diversity to be represented by that core collection. To describe the structure, a genetic diversity tree can be constructed by branching (based on a knowledge of natural evolution, domestication, distribution and utilization) or by clustering (based on a phonetic analysis of individual accessions). For clustering, data obtained as close to the DNA as possible should be used [that is, Restriction Fragment Length Polymorphism (RFLP)]. If a

diversity tree is available, a core collection can be selected by deciding on the number of accessions per sample and subsequently, choosing accessions with maximal diversity within each group (van Hintum, 1995).

The objective of this paper therefore is to outline approaches or models which may be employed to classify plants into hierarchies upon analysis of their genetic diversity. To improve on crop breeding programs and to facilitate reliable classification of accessions and identification of subsets of core accessions with possible utility for specific breeding purposes. To review the molecular techniques of genetic variability and their application to plant sciences.

2. GENETIC DIVERSITY IN PLANTS

Biological diversity thus refers to the range of differencies among set of entities; within the living world, and includes all living organisms both individuals and their relationship with one another (Emma-Okafor *et al*., 2009). This may be discussed from three levels: genetic diversity, species diversity and ecosystem diversity (Ramanatha & Hodgkin, 2002).

Genetic diversity is the variation of heritable characteristics present in a population of the same species (Swingland, 2001). The result of genetic differences may manifest as differences in DNA sequence, biochemical characteristics (e.g in protein structure or isoenzyme properties), physiological properties (e.g abiotic stress resistance or growth rate) and/or morphological characters such as flower colour or plant form and function. In line with Ramanatha & Hodgkin, (2002) and van Hintum, (1995) genetic diversity are usually highlighted through; the number of different forms (alleles), their distribution(genotype), the effect on their performance (phenotype) and the overall sum of genome.

Ecosystem diversity refers to the variety of habitats, living communities, and ecological processes in the living world (Swingland, 2001). Each ecosystem provides different kinds of habitats (niche), both living and non-living

environments interact constantly and in complex ways that may change over time. No two ecosystems are the same in composition and diversity. Its diversity is determined by the types of plants, animals and microorganisms, as well as by the physical characteristics (e.g. substrates, light, nutrients, e.t.c.) and the interactions (e.g. predators-prey relationships).

Species diversity is a measure of the diversity within population that incorporates both species richness and their evenness. Species diversity is influenced by species richness (Swingland, 2001). Communities with more species are considered to be more diverse. Evenness measure the variation in the abundance of individual per species within a community. Communities with less variation in the relative abundance of species are considered to be more even than a community with more variation in relative abundance. These components of species diversity respond differently to various environmental conditions. The species level is generally regarded as the most natural one at which to consider whole-organism diversity (Swingland, 2001). The ecological importance of species can have a direct effect on community structure, and thus on over-all biological diversity.

Even though genetic diversity is at the lowest hierarchy, without genetic diversity, a population cannot evolve and adapt to environmental changes (Paulauskas *et. al*., 2013). The genetic diversity has an impact on the higher levels of biodiversity (Templeton 1991, 1993). Genetic biodiversity finds its natural resources in wild species for which it is important to find out the amount of genetic variability by the way of morphological, biochemical and molecular markers, besides some interesting physiological turns. Genetic diversity is the basis for survival and adaption and makes it possible to continue and advance the adaptation processes on which evolutionary success and to some extent human survival depends.It is generally accepted that the genetic variation in plant populations is structured in space and time (Loveless and Hamrick, 1984).

Plate 1: Genetic Diversity as exhibited by seed colours in African Yam Bean (S*phenostylis stenocarpa*) held in IITA seed bank, Ibadan Nigeria. Photo credit: Aiwansoba, R. O

SOME CAUSES OF GENETIC DIVERSITY

Genetic diversity is affected by several ongoing processes both natural and artificial. These processes include one or more of the following;

Evolution: Evolution is a population phenomenon. It is concerned with the effect of changes in the frequency of alleles within a gene pool of a population, such changes leading to changes in genetic diversity and ability of the population to undergo evolutionary divergence (Acquaah, 2007). Evolutionary divergence resulted in association of characters that makes it possible to distinguish well-defined groups. Often, it will be possible to divide the genetic diversity in such a group into new or smaller groups on the same basis (van Hintum, 1995). This principle of subsequent divisions implies that species diversity has a basically hierarchical structure. Taxonomic nomenclature is based on this hierarchy. There

are numerous natural factors, such as hybridization and clinical variation, which can complicate the structure. In general, however, the hierarchical model will suffice.

Domestication: Primitively, man selected the plant types that were better suited for his needs in an unsystematic way with domestication as the first outcome of this selection (Rauf *et al*., 2010). McCouch (2004) indicated that during the process of domestication, selection of unusual phenotypes, such as large fruit or seed size, intense colour, sweet flavor or pleasant aroma was carried out. Domestication can be defined as the process by which genetic changes (or shifts) in wild plants are brought about through a selection process imposed by humans. It is an evolutionary process in which selection (both natural and artificial) operates to change plants genetically, morphologically and physiologically (Acquaah, 2007).

The second level of genetic differentiation results from infraspecific divergence as a result of domestication. The processes causing genetic changes under domestication are basically the same as those occurring under evolution in nature, except that human selection, partly conscious, partly unconscious, is added. The major difference, apart from being centered on human activity, is that the rate of change under domestication is much higher than the relative slower process under natural evolution (Pickersgill, 1984).

Domestication usually causes a reduction of genetic diversity compared with the diversity of the species in the wild. Several reports are available regarding the reduction of genetic diversity during domestication. Reif *et al*. (2005b) showed that loss of genetic diversity occurred during wheat domestication. The loss of genetic diversity occurred from wild species *Triticum tauschii* to the land races and from land races to elite cultivars. Similarly, Haudry *et al*. (2007) also noted that wheat genetic diversity was reduced in the cultivated form by 69% in bread wheat and 8% in durum wheat during the process of domestication. Sonante *et al*. (1994) also noted losses in genetic diversity during the process of domestication in common beans (*Phaesolus vulgaris).* Loss of genetic diversity during domestication has also been reported in other species (Hollingsworth *et al.*,

2005). The extent to which the hierarchical structure applies to the genetic diversity within a domesticated crop will generally be less n at the species level. The association between characters within a crop species will generally be much smaller than between species since the reproductive isolation at the species level has occurred over a much longer period of time, resulting in a cleaner delimitation of variation. However the association that exist in landraces, can like species diversity,

Plant Breeding: Plant breeding is a technology that deals with the evolution of crop varieties using the principles of various sciences and skills of the plant breeder gained over the years. The pre green revolution era is marked by nobalization of sugar cane, utilization of commercial Heterosis and development of plant breeding and biometric techniques (Troyer, 1996). This resulted in the evolution of crop varieties more uniform in yield and growth. A major leap in yield was achieved on the onset of the green revolution era, specifically in the Indian sub-continent. After the green revolution, the per capita availability of cereals increased from 275 to 370kg (Ortiz, 1998). The green revolution was achieved when dwarfing genes were exploited in plant species. The genes characteristically reduced induced plant heights, induced early maturity, increased harvest index (HI), stomatal conductance and defective plant growth regulators (PGRs) such as auxins and gibberellins (Reynolds *et al*., 1999; Salamni, 2003; Routray *et al*., 2007; Rauf and Sadaqat, 2008a). The induction of dwarfing genes also helps to maintain high density populations and a double cropping system (Neelu and Rajbir, 2009).

Some plant breeding techniques increase crop diversity (for example induced mutations, hybridization between previously incompatible populations and introgression from previously isolated populations) (van Hintum, 1995). Others cause a reduction in diversity (Rauf *et al*., 2010). In many cases, intra specific hybridization was found to lower the genetic diversity when it was attempted to explore the transgressive segregation (Van Esbroeck & Bowman, 1998; Jordan *et al*., 1998; Cui *et al*., 2000). Van Esbroech and Bowman (1998) concluded that plant breeders have lowered the genetic diversity due to preference for the few cultivars to be used as parents in the establishment of transgressive generation.

Cui *et al*. (2000) showed that genetic diversity of Chinese soya bean (*Glycine max*) was high due to careful selection of unrelated parents while constituting transgressive segregation.

Mutation: Mutation is the origin of all new genetic diversity, occurring when there are occasional errors in the replication of DNA or other elements of the production and packaging of genetic information within the cells. Although it implies something negative, mutations can have positive, neutral, or deleterious impacts. Mutations occur rather slowly but continuously. Mutations at one level, for example, in the nucleotides that are the basis of DNA, may not all be expressed at other levels such as protein differences or observable changes in the appearance of a plant. The rate of mutation is useful in determining evolutionary relationships.

Migration: Migration is the movement of genetic diversity, usually within a species. In plants, this occurs through pollen dispersal, seed dispersal, and movement of vegetative propagules, such as suckers or rhizomes, in species that can reproduce asexually. Migration, also called gene flow, occurs both with the advancing front of a population when it is colonizing new areas, and when genes of two or more populations mix through pollen and seed dispersal. The rate of migration is obviously related to the frequency of reproduction and the distances over which pollen and seeds typically disperse.

Selection: Selection is perhaps the best known of the processes affecting genetic diversity and is the only process that directly results in populations becoming better adapted to their environment. For natural selection to occur, there must be differences in fitness and survival among individuals and a genetic basis for those differences. Over time (generations), those individuals that are better suited to the environment live, or live longer, and produce more offspring — those offspring having inherited the more adaptive traits (or rather, have a higher frequency of the alleles that confer better adaptation). The above processes also regulate human interactions with plants which might reduce or increase genetic diversity.

3. SAMPLING STRATEGIES AND GENETIC DISTANCE

SAMPLING STRATEGIES: Sampling is the selection of subset of individuals from within a population. Genetic diversity in plants may be analyzed at different levels, which range from one another due to nature of genetic materials viz; individual genotypes such as inbred lines or pure lines or clones, populations, germplasm accessions and species (David & George, 2005). According to David & George, (2005) the sampling process for analyzing genetic diversity includes;

1. Defining the population of interest
2. Specifying a sampling frame
3. Specifying a sampling method
4. Determine the sample size
5. Implementing the sampling plan
6. Sampling and data collecting

These strategies may either be non-probabilistic or probabilistic. Probability sampling is one in which every unit in the population has a chance of being selected in the sample while non-probability sampling is any sampling method where some elements of the population have no chance of selection. The tables below summarize various class of sampling strategies;

Table 1: Probability sampling approach

Method	Strategy for Selecting Sample
Simple random sampling	Every member of population being studied has an equal chance of being selected.
Interval sampling	This method is used when a stream of representative population are available.
Stratified sampling	In this method the population is divided into non-overlapping groups and samples are taken from within these groups.
Clustered sampling	This method is used when the population of interest is large and widely geographically dispersed. Clusters within the population are randomly selected.

(Source: Adapted from David and George, 2005)

Table 2: Nonprobability sampling approach

Method	Strategy for Selecting Sample
Convenience sampling	Participants will be those that the researcher has relatively easy access to.
Snowball sampling	Participants meeting the study requirements will recommend others with the same characteristics. The method is used when trying to access difficult to reach population.
Quota sampling	Participants are non-randomly selected according to pre-defined fixed quota.
Typical case sampling	Participants are selected to be typical, normal or average for a particular phenomenon
Theoretical sampling	Participants are selected on the basis of the results of the data collected to date. The goal is to develop a deeper understanding of the topic and to develop theory.

(Source: Adapted from David and George, 2005)

GENETIC DISTANCE

Genetic distance is a measure of the genetic divergence between species or between populations within a species. Populations with many similar genes have small genetic distances. This suggests that they are closely related and/or have a recent common ancestor. Genetic distance is useful for reconstructing the history of populations and is also useful in understanding the origin of biodiversity.
There are several different statistical measures that have been proposed. Most commonly used distance measures are as follows;

1. Nei's genetic distance (Nei, 1972)
2. Cavalli-Sforza chord measure (cavalli-Sforza and Edwards, 1967)
3. Reynolds genetic distance (Reynolds weir and Cockerhan's, 1983)

In all the formulae, X and Y represent two different populations for which L loci have been studied. Let X_u represent the u^{th} allele at the l^{th} locus.

Neis genetic distance: The distance is expressed by multiple loci or the use of arithmetic means across all loci and it can also be interpreted as a mean number of codon substitutions per locus

$$D = -\ln I$$

$$where\ I = \frac{\sum x_i y_i}{\sqrt{\sum x_i^2 \sum y_i^2}}$$

Where x_i and y_i are the frequencies of the ith allele in population x and y
Nei's assume that;

- All loci have same rate of neutral mutation
- Mutation-genetic drift equilibrium
- Stable effective population size

Cavalli-Sforza Chord Distance: Populations are conceptualised as existing as points in **m**-dimensional Euclidean space which are specified by **m** allele

frequencies (i.e. **m** equals the total number of alleles in both populations). The distance is the angle between these points:

$$D_{CH} = \frac{2}{\pi}\sqrt{2\left(1 - \sum_{l}\sum_{u}\sqrt{X_u Y_u}\right)}$$

- Assumes genetic drift only (no mutation)
- Geometric distance between points in multi-dimensional space

Reynold Distance: This measure assumes that genetic differentiation occurs only by genetic drift without mutations. It estimates the coancestry coefficient Θ, which provides a measure of the genetic divergence by

$$D = -\ln(1-\theta)$$

Other distance measures include Euclidean distance, Share allele distance, Roger's distance, Goldstein distance (for microsatallite).

4. METHODS OF ANALYZING GENETIC DIVERSITY

The assessment of genetic diversity in crop plants may be carried out in two major ways namely

I. Morphological characterization techniques
II. Molecular techniques

MORPHOLOGICAL CHARACTERIZATION: Characterization is the exercise involved in the description of plant germplasm. It determines the expression of highly heritable characters ranging from morphological or agronomical features. Morphological characterization of crop plants is essential to provide information on traits of accessions assuring the maximum utilization of germplasm collection to the final user (Engel & Visser, 2003). This method has been applied by Osawaru *et al.* (2013) for West African okra (*Abelmoschus cailei*). A minimum of three replicates and preferably four with data taken from

at least 10 plants across replications has been shown to be statistically acceptable for some crops. Care should be taken to select a site where the species is adapted and traits will be expressed. Morphological characterization may include one or combination of the following:

A. **Morphological Descriptors:** A set of morphological descriptors can be used to describe the phenotype. International plant genetic resource institute (IPGRI) descriptors are widely available in the internet and these have been applied by Osawaru *et al*., (2009).

B. **Herbarium Samples:** Herbarium samples are a good record of variation and with care can be kept for many years and used for rechecking trait after the material is removed from the field. Through visits and request of type samples the genetic diversity of a particular sample can be analyzed. This has been applied by Borokini *et al*, (2010).

C. **Pictures:** Sufficient details should be captured in images to taxonomically identify the plant and demonstrate the traits that show variation. The images can be stored in a database linked with the morphological data. Web portals are available (like IPNI [ipni.org], PNAS [pnas.org], PROTA [prota.org]) were digital images of plant can be pooled and used for characterization. This method has been employed by Osawaru *et al*., (2013b).

D. **Nutritional Traits:** Nutritional traits which include food or feed value are important for many vegetables, fruits and forages. Nutrition trait can be characterized, this method has be applied by Osawaru *et al*., (2012 & 2013).

Morphological characterization does not require expensive technology but large tracts of land are often required for these experiments, making it possibly more expensive than molecular assessment. These traits are often susceptible to phenotypic plasticity; conversely, this allows assessment of diversity in the presence of environmental variations (Pagnotta *et al*., 2009)

MOLECULAR TECHNIQUES: Analyses of genetic diversity are usually based on assessing the diversity of an individual using either allozymes (i.e.,

variant forms of an enzyme that are coded for by different alleles at the same locus) or molecular markers, which tend to be selectively neutral (Pagnotta *et al*., 2009). Genetic variability within a population can be assessed through: the number (and percentage) of polymorphic genes in the population, the number of alleles for each polymorphic gene and the proportion of heterozygous loci per individual. However, they have the disadvantage of being relatively expensive, time consuming and require high levels of expertise and materials in analysis. There are credited for higher precision when compared with morphological markers. To assess genetic variability different markers are used and this markers are categorized into; biochemical markers and molecular markers.

Biochemical Markers: The use of biochemical markers largely involves the analysis of seed storage proteins and isozymes. This technique utilizes enzymatic functions and is a comparatively inexpensive yet powerful method of measuring allele frequencies for specific genes. Allozymes, being allelic variants of enzymes, provide an estimate of gene and genotypic frequencies within and between populations (Pagnotta *et al*., 2009).

Major advantages of these types of markers consist in assessing co-dominance, absence of epistatic and pleiotrophic effects, ease of use, and low costs. Disadvantages of isozymes include: (i) there are only few isozyme systems per species (no more than 30) with correspondingly few markers; (ii) the number of polymorphic enzymatic systems available is limited and the enzymatic loci represent only a small and not random part of the genome (the expressed part) - therefore, the observed variability may be not representative of the entire genome; (iii) although these markers allow large numbers of samples to be analyzed, comparisons of samples from different species, loci, and laboratories are problematic, since they are affected by extraction methodology, plant tissue, and plant stage.

Molecular Markers: Recently, a variety of different genetic markers have been proposed to assess genetic variability as a complementary strategy to more traditional approaches in genetic resources management. Different markers have different genetic qualities which may be dominant or co-dominant, can amplify

anonymous or characterized loci, can contain expressed or non-expressed sequences, e. t. c (Pagnotta *et al*., 2009).

A molecular markers can be defined as a genomic locus, detected through probe or specific starters (primers) which in virtue of its presence, distinguishes unequivocally the chromosomic trait which it represent as well as the flanky regions at the 3' and 5' extremity.

Molecular markers may or may not correlate with phenotypic expression of a genomic trait. They offer numerous advantages over conventional, phenotype-based alternatives as they are stable and detectable in all tissues regardless of growth, differentiation, development, or defense status of the cell. Additionally, they are not confounded by environmental, pleiotropic and epistatic effects.

An ideal molecular marker should possesses the following features: (1) be polymorphic and evenly distributed throughout the genome; (2) provide adequate resolution of genetic differences; (3) generate multiple, independent and reliable markers; (4) be simple, quick and inexpensive; (5) need small amounts of tissue and DNA samples; (6) link to distinct phenotypes; and, (7) require no prior information about the genome of an organism. Nevertheless, no molecular marker (Pagnotta *et al*., 2009).

Genetic or DNA based marker techniques such as Restriction Fragment Length Polymorphism (RFLP), Random Amplified Polymorphic DNA (RAPD), Simple Sequence Repeats (SSR) and Amplified Fragment Length Polymorphism (AFLP) are now in common use for ecological, evolutionary, taxonomical, phylogenic and genetic studies of plant sciences.

Molecular markers work by highlighting differences (polymorphisms) within a nucleic sequence between different individuals. These differences include insertions, deletions, translocations, duplications and point mutations. They do not, however, encompass the activity of specific genes.

In addition to being relatively impervious to environmental factor, molecular markers have the advantage of: (i) being applicable to any part of the genome (introns, exons and regulation regions); (ii) not possessing pleiotrophic or

epistatic effects; (iii) being able to distinguish polymorphisms which not produce phenotypic variation and finally, (iv) being some of them co-dominant.

The different techniques employed are based either on restriction-hybridization of nucleic acids or techniques based on Polymerase Chain Reaction (PCR), or both. In addition, the different techniques can assess either multi-locus or single-locus markers. Multi-locus markers allow simultaneous analyses of several genomic loci, which are based on the amplification of casual chromosomic traits through oligonucleic primers with arbitrary sequences. These types of markers are also defined as *dominant* since it is possible to observe the presence or the absence of a band for any locus, but it is not possible to distinguish between heterozygote (a/-) conditions and homozygote for the same allele (a/a). By contrast, single-locus markers employ probes or primers specific to genomic loci, and are able to hybridize or amplify chromosome traits with well-known sequences. They are defined as co-dominant since they allow discrimination between homozygote and heterozygote loci.

Basic marker techniques can be classified into two categories: (1) non-PCR-based techniques or hybridization based techniques; and (2) PCR-based techniques.

Restriction Fragment Length Polymorphism: All organisms have numerous differences in their genomic DNA sequence and therefore are genotypically distinct. This difference results in a restriction fragment length polymorphism. This can be detected by southern hybridization after running agarose gel electrophoresis. Hybridization analysis is carried out using probe that spans the region of interest. The probe hybridizes to the relevant region, 'lighting up' the appropriate restriction fragments on the resulting autoradiograph. If an RFLP is present then it will be clearly visible on the autoradiograph. Thus RFLP is used as a major tool to identify the genetic diversity within and between species (Somasundaram & Kalaiselvam 2011).

Random Amplified Polymorphic DNA (RAPD): The invention of PCR (polymerase chain reaction) is a milestone in the development of molecular techniques. PCR results in the selective amplification of a chosen region of a DNA molecule. Random amplification of DNA with short primer by PCR is a useful technique in phylogenetics. The important point is the banding pattern seen, when the products of PCR with random primers are 2electrophoresed in a reflection of the overall structure of the DNA molecule used as the template. If the starting material is total cell DNA then the banding pattern represents the organization of the cell's genome. Differences between the genomes of two organisms can be measured with RAPD. Two closely related organisms would be expected to yield more similar banding patterns than two organisms that are distant in evolutionary terms (Miesfeld, 1999).

Amplified Polymorphic Length Polymorphism (AFLP): AFLP analysis is able to detect high levels of polymorphism and has high repeatability and speed of analysis. These markers have a very high diversity index, resulting in a limited number of primer combinations required to screen a whole genome and has been applied to develop a system for the fingerprinting of an organism (Faccioli *et al*., 1999) and for map expansion (Castiglioni *et al*., 1998). Vos *et al*. (1995) described the AFLP technique as being based on the detection of restriction fragments by PCR amplification and argued that ′the reliability of the RFLP technique is combined with the power of the PCR technique′. AFLPs provide high levels of resolution to allow delineation of complex genetic structures. AFLPs are fragments of DNA that have been amplified using directed primers from restriction digested genomic DNA (Matthes *et al*., 1998; Karp *et al*., 1997).

The major advantage of the AFLP technique is the large number of polymorphisms that the method generates. Its ability to differentiate individuals in a population makes the technique useful for paternity analyses (Krauss, 1999), gene-flow experiments, and also for Plant variety registration (Law *et al*., 1998). Other advantageous features of the AFLP technique are: i) no sequence information is required; ii) the AFLP technique are: i) no sequence information is required; ii) the PCR technique is fast; and iii) a high multiplex ratio is possible (Rafalski *et al*., 1996).

Microsatellites: Microsatellites, alternatively known as simple sequence repeats (SSRs), short tandem repeats (STRs) or simple sequence length polymorphisms (SSLPs), are tandem repeats of sequence units generally less than 5 bp in length (Bruford and Wayne, 1993). These markers appear to be hypervariable, in addition to which their co-dominance and reproducibility make them ideal for genome mapping, as well as for population genetic studies (Dayanandan *et al.*, 1998). One common example of a microsatellite is a (CA)n repeat, where n is variable between alleles. These markers often present high levels of inter- and intra-specific polymorphism, particularly when tandem repeats number ten or greater. CA nucleotide repeats are very frequent in human and other genomes, and present every few thousand base pairs. Inter-SSRs are a variant of the RAPD technique, although the higher annealing temperatures probably mean that they are more rigorous than RAPDs. Chloroplast microsatellites (cpSSRs) are similar to nuclear microsatellites but the repeat is usually only 1 bp, i.e. (T) n (Provan *et al.*, 1999). Microsatellites owe their variability to an increased rate of mutation compared to other neutral regions of DNA.

Recently, a new class of advanced techniques has emerged, primarily derived from combination of earlier, more basic techniques. These advanced marker techniques combine advantageous aspects of several basic techniques. In particular, the newer methods incorporate modifications in the basic techniques, thereby increasing the sensitivity and resolution in detecting genetic discontinuity and distinctiveness. The advanced marker techniques also utilize newer classes of DNA elements such as retrotransposons, mitochondrial and chloroplast based microsatellites, allowing increased genome coverage. Techniques such as RAPD and AFLP are also being applied to cDNA- based templates (i.e., sequences of complementary DNA obtained by mRNA retrotranscription) to study patterns of gene expression and uncover the genetic basis of biological responses. The recent development of high-throughput sequencing technology provides the possibility of analysing high numbers of samples over smaller periods of time.

Table 3. Comparison of different characteristics of most frequently used molecular markers techniques.

MOLECULAR MARKERS	RFLP	RAPD	AFLP	SSR
Degree of polymorphism	M	M	M	M
Locus specificity	Y	N	N	N
Dominance (D)/Co-dom. (C)	C	D	D	C
Ease of Replication	H	L	H	M
Abundance	H	H	H	M
Sequence information required	Y	N	N	N
Quantity of DNA required	H	L	M	L
Automation	N	Y	Y	Y
Costs per assay	H	L	M	L/M
Technical requirement	H	L	M	L/M

Key: H = High; M= Medium; L = Low; Y = Yes; N= No

(Source: Adapted from Pagnotta *et al*., 2009)

5. DATA ANALYSIS

The genetic data generated using various analytical techniques can be analyzed using either specific statistical methods or combination of methods. The choice of statistical method to be used is dependent on the achievable objectives laid out in the studies. It is important to take the measurement levels of the variables into account for the analysis, as special statistical techniques are

available for each level. The data often involve numerical measurement and in many cases, combinations of different types of variables. Base on the number of variables used data analysis can be viewed from two methods namely;

- Univariate and
- Multivariate

Univariate Methods: Univariate analysis involves the examination across cases of one variable at a time. The analysis of this type of data can discussed from the distribution, the central tendency and dispersion. The distribution is a summary of frequency of individual values or range of values for a variable. One of the most common ways to describe a single variable is with a frequency distribution. Frequency distribution can be depicted in two ways, as a table or as a graph.

The central tendency of a distribution is estimated of the "center" of a distribution of values. There are three major types of estimates of central tendency namely; mean, median and mode. The mean is probably the most commonly used method of describing central tendency. To compute the mean, all you do is add up all the values and divide by the number of samples. The median is the score found at the exact middle of set of values. One way to compute the median is to list all scores in numerical order and then locate the score in the center of the sample. The mode is the most frequently occurring value in the set of scores.

Dispersion refers to the spread of the values around the central tendency. There are two common measures of dispersion, the range and the standard deviation. The range is simply the highest values minus the lowest values. The standard deviation is a more accurate and detailed estimate of dispersion because an outlier can greatly exaggerate the range. The standard deviation shows the relation that set of scores has to the mean of the sample. To compute the standard deviation we apply the formula below;

Mathematically:

$$\sqrt{\frac{\sum(x-\bar{x})^2}{(n-1)}}$$

Where: $\boldsymbol{x}$ = each score, $\bar{x}$ = the mean, n = the number of value, $\sum$ = sum across the values

Although it may be calculated manually, most statistical programs is capable of calculating them easily which include Ms-Excel, SPSS, PAST e.t.c.

Multivariate Methods: With increases in the sample size of germplasm accession used in crop improvement programs, classification and to create order in genetic variability assuming considerable significance. The use of established multivariate statistical algorithms is an important strategy for classifying germplasm, ordering variability for a large number of accessions, or analyzing genetic relationships among breeding materials (Mohammadi & Prasanna, 2003). Multivariate analytical techniques, which simultaneously analyze multiple measurements on each individual under investigation, are widely used in analysis of genetic diversity irrespective of the dataset (morphological, biochemical, or molecular marker data). Among these are algorithms, cluster analysis, principal component analysis (PCA), principal coordinate analysis (PCoA), and multidimensional scaling (MDS) are, at present, most commonly employed and appear particularly useful (Brown-Guedira *et al*., 2000).

Cluster Analysis: Cluster analysis refers to "a group of multivariate techniques whose primary purpose is to group individuals or objects based on the characteristics they possess, so that individuals with similar descriptions are mathematically gathered into the same cluster" (Hair *et al*., 1995). The resulting clusters of individuals should then exhibit high internal (within cluster) homogeneity and high external (between clusters) heterogeneity. Thus, if the classification is successful, individuals within a cluster shall be closer when plotted geometrically and different clusters shall be farther apart (Hair *et*

al., 1995). There are broadly two types of clustering methods:

(i) Distance-based methods, in which a pair-wise distance matrix is used as an input for analysis by a specific clustering algorithm (Johnson & Wichern, 1992), leading to a graphical representation (such a tree or dendrogram) in which clusters may be visually identified.

(ii) Model-based methods, in which observations from each cluster are assumed to be random draws from some parametric model, and inferences about parameters corresponding to each cluster and cluster membership of each individual are performed jointly using standard statistical methods such as maximum-likelihood or Bayesian methods.

Thornsberry *et. al*. (2001) provided the first empirical demonstration of the utility of the "structured association" method in plant genetics, identifying a gene associated with variation for flowering time in maize. At present, distance-based methods are most frequently applied.

Distance-based clustering methods can be categorized into two groups: hierarchical and nonhierarchical. Hierarchical clustering methods are more commonly employed in analysis of genetic diversity in crop species. These methods proceed either by a series of successive mergers or by a series of successive divisions of group of individuals. The former, known as "agglomerative hierarchical" methods, start with a single individual. Thus, there are initially as many clusters as individuals. The most similar individuals are first grouped and these initial groups are merged according to their similarities.

The nonhierarchical clustering procedures do not involve construction of dendrograms or trees. These procedures, also frequently referred to as "K-means clustering," are based on "sequential threshold," "parallel threshold," or "optimizing" approaches for assigning individuals to specific clusters, once the number of clusters to be formed is specified.

Dendrogram: The resulting cluster of individuals should then exhibit high internal (within cluster) homogeneity and high external (between cluster) homogeneity. Thus, if the classification is successful, individuals within a cluster shall be closer when plotted geometrically and different clusters shall be farther apart (Hair *et al.*, 1995).

The results of a cluster analysis are shown by a dendrogram which lists all the samples and indicates what level of similarity any two clusters are joined (Holland, 2006). A dendrogram could also be seen as a structure (a graph) able to be decomposed into subgroups (sub graphs) (Restrepo *et al.*, 2007).

A dendrogram is represented by its clusters. Perhaps the most important purpose of a cluster is its ability to show similarities between the elements of a particular plant species which is the main reason why cluster analysis is broadly used in drug discovery processes and molecular diversity studies (Restrepo *et al.*, 2007).
Based on the data type used, different types of dendrograms are produced. They include

- Phenograms
- Cladograms, and
- Phylograms

Phenogram: This is a diagram depicting taxonomic relationships among organisms based on overall similarity of many characteristics without regard to evolutionary history or assumed significance of specific character.

Cladogram: This is a branching tree-like diagram used to illustrate the evolutionary (phylogenetic) relationship among organisms. Each node or point of divergence has two branching lines of descendance, indicating evolutionary divergence from a common ancestor. The end point of the tree represents individual species, and any node, together with its descendant braches and sub-branches constitute a node.

Phylogram: A phylogenetic tree, or phylogram, sometimes called 'the tree of life' shows the evolutionary relationship among various organisms that are believed to have a common ancestor. Every node with descendants represent the most recent common ancestor of the descendant. Each node in a phylogenetic tree is called a taxonomic unit.

Although phylogenetic trees can provide evolutionary insight, they do have important limitations. Phylogenetic trees do not necessarily (and likely do not) represent actual evolutionary history. The data on which they are based are noisy and recombination and back mutations can all confound the analysis (Wright, 2011).

Principal components, canonical and multidimensional analyses are used to derive a 2-or 3- dimensional scatter plot of individuals such that the geometrical distances among individual genotypes reflect the genetic distances among them. Principal component as a reduced data form which clarify the relationship between breeding materials into interpretable fewer dimensions to form new variables. These new variables are visualized as different non correlating groups.

Principal components analysis first determines Eigen values which explain the amount of total variation displayed on the component axes. It is expected that the first 3 axes will explain a large sum of the variations captured by the genotypes. Cluster and principal component analysis can be jointly used to explain the variations in breeding materials in genetic diversity studies.

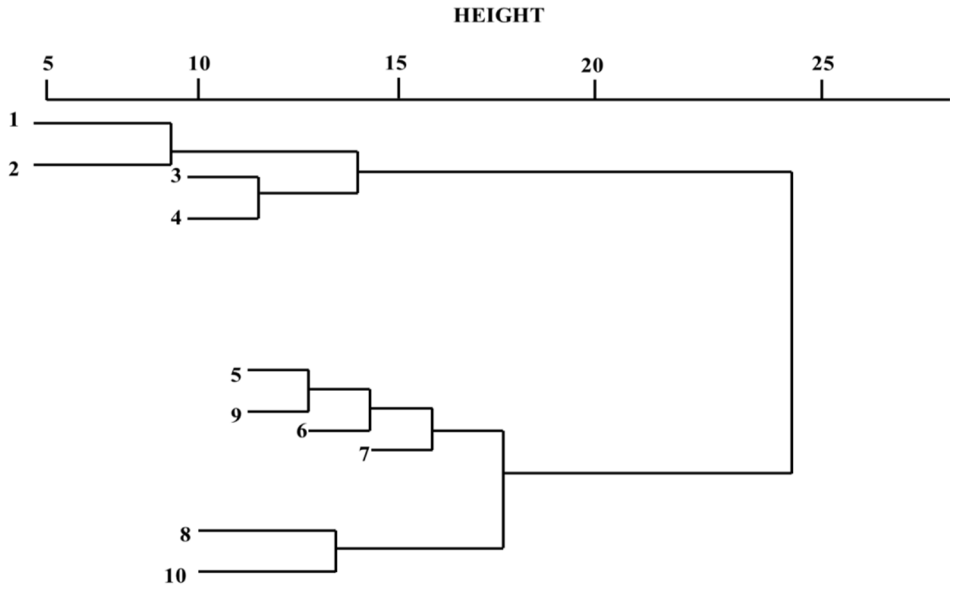

Figure 3: A Hypothetical Dendrogram.

(Source: Holland, 2006)

6. RECOMMENDATIONS AND CONCLUSION

The following are thereby recommended:

1. Support for policies and principles regarding plant genetic resources (PGR);
2. Encourage establishment of structure and international treaties to ensure sustainable management of PGR;
3. Sampling strategy to be adopted for hierarchical analysis of PGR should be designed to confirm with the needs of germplasm and location;
4. Support collection, characterization and evaluation of germplasm for the basis of conservation;
5. Causes of loss in genetic diversity should be studied to reduce their impacts;
6. Modern approaches to studying genetic diversity should be adopted by

scientist in their studies and

7. Clustering techniques for analysis plant genetic diversity should be adopted by gene bank curators to enable them check for redundancy in their collections.

Analysis of genetic diversity in plants may play roles in uncovering evolutionary history, biogeography and estimating their diversity, distinctiveness and population structure. Awareness of the level of genetic diversity and the proper management of genetic resources are important issues in modern era for which hierarchical analysis is relevant. Application of advanced techniques is a valuable tool to study genetic variability for the basis of conservation and classification. In the near future, the advent of more advanced techniques will be an impressive tool for evaluation of plant genetic resources. Many computer software packages are available for analyzing genetic diversity. There are two primary considerations when choosing software packages, which are: (i) statistical packages that offer analyses on the basis of relevant evolutionary models and, (ii) user-friendliness of the packages and output.

7. REFERENCES

ACQUAAH, G. (2007). Principles of Plant Genetics and Breeding. Wiley-Blackwell Publishers. 569pp.

BAR-HEN, A. A., CHARCOSSET, M. B. & CUIARD J. (1995). Relationships Between Genetic Markers and Morphological Traits in a Maize Inbred Lines Collection. *Euphytica,* 84:145–154.

BERNARDO, R. (1993). Estimation of Coefficient of Co ancestry Using Molecular Markers in Maize. *Theoretical Application Genetic* 85, 1055–1062

BOROKINI, T.I, OKERE, A.U, DARAMOLA. B.O. & ODOFIN, W.T. (2010). Biodiversity and Conservation of Plant Genetic Resources in the Field Genebank of the National Centre for Genetic Resources and Biotechnology, Ibadan, Nigeria. *International Journal of Biodiversity and Conservation* 2(3), 37-50.

BRUFORD, M .W. & WAYNE, R. K. (1993). Microsatellites and their Application to Population Genetic Studies. *Current Opinion in Genetics and Development* 3, 939-943.

CASTIGLIONI, P., POZZI, C., HEUN, M., TERZI, V., MULLER, K. J., ROHDE, W. & SALAMINI, F. (1998). An AFLP-Based Procedure for the Efficient Mapping of Mutations and DNA Probes in Barley. *Genetics* 149, 2039-2056.

CAVALLA-SFORZA, L. L. & EDWARDS, A. W. F. (1967). Phylogenetic Analysis-models and Estimation Procedures. *The American Journal of Human Genetics* 19, 1-3

CUI, Z., CANTER, T.E. & BURTON, J.W. (2000). Genetic Diversity Patterns in Chinese Soybean Cultivars based on Coefficient of Parentage. *Crop Science* 40, 1780-179.

DAVID, S. M. & GEORGE, P. M. (2005). Introduction to the Practice of Statistics 5th Edition. W. H. Freeman and Company, United Kingdom. 896p.

DAYANANDAN, S. RAJORA, O. P. & BAWA, K. S. (1998). Isolation and Characterisation of Microsatellites in Trembling Aspen (*Populus tremuloides*). *Theoretical and Applied Genetics* 96, 950 -956.

EMMA-OKAFOR, L. C., IBEAWUCHI, I. I. & OBIEFUNA, J. C. (2009). Biodiversity Conservation for Sustainable Agriculture in Tropical Rainforest of Nigeria. *New York Journal Science* 2(7), 81-88.

ENGELS, J.M.M. & VISSER, L. (2003). IPGRL Hand Book for Genebank No. 6. A guide to effective management of germplasm collections. IPGRI, Rome, Italy. 182p.

FACCIOLI P., PECCHIONI, N., STANCA, A.M. & TERZI, V. (1999). Amplified Fragment Length
Polymorphism (AFLP) Markers for Barley
Malt Fingerprinting. *Journal of Cereal Science* 29, 257-260.

HAIR, J.R., ANDERSON, R. E., TATHAM, R. L. & BLACK, W. C. (1995). Multivariate Data Analysis with Readings. 4th Edition, Prentice-Hall, Englewood Cliffs, New Jersey. Pp178-255.

HAMRICK, J. L. & GODT, M. J. W. (1997). Allozyme Diversity in Cultivated Crops. *Crop Science* 37, 26–30.

HAUDRY, A., CENCI, A., RAVEL, C., BATAILLON, T., BRUNEL, D., PONCECT, C., HONCHU, I., PRORIER, S., SANTONI, S., GLENN, S. & DAVID, J. (2007). Grinding up Wheat. A Massive Loss of nucleotide Diversity since Domestication. *Molecular Biology and Evolution* 24(7), 1506-1517

HOLLAND, S.M. (2006). Cluster analysis. Department of Geology, University of Georgia, Athens, GA 30602-2501. 45p

HOLLINGSWORTH, P.M., DAWSON, I.K., GOODALL-COPESTAKE, W.P., RICHARDSON, J.E., WEBER, J.C., STOOL MONTES, C & PENNINGTON, R.T. (2005). Do farmers reduce genetic diversity when they domesticate tropical trees? A case study from Amazonia. *Molecular Ecology* 14(2), 497-501.

JOHNSON, A. R. & WICHERN, D. W. (1992). Applied Multivariate Statistical analysis. 3rd Edition, Prentice-Hall, Englewood Cliffs, New Jersey. 642p

JORDAN, D.R., TAO, Y.Z., GODWIN, I.D., HENZELL, R.G., COOPER, M. & MCLINTYRE, C.L. (1998). Loss of genetic diversity associated with selection for resistance to sorghum midge in Australian sorghum. *Euphytica* 102 (1), 1-7.

KARP, A., KRESOVICH, S., BHAT, K. V., AYAND, W. G. & HODGKIN, T. (1997). Molecular Tools in Plant Genetic Resources Conservation: A Guide to the Technologies. IPGRI Technical Bulletin No. 2, International Plant Genetic Resources Institute, Rome, Italy. [Available at http:// 198.93.227.125/publicat/ techbull/TB2.pdf]. [Accessed 18/08/2014].

KRAUSS, S. L. (1999) Complete Exclusion of Nonsires in an Analysis of Paternity in a Natural Plant Population Using Amplified Fragment Length Polymorphism (AFLP). *Molecular Ecology* 8, 217-226.

LAW, J.R., DONINI, P., KOEBNER, R.M.D., JONES, C.R. & COOK E, R.J. (1998). DNA profiling and plant variety registration III:

The statistical assessment of distinctness in wheat using ampli fied fragment length polymorphisms. *Euphytica*, 102, 335 -342.

LOVELESS, M. D. & HAMRICK, J. L. (1984). Ecological Determinants of Genetic Structure in Plant Population. *Annual Reviews of Ecology and Systematics* 15, 65-95.

MATTHES, M. C., DALY, A. & EDWARDS, K.J. (1998). Amplified Fragment Length Polymorphism (AFLP). *In*: Karp, A. Isaac, P. G., and Ingram, D. S. (Eds.). Molecular Tools for Screening Biodiversity. Chapman and Hall, London. pp. 183-190.

MCCOUCH, S. (2004). Diversifying Selection in Plant Breeding. *PloS Biology* 2(10): 347.

MESSMER, M. M., MELCHINGER A. E., HERRMANN, R. G., & BOPPEN-MAIER, J. (1993). Relationships Among Early European Maize Inbreds: II. Comparison of Pedigree and RFLP Data. *Crop Science* 33, 944–950.

MIESFELD, R. L. (1999). Rapid Amplification of DNA. In: Applied Molecular Genetics. A John Wiley and Sons, UK. 312p.

MOHAMMADI, S. A. & PRASANNA, B. M. (2003). Analysis of Genetic Diversity in Crop Plants-Salient Statistical Tools and Considerations. *Crop Science* 43, 1235–1248

NEELU, J. & RAJBIR, Y. (2009). Genetic Diversity Analysis among Pre-Green Revolution, Post-Green Revolution era Cultivars and Wheat Landraces as Revealed by Microsatellite Markers. *Crop and Pasture Science* 60(4), 373-381.

NEI, M. (1972). Genetic Distance Between Population. *American Nature* 106, 283-292.

ORITZ, R (1998). Critical Role of Plant Biotechnology for the Genetic Improvement of Food Crop Preservatives for the next Millennium. *Electronic Journal of Biotechnology,* 1(3), 67-74

OSAWARU, M. E., OGWU, M. C. & DAINA-OGBE F.M. (2013a). Morphological assessment of genetic variability among 53 accessions of West African Okro [*Abelmoschus caillei* (A. Chev) Stevels.] from South Western Nigeria. *Nigeria Journal of Basic and Applied Sciences* 21(3), 227-258.

OSAWARU, M. E., OGWU, M. C. & IMARHIAGBE, O. (2013b). Biochemical characterization of some Nigerian *Corchorus* L. *Bayero Journal of Pure and Applied Sciences* 6, 69-75.

OSAWARU, M. E., OGWU, M. C., CHIME, A. O. & AMORIGOYE, A. R. (2012). Morphological evaluation and protein profiling of three accessions of Nigeria *Corchorus* L. species. *Bayero Journal of Pure and Applied Sciences* 5(1), 26-32.

PAGNOTTA, M. A., MONDINI, L. & NOORANI, A. (2009). Assessing Plant Genetic Diversity by Molecular Tools. *Diversity* 1, 19-35.

PAULAUSKAS, A., JODINSKIENE, M., GRICIUVIENE L., ZUKAUSKIENE J., PETRAITIENE E. & BRAZAUSKIENE I. (2013). Morphological Traits and Genetic Diversity of Differently Overwintered Oil Seed Rape (*Brassica napus* L.) Cultivars. *Agrivulture* 100(4), 409-411.

PICKERSGILL, B. (1984). Evolution of hierarchical variation patterns under domestication and their taxonomic treatment. In Styles, BT (ed.) Infraspecific classification of wild and cultivated plants. Oxford, UK : Clarenton. Pp 191-209.

PROVAN, J., RUSSELL, J. R., BOOTH, A. & POWELL, W. (1999). Polymorphic Chloroplast Simple Sequence Repeat Primers for Systematic and Population Studies in the Genus *Hordeum*. *Molecular Ecology* 8, 505 -511.

RAFALSKI, J. A., VOGEL, J. M., MORGANTE, M., POWELL, W., ANDRE, C. & TINGEY, S. V. (1996). Generating and Using DNA Markers in Plants. In: Birren, B, Lai E. (Eds.), Non-Mammalian Genomic Analysis: A practical guide. Academic Press, London, Pp. 75-134.

RAMANATHA, V. R. & HODGKIN, T. (2002). Genetic Diversity and Conservation and Utilization of Plant Genetic Resources. *Plant Cell, Tissue and Organ Culture* 68, 1-19.

RAUF, S. and SADAQAT, H.A. (2008). Identification of physiological traits and genotype combined to high achene yield in sunflower (*Helianthus annus* L.) under contrasting water regions. *Australian Journal of Crop Science* 1(1), 23-30.

RAUF, S., TEIXEIRA DA SILVA, J.A., KHAN, A.A. & NAVEED, A. (2010). Consequences of plant breeding on genetic diversity. *International Journal of Plant Breeding*, 4(1), 1-21.

REIF, J.P., ZHANG, P., DREISGACKER, S., WAIBURTON, M.L., GUKEL MVAN, HOISINGTON, D., MOHN, .M & MELCHYER, A.E. (2005). Wheat Genetic Diversity Traits during Domestication and Breeding. *Theoretical and Applied Genetics* 110(5), 859-864.

RESTREPO, G., MESA, H & LLANOS, E.J. (2007). Three Dissimilarity Features to Contrast Dendrograms. *J. Chem. Info Model* 47, 765-770.

REYNOLDS, J. B. S. & COCKERHAN, W. C. C. (1983). Estimation of the Coancestry Coefficient: Basis for Short-term Genetic Distance. *Genetic* 105, 767-779.

REYNOLDS, M.P., RAJARAM, S. & SAYRE, K.D. (1999). Physiological and Genetic Changes of Irrigated Wheat in the Post Green Revoltion Perod and Approaches to Meeting Projected Global Demand. *Crop Science* 3, 1611-1621.

ROUTRAY, P., BASHA O., GANG, M., SINGH, N.K. & DHAILIWAL, H.S. (2007). Genetic Diversity of Landraces of Wheat (*Triticum aestivium*. L) from hilly areas of Uttaranchal India. *Genetic Resources and Crop Evolution* 54(6), 1315-1326.

SALAMNI, F. (2003). Enhanced Hormone and the Green Revolution. *Science* 302 (5642), 71-72.

SMITH, J. S. C., PASZKIEWICS S., SMITH, O.S. & SCHAEFFER, J. (1987). Electrophoretic, Chromatograhic and Genetic Techniques for Identifying Associations and Measuring Genetic Diversity Among Corn Hybrids. Proceedings, 42nd Annual. Corn Sorghum Research Conference, Chicago, Seed Trade Association, Washington, DC. Pp. 187–203.

SMITH, O. S. and SMITH, J. S. C. (1992). Measurement of Genetic Diversity Among Maize Hybrids; A Comparison of Isozymic, RFLP, Pedigree, and Heterosis Data. *Maydica* 37, 53–60.

SOMASUNDARAM, S. T. & KALAISELVAM, M. (2011). Molecular Tools for Assessing Genetic Diversity. International Taining Course on Mangroves and Biodiversity, Annamalai University, Centre of Advanced Study in Marine Biology, India. Pp 82-91.

SONANTE, G., STOCKTON, T., NODARI, R.O., BECERRA, V.L., VELASQUEZ & GEPTS, P. (1994). Evaluation of Genetic Diversity during the Domesticaion of Common bean (*Phaseolus vulgaris*. L.).

Theoretical and Applied Genetics 89 (5), 629-635.

SWINGLAND, I. R. (2001). Biodiversity, Definition of. *Encyclopedia of Biodiversity* 1, 377-390.

TEMPLETON, A. R. (1991). Genetics and Conservation Biology. *In*: Seitz, A. and Loeschche V. (Ed.), Species Conservation a Population Biological Approach. Basel, Birkhauser Verlag., Pp15-29.

TEMPLETON, A. R. (1993). Translocation as Conservation Tool. *In*: Szaro, R. (Ed.), Biodiversity in Mangrove Landscapes, Theory and Practice. Oxford University Press. 42p

THORNSBERRY, J. M., GOODMAN, M. M., DOEBLEY, J., KRESOVICH, S., NIELSEN, D. & BUCKLER. E. S. (2001). Dwarf Polymorphisms Associate with Variation in Flowering Time. *Nature Genetic* 28, 286–289.

TROYER, A.F. (1996). Breeding Widely Adapted Popular Maize Hybrid. *Euphytica* 92, 163-174.

VAN ESBROECK, C.A. & BOWMAN, D.T. (1998).Changes in the Genetic Diversity of Cotton in the USA from 1970 to 1995. *Crop Science* 38, 33-37.

VAN HINTUM, T. J. L., & HAALMAN, D. (1994). Pedigree Analysis for Composing a Core Collection of Modern Cultivars, with Examples from Barley (*Hordeum vulgare*). *Theoretical Application Genetic* 88, 70–74

VAN HINTUM, T.J.L (1995). Hierarchical approaches to the analysis of genetic diversity in crop plants. *In*: Hodgkin, T., Brown, A.H.D., van Hintum, T.J.L. and Morales, E.A.V. (Eds.). Core Collection of Plant Genetic Resources. A Wiley Sayce Publication, pp23-34.

VOS. P, HOGERS. R, BLEEKER, M., REIJANS, M., VANDE-LEE, T., HORNES, M., FRIJTERS, A., POT, J., POLEMAN, J., KUIPER, M. & ZABEAU, M. (1995). AFLP: A New Technique for DNA Fingerprinting. *Nucleic Acids Research* 23, 4407-4414.

WRIGHT. D. (2011). The Irish type III website. [Online] [Available at http://irishtype3dna.org/phylogram/php] [Accessed: 02/08/2014].