CHAPTER I

GENERAL INTRODUCTION

Oil palm is a tropical crop grown mainly for its industrial production of vegetable oil. According to the World Bank report (2010), oil palm provides one of the leading vegetable oils produced globally and it accounts for one quarter of global consumption and 60 % of international trade in oils and fats. The oil extracted from the oil palm is of two types, namely the palm oil and palm kernel oil, 90 % of which is used mainly for food purposes and the other 10 % for non-food purposes (Jalani, 2012). Malaysia and Indonesia presently account for about 85 % of the world's production of palm oil with their main markets in China, European Union, Indian and the United States (Oil World, 2012).

The modern oil palm industry of the world was sown from only four oil palm trees that were introduced from Amsterdam Botanical Garden and Bogor. These four palms were planted 1848 in Java as ornamental plants (Jalani, 2012). This shows that the oil palm planting materials have a very narrow genetic base and a lot of concern was generated amongst researchers in Malaysia and thus, provided the initial impulsion for the collection of oil palm genetic materials from its area of endemism (centre of origin such as the *Elaeis guineensis* from West Africa and *E. oleifera* from South America), as genetic diversity is a crucial factor of any agricultural production (Rajanaidu, 1994; Rajanaidu & Jalani, 1994).

The first oil palm collection expedition was however made first in 1973 to Nigeria and 180,000 oil palm seeds were brought to Malaysia after a strict quarantine routine. The collection expedition progressed to other countries in Africa (Cameroon, Zaire, Tanzania, Madagascar, Angola, Senegal, etc.) and South America to collect extensive germplasm samples of *E. guineensis*, *E. oleifera* and other economic palms (Marhalil & Rajanaidu, 2011). All the prospected oil palm germplasm were field planted at the Malaysian Palm Oil Board (MPOB) research station, thus making it the world's largest ex-situ oil palm germplasm conservation programme (Kushairi, 2009).

Collection, conservation and evaluation of germplasm materials provide opportunity for tapping advantages of natural diversity for improvement (Sarkar et al., 2012) and these have been accelerated in the past decades to prevent the extinction of landraces and wild relatives (Sapey et al., 2012). The concept of germplasm conservation therefore, requires that collection methods initially seize maximum variation and to this essence, plant genetic resources conservation entails collecting, conservation and management, identification of potentially valuable materials by characterization and evaluation for subsequent utilization (Rao, 2004).

Germplasm collecting involves the assembling of samples of a species from populations in the field or natural habitats for conservation and subsequent utilization (Maji & Shaibu, 2012; Rao, 2004). Conservation of germplasm materials on the other hand ensures the sustenance of plant genetic resources and can be done in the natural habitat they occur (in-situ) or outside the native environment (ex-situ) and this constitutes the most important means of conserving plant genetic resources (Gonzalez-Benito et al., 2004; Rao, 2004).

Characterization and accurate assessment of plant genetic resources make up the first measure towards efficient usage of plants germplasm, because the power to key out genetic variation is essential for effective management and the use of genetic resources (Tahir, 2013; Rao, 2004). Adequate characterization for agronomic and morphological traits is necessary to ease utilization of germplasm materials by breeders (Upadhayaya, 2008), and the utilization of grounded multivariate statistical algorithms have become significant in classifying germplasm, ordering variability for magnanimous number of accessions and studying genetic relationships in any breeding materials (Mohammadi & Prasanna, 2003).

Multivariate analysis consists of a collection of methods than can be utilized when several measurements are made on each individual or objects in one or more samples (Rencher, 2002). Out of all multivariate techniques, principal components analysis and clustering have been extensively used to classify germplasm materials. PCA generates the general relationship between variables while clustering is an exploratory data analysis tool for grouping accessions (Richard & Wichern, 2007). Principal component analysis (PCA) and Cluster analysis (CA) have been successfully used in germplasm evaluation for years as they reveal the relationship and correlation among variables studied (Zafar et al., 2008; Maji & Shaibu, 2012; Odewale et al., 2012).

This research thereby aims to use multivariate statistical analysis (PCA and clustering) to unveil novel traits in the oil palm germplasm that will account for high variation among the accessions and also suggest individuals that show high level of diversity in order to maximize heterosis for oil palm improvement and germplasm enrichment.

1.1 Problem Statement

The oil palm stems from a very narrow genetic base and the effective use of crop genetic resources stored in gene banks by breeding programmes is limited. The number of accessions deposited in gene banks, however, is continuously growing.

Slow characterization of germplasm has been pointed out as a major cause of this and the use of principal component analysis and clustering in characterizing oil palm germplasm is limited.

1.2 Objectives

The intention of the research is to find answers to the following questions:

a. What is the extent of variation in the Nigerian-MPOB oil palm germplasm?
b. What are the main characters contributing to the variation in the oil palm germplasm collections?
c. How similar are the oil palm accessions?

Therefore, the objectives of the study were to;

i. Explore the extent of variability in Nigerian-MPOB based oil palm germplasm collections using univariate methods.
ii. Explore the pattern of variability in the oil palm materials and identify the main characters contributing to the overall variability using principal component analysis.
iii. Classify the variation pattern within the germplasm accessions based on their overall similarity using clustering analysis.

1.3 Scope of Research

Analyzing multivariate data set of MPOB-Nigerian oil palm germplasm; using PCA to find pattern of variation and cluster analysis to reveal the relative positions of the germplasm materials. However, PCA is poorly understood as some of its output cannot be easily interpreted. The problem also encountered with clustering analysis is determining the amount of clusters to generate.

1.4 Organization of Thesis

The thesis is structured into six chapters, with introduction as chapter I. Chapter II highlights the background to the study with a review of current literature into the question asked. Chapter III reviews the extent of variation in the Nigerian-MPOB based oil palm germpalsm. Chapter IV highlights the pattern of variation in the oil palm materials. Chapter V reveals the clustering of the oil palm materials into groups while chapter VI highlights the general conclusions of the study as well as recommendation.