KWAME NKRUMAH UNIVERSITY OF SCIENCE AND TECHNOLOGY

SCHOOL OF GRADUATE STUDIES

DEPARTMENT OF CROP AND SOIL SCIENCES

DIVERSITY STUDIES IN SOYBEAN (*Glycine max* (L.) Merrill) AND VALIDATION OF SHATTERING RESISTANT MARKERS FOR MARKER ASSISTED SELECTION

DAVID APPIAH-KUBI (BSc. AGRICULTURE)

AUGUST, 2012

DIVERSITY STUDIES IN SOYBEAN (*Glycine max* (L.)Merill) AND VALIDATION OF SHATTERING RESISTANT MARKERS FOR MARKER ASSISTED SELECTION

BY

DAVID APPIAH-KUBI BSc. AGRIC (HONS.)

A THESIS SUBMITTED TO THE DEPARTMENT OF CROP AND SOIL SCIENCES, FACULTY OF AGRICULTURE AND NATURAL RESOURCES, KWAME NKRUMAH UNIVERSITY OF SCIENCE AND TECHNOLOGY, KUMASI, IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE AWARD OF A MSc. DEGREE IN AGRONOMY (PLANT BREEDING OPTION)

AUGUST, 2012

DECLARATION

I, David Appiah-Kubi, hereby declare that this thesis is my own, produced from research work during my MSc programme. To the best of my knowledge, it contains no material previously published by any other person or material which has been accepted for the award of any other degree of the University. References to other published works have been duly acknowledged.

DAVID APPIAH –KUBI ....

PG3989409 Signature Date

Student

Certified by:

PROF. RICHARD AKROMAH ....

(Main Supervisor) Signature Date

DR. JAMES YAW ASIBUO ....

(Co-Supervisor) Signature Date

Certified by:

DR. J. SARKODIE-ADDO ....

(Head of Dept) Signature Date

DEDICATION

This work is dedicated to my wife, Mrs Zippora Appiah-Kubi.

ACKNOWLEDGEMENT

I thank the Almighty God for seeing me through this programme. My special thanks go to my able Supervisors; Prof. Richard Akromah, Dean, Faculty of Agriculture and of Crop and Soil Sciences Department, KNUST, for his pieces of advice and special interest shown in this work and Dr James Y. Asibuo of the CSIR-Crops Research Institute (CSIR-CRI) for the technical support.

I wish to express my heartfelt gratitude to the Director-General of Council for Scientific and Industrial Research (CSIR), Dr. A.B. Salifu, Deputy Director General CSIR, Dr. Mrs. RoseEmma Mamaa Enstua-Mensah and the Director of CSIR-Crops Research Institute, Dr. Hans Adu-Dapaah, and the entire management of the CSIR who made it possible for me to pursue this Master's programme.

I also wish to acknowledge the support, guidance and encouragement I received from Dr. (Mrs) Elizabeth Yaa Parkes, Dr. (Mrs.) Marian D. Quain, Mr. Albert Aubyn, Mrs. Ruth N.A Prempeh, Mrs. Priscilla Francisco Ribeiro, Mrs. Linda Appianimaa Oppong and Mr. Charles Debrah all from Biotechnology Unit of the CSIR-Crops Research Institute. I cannot forget to express my appreciation to Mr Emmanuel Owusu Kyere and Mr Godwin Essiaw-Quayson for their useful support in ensuring early completion of this work. Special mention is made of the following people who in one way or the other contributed to the success of this work; Mr. Allen Oppong, Mrs. Rose Osei-Kofi, Mr. Lord Ohemeng-Appiah and the entire staff of the legumes division CSIR-CRI, especially, Mr Michael Adu, Samuel Akuffo and of course all my friends and colleagues at the CSIR-CRI.

I wish also to express my sincere thanks to Mr Sam Korie of IITA-Ibadan, Nigeria and Mrs Adelaide Agyeman of CSIR-CRI, Kumasi-Ghana, for helping me in analysing the research data.

My acknowledgement would not have been complete without mentioning Dr. Hideyuki Funatsuki and his entire soybean research team at National Institute of Crop Science, Kannondai – Tsukuba, Japan, from whose generosity that shattering resistance markers and publications were made available to me for free. I will forever be grateful for this magnanimity.

Finally, I express my deepest appreciation for the support, encouragement and inspiration I received from my dear wife Zippora, my parents and all other family members and everybody, who in one way or the other, has contributed to my studies.

ABSTRACT

Experiments were conducted to study the genetic diversity among soybean accessions from three different countries and also to validate the usefulness of SSR shattering resistant markers for marker-assisted selection in breeding for soybean shattering resistance in Ghana. In all, 36 soybean accessions from Ghana, Nigeria and Brazil were studied. A genetic diversity study was conducted using 20 soybean SSR markers via PCR and PAGE electrophoresis. Molecular data was scored and analysed using CLUSTER procedure in SAS version 9.2 (2007). Morphological characterisation based on qualitative and quantitative traits, were analysed using hierarchical cluster analysis and principal component analysis (PCA) procedures of SAS (2007). Validation of the three novel QTL (*qPDH1*) soybean SSR shattering resistant markers (SRM0, SRM1 and SRM2) was done by evaluating known shattering resistant and shattering susceptible parents and their F_1 progenies with the markers on agarose gel. The 20 SSR markers grouped the germplasm into six clusters based on Jaccard similarity coefficient. Three morphological traits namely: plant height, number of seeds per pod and days to maturity were the most important traits that discriminated the germplasm into three clusters based on PCA biplot, which corresponded to their country of origin. The three resistant markers were successfully validated for their usefulness for breeding shattering resistance in the germplasm studied. It is recommended that these SSR shattering resistant markers be used as integral part for breeding for soybean shattering resistance in Ghana.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF PLATES

CHAPTER ONE

1.0 Introduction

Soybean [*Glycine max.* (L.) Merrill] is a leguminous annual crop belonging to the family Fabaceae. It is an erect bushy plant with a well-defined main stem and branches, with numerous leaves. According to FAO (2012), total world production at 2010 was 261.6 million metric tonnes. The three major world producing countries are U.S.A (90.6 million metric tonnes), Brazil (68.5 million metric tonnes) and Argentina (52.6 million metric tonnes). The total production in Africa was 1.5 million tonnes with West Africa producing 437,115 metric tonnes. Nigeria is the leading producer in West Africa with 393,860 metric tonnes (FAO, 2012). According to the Ministry of Food and Agriculture (MoFA), 124,045 metric tonnes of soybean was produced in Ghana in 2010, with an additional import of 200 metric tonnes in that same year (MoFA, 2011).

Soybean is grown primarily for the production of seed and has several uses in the food and industrial sectors. It represents one of the major sources of edible vegetable oil and proteins for livestock feed. Soybean is important in Ghana for several reasons. In Ghana, soybean is consumed by humans, animals and it improves soil fertility. It is one of the few legume crops that have the greatest potential to contribute to employment and income generation in rural communities especially if its agro-industrial potentials are exploited. Among the grain legumes, soybean currently ranks third after groundnut and cowpea in terms of production and utilization in Ghana (MoFA, 2011).

The demand for soybean as a raw material for the oil and poultry industries has been on the ascendency over the years. Its nutritional benefits for the consumer have been well documented (Asalm *et al*., 1995; Asafo-Adjei *et al*., 2005, 2007a; Mebrahtu,

2008). Soybean research and production in Ghana are besieged with a lot of constraints. These include marketing, seed viability, low research effort, pest and diseases, narrow genetic base, and pod shattering among others. However, narrow genetic base and pod shattering are two of the major problems that affect soybean production in Ghana.

The exiting soybean varieties in the country were introductions from International Institute of Tropical Agriculture (IITA); which were evaluated by CSIR-Crops Research Institute and CSIR-Savannah Agriculture Research Institute and released as varieties to farmers (Asafo Adjei and Adekunle, 2001; Tefera *et al.,* 2010). Eight commercial soybean varieties have so far been released. Since the soybean varieties Ahoto and Nangbaar were released in 2005, no new variety has been reported.

Introducing new soybean lines can increase genetic diversity, thereby facilitating the development of new varieties that can address some of the constraints in soybean production. There is the need to assemble new germplasm from different geographical origins in order to assess their genetic diversity at both morphological and molecular levels. In this way, new parental lines with desirable traits could be obtained to initiate hybridisation to produce new hybrids to meet both industrial and consumer needs.

Among soybean growing regions in Ghana, the largest production occurs in the northern part which lies within the Guinea-Savanna and Sahel agro-ecological zones (Lawson *et a*l., 2008). These regions are characterised by high temperature, low rainfall and occasional drought which induce pod shattering leading to yield loss (Philbrook and Oplinger, 1989). Yield losses due to shattering have been reported to be 422 kg/ha (Shirota *et al*., 2001) and seed losses of 50-100 % have been reported (IITA, 1986). In the advent of climate change and its associated high temperatures

and prolonged drought, identifying pod shattering-resistant cultivars will minimise yield loss at harvest and, hence, increase yield.

The availability of biotechnology tools, such as molecular markers, can facilitate the breeding process. Shattering markers have been developed by Suzuki *et al.* (2010). Validating these novel markers for their usefulness in detecting pod shattering-resistant cultivars could enhance breeding activities and thereby improve soybean production.

There is narrow genetic base of soybean in Ghana (MoFA, 2003b). The need to widen the genetic base so that suitable varieties that possess required agronomic traits can be developed is necessary. Most of the varieties currently available to farmers are susceptible to shattering (Mohammed, 2010) which require research attention. The shattering molecular markers developed have not been validated with the accessions in Ghana. This requires research effort to assess their usefulness for breeding shattering- resistant cultivars.

The narrow genetic base of soybean in Ghana is a major constraint to genetic improvement, therefore broadening the genetic base will ensure development in the country. Several soybean processing factories have sprung up and are looking for varieties with high yields and high oil content. Soybean varieties developed in IITA may not meet the specific needs of end users in Ghana; there is therefore the need for crosses to be initiated to develop varieties suitable for the local consumers. To address this problem requires the introduction of new germplasm from different geographical locations to enhance the genetic diversity in Ghana.

Similarly, if efforts to increase productivity could be realised, then shattering, this is a major constraint to yield and needs to be addressed. Therefore, this work was

designed to find solutions to soybean pod shattering problem in Ghana. Soybean shattering-resistant markers have been developed and these markers have to be validated for their usefulness in breeding and selection for shattering-resistant genotypes in Ghana.

The main objective of this study was to evaluate the genetic diversity existing among local and exotic germplasm (from IITA, Nigeria and Brazilian Agricultural Research Corporation (EMBRAPA), Brazil) and to validate SSR shattering-resistant markers for marker assisted breeding.

The specific objectives were to:

i. Study the relationship between soybeans from EMBRAPA, IITA and released varieties currently grown by farmers.
ii. Correlate the cluster generated from both SSR molecular markers and morphological traits.
iii. Validate the usefulness of shattering resistance markers on 36 soybean germplasm and F_1 hybrids from known susceptible and resistant genotypes.

CHAPTER TWO

2.0 Literature Review

2.1 Biology of Soybean

The genus *Glycine* (Wild) is composed of two subgenera: *Glycine* and *Soja* (Moench) F.J. Herm. The wild perennial soybeans belong to the subgenus *Glycine* and composed of 16 wild perennial species (Newell and Hymowitz, 1980; Tindale, 1984, 1986a, 1986b; Singh and Hymowitz, 1985; Tindale and Craven, 1988, 1993). Wild perennials are diverse in morphology, cytology and genome (Singh *et al*., 1988; 1998; 1992a; b).

The cultivated soybean and its wild annual progenitor *G. soja* (Sieb. and Zucc.) belong to the subgenus *Soja*, which are diploid (2n=2x=40), are cross compatible, usually produce vigorous fertile F_1 hybrids, and carry similar genomes (Palmer *et al*., 1996; Acquaah, 2007). There are more than 100,000 *G. max* accessions, less than 10,000 *G. soja* accessions, and approximately 3,500 accessions of perennial *Glycine* species in germplasm collections throughout the world (Juvik *et al*., 1985).

In ideal soil conditions, soybean is infected by *Rhizobium* bacterium, resulting in roundish nodules on the roots in which the nitrogen-fixing bacteria live for N-fixation. The nodulated root system consists of a taproot from which emerges a lateral root system. The plants of most cultivars are covered with fine trichomes, but glabrous types also exist.

The primary leaves of soybean are unifoliate, opposite and ovate. The secondary leaves are trifoliate and alternate; the leaves are compound with four or more leaflets occasionally present.

The flowers are either purple or white, and are borne in axillary racemes on penduncles at the nodes. The papilionaceous flower consists of a tubular calyx of five sepals, a corolla of five petals (one banner, two wings and two keels), one pistil and nine stamens with a single separate posterior stamen. The stamens form a ring at the base of the stigma and elongate one day before pollination, at which time the elevated anthers form a ring around the stigma and are self pollinated (Acquaah, 2007).

The plant produces a large number of flowers, but only about two-thirds to three-quarters of them produce pods (Acquaah, 2007). The pods are also pubescent and range in colour from light-yellow to black. They are usually straight or slightly curved in shape, vary in length from two to seven centimetres, and consist of two halves of a single carpel which are joined by a dorsal and ventral suture.

The pod usually contains one to three seeds (occasionally four) (Asafo-Adjei, 2005). The shape of the seed, usually oval, can vary amongst cultivars from almost spherical to elongated and flattened. The seeds are usually uncoloured and may be straw-yellow, greenish-yellow green, brown, or black (Acquaah, 2007). Bicoloured seeds exist, such as yellow with a saddle of black or brown. The hilum is also coloured with various patterns such as yellow, buff, brown or black (Acquaah, 2007).

2.2 Soybean Improvement in Ghana

Soybean is a relatively new crop in Ghana compared to other leguminous crops such as cowpea and groundnuts (Asafo-Adjei *et al*., 2005). The soybean breeding programme in Ghana is spearheaded by the Crops Research Institute, Kumasi and Savannah Agricultural Research Institute, Tamale; of the Council for Scientific and

Industrial Research (CSIR) with the support of other international institutions, such as IITA and EMBRAPA.

The released soybean in Ghana varieties are Salintuya-1, Bengbie, Anidaso, Quarshie, Jenguma, Nangbaar, Ahoto and Salintuya-2; and were original breeding lines developed in IITA (Asafo Adjei and Adekunle, 2001; Asafo-Adjei *et al*., 2005). These varieties had some desired traits such as high stable yields, disease resistance and shattering resistance (Mohammed, 2010). Grain yields of these released varieties range from 1.2-2.8 t/ha depending on the cultivar, environment and management practices employed (Lawson *et al*., 2008).

Soybean is well adapted to wide range of climate and soil conditions. In Ghana, the best environments for soybean cultivation are the Forest-savanna transition, Guinea-savanna and Sahel agro-ecological zones with well-drained fertile soils and annual rainfall not less than 700 mm distributed throughout the growing period (Asafo-Adjei *et al*., 2005; Lawson *et al*., 2008).

The demand for soybean has increased after an initial production of about 1,000 metric tonnes in 1990 to an estimated production of about 15,000 metric tonnes in 2004 (Asafo-Adjei *et al*., 2005). Ministry of Food and Agriculture Draft Policy Report (2006) projected an annual production of 40,000 metric tonnes by 2008. While this is far below the demand of over 90, 000 metric tonnes required by the existing 20 and smaller to medium scale industries (MoFA, 2002) that process soybean.

The Ministry of Food and Agriculture projects that Ghana has a potential of 750, 000 MT annually from an estimated 500,000 ha of arable land that is suitable for soybean cultivation in Ghana (MoFA, 2003a). According to the Statistics, Research and

Information Directorate (SRID) of Ministry of Food and Agriculture, total domestic production in 2010 was 124,045 MT (MoFA, 2011).

World production of soybean for 2010 indicated that the USA is leading producer; produced about 90.6 million metric tonnes followed by Brazil with 68.6 million MT; and China being the centre of origin produced 15.1 million MT (FAO 2012). In terms productivity in that same year, USA and Brazil produced almost the same with 2.92 ton/ha and 2.94 ton/ha respectively; whilst China recorded 1.77 ton/ha (FAO 2012).

In Africa, South Africa is the leading producer with 566,000 MT; followed by Nigeria, 393,860 MT and Uganda, 175,000 MT (FAO 2012). Soybean productivity figures in African have been relatively low. According FAO (2012), South Africa produced 1.82 ton/ha, followed by Nigeria with 1.40 ton/ha then Uganda producing 1.13 ton/ha as at 2010.

2.3 Economic Importance of Soybean

Soybean, grown primarily for the production of seed, has several uses in the food and industrial sectors, it represents one of the major sources of edible vegetable oil and proteins for livestock feed (Asafo-Adjei, 2005).

In Ghana the cereals, root and tuber crops and plantains supply mainly carbohydrates, whilst the grain legumes are the major source of proteins, fat and essential micronutrients in the diets. Among the grain legumes, soybean currently ranks third after groundnut and cowpea in terms of production and utilization (Asafo-Adjei et *al*., 2005). Soybean seed contains about 38.5 - 45.8 % protein, 15.84 – 30 % carbohydrate and 17.4 – 20 % oil (Asalm *et al*., 1995; Asafo-Adjei, *et al*., 2005, 2007a; Mebrahtu, 2008). It is also rich in minerals particularly calcium, phosphorus, iron and vitamins

(thiamin, riboflavin and niacin) (Messina, 1999; Nti *et al*., 2005). Soybean oil is of high quality and has no cholesterol. Its protein is also of high quality and comparable to those from animal sources such as meat, egg and milk (Carroll *et al*., 1978).

In Ghana, soybean cake/meal is a preferred protein source in the animal feed industry. Currently, it is used in preparing weaning foods for infants to prevent kwashiorkor (protein malnutrition) in children (Asafo-Adjei *et al*., 2005). It is used to fortify various traditional foods such as gari, stew, sauces, banku, and kenkey to improve their nutritional levels without changing their taste or cooking time (Asafo-Adjei *et al*., 2005).

The soybean crop is able to fix atmospheric nitrogen for its own use and a residual for the succeeding crop in rotation. The use of soybean in rotation with cereals results in drastic reduction in striga seed bank in soils (Kroschel and Sauerborn, 1988; Alabi *et al*., 1994; Berner *et al*., 1994; Denwar and Ofori, 2003) thus, making it possible for such cereals to be grown with minimal or no striga attack.

In Brazil, which accounts for about 27% of the world soybean production, harvested raw produce are intrinsically related to its agribusiness. In 2004, soybean exports from Brazil were US$10.0 billion (EMBRAPA 2003; MAPA 2005).

2.4 Shattering in soybean

Shattering (dehiscence) is simply the splitting of dry fruit to release its seeds prior to harvesting. Matured dried pods of soybean burst open along the dorsal and ventral sutures and scatter their seeds under less humid conditions. The dehiscence of soybean pods is one of the major obstacles to mechanical harvesting, because of losses due to scattering seeds. Therefore, a high degree of resistance to pod shattering is essential in commercial soybean cultivation. There are a lot of factors influencing

pod dehiscence in soybean; these include anatomical structure of pod, environmental factors, and genotype by environment interactions (Tukamuhabwa *et al*., 2002)). Also, low humidity, high temperature, rapid temperature changes, wetting and drying are some of the environmental conditions. Calson (1973) stated that the direct cause of pod dehiscence must be the difference in tension developed in cells of the inner sclerenchyma layer due to moisture loss.

Evaluation methods for pod shattering have been established (Tsuchiya, 1986; Jiang *et al*., 1991; Tukamuhabwa *et al*., 2002) and have proven usable in breeding programmes (Tsuchiya, 1986). This involves heated air drying treatment of pods at 60^0C for three hours in the laboratory. In the field, pod shattering is indicated by the date of incipient shattering (number of days from maturity to the date when pod shattering began in each plant) and the degree of shattering (percentage of shattered pods to all pods of each plant on the 40^{th} day after maturity).

2.5 Genetic Diversity of soybean

Information on genetic diversity and relationships in crop plants are important for efficient selection of parental lines for new crosses and preservation of germplasm by plant breeders (Tatineni *et al*., 1996). Traditionally, morphological traits have been used to distinguish crop varieties (Chowdhury *et al*., 2001). Morpholgical traits are highly influenced by environmental factors. For plants with narrow genetic base, such as soybean, molecular characterisation can provide additional information on their degree of genetic diversity (Acquaah, 2007).

2.5.1 Genetic diversity in soybean germplasm based on morphological characters

Success of a crop breeding programme depends on the extent of variability present in the available germplasm, choice of the parents and the selection procedure. Morphological traits or characters reflect not only on the genetic composition of a cultivar, but also the interaction of the genotype with the environment in which it is expressed (Smith and Smith, 1992).

According to Shadakshari *et al*. (2011), among the 12 morphological characters used to analyse the genetic diversity of 50 soybean germplasm, number of seeds per plant accounted for 40.24 % in assessing the diversity; followed by seed yield per plant contributing 20.12 %. Dayaman *et al*. (2009) also reported that among the 22 morphological traits used to investigate diversity of selected Indian soybean accessions, seed yield recorded highest coefficient of variation of 40.06 followed by number of branches per plant.

Mebatsion *et al.* (2012) evaluated grain shape variability using principal component analysis (PCA) and 99 % of the variation in the shape of grains was captured by the first two principal components. Similarly, Bhartiya *et al*. (2011) used PCA to determine the variability of both indigenous and exotic black soybean from different eco-geographic regions of the world for which the first four principal components together accounted for 70.28 % of the total variance. Feng *et al*., (2008) also used PCA to evaluate the genetic relationship among 40 soybean genotypes of Southern United States of America. They found that the first and second principal components accounted for 69 % and 3 %, respectively, of the total variation.

In using morphological data for cluster analysis, Dayamann *et al*. (2009) used 45 soybean accessions and grouped them into six different clusters based on morphology.

Griffin and Palmer (1995) grouped 68 genotypes of soybean into seven clusters based on morphology. Ojo *et al*., (2012) also reported that phenotypic diversity among 40 soybean genotypes using cluster analysis generated seven clusters.

2.6 Use of molecular markers for breeding

Traditionally, genetic diversity of cultivars of *Glycine max* is determined by a combination of morphological or agronomic traits and biochemical tests/assays (Chowdhury *et al*., 2001: Dayaman *et al*., 2009). Most commercial and released soybean cultivars arose from hybridization between members of an elite group of genotypes; hence the amount of genetic variability among those cultivars is small (Chowdhury *et al*., 2001). Such cultivars are often indistinguishable based on agro-morphological traits or biochemical tests which are often subjected to environmental influence interplaying with a number of genes and thus may not represent genetic divergence in the entire genome (Diwan and Cregan, 1997; Brown-Guidira *et al*., 2000). A large number of polymorphic markers are required to measure genetic relationships and genetic diversity; as a result, it is now widely accepted that information generated from DNA-based analyses using Restricted Fragment Length Polymorphism (RFLP), Random Amplified Polymorphic DNA (RAPDs), Simple Sequence Repeats (SSR) and Amplified Fragment Length Polymorphism (AFLP) alone, or with morphological analyses provide the best estimate of genetic diversity (Chowdhury *et al*., 2001).

2.6.1 Use of Microsatellites as molecular marker

Microsatellites also known as simple sequence repeats (SSR) consist of tandemly repeated, short DNA sequence motifs (Maughan *et al*., 1995). The popularity of microsatellites stems from a unique combination of several important advantages;

they are codominant markers, high genomic abundance in a population, and random distribution throughout the genome (Morgante *et al*., 2002). They exhibit allelic diversity. Their reproducibility is much higher than RAPDs (Demeke *et al*., 1997; Karp *et al.*, 1997). The flanking sequences of microsatellites are usually highly conserved, making it possible to design universal primers for their study across genomes (Akkaya *et al*., 1992; Diwan and Cregan, 1997).

Although microsatellites are very useful in general, they also have certain disadvantages; including relatively high cost of marker development, occasional occurrence of artefacts, such as stutter bands (Walsh *et al*., 1996).

In general, microsatellites show a high level of polymorphism, so they are very informative markers. They can be used for population genetic studies and gene mapping, ranging from the individual level (e.g. clone and strain identification) to that of closely related species (Jarne and Lagoda, 1996).

2.6.2 Assessing genetic diversity in soybean germplasm using SSR markers.

Molecular markers are frequently used in the analysis of soybean germplasm. Simple sequence repeats markers have been shown to be highly polymorphic in soybean (Akkaya *et al*., 1992; Diwan and Cregan, 1997). The analysis of the polymorphism in DNA sequences allow for a more accurate genetic characterization.

Abe *et al*., (2003) used 20 SSR loci in 131 accessions introduced from 14 Asian countries to detect genetic diversity among them. Morgante and Olivieri (1994) detected similar levels of polymorphism in seven SSR loci in a group of soybean genotypes. Akkaya *et al*. (1992) used several types of SSRs to analyze the diversity of 43 soybean genotypes including ancestral and domestic cultivars representing the northern and southern U.S germplasm. Doldi *et al*. (1997) found two to six alleles per

locus in a group of 18 soybean cultivars using 12 microsatellite loci. Diwan and Cregan (1997) observed an average of 10.1 alleles per locus in a total of 20 loci studied in soybean genotypes that represented 95% of all alleles of the germplasm cultivated in the north of the United States. In a study on 186 Brazilian soybean cultivars, Priolli *et al*. (2002) found four to eight alleles per loci using 12 SSR loci studied. They determined that SSR with (AT) and (ATT) repeat motifs were highly polymorphic in soybean and identified up to eight alleles at each locus.

Rongwen *et al*. (1995) identified 11 to 26 alleles at each of seven SSR loci in a diverse sample of soybean genotypes that included U.S. cultivars, *G. max* and *G. soja* plant introductions and Chinese landraces. Maughan *et al*. (1995) detected 79 alleles across five SSR loci in a sample of 94 soybean accessions of *G. max* and *G. soja* genotypes.

Tantasawat *et al*. (2011) used 11 SSR primers to analyse genetic relationships among 25 soybean genotypes. They reported that genetic similarity between genotypes and that the 25 genotypes formed four major clusters. Singh *et al*. (2010) also reported a cluster analysis based on coefficient of similarity classified 44 soybean genotypes into four major clusters derived from 120 SSR makers. Dayamann *et al*. (2009) used 11 SSR markers to analyse genetic diversity of 45 soybean genotypes and these accessions were grouped into 14 different clusters.

2.6.3 Use of shattering markers in marker-assisted selection (MAS)

Marker assisted selection (MAS) involves the use of genetic markers to follow regions of the genome that encode specific characteristics of a plant. The reliability of a marker to genetically predict a trait to a target locus depends upon the closeness of the genetic linkage (Barr, 2009). According to Barr, (2009), markers that co-segregate

(presence of resistant or susceptible allele) with the target trait are reliable and are regarded as diagnostic.

Although conventional heat and drying methods have been established in evaluating pod shattering in soybean, they are not convenient (Tsuchiya, 1986). However, these methods are not suitable for backcross breeding because pod shattering resistance has proven partially recessive (Tsuchiya, 1986, 1987; Tukamuhabwa *et al*., 2002), which implies the need for progeny testing for selection. Thus, a more efficient method such as marker-assisted selection is desirable.

Funatsuki *et al*. (2006) reported that the major gene conferring pod shattering resistance derived from soybean cultivar Hayahikari was located between Sat-093 and Sat-366 on linkage group J. The distance between these markers was estimated to be 2.9 cM and 6.7 cM by Funatsuki *et al*. (2006) and Song *et al*. (2004) respectively, which may be at usable level for applying marker-assisted selection. Soybean pod shattering has been found to be a quantitative trait locus (QTL) and was designated as *qPDH1* (Bailey *et al*., 1997; Funatsuki *et al*., 2008).

In the development of a useful QTL selection marker, progeny of residual heterozygous line segregating at genomic region around *qPDH1* was screened for flanking markers. Analysis of the relationship between degree of pod dehiscence and graphical genotype of the lines confined the location of *qPDH1* to a 134 kb region on chromosome 16, formerly linkage group J (Suzuki *et al*., 2010). Sequencing analysis of the parental shattering-resistant (Hayahikari) and shattering-susceptible (Toyosmusume) cultivars for the candidate genes revealed a high-frequency nucleotide polymorphism at this genomic region; for which three markers namely

SRM0, SRM1 and SRM2 have been developed using insertion/deletion (In/Del) variations of relatively large size (Suzuki *et al*., 2010).

SRM0 marker, associated with shattering-resistance, was designed to amplify at genomic regions with an 18 bp In/ Del variation with resistant and susceptible cultivars amplifying at 213 bp and 231 bp, respectively. Similarly, SRM1 marker distinguished the alleles at 12 bp In/ Del variation with resistant and susceptible cultivars amplified at 234 bp and 222 bp, respectively. The SRM2 marker also yielded two distinct alleles with more than 20 bp size difference with resistant and susceptible cultivars amplifying at 185 bp and 157 bp, respectively. These three markers were tightly linked to the major QTL, *qPDH1*, as the genetic distances between the makers and *qPDH1* were estimated to be less than 0.2 cM (Funatsuki *et al*., 2006; Suzuki *et al*., 2010).

The shattering-resistant cultivars that harbour the resistance allele at *qPDH1* exhibited the same genotype at these marker loci, which were clearly distinguished from those for shattering-susceptible cultivars with two exceptions for SRM2 (Suzuki *et al*., 2010).

The polymorphism at these marker loci were conserved between diverse shattering-resistant and susceptible cultivars. Hence, suggesting the versatility and usefulness of these markers in marker-assisted selection

CHAPTER THREE

3.0 Materials and Methods

3.1 Experiment 1: Diversity of soybean accessions based on morphological traits

3.1.1 Location and field establishment

The field experiment was conducted at the CSIR-Crops Research Institute at Fumesua during the major season of 2010. Fumesua ($6^0$43'N, $1^0$31'W) is located in the semi-deciduous forest in Ashanti Region of Ghana and characterised by sandy-loam soils (Parkes *et al*., 2012). The experimental field was ploughed and harrowed. Thirty-six soybean accessions made up of 20 lines from IITA, Nigeria; 10 lines from EMBRAPA, Brazil and six released varieties in Ghana by CSIR-SARI and CRI were planted in June 2010. The soybean accessions used for the study are listed in Table 3.1. Four row plots measuring 4 m long and a spacing of 60 cm (between rows) by 5 cm (within rows) with three replications were used (Plate 1). Weeds were controlled by hand hoeing at three and six weeks after planting. Pre-flowering insects were controlled using karate (2.5 % lamda-cyhalothrin) at 50 ml karate in 15 L of water in spraying, and against post-flowering insects using cydimethoate (36g cypermethrin plus 400 g dimethoate per litre) at 100 ml cydimethoate in 15 L of water for application.

PLATE 1: Soybean field showing immature green pods

3.1.2 Morphological data collection

Nine qualitative and sixteen quantitative data were collected using morphological and phenological descriptors based on United States Department of Agriculture (USDA), Agricultural Research Service (ARS), Germplasm Resources Information Network (GRIN) (USDA-ARS, 2010). Field data collection was based on the two central rows of each plot as harvested soybean accessions are shown in Plate 2.

PLATE 2: Labelled dried - harvested soybean genotypes for morphological data collection.

Table 3.1 Soybean accessions used for the study

Soybean cultivar/Lines	Source /Institution	Country
TGX 1805-31F	IITA	Nigeria
TGX 1903-7F	IITA	Nigeria
TGX 1903-8F	IITA	Nigeria
TGX 1904-2F	IITA	Nigeria
TGX 1904-3 F	IITA	Nigeria
TGX 1904-6F	IITA	Nigeria
TGX 1909-3F	IITA	Nigeria
TGX 1910-16F	IITA	Nigeria
TGX 1835-10E	IITA	Nigeria
TGX1903-2 F	IITA	Nigeria
TGX 1903-1 F	IITA	Nigeria
TGX 1842-18E	IITA	Nigeria
TGX 1843-29E	IITA	Nigeria
TGX 1844-4 E	IITA	Nigeria
TGX 1844-18E	IITA	Nigeria
TGX 1910-2F	IITA	Nigeria
TGX 1910-3F	IITA	Nigeria
TGX 1910-6F	IITA	Nigeria
TGX 1910-14F	IITA	Nigeria
GMX 92-6-10M	IITA	Nigeria
Boavista	EMBRAPA	Brazil
Celeste	EMBRAPA	Brazil
Tracaja	EMBRAPA	Brazil
Pirarara	EMBRAPA	Brazil
Flora	EMBRAPA	Brazil
Brazillia	EMBRAPA	Brazil
MG 68	EMBRAPA	Brazil
Raimunda	EMBRAPA	Brazil
MG/BR 46	EMBRAPA	Brazil
Sambaiba	EMBRAPA	Brazil
Jenguma	CSIR-SARI	Ghana
Salintuya-1	CSIR-SARI	Ghana
Sanlintuya-2	CSIR-SARI	Ghana
Quarshie	CSIR-CRI	Ghana
Nangbaar	CSIR-SARI	Ghana
Anidaso	CSIR-CRI	Ghana

3.1.3 Scoring of Qualitative traits soybean

3.1.3.1 Flower Colour

Flower colour was scored using the code in Table 3.2

Table 3.2 Flower code

Code	Definition
B	Blue
Dp	Dark purple
Lp	Light purple
M	Magenta
Nw	Near white
P	Purple
Pth	Dilute purple (purple throat)
W	White

3.1.3.2 Pod Colour

Pod colour was recorded when pods had dried and were scored as in Table 3.3.

Table 3.3 Pod Colour

Code	Definition
Bl	Black
Br	Brown
Dbr	Dark brown
Lbr	Light brown
Tn	Tan

3.1.3.3 Pubescence colour

Pubescence colour was recorded when plants have dried and was scored as in Table 3.4

Table 3.4 Pubescence colour

Code	Definition
G	Gray
Lt	Light tawny
Ng	Near gray
T	Tawny

3.1.3.4 Pubescence density

Pubescence density was recorded when plants were fully matured and scored using Table 3.5

Table 3.5 Pubescence density

Code	Definition
Dn	Dense (increased density most noticeable on stem)
G	Glabrous (no pubescence)
N	Normal density
Sdn	Semi-dense (slightly increased density)
Sp	Sparse (greatly reduced density, most noticeable on stem)
Ssp	Semi-sparse (slightly reduced density, especially on pulvinus)

3.1.3.5 Pubescence form

Pubescence form was scored as in Table 3.6.

Table 3.6 Pubescence form

Code	Definition
A	Appressed (most hairs flat on leaf surface)
C	Curly (twisted and appressed)
E	Erect on leaf surface
I	Irregular (slightly curly or twisted)
Pt	Puberulent (minute)
Sa	Semi-appressed (between erect and appressed)
Va	Very appressed (all hairs flat on leaf surface)

3.1.3.6 General shape of seed

General shape of seed was scored using as in Table 3.7

Table 3.7 General shape of seed

Code	Definition
Nr	Near round
Ob	Oblong (near rectangular in lateral outline)
Ov	Oval (elliptical in lateral outline)

3.1.3.7 Seed Coat Colour and Hilum Colour

Seed coat colour and observed hilum colour characters were recorded using the colour code in Table 3.8.

Table 3.8 Colour code for seed coat and hilum

Colour Code	Definition
Bf	Buff
Bfib	Buff/imperfect black
Bl	Black
Blbr	Black hilum w/ brown
Br	Brown
Brbl	Brown w/ black
Dbf	Dark buff
Dbr	Dark brown
Dg	Dark gray
Dib	Dark imperfect black
Drbr	Dark red brown
G	Gray
Ggn	Grayish green
Gn	Green
Gnbl	Greenish black
Gnbr	Greenish brown
H	Heterogeneous
Ib	Imperfect black
Ig	Imperfect gray
Lbf	Light buff
Lbl	Light black
Lbr	Light brown
Lg	Light gray
Lggn	Light gray green
Lgn	Light green
Lib	Light imperfect black
Rbf	Reddish buff
Rbl	Reddish black
Rbr	Reddish brown
Tn	Tan
Y	Yellow

3.1.3.8 Maturity group

Maturity group was evaluated based on the number of days from germination to when senescence of the leaf begun and pod colour changed from green to light brown or tan. Three categories of maturity period namely: early, medium, and late were assessed as indicated in Table 3.9.

Table 3.9 Maturity group

Days	Maturity group	code
≤ 100	Early	E
101-114	Medium	M
≥ 115	Late	L

3.1.3.9 Statistical analysis of qualitative data collected

The qualitative data was analysed using Statistical Analysis System (SAS) version 9.2 (2007). Qualitative traits were considered as nominal and were transformed into dummy binary-valued variables as presence or the opposite as applied to the attribute and were indicated as "1" or "0", respectively. Jaccard coefficient was computed using DISTANCE procedure in SAS (SAS, 2008). The CENTROID method was also used to perform the hierarchical cluster analysis using the distance matrix produced by the DISTANCE procedure. Distance matrix was then converted into similarity measure using PROC CLUSTER method as in SAS (2008).

3.1.4 Scoring of quantitative traits

3.1.4.1 Nodulation count

Visual nodulation was counted at 52 days after planting (when the number of nodules had increased till pod filling began) and was scored on a scale from 1= no nodule to

5 = ample large, active nodules.

3.1.4.2 Days to 50% flowering

Days after planting to the day that 50 % of a soybean genotype flowers was scored visually.

3.1.4.3 Plant height at maturity

Three plants was randomly selected and plant height measured using measuring rule, from ground level to stem tip in centimetres at maturity, for each replication and mean plant height recorded.

3.1.4.4 Scoring of stem lodging

The tendency of stem to lodge was measured at maturity. Lodging was scored visually on the scale 1 (erect) to 5 (prostrate).

3.1.4.5 Pod length of soybean

Average length of Five pods from each genotype was randomly selected was measured in millimetres using vernier calliper (Tricle brand-name) and average length recorded.

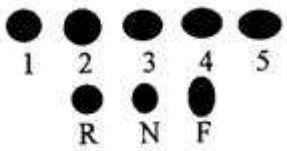
3.1.4.6 Seed Shape

Seed shape in USDA-ARS descriptor (2010) generated data on seed height, seed length and seed thickness. Five seeds from each of the selected plants (in section 3.1.4.7) were selected and measurements taken using vernier caliper (Tricle brand) and their mean measurements recorded. The seed shape based on height/length ratio and height/thickness ratio was also measured using vernier caliper. These ratios were compared based on the USDA-ARS Soybean descriptors as in Tables 3.10 and 3.11.

Table 3.10 Distribution of values for seed shape

Code	Definition
1F	Height-length ratio ≥ 0.95 and height-thickness ratio ≥ 1.40
1N	Height-length ratio ≥ 0.95 and height-thickness ratio 1.11 - 1.39
1R	Height-length ratio ≥ 0.95 and height-thickness ratio ≤ 1.10
2F	Height-length ratio 0.87 - 0.94 and height-thickness ratio ≥ 1.40
2N	Height-length ratio 0.87 - 0.94 and height-thickness ratio 1.11 - 1.39
2R	Height-length ratio 0.87 - 0.94 height-thickness ratio ≤ 1.10
3F	Height-length ratio 0.79 - 0.86 and height-thickness ratio ≥ 1.40
3N	Height-length ratio 0.79 - 0.86 and height-thickness ratio 1.11-1.39
3R	Height-length ratio 0.79 - 0.86 height-thickness ratio ≤ 1.10
4F	Height-length ratio 0.71 - 0.78 and height-thickness ratio ≥ 1.40
4N	Height-length ratio 0.71 - 0.78 and height-thickness ratio 1.11 - 1.39
4	Height-length ratio 0.71 - 0.78 height-thickness ratio ≤ 1.10
5F	Height-length ratio ≤ 0.70 and height-thickness ratio ≥1.40
5N	Height-length ratio ≤ 0.70 and height-thickness ratio 1.11 - 1.39
5R	Height-length ratio ≤ 0.70 height-thickness ratio ≤ 1.10

Side view: 1 (Round) – 5 (Very elongated)



End view: R (Round), N (Normal), and F (Flat)

Table 3.11 Longitudinal section of seed shape

Height/length	Class
≥0.95	1
0.87- 094	2
0.79 - 0.86	3
0.71- 0.78	4
≤0.70	5

Height/thickness	Class
≥1.40	F (flat)
1.11- 1.39	N (normal)
≤1.10	R (round)

3.1.4.7 Number of branching per plant

Five plants were randomly selected at harvest and branches that had emerged from the main stem were counted using a tally-counter (Brannan brand-name).

3.1.4.8 Biomass per plant

Five plants selected (same plants as in section 3.1.4.7), each whole plant with stem, branches, and pods was weighed; and a mean weight in grams recorded using analytical balance (Model: aeADAM AEP-6000G).

3.1.4.9 Number of pods per plant

Selected plants (same plants as in 3.1.4.7) and the number of pods on each plant were counted using a tally-counter (Brannan brand-name).

3.1.4.10 Number of seeds per plant

The randomly selected five plants (Section 3.1.4.7) were threshed and seeds from each plant were counted using a tally-counter (Brannan brand-name).

3.1.4.11 Hundred-seed weight

Hundred seeds from each of the five plants selected (in section 3.1.4.7) was counted and their mean weight (100 seeds) in grams (g) recorded using an analytical balance (Model: aeADAM AEP-6000G).

3.1.4.12 Seed weight per plant

Threshed seeds from each of the five plants were weighed and a mean weight in grams recorded using an analytical balance (Model: aeADAM AEP-6000G).

3.1.4.13 Days to maturity

The date that 95 % of the pods reached final colour (when pod colour changes from green to tawny) on the field was expressed in days that a soybean genotype had reached its maturity to be harvested.

3.1.4.14 Statistical analysis of quantitative data collected

For the quantitative traits, genetic similarity between pairs was calculated according to Euclidian coefficient and also centroid method of hierarchical cluster analysis was performed (SAS, 2008). The morphological data were subjected to analysis of variance (ANOVA) using the general linear model (GLM) procedure and mean separations were done using the Student Newman Keuls test under SAS (2007). Log(x+1) transformation was used on the morphological count data to stabilize the variance and normalize the data before the analysis. Also, using the correlation

procedure of SAS (2007), Pearson correlation coefficients matrix of quantitative traits was calculated.

A stepwise discriminant analysis was also conducted with the quantitative trait variables to rank them in order of importance in terms of phenotypic variance among the soybean accessions. Traits that did not have significant correlation with principal component scores were considered as redundant and unimportant. Again using the principal component analysis procedure of SAS (2007), PCA was performed on the mean values of the important traits from the stepwise discriminant analysis that accounted for most of the variations within the accessions. The biplot analysis and its graphical output were then used to understand and interpret the relationship and association between the soybean accessions and the observed quantitative trait.

3.2 Experiment 2: Diversity studies in soybean using SSR molecular markers

This study was conducted at CSIR-CRI Molecular Biology Laboratory, Fumesua in Kumasi.

3.2.1 DNA extraction:

Genomic DNA extraction was done using young but opened leaves from three-week-old soybean seedlings, Qiagen extraction kit (Qiagen sciences) was used for the extraction following the manufacturer's instructions.

3.2.2 Determination of quality and quantity of DNA extracted

Quality of DNA was assessed using gel electrophoresis and UV-adaptable spectrophotometer (Biochrom Libra S12). DNA quantification was performed by taking absorbance measurements at 260 nm and 280 nm.

A volume of 5 μl genomic DNA of each soybean genotype was added to 495 μl sterilised water in 0.5 ml eppendorf tube and were thoroughly mixed by vortexing. Resultant solution was carefully pipetted into a cuvette (*UVette*® 220-1600) and absorbance at 260 nm and 280 nm were read using spectrophotometer. The ratio of the readings obtained at 260 nm and 280 nm were calculated to assess the quality. Pure DNA free of protein contamination had A_{260}/A_{280} ratio close to 1.8. Phenol or protein contamination in the DNA prep gave less than 1.8 for the ratio A_{260}/A_{280}.

A standard solution of DNA with a concentration of 50µg/ml had an absorbance at 260 nm equal 1.0. (Stephenson H. F., 2003)

An absorbance of 1 at 260 nm = 50 µg/ml dsDNA

$$\text{DNA concentration } (\mu g/\mu l) = \frac{A_{260} \text{ X dilution factor X 50 } \mu g/ml}{1000}$$

Dilution factor = 100

3.2.3 Genetic characterisation of Soybean using SSR markers/primers.

The genetic diversity of soybean germplasm obtained from Ghana, Nigeria and Brazil were studied using 20 SSR markers (Appendix 11). The germplasm from Nigeria were developed at IITA, the Brazilian germplasm were obtained from EMBRAPA and varieties from Ghana were obtained from CSIR- Crops Research Institute. DNA extraction and purity were done as indicated in sections 3.2.1 and 3.2.2, respectively. PCR amplification was performed using the 20 soybean SSR primers.

3.2.4 PCR amplification of soybean SSRs

The soybean SSR primers were used to amplify the genomic DNA extracted from soybean (Table 3.1).1X PCR reaction mixture (*Biolabs* regents) comprised of 6.15 μl

nuclease free water, 1 μl of 10X PCR buffer, 0.9 μl of 25 mM $MgCl_2$, 0.4 μl of 10 mM dNTPs, 0.25 μl of each 5 μM primer (F/R), 0.05 μl of 5 U/μl *Taq* Polymerase and 1 μl of 30 ng/μl genomic DNA were pipetted into a PCR tube. In all, 38 X PCR reaction mixtures were prepared for one primer for the 36 samples; this was subsequently repeated for the twenty (20) SSR primers.

The amplification processes were carried out using thermocyler (Gene Amp® PCR system 9700 version 3.09) at 95 ^{0}C / 2 min as initial denaturation; followed by 92 ^{0}C/ 1min, 47 ^{0}C/ 1 min, 72 ^{0}C / 1 min for 33cycles; and 72 ^{0}C / 10 min. as the final extension. After the amplification process 2 μl of 6X loading dye (Fermentas) was added to each PCR product; and electrophoresed on 6 % non-denaturing polyacrylamide gel (Appendix 8). The extreme end wells were loaded with 4 μl 100 bp ladder to estimate band size. Each well was loaded with 12 μl PCR-dyed products and were run at 120 volts for 45 minutes after which it was silver-stained.

3.2.5 Staining of the polyacrylamide gel (PAGE)

The electrophoresed gel was fixed for 10 minutes in fixation solution (10 % v/v acetic acid) with gentle shaking and then washed in distilled water for two minutes. The fixation step was followed with oxidation with 1.5 % v/v nitric acid for three minutes. After incubating in silver staining solution (0.1% w/v silver nitrate, 750 μl formaldehyde), the gel was washed in distilled water for 10 seconds, and then transferred to cold developing solution (3 % w/v sodium carbonate, 3 ml formaldehyde, 250 μl 1X sodium thiosulphate) to develop the silver-stained DNA bands. The process was stopped using a stop solution (10 % v/v acetic acid). The gel was washed in de-ionized autoclaved water and photographed using digital camera (Rollei FLexline 202). The photographed gels were downloaded onto a computer for scoring. The scores were analysed using SAS for Windows, version 9.2 (SAS, 2007).

3.2.6 Statistical analysis of SSR molecular marker data

The photographed gels were downloaded onto a computer and weighted bands were scored as presence or absence of band using DNA ladder as the reference (1 kb Invitrogen and 100 bp Fermentas). Using the binary scored data (presence or absence of DNA bands), genetic similarity between pairs of soybean genotypes was calculated according to Jaccard's similarity coefficient, followed by centroid method of hierarchical cluster analysis using CLUSTER procedure in SAS statistical software (SAS, 2007).

3.3 Experiment 3: Validation of SSR shattering resistance markers for marker-assisted selection in soybean

The usefulness of available soybean shattering markers was validated on the 36 soybean accessions used in this study. First, the three SSR shattering resistance markers were analyzed using all the accessions. Secondly, crosses were made between susceptible and resistant (Salintuya-2 X Jenguma) cultivars, resistant by resistant (Flora X Jenguma) cultivars and another resistant by resistant (Tracaja X Flora) cultivars. The markers used are listed in Table 3.12. DNA was extracted from the leaves of two plants each of the Salintuya-2 X Jenguma and Flora X Jenguma F_1 hybrids and one plant of the Tracaja X Flora hybrid as was done in section 3.2.1. PCR and agarose gel electrophoresis was also done as in section 3.2.4. DNA bands were scored according to the reported base pair sizes of their respective resistance/ susceptible shattering markers.

3.3.1 Generation of F_1 Hybrids

Parents for the generation of F_1 hybrids were planted in sterilized soils (sandy loam) soils. The soil was sieved to obtain loose, aerated and uniform soil structure. Thirty-two pots with drainage holes beneath them; each measuring 19.0 x 15.5 x 15.5 cm were filled with the treated soil. Four seeds were planted per pot and thinned to two stands per pot after emergence. Each of the four soybean parental lines had eight pots. Planting dates were staggered (10 days difference) to ensure synchronization of flowering for crossings. Watering was done as and when necessary. DNA was extracted after three weeks of leaf emergence.

At flowering, crosses were done by pulling off buds of female parents gently and emasculating using a pair of forceps and with the aid of optical visor (magnifying lens). All other axillary shoots were removed from the axil to avoid misidentification of hybridized-pod at maturity. Desired pollen from a male flower (fully matured opened flower) was transferred to the stigma of emasculated female flower (Plate 3).

PLATE 3: Artificial of pollination soybean plant

PLATE 4: Labelled pollinated soybean in potted plants

Each cross was well-labelled with names of the two parents and date of pollination (Plate 4). Pollination/hybridization was done in the morning to ensure high success rate.

Successful crosses at F_1 (Salintuya-2 X Jenguma, Tracaja X Jenguma and Flora X Jenguma) were also planted in a sterilized potted- sandy-loam soils.

DNA extractions from these successful F_1 progenies were also done as described section in 3.2.1 after which PCR amplification was conducted with the soybean SSR shattering resistance markers.

3.3.2 Soybean SSR markers controlling pod shattering

The SSR shattering markers with their sequence and their order of closeness to quantitative trait locus (QTL) controlling pod shattering designated as *qPDH1* are presented in Table 3.12

Table 3.12 List of soybean SSR shattering resistance markers

Order of closeness to *qPDH1*	Name of Marker	Forward primer sequence (5'—3')	Reverse primer sequence (5'—3')
1	SRM0	GCCAGCCTTGTCTGTCATTT	TGATGATCAATGGTCAGATTCA
2	SRM1	AGAGCAAGAAATCACGTTGCA	CACCTCACCCCTTTTTCTCA
3	SRM2	AATCGTATTAAAATTGAAGGCATGT	AGGGGTTGAGGATGAGGAGT

3.3.3 PCR amplification of pod shattering markers

The soybean SSR primers were used to amplify the genomic DNA extracted from the four soybean parental cultivars (Salintuya-2, Jenguma, Tracaja and Flora) and their successful F_1 progenies. 1X PCR reaction mixture (*Biolabs* regents) comprise of 6.15 μl nuclease free water, 1 μl of 10X PCR buffer, 0.9 μl of 25 mM $MgCl_2$, 0.4 μl of 10 mM dNTPs, 0.5 μl of 5 μM Primer (F/R), 0.05 μl of 5 U/μl *Taq* Polymerase and 1μl of 30 ng/μl genomic DNA were pipetted into a PCR tube. In all, 10X PCR reaction

mixtures were prepared for one primer/marker; this was subsequently repeated for the three pod shattering SSR primers.

The amplification process was carried out using thermocycler (Gene Amp® PCR system 9700, version 3.09) at 95 ^{0}C/ 2 min as initial denaturation; followed by 92 ^{0}C/ 1 min, 55 ^{0}C/ 1 min, 72 ^{0}C/ 1 min for 33 cycles; and 72 ^{0}C/ 10 mins as the final extension.

After the amplification process, 2 μl of 6X loading dye (Fermentas) was added to each PCR product; and electrophoresed on 1.8 % agarose gel with 7.5 μl ethidium bromide (to fluorescence under UV-light) at 120 volts for 45 minutes. The extreme end wells were loaded with 4 ul of 100 bp ladder to estimate band size. The gel was then captured in a DNR Bio-imaging system with an in-built camera and connected to a computer and the bands scored for analyses.

3.3.4 Statistical analysis of shattering marker data

The DNA bands for shattering resistance and susceptible scored alleles were transformed to dummy binary valued variables. The presence of resistance allele was indicated as "1" whilst susceptible allele (the opposite) was indicated as "0".

Jaccard coefficient between each pair of soybean genotype was computed using DISTANCE procedure in SAS (2008). Again, the CENTROID method was used to perform the hierarchical cluster analysis using the distance matrix. Distance matrix was converted into similarity measure using PROC CLUSTER as in SAS (2007).

CHAPTER FOUR

4.0 RESULTS

4.1 Genetic Diversity of Soybean accessions based on morphological traits

Genetic diversity of soybean was assessed using both quantitative and qualitative traits and their cluster analyses are shown in dendrograms presented in figures 4.1 and 4.2 respectively below.

Table 4.1 shows descriptive statistics of quantitative traits with the mean values, standard deviation, maximum and minimum values of 36 soybean accessions. These quantitative traits can be grouped into three namely: yield performance, seed characteristics and other agronomic characteristics. Yield performance attributes include seed weight per plant, 100-seed weight and shoot biomass per plant. Seed characteristics variables also include number of seeds per pod, seeds per plant, pod length, number of pods per plant, seed height, seed length and seed thickness. Other agronomic characteristics which include lodging count, nodule count, number of branches per plant, plant height, days to 50 % flowering and days to maturity.

The yield performance of soybean accessions, for example had seed weight per plant ranging from 3.06 to 44.06 g with a mean seed weight per plant of 20.82 g; 100-seed weight also ranged from 6.28 to 29.29 g with a mean value of 13.25 g; biomass per plant ranged from 7.70 to 104.14 g with a mean of 49.03 g.

Table 4.1: Descriptive Statistics on 18 quantitative traits of Soybean accessions

Variable	N	Mean	Std Dev	Sum	Minimum	Maximum
Seed wt per plt	36	20.82	9.30	749.42	3.06	44.06
100-seed wt	36	13.25	3.96	476.97	6.28	29.29
Bmas per Plt	36	49.03	21.79	1765.00	7.70	104.14
Seeds per pod	36	2.25	0.44	81.00	2.00	3.00
Seeds per plt	36	164.49	63.31	5922.00	37.10	286.00
Pod Lt	36	3.46	0.44	124.65	2.07	4.20
Pods per plt	36	108.74	31.86	3914.00	47.00	178.00
Seed Ht	36	0.57	0.05	20.66	0.47	0.66
Seed Lt	36	0.69	0.05	25.00	0.59	0.80
Seed thickness	36	0.49	0.07	17.70	0.36	0.63
Lodging count	36	1.53	0.39	55.03	1.07	2.50
Nodule count	36	2.25	0.38	81.07	1.63	3.13
Brchs per plt	36	16.37	9.59	589.16	5.60	56.40
Plt Ht	36	61.19	16.17	2203.00	36 00	86.33
Days to 50% flower	36	42.06	2.96	1514.00	38.00	52.00
Days to maturity	36	105.89	7.23	3812.00	93.00	124.00
*SdHt By SdLt	36	0.83	0.05	29.79	0.68	0.93
*SdHt By SdTk	36	1.18	0.11	42.45	1.02	1.47

*Derived variable. Key: Lt=length Plt=plant Ht=height wt=weight Sd=seed Tk=thickness Brchs=branches Bmas=biomass

Pod length also ranged from 2.07 to 4.20 cm and the number of seeds per pod varied between 2.07 and 3.00.

The number of seeds per plant varied from 37 to 286. Again, seed height ranged from 0.47 to 0.66 cm; also seed length ranged from 0.59 to 0.80 cm and seed thickness ranged from 0.36 to 0.63 cm. Similarly, the number of pods per plant ranged from 47 to 178. Based on standard deviation values, seeds per plant were more variable.

For other agronomic characteristics, days to maturity ranged from 93 to 124, whilst days to 50 % flower ranged from 38 to 52.

The numbers of branches per plant also ranged from 6 to 56. The other quantitative traits variables were as shown in the Table 4.1.

The dendrogram in Fig. 4.1 was derived from eighteen quantitative trait (including derived variables) variables, with an indicated percentage similarity range of 71-99 %. At 72 % similarity, three clusters were identified namely I, II, and III. The major cluster among them was cluster II comprising of 32 cultivars of mixed accessions from all the three countries and had different shades of sub-clusters within it. A minor cluster, I comprised of three Brazilian cultivars namely: Sambaiba, Flora and Brazillia. Cluster III had an isolated cultivar (Salituya-2) from Ghana. Genotypes Tracaja and Pirarara (both Brazilian lines) were 98 % similar quantitatively.

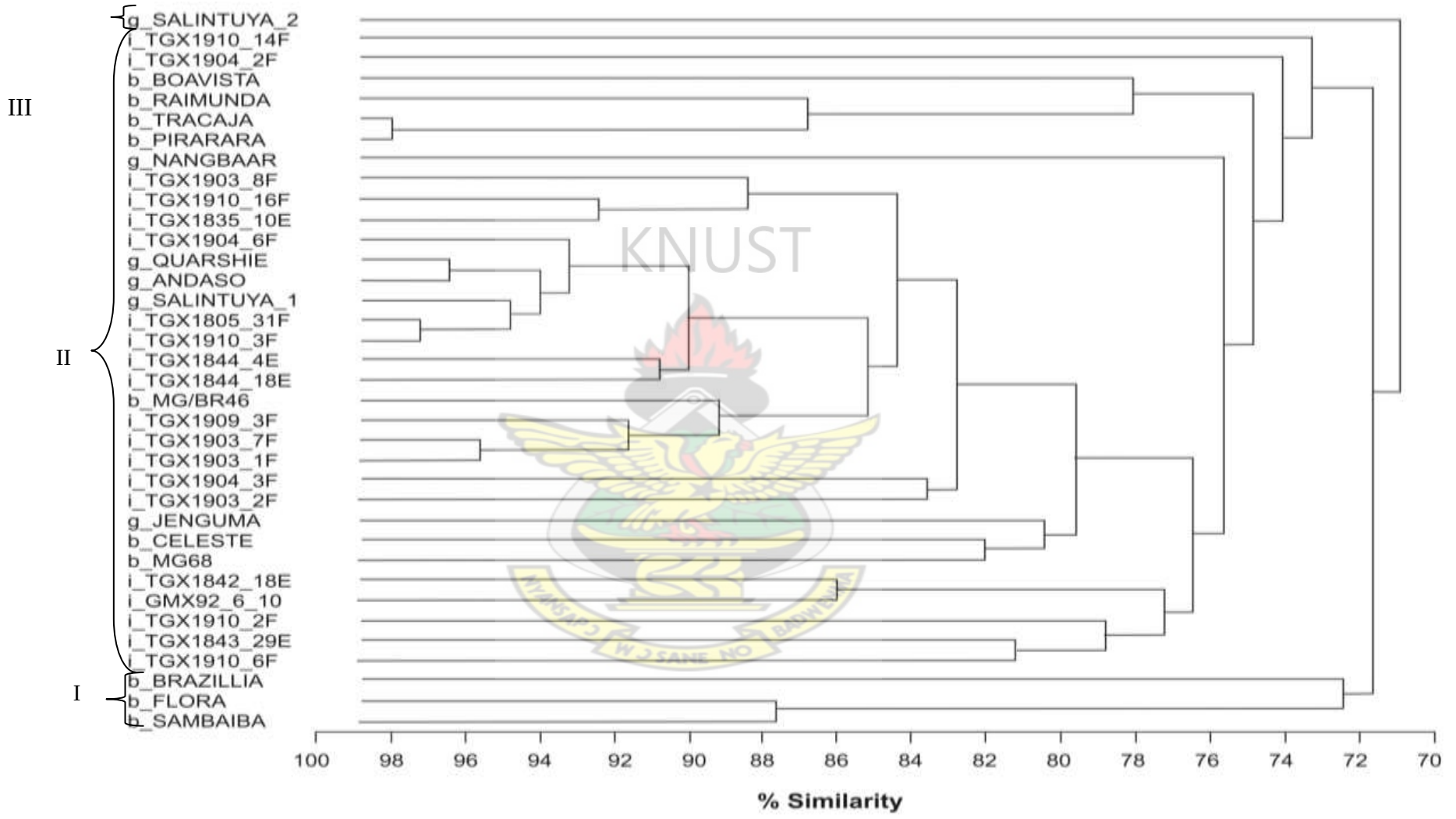


Figure 4.1: Dendrogram of soybean accessions using 18 quantitative traits. Key: Prefix; b=Brazilian line g=Ghanaian line i=IITA line

Nine qualitative traits which include flower colour, pod colour, pubescence colour, pubescence density, pubescence form, seed shape, seed coat colour, hilum colour and maturity group were also used to assess diversity.

Among the flower colours, two colours namely purple and white were identified in all the 36 soybean accessions. About 80.6 % was purple whilst 19.4 % being white. For pod colour, 80.5 % were light brown, 13.9 % brown and 5.6 % dark brown. The light tawny and tawny colours constituted 13.9 % each of the pubescence colour whilst near gray and gray were 52.7 % and 19.4 % respectively. The pubescence densities among the accessions were that, 2.8 % was dense, 13.9 % semi-dense, 5.5 % sparsely dense and 77.7 % normal density. With the levels of pubescence form, 2.8 % were erect, 86.1 % were irregular and 11.1 % being semi-appressed. For seed shape of grains, 5.6 % was near round, 58.3 % oblong and 36.1 % being oval. Among the various seed coat colours 11.1 % of the accessions were brown, 8.3 % were light gray, 22.2 % were greenish brown and 58.3 % light brown. Also for the hilum colour, 2.8 % was black, brown and dark brown had 19.4 % each; 22.7 % were brown with black; 30.5 % were light brown. Majority of the maturity grouping were medium constituting 77.8 %, whilst the early and late maturing were 19.4 % and 2.8 %, respectively. Detailed frequencies of respective qualitative traits has been summarised in Appendix 9.

From the dendrogram (Figure 4.2), percentage similarity scale ranged from 25.3-95 %. At 33 % similarilty, seven different clusters were identified. The major cluster was cluster II comprising 16 heterogeneous accessions from Ghana (Quarshie, Jenguma, Anidaso, Nangbaar and Salintuya-1) and IITA (TGX 1904-3F, TGX 1904-6F, TGX 1910-6F, TGX 1910-14F, TGX 1910-3F, TGX 1904-2F, TGX 1910-2F, TGX 1909-

3F, TGX 1910-16F, TGX 1844-4E, TGX 1844-18). The second major cluster was VI comprising seven accessions from Brazil (Flora, Brazillia, Boavista, Pirarara, Sambaiba, Celeste and Ramunda) with the remaining four clusters (I, III, IV, V) being minor having accessions ranging from two-four as shown in Figure 4.2. Clusters IV and V had three IITA lines each and appeared to be a disjointed group. There was one outlier, which was cluster VII, a Brazilian line, MG68.

Genotypes 'TGX 1909-3F' and 'TGX 1910-16F' are IITA lines which were about 95 % similar qualitatively. Genotypes 'TGX 1904-2F' and 'TGX 1910-2F' also from IITA were about 90 % similar.

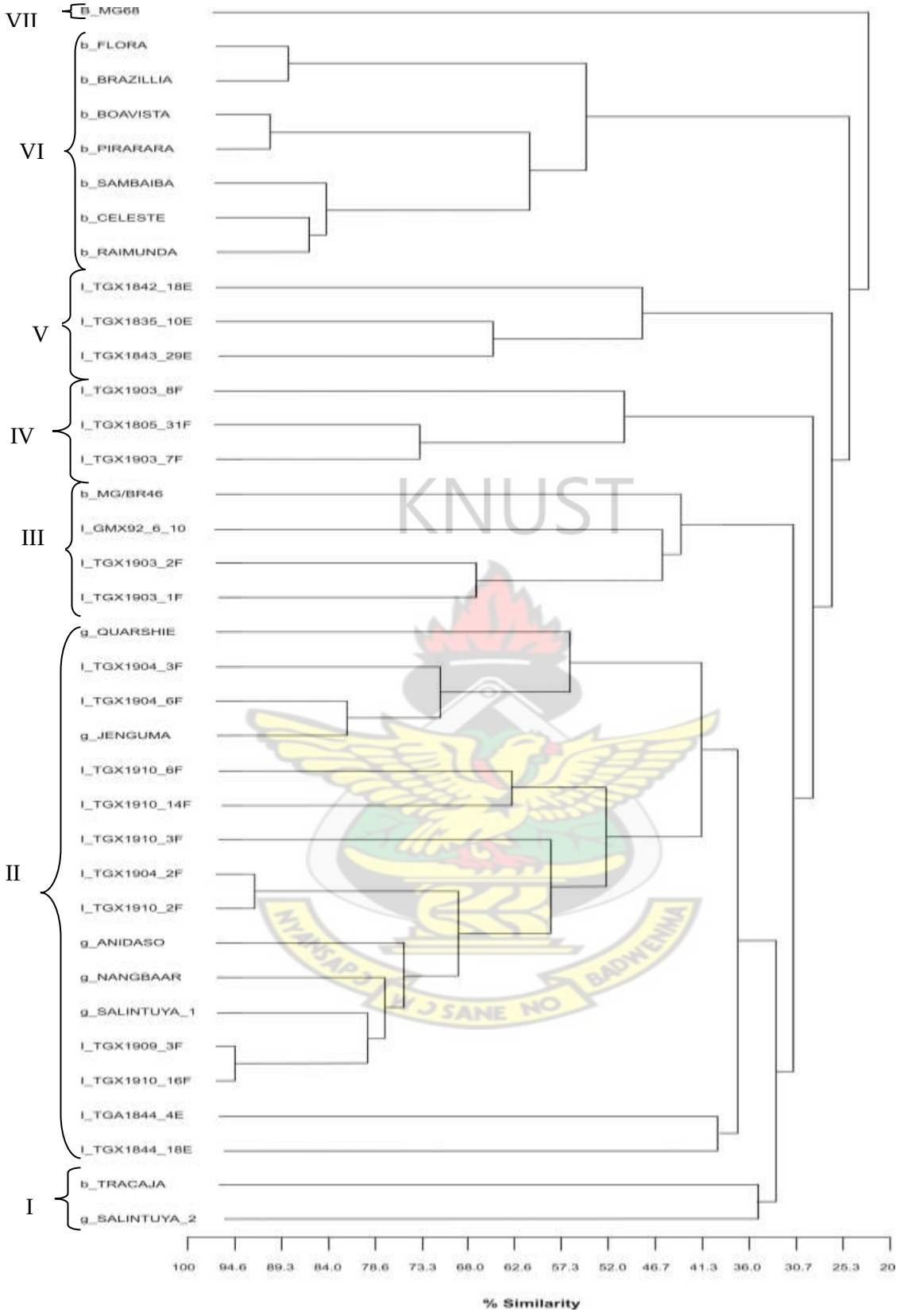


Figure 4.2: Dendrogram of 36 soybean accessions using nine qualitative traits
Key: Prefix; b=Brazilian, g=Ghanaian, i=IITA

4.1.1 Multivariate analysis of sixteen agro-morphological traits of soybean accessions

A correlation matrix of 16 quantitative trait variables (excluding the derived variables) was computed using Pearson correlation coefficient (Table 4.2). There were 136 possible correlation coefficients among the traits evaluated and only 48 reached statistically significant levels at probability of 95 % and above ($P \leq 0.05$); with not less than 0.3 regression value ($r \geq 0.3$).

Pods per plant was negatively correlated with pod length (r = -0.04) and seed per pod (r = -0.17) but positively correlated with number of branches per plant (r = 0.38). Plant height was negatively correlated with number of pods per plant (r = -0.23) but positively correlated to biomass per plant (r = 0.55), seed weight per plant (r = 0.55) and lodging (r = 0.41). Biomass per plant was positively correlated to seed per pod (r= 0.66; P < 0.01), pod length, branches per plant and seed weight per plant. Days to maturity was positively correlated to days to 50 % flowering (r = 0.64; P < 0.01); and also with seeds per plant, biomass per plant, seed weight per plant and plant height. Lodging was positively associated with seeds per plant (r = 0.33), biomass per plant (r = 0.35), seed weight per plant (r = 0.37) and plant height (0.41).

Table 4.2: Correlation matrix of 16 morphological traits of soybean accessions

	Pod Lt	Sds/ Pod	Brches /Plt	Pods/ Plt	Sds /Plt	Sd Wt 100	Bmas /Plt	Sd Wt / Plt	Ldgn cnt	Nodule cnt	Plt Ht	Days to 50% Fl	Days to Maty	Sd Ht	Sd Lt	Sd Tk
Pod Lt	1.00															
Sds/Pod	0.63**	1.00														
Brches/ Plt	0.13	0.10	1.00													
Pods/Plt	-0.04	-0.17	0.38*	1.00												
Sds/Plnt	0.30	0.17	0.35*	0.65**	1.00											
Sd Wt 100	0.56**	0.44*	0.17	-0.06	0.23	1.00										
Bmas/Plt	0.47**	0.48**	0.49**	0.48**	0.66**	0.32	1.00									
Sd Wt /Plt	0.63**	0.53**	0.33*	0.32	0.71**	0.46*	0.83**	1.00								
Ldgn cnt	0.07	-0.07	0.31	0.25	0.33*	0.08	0.35*	0.37*	1.00							
Nodule cnt	0.03	0.15	-0.01	0.03	0.21	0.20	0.12	0.27	-0.24	1.00						
Plt Ht	0.52**	0.63**	0.24	-0.23	0.11	0.30	0.55**	0.55**	0.41**	-0.01	1.00					
Days to 50% Flower	-0.12	0.08	0.24	0.07	0.06	-0.05	0.32	-0.04	0.04	-0.30	0.22	1.00				
Days to Maty	0.20	0.20	0.31	0.20	0.35*	0.05	0.56**	0.38*	0.28	-0.27	0.37*	0.64**	1.00			
Sd Ht	0.71**	0.64**	0.00	-0.18	0.31	0.57**	0.35*	0.61**	0.12	0.16	0.44*	-0.32	0.11	1.00		
Sd Lt	0.47**	0.39	-0.04	-0.33	0.00	0.48**	0.10	0.26	-0.03	0.31	0.26	-0.26	-0.05	0.66**	1.00	
Sd Tk	0.55**	0.65**	0.12	-0.24	0.24	0.58**	0.49*	0.65**	0.16	0.19	0.65**	-0.05	0.20	0.78**	0.56**	1.00

Key: Lt=length Ht=Height cnt=count Fl=flowering Bmas=biomass Maty=maturity Tk=thickness Ldgn=lodging Plt=plant Sd=seed Brches=branches. Significant coefficient: ($P \leq 0.05$)* ($P \leq 0.01$) **

The 18 quantitative trait variables were subjected to stepwise discriminant analysis and seven of the traits were identified as important based on cumulative average squared canonical correlation (CASCCC) values.

These seven traits contributed 84.3 % to phenotypic variation among the 36 soybean genotypes. The seven traits were plant height, which contributed 44 % variation, seeds per pod 24 %, days to maturity 7.9 %, nodule count 2.9 %, seed thickness 2.3 %, pod length 1.9 % and days to 50 % flowering 1.3 % (Table 4.3).

Plant height, seed per pod, and days to maturity accounted for 76 % of the total phenotypic variation among the 36 soybean accessions and they were used for the principal component analyses shown in Table 4.4.

Table 4.3: Summary of stepwise selection of seven quantitative trait variables based on discriminant analysis of 36 soybean accessions

Trait	Partial R^2	F Value	Pr > F	Cumulative average sq. canonical correlation (CASCC)	P>CASCC
Plant height	0.880	121.84	< 0.0001	0.440	< 0.0001
Seeds per pod	0.484	15.03	< 0.0001	0.681	< 0.0001
Days to maturity	0.321	7.31	0.0025	0.760	< 0.0001
Seed thickness	0.322	7.13	0.0029	0.783	< 0.0001
Pod length	0.308	6.48	0.0047	0.802	< 0.0001
Nodule count	0.168	2.82	0.0764	0.831	< 0.0001
Days to 50% flowering	0.158	2.53	0.0981	0.843	< 0.0001

Principal component one (PC1) accounted for 61 % of the total variance; principal component two (PC2) and principal component three (PC3) contributed 28 % and 11 %, respectively, thus making a total of 100 % (Table 4.4).

Plant height, seeds per pod, and days to maturity had positive loadings on PC1; however, plant height and seed per pod had negative loadings on PC2, but days to maturity were positive on the axis. The first two PC axes were the most important which together accounted for 89 % (critical value) of the phenotypic variation among the soybean accessions. The traits were loaded on these axes to group the accessions (Appendix 10) from the different countries into closely related clusters. Three clusters were identified in the biplot as shown in Figure 4.3 suggesting the presence of distinct groups on the basis of quantitative characters.

Table 4.4: Eigenvectors of the first three principal component axes (PC1, PC2 and PC3) for 36 soybean accessions using three most important quantitative traits.

	Eigenvectors		
Variable trait	PC1	PC2	PC3
Plant height	0.66	-0.15	-0.74
Seeds per pod	0.61	-0.47	0.64
Days to maturity	0.44	0.87	0.22
Proportion of variance phenotypic accounted for by PC	0.61	0.28	0.11
Cumulative proportion of phenotpic variance accounted for by PC	0.61	0.89	1.00

The graph of PC biplot apparently grouped the accessions into three different countries from which they originated, namely: Brazil, IITA-Nigeria, and Ghana. The results also revealed that quantitative traits namely: plant height, seeds per pod and days to maturity were able to identify the diversity within soybean accessions from the countries.

Cultivar g03 (Salintuya-2, see Appendix 10), which was part of the Ghanaian accessions; a known late maturing variety isolated from the Ghanaian group in ‘days to maturity’ coordinates.

The Brazilian cultivars were identified by pod with high number of seeds with relatively tall plants. There were two Brazilian cultivars, b07 and b09 (i.e., MG68 and MG/BR46) which were early maturing and associated with the IITA accessions cluster. Genotype “i19” (i.e., TGX 1910-14F), an IITA line also associated with the Brazilian cluster. From the graph of PCA biplot, it could be inferred that early maturing cultivars were relatively shorter in height and had fewer seeds per pod (Figure 4.3).

From the PCA biplot, there were two Brazilian cultivars “b07” and “b09” (i.e., MG68 and MG/BR46) that had overlapped into the IITA cluster of lines, giving an indication of earliness, short plants and bearing smaller number of seeds per pod. There was also one IITA line “i19” (i.e., TGX 1910-14F), that had overlapped with the Brazilian cluster of cultivars signifying that it had high number of seeds per pod and was relatively taller in height.

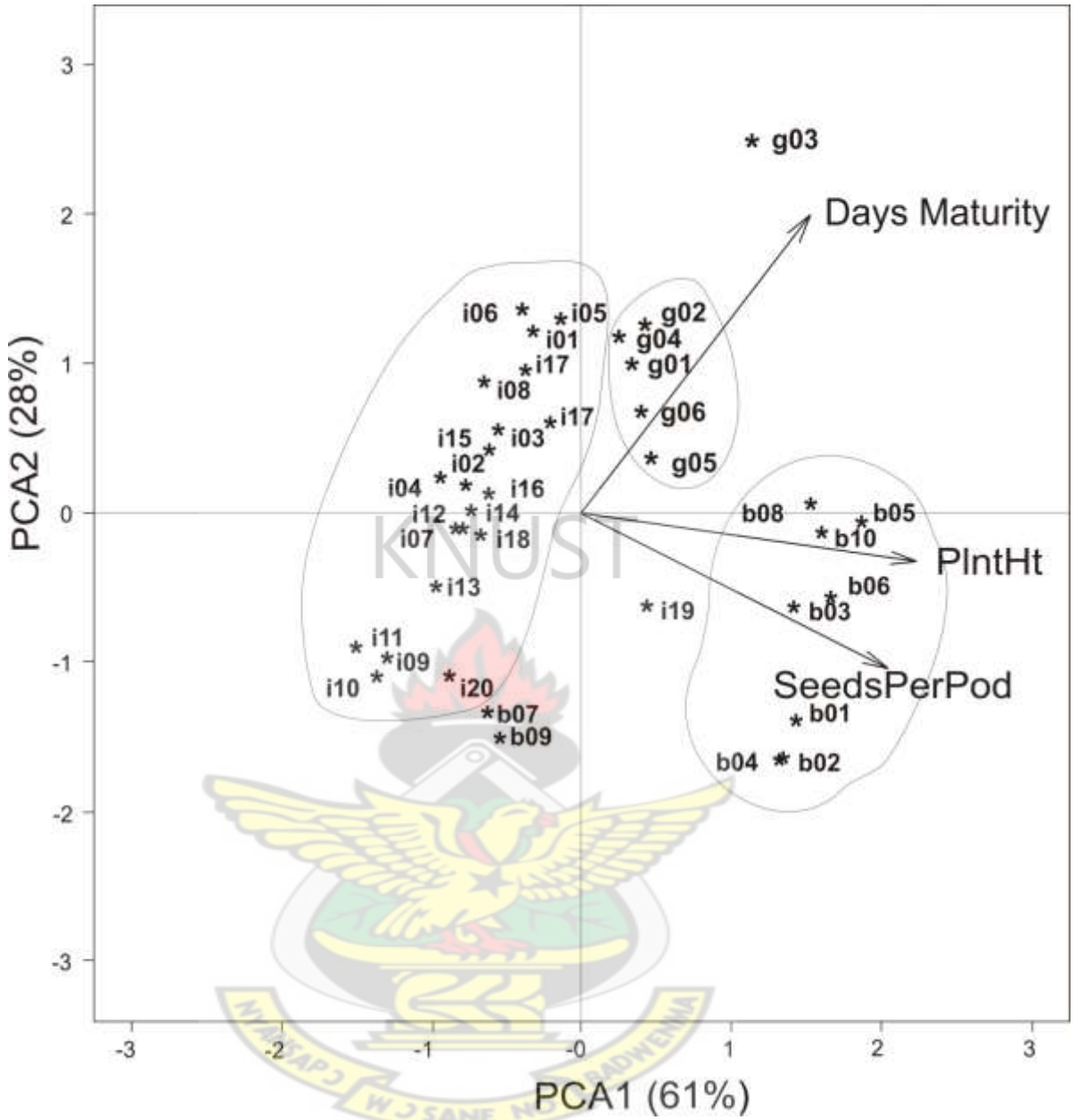


Figure 4.3: PC 1 and PC 2 biplot using quantitative trait scores of 36 soybean accessions from three different countries. Key to prefix: g = Ghana; i = IITA; b = Brazil

4.2 Diversity studies in soybean using SSR molecular markers

Figure 4.4 is a dendrogram of soybean accessions analyzed using PCR products of the 20 SSR primers/markers. A summary of all the SSR primers and their amplification output are also shown in Table 4.5 The cluster divergence of accessions were spread

out at percent similarity range of 15-96 % based on Jaccard's similarity coefficients. At a reference point of 23 % similarity level, six clusters were derived.

Clusters I, II and III were all IITA lines with an exception of Boavista in cluster I which was a Brazilian cultivar. Cluster II consisted of 17 IITA lines. Cluster III comprised only one cultivar, TGX 1842-18E.

Cluster IV, also a major cluster, consisted of a heterogenous group of cultivars composed of six varieties (Anidaso, Nangbaar, Quarshie, Salintuya-1, Jenguma and Salintuya-2) from Ghana and five cultivars, (Sambaiba, Ramunda, MG68, MG/Br 46 and Brazilia) all from Brazil.

Cluster V, consisted of three cultivars (Flora, Pirarara and Tracaja) from Brazil. Cluster VI, consisted of one Brazilian cultivar (Celeste) which was an outlier. The cluster analysis from the dendrogram indicated that two of the IITA lines TGX1910-2F and TGX1910-3F were about 94% similar at the molecular level.

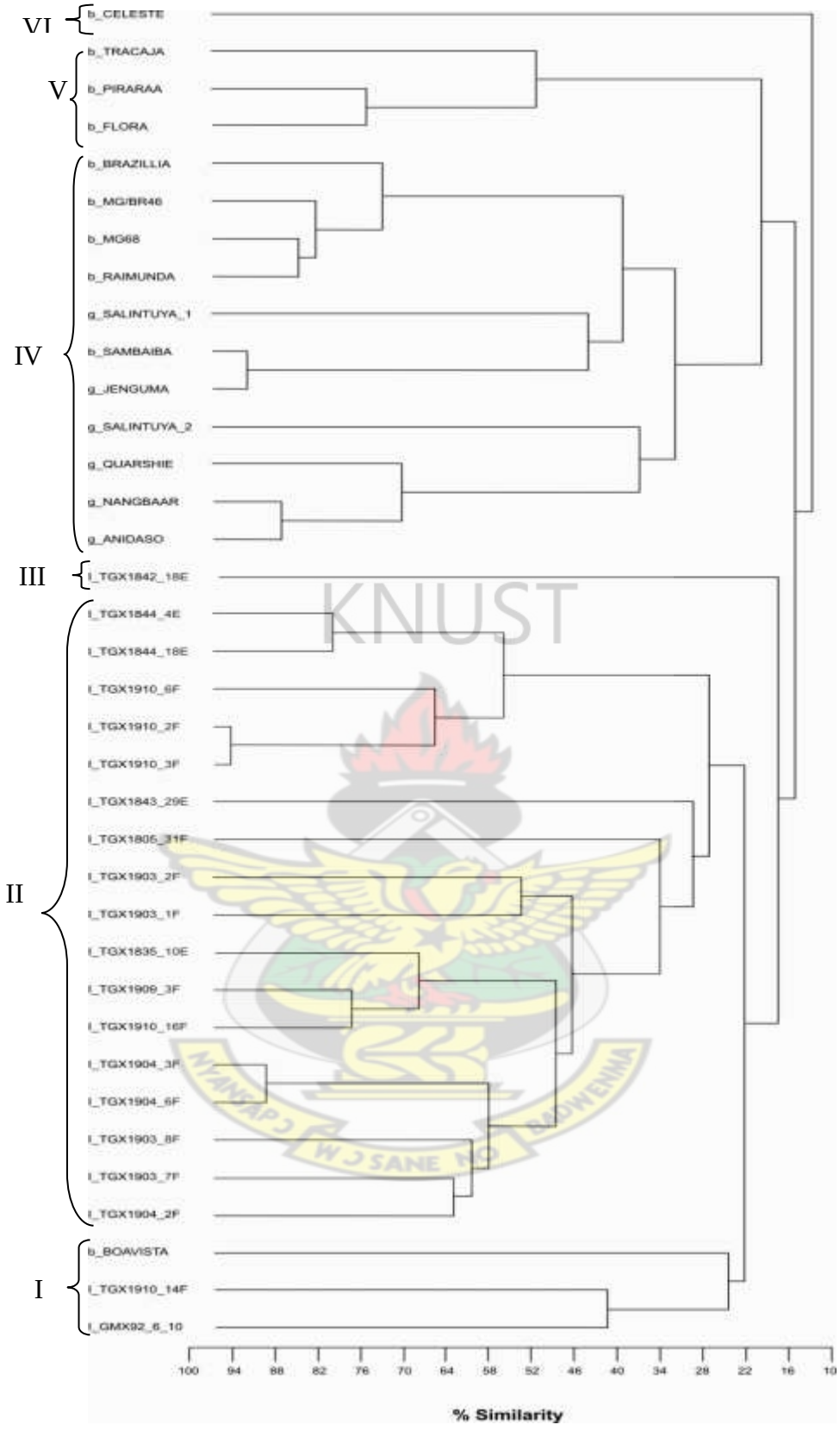


Figure 4.4: Dendrogram of 36 soybean genotypes generated from Jaccard similarity coefficient using 20 SSR markers

The genotypes “Jenguma (from Ghana) and Sambaiba (from Brazil) were about 90 % molecularly similar. Also Anidaso and Nangbaar (both from Ghana) were about 88 % similar.

4.2.1 Analysis of Variance (ANOVA) of PCR products of 20 Soybean SSR Primers

Twenty (20) SSR markers/primers were used to analyse 36 soybean accessions to determine genetic diversity within the accessions. All the primers were polymorphic expressing maximum number of allele ranging from two - seven at each locus. The primer Sat_119 was highly polymorphic with total sum of 158 alleles across the 36 soybean accessions. The mean alleles of the 20 SSR primers across the 36 soybean accession range from 0.72 – 4.39 (Table 4.5).

Based on means separation values, that is means with the same letters are not significantly different; nine (9) primer groups could be identified. Group 1 comprised one primer Sat_119. Group 2, comprised of three primers namely Sat_143, BE _4753, and Sat_219. Group 3, had one primer, Satt_593. Group 4, consisted of one primer; again, Sat_260 also had one primer Satt_522. Group 6, consisted of two primers, Sat_192 and Sat_229; again, group 7 consisted of two primers, Satt_531 and Satt_366. Group 8 comprised of four primers namely, Sat_147, Satt_631, Sat_151 and sat_299. Group 9 also consisted of five primers namely, Sat_172, Sat_250, Sat_150 and Satt_597. Within the nine groups, a primer representative from each group could be selected and used to assess the diversity across the 36 accessions. Instead of using the 20 SSR markers, nine representative markers could have been used.

Table 4.5: Summary Statistics of Soybean SSR alleles

Primer	N	No. of alleles	Mean alleles	Sum of alleles	S.E alleles	*Means
Sat_119	36	7	4.39	158	0.230	1.653^{a}
Sat_143	36	6	2.64	95	0.296	1.122^{b}
BE_4753	36	5	2.58	93	0.227	1.185^{b}
Sat_219	36	5	2.33	84	0.225	1.123^{b}
Satt_593	36	3	1.83	66	0.102	1.018^{bc}
Sat_260	36	4	1.81	65	0.173	0.956^{bcd}
Satt_522	36	4	1.64	59	0.165	0.894^{bcde}
Sat_192	36	5	1.56	56	0.317	0.661^{def}
Sat_229	36	7	1.56	56	0.366	0.622^{def}
Satt_531	36	2	1.56	56	0.101	0.902^{bcde}
Sat_366	36	4	1.53	55	0.167	0.869^{bcde}
Sat_147	36	4	1.36	49	0.204	0.712^{cdef}
Satt_631	36	4	1.36	49	0.174	0.754^{cdef}
Sat_151	36	2	1.22	44	0.070	0.783^{cdef}
Sat_299	36	2	1.03	37	0.049	0.696^{cdef}
Sat_172	36	6	0.92	33	0.220	0.478^{f}
Sat_237	36	2	0.89	32	0.096	0.584^{ef}
Sat_250	36	4	0.86	31	0.179	0.494^{f}
Sat_150	36	3	0.78	28	0.127	0.488^{f}
Satt_597	36	2	0.72	26	0.117	0.461^{f}

*Means with the same letters are not significantly different. S.E=Standard Error

Figure 4.5 is the representative electrophoregram of primer Sat_192 showing pattern of allelic bands across the 36 soybean accessions on polyacrylamide gel (PAGE). The marker/primer was polymorphic with five alleles.

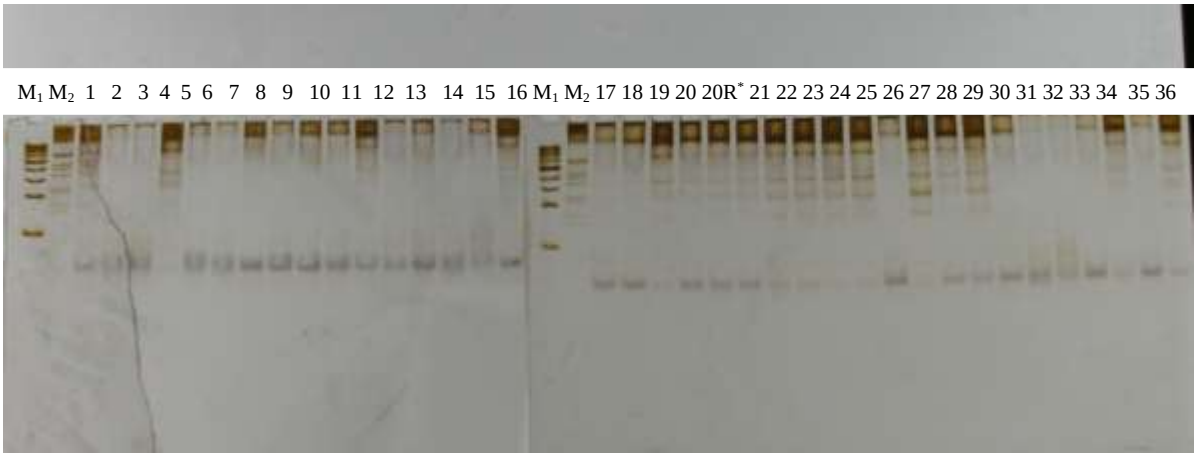


Figure 4.5: PCR Amplifications of Sat_192 using 36 soybean genotypes on PAGE. M_1=100bp ladder, M2=1kb ladder. *Lane 20R is a repetition of lane 20.

Table 4.6 Analysis of variance of SSR alleles scored

Source of variation	Df	Sum of squares	Mean squares	F-value	P > F
Soybean SSR Primers	19	61.12	3.22	15.04	< 0.001
Error	700	149.71	0.21		
Corrected Total	719	210.83			

The purpose of analysis of variance of primers was to test for statistical significance difference between means. From Table 4.6 above, there was significant difference between the primer means with F-value of 15.04 and P < 0.0001 probability (P< 0.01 i.e. 99% confidence that the means were significantly different). The null hypothesis was rejected and the alternative hypothesis was accepted that there were differences between the means. It was therefore concluded that the alleles scored for the SSR was efficient in identifying variation among the accessions.

4.3 SSR shattering resistance markers for marker-assisted selection (MAS) in soybean

Soybean pod shattering resistance markers SRM0, SRM1 and SRM2 (SRM0-2) have been found to be tightly linked to the *qPDH1* locus, a major quantitative trait locus (QTL) controlling pod dehiscence. These were used to screen the 36 soybean accessions to identify those resistant or susceptible to pod shattering. The genetic distances between the markers and the *qPDH1* are estimated to be less than 0.2cM (Suzuki *et al*., 2010).

The accessions consisted of six released varieties in Ghana, twenty lines from IITA and ten from EMBRAPA-Brazil (Table 4.7). The results showed that all the shattering markers amplified distinct allelic bands which were within the expected base pair sizes in most of the accessions. However, there were other amplifications which did

not correspond to any of the reported/expected allelic base pair size and thus were classified as outliers. There were instances where no apparent amplification was seen and it was considered as missing band as indicated in Table 4.7.

The marker SRM0 had amplified bands for 34 genotypes of the 36 accessions as shown in Figure 4.6. Two distinct allelic bands, 213 bp for resistant allele and 231bp for susceptible allele were detected. Six genotypes had susceptible alleles while 28 genotypes had resistant alleles. There were two missing bands and no outlier.

The marker SRM1 also had amplified bands for 34 of the 36 accessions, with two distinct bands; 234 bp for the resistant allele and 222 bp for the susceptible. Five genotypes had susceptible alleles and 28 genotypes had resistant ones. There were two missing bands and one outlier.

Thirty-three genotypes had amplified bands for SRM2 with two distinct bands; 185 bp for the resistant allele and 157 bp for the susceptible. Four genotypes had the susceptible allele while 26 genotypes had the resistant one. There were three missing bands and three outliers as indicated in Table 4.7.

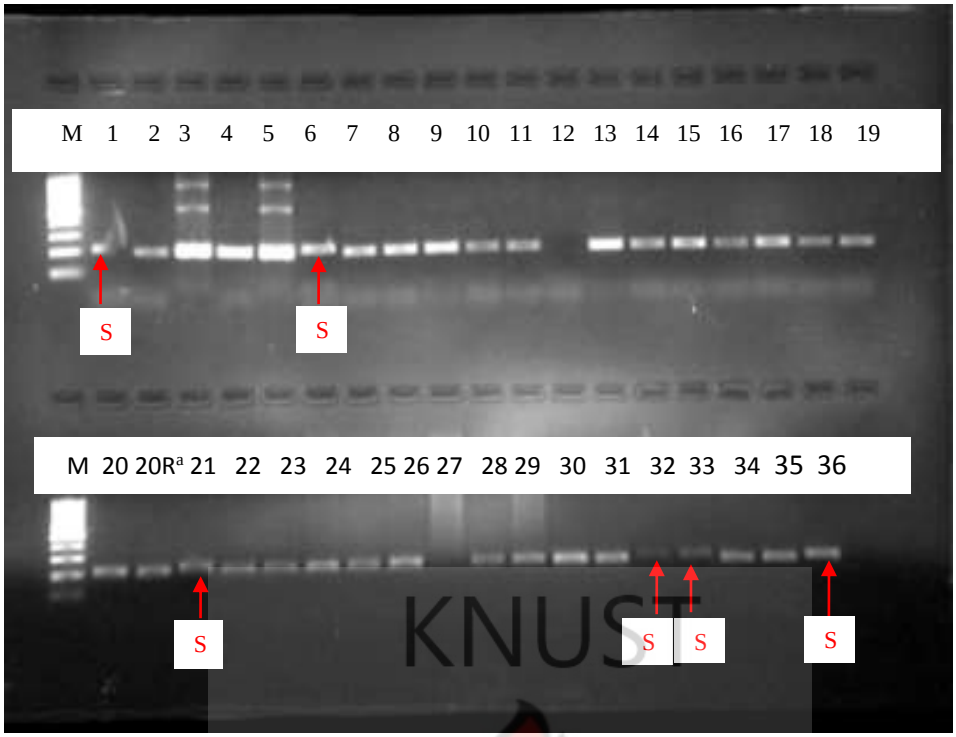


Figure 4.6: PCR Amplifications of SRM0 marker among 36 soybean genotypes on agarose gel showing the susceptible allele [(S) in red arrow] =231bp; and bands of resistant alleles (R) =213bp. Legend: M=100bp ladder. NB: [a]Repetition of lane 20

Based on the result in Table 4.7, the dendrogram in Figure 4.7 was generated. At the reference point of 17% similarity level, five clusters namely: I, II, III, IV, and V were expressed.

Cluster I comprised highly shattering-susceptible cultivars (Salintuya-1, TGX 1904-6F, Salintuya-2 and Anidaso), where almost all cultivars had the susceptible allele of all the three markers; except Anidaso which showed resistance with marker SRM1.

Cluster II had only one cultivar, TGX 1805-31F; which had susceptibility allele for marker SRM0 and SRM1, but the allele band size (less than 100bp) did not conform to any of the reported susceptible or resistant alleles of SRM2, therefore, was considered as outlier to SRM2 marker.

Cluster III comprised of two Brazilian cultivars, Boavista and Celeste that expressed two resistance alleles each to SRM1/SRM2 and SRM0/SRM2, respectively; but had susceptible alleles for SRM0 and SRM1 makers accordingly.

Cluster IV, a major cluster, comprised 28 accessions. None of them had a susceptible allele. They registered at least two resistance alleles except MG68, which had no PCR product in two accessions. Cultivars TGX 190-3F, TGX 1910-16F and Ramunda had no PCR product for markers SRM1 while cultivars TGX 1909-3F and TGX 1835-10E had no PCR product for SRM2.

Cluster V, consisted of TGX 1842-18E as the only isolated cultivar from the accessions. It had no PCR product for SRM0 and SRM2. The only band registered was an outlier with SRM1 marker.

Table 4.7: Results of scoring on 36 soybean accessions using shattering resistant markers

Cultivar/Line	Origin	[b]Lane	Marker alleles for pod shattering *qPDHI*		
			SRM0: R=213bp S=231bp	SRM1: R=234bp S=222bp	SRM2: R=185bp S=157bp
TGX1805_31F	IITA	1	S	S	Outlier
TGX1903_7F	IITA	2	R	R	R
TGX1903_8F	IITA	3	R	R	R
TGX1904_2F	IITA	4	R	R	R
TGX1904_3F	IITA	5	R	R	R
TGX1904_6F	IITA	6	S	S	S
TGX1909_3F	IITA	7	R	R	Outlier
TGX1910_16F	IITA	8	R	NP	R
TGX1835_10E	IITA	9	R	R	Outlier
TGX1903_2F	IITA	10	R	R	R
TGX1903_1F	IITA	11	R	R	R
TGX1842_18E	IITA	12	NP	Outlier	NP
TGX1843_29E	IITA	13	R	R	R
TGX1844_4E	IITA	14	R	R	R
TGX1844_18E	IITA	15	R	R	R
TGX1910_2F	IITA	16	R	R	R
TGX1910_3F	IITA	17	R	NP	R
TGX1910_6F	IITA	18	R	R	R
TGX1910_14F	IITA	19	R	R	R
GMX92_6_10	IITA	20	R	R	R
BOAVISTA	Brazil	21	S	R	R
CELESTE	Brazil	22	R	S	R
TRACAJA	Brazil	23	R	R	R
PIRARARA	Brazil	24	R	R	R
FLORA	Brazil	25	R	R	R
BRAZILLIA	Brazil	26	R	R	R
MG68	Brazil	27	NP	R	NP
RAMUNDA	Brazil	28	R	R	NP
MG/BR46	Brazil	29	R	R	R
SAMBAIBA	Brazil	30	R	R	R
JENGUMA	Ghana	31	R	R	R
SALINTUYA_1	Ghana	32	S	S	S
SALINTUYA_2	Ghana	33	S	S	S
QUARSHIE	Ghana	34	R	R	R
NANGBAAR	Ghana	35	R	R	R
ANIDASO	Ghana	36	S	R	S

R=resistance to pod shattering S=susceptible to pod shattering NP=no PCR product [b]Lane number in fig. 4.6

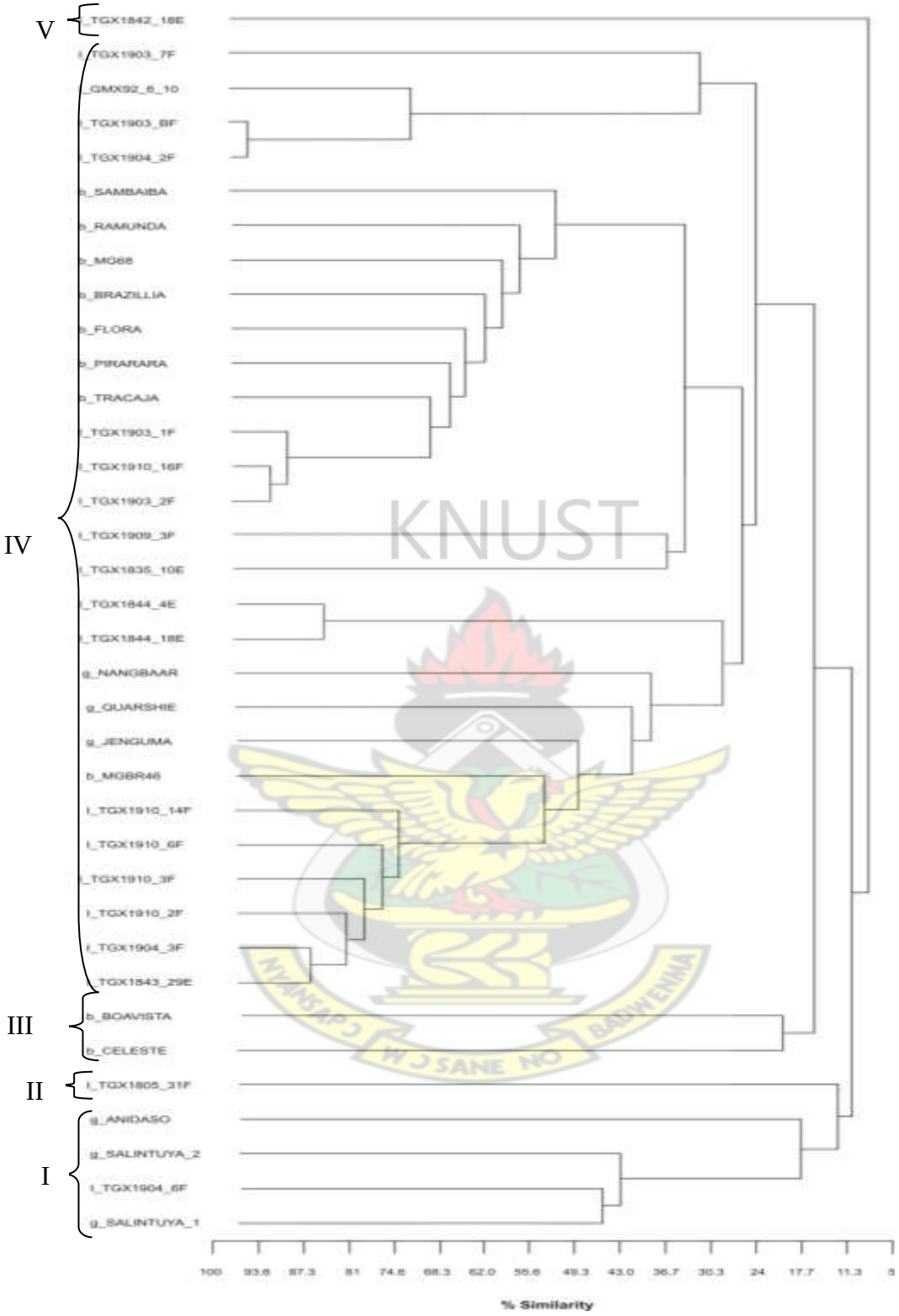


Figure 4.7: Dendrogram of 36 soybean accessions using three SSR shattering markers

The PCR amplifications of a cross between known resistant (Jenguma) and known susceptible (Salintuya-2) soybean genotypes with SRM0-2 allele markers detected at

F_1 hybrids in heterozygous state that had inherited both parental alleles as indicated in Figures 4.8 and 4.10. This invariably established the successfulness of F_1 crosses from two genotypes with contrasting alleles and that the alleles were true heritable.

On the other hand, a cross between two resistant genotypes (Jenguma X Flora and Jenguma X Tracaja) having same base pair of alleles produced F_1 progenies with the same pair of alleles as the parents ; thus, expressing homozygous state in F_1 as shown in figure 4.9.

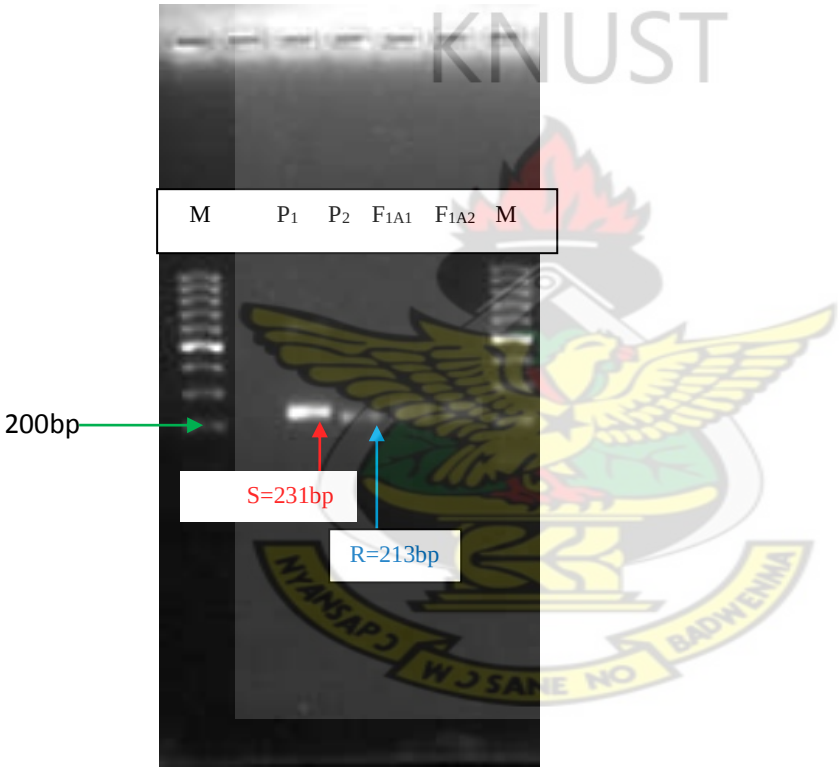


Figure 4.8: PCR Amplification of known shattering resistant and susceptible soybean genotypes and its F_1 hybrid with SRM0 marker on agarose gel. Legend: M-100bp ladder, P_1- Salintuya-2(S = susceptible genotype- 231bp in red arrow), P2- Jenguma (R = resistant genotype-213bp in blue arrow), F_{1A1}+ F_{1A2} = F_1 progenies (Salintuya-2 by Jenguma).

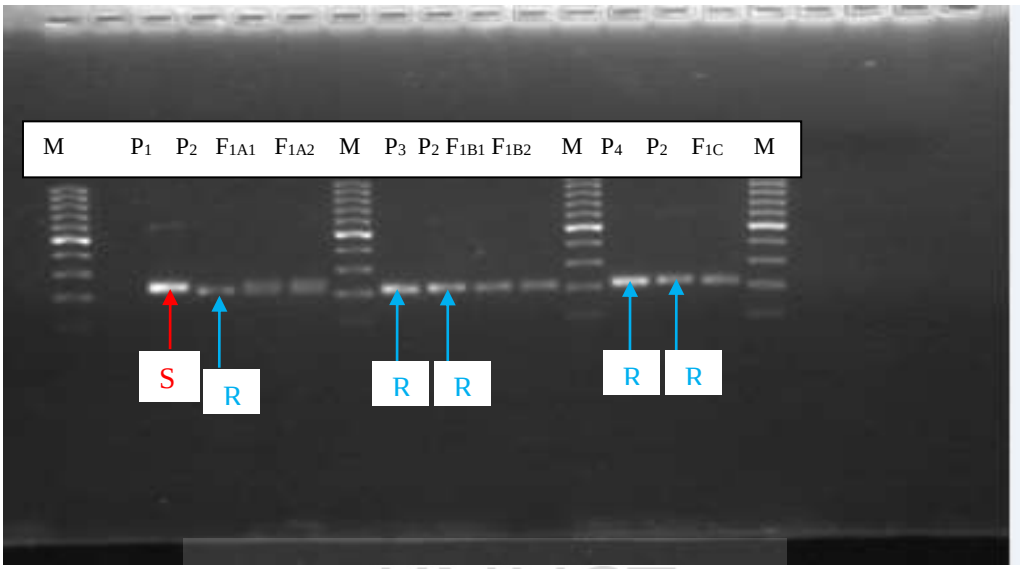


Figure 4.9: PCR Amplification using SRM0 Marker on agarose gel Legend: M-100bp ladder, P_1- Salintuya-2 (S= susceptible genotype=231bp), P_2- Jenguma (R=resistant genotype=213bp), P_3-Flora, P_4 –Tracaja (R= resistant genotypes=213bp); F_{1A}, F_{1B}, F_{1C} are the respective F_1 progenies; P_1X P_2 =F_{1A1}+ F_{1A2}= F_{1A};
P_3 X P_2 =F_{1B1}+F_{1B2}=F_{1B}; P4 X P_2= F_{1C}

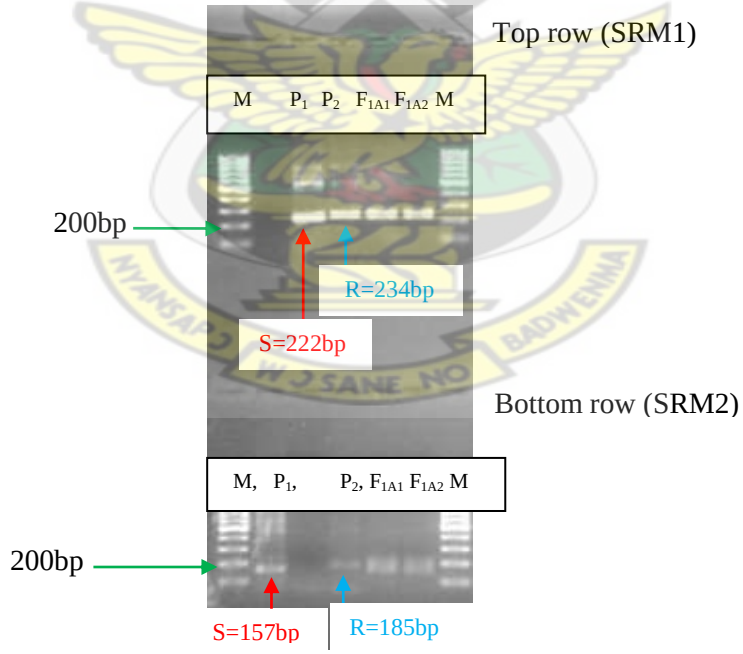


Figure 4.10: PCR Amplifications of SRM1 marker (top row) with resistant allele=234bp and susceptible allele=222bp and SRM2 marker (bottom row) with resistant allele(R) =185 and susceptible allele(S) =157 soybean genotypes on agarose gel. Legend: M-100bp ladder, P_1- Salintuya-2 (susceptible genotype in red arrow), P_2- Jenguma (resistant genotype in blue arrow) F_{1A1+} F_{1A2} =F_{1A}= F_1 progenies (Salintuya-2 X Jenguma).

CHAPTER FIVE

5.0 Discussion, Conclusions and Recommendations

5.1 DISCUSSION

The main objective of this study was to assess genetic diversity and to establish relationship among 36 soybean germplasm using both morphological and molecular data; and also to validate the usefulness of soybean shattering resistance markers for marker-assisted selection (MAS) in soybean breeding programmes. Because morphological traits that will be useful to assess diversity was not known, step-wise discriminate analysis was conducted on the quantitative trait variables to identify the order of importance in terms of phenotypic variance. Principal component analysis (PCA) was performed on the mean values of the important traits that accounted for most variation within the accessions. Again, DNA molecular analysis of the accessions was conducted using soybean SSR primers; and based on the dendrogram generated, discernible clusters of accessions were established.

The relationships among the 36 accessions based on morphological characterisation show that suitable parental lines could be recommended and used for hybridisation. Analysis of plant morphology based on quantitative and qualitative traits revealed three and seven clusters, respectively. The quantitative trait revealed less variation as percent similarity among accessions range between 71-99 % compared to the qualitative trait which had percent similarity of between 25.3-95 %. Chowdhury *et al*. (2001) reported that analysis of eight morphological traits (qualitative) resulted in differentiation of 47 soybean varieties into four main clusters at about 22 distance scale level.

Among the 18 quantitative trait variables, seven were identified as important from stepwise-discriminant analysis (Table 4.1). This decision was based on partial r-squared values and the F-values. R-square measures the adequacy or strength of a variable. The higher the R-square the better the regression and is expressed as a percentage or fraction. The higher the R-square value of a variable the better the discriminating power and greater the chances to be loaded onto principal component axis (SAS online). The R-square value for seed thickness (0.322) was similar to that of days to maturity (0.321). A higher F-value of significance for days to maturity (7.31) makes it a preferred trait to seed thickness (7.13) in statistical discriminant analysis. Plant height had the highest partial r-square value of 0.880 followed by seeds per pod with 0.484; days to 50 % flowering had the least (0.158).

Cumulative average squared canonical correlation (CASCC) measures the contribution of each variable plus the preceding variables to the discriminant analysis (SAS online). As shown in Table 4.3, all the seven trait variables contributed approximately 84.3 % (CASCC = 0.843) phenotypic variation of the soybean accessions. The three most important traits from discriminant analysis were plant height, seed per pod and days to maturity. These three traits were used for the principal component analysis.

From the graph of PCA biplot, IITA lines and Ghanaian cultivars showed contrasting association with respect to days to maturity.

The cultivars from Ghana showed positive association with days to maturity suggesting cultivars require longer days to mature. The IITA lines, however, showed negative association with days to maturity. The IITA lines were early-maturing compared to those from Ghana and Brazil.

In determining the relationships among the traits, simple linear correlation coefficient and principal component analysis were employed. Correlation analysis results alone could not give a complete picture of interrelationships because it considers only two traits at a time, irrespective of interrelationship with other traits (SAS online). PCA on the other hand considers the underlying interrelationships and selects the best linear combination of traits that explains the largest proportion of variance. The three most important traits were plant height, number of seeds per pod and days to maturity. These traits were able to discriminate the soybean genotypes from the three countries.

Soybean cultivars from Ghana were distinguished from Brazilian and IITA lines on the fact that they require more days to mature. Within the coordinate of "days to maturity" was an isolated cultivar "g03" (Salintuya-2) and is a known late maturing variety that requires 120 days to mature; similar to what was reported by Mohammed, (2010). Brazilian lines were identified by pods with higher number of seeds; they poduced three-seeded pods and were relatively taller in height. The IITA lines were found at the negative coordinate to these three trait variables; which showed a direct contrast to their Ghanaian and Brazilian genotypes in their morphology.

The first two principal component axes accounted for 89 % of the phenotypic variation among the 36 soybean accessions studied. This compares well with Mardia *et al*. (1979), who indicated that the total variance accumulated by principal components close to 80 % explained satisfactorily the variability manifested among individuals. Thus, the 89 % phenotypic variation achieved could explain the genetic diversity of soybean accessions. Also the PCA results in assessing diversity were consistent with the results obtained by Bhartiya *et al*. (2011), where they used PCA to determine the variability of both indigenous and exotic black soybean from different eco-geographic regions of the world. The results also conformed to earlier work by

Mebatsion *et al.* (2012). They evaluated grain shape variability using Principal component analysis (PCA); where 99 % of the variation in the shape of grains was captured by the first two principal components.

Analysis of variance of the 20 SSR primers used showed a high level of significant difference with F-value of 15.04 ($P < 0.0001$). It was concluded that the primers were efficient in identifying variation among the accessions. Genetic diversity analysis also revealed that the 36 soybean genotypes could be clustered into six major groups. This result was consistent with the results reported by Tantasawat *et al*. (2011); where four major clusters were revealed when they used 11 SSR markers to assess genetic relationships among 25 soybean genotypes. Similarly, the results obtained were in agreement with the one reported by Singh *et al*. (2010) who classified 44 soybean genotypes into four clusters based on coefficient of similarity derived from 120 SSR markers.

The results of molecular diversity of the accessions revealed that, the IITA lines appeared to have been bred from a common or closely related ancestry. Such a narrow genetic base needs to be expanded with exotic introductions from distant areas such as Brazil and China. The variety 'Jenguma' which was about 90 % similar with 'Sambaiba' at the molecular level may have been an introduction from Brazil, via IITA, that has been given a local name in Ghana. Similarly, 'Anidaso' and 'Nangbaar' could have been derived from common ancestry.

From the dendrogram (Fig. 4.7) the genetic diversity of pod shattering resistance in soybean conforms to earlier work done by Caviness (1965), Tsuchiya (1986), Helms (1994), Romkaew and Umezaki (2006) and Yamada (2009), who reported

considerable genetic diversity among soybean cultivars with regard to shattering resistance.

Asafo-Adjei, (2005; 2007a, b) scored Anidaso as resistant to pod shattering. However, from the current study at molecular level, it had expressed two alleles for susceptibility to SRM0 and SRM2. It was therefore inclined more onto susceptibility than being resistant to pod shattering. The field conditions under which Asafo-Adjei (2005; 2007), assessed Anidaso may have influenced his rating (evaluation), as temperature and relative humidity do affect pod dehiscence. The current assessment with DNA markers is therefore more credible.

The genotypes Boavista and Celeste all from Brazil also expressed one allele with SRM0 and SRM1, respectively.

Three soybean genotypes: Salintuya-1, Salintuya-2 (from Ghana) and TGX 1904_6F (from IITA) also expressed all the three susceptible alleles for the SRM0-2 markers and were considered highly susceptible to pod shattering. These results confirm a genetic analysis on pod shattering studies by Mohammed (2010) that Salintuya-2 and Salintuya-1 were very susceptible and moderately susceptible to pod shattering respectively.

Genotype TGX1842-18E (IITA line) was isolated from all other clusters, and hence the current findings could not determine its pod dehiscence status with the shattering markers SRM0, SRM1 and SRM2; and may have a different genome. The current study with SSR molecular diversity (Figure 4.4) with these soybean accessions also isolated TGX 1842-18E from the IITA clusters of soybean accession.

The order of closeness of the three resistance shattering markers to the target locus (*qPDH1*) differed among them (Personal communication, Funatsuki, H.). Marker SRM0 the closest showed no outlier with the 36 accessions whilst markers SRM1 and SRM2 followed in that order of closeness recorded one and three outliers, respectively. This seems to suggest the farther away the marker from the target locus the more it registers outlier bands.

The polymorphic nature of these SSR shattering resistance markers (SRM0-2) had helped to distinguish the known susceptible pod shattering cultivar (Salintuya-2) from known resistant pod shattering cultivar (Jenguma). The markers were able to detect F_1 hybrids from a cross between these two cultivars as shown in Figures 4.8 and 4.10.

The stable effect of *qPDH1* allele had been demonstrated in progeny derived from various cross combinations between shattering-resistant and shattering-susceptible cultivars as reported by Funatsuki *et al*. (2006, 2008). The results are in agreement with Hwang *et al*. (2008) who also have reported that the alleles for shattering susceptibility at SRM0-2 are not specific to small population of Japanese origin, but can be used as universal markers to assess shattering resistance and susceptibility in soybean worldwide. Both shattering-resistant and -susceptible cultivars possess resistant alleles and susceptible alleles, respectively, at *qPDH1* and exhibit the same genotype at the marker locus.

5.2 Conclusion and recommendations

The 20 SSR soybean primers were able to detect genetic diversity among the soybean germplasm from the three countries (Ghana, Brazil and Nigeria).

The use of principal component analysis (PCA) revealed that days to maturity, seeds per pod and plant height were important traits for estimating the relationship among

accessions from the respective countries. The PCA also showed a clear picture for selecting suitable parental lines for crosses.

The results clearly confirmed that SRM0, SRM1 and SRM2 markers could be very useful in marker-assisted selection (MAS) for selecting shattering resistant soybean genotypes. The heritability of these markers was demonstrated in the F_1 hybrids obtained.

However, it is recommended that a larger number of soybean genotypes should be tested to find out if these markers can detect shattering resistance. This can lead to the detection of novel alleles also conferring resistance to shattering in soybean.

It is also recommended that F_2 seeds from Salintuya-2 X Jenguma should be used for further genetic analysis to validate the inheritance of shattering alleles. Further work can also be done in future by assessing correlation between SRM0 genotypes and degree of shattering resistance in the F_2 segregation population.

REFFERENCES

Abe, J., Xu, D. H., Susuki Y, Kanazawa, A. and Shimomoto, Y., (2003). Soybean germplasm pools in Asia revealed by SSRs. *Theoretical and Applied Genetics* 106: 445-453.

Acquaah, G. (2007). Principle of Plant Genetics and Breeding. Blackwell Publishing Ltd., U.S.A. (Pages 520-21).

Akkaya, M.S., Bhagwat A.A. and Cregan P. B., (1992). Length polymorphism of simple sequence repeat DNA in soybean. *Genetics* 132: 1131- 1139.

Alabi, M. O., Berner, D. K. and Olaniyan, G. O., (1994). Characterisation of soybean cultivars for *Striga hermothica* control. *Phytopathol*.84:1151

Asafo-Adjei, B. and Adekunle, A. A., (2001) Characteristics of some released IITA soybean varieties and promising advanced breeding lines. International Institute of Tropical Agriculture. Ibadan, Nigeria.

Asafo-Adjei, B., Adu-Dapaah, H. and Oti-Boateng, C., (2007a). *CSIR-CRI/CSIR-SARI/MOFA FCDP Fact Sheet*: Nangbaar and Ahoto early Maturing Soybean varieties for Ghana.

Asafo-Adjei, B., Ansah,I. O. O., Asuboah, R.A., Dapaah,H., Haruna,M. and Oti-Boateng, C. (2005). *Soybean Production Guide*. CSIR-CRI, Kumasi, Ghana.

Asafo-Adjei, B., Denwar, N. and Oti-Boateng, C., (2007b). *CSIR-CRI/CSIR-SARI/MOFA FCDP Fact Sheet*: Quarshie and Jenguma: Shattering Resistant Soybean varieties for Northern Ghana.

Asalm, M., Mirza, M. S., Shah, S.M., Shafeeq, S., Naeem Ullah (1995). Crop Production Bulletin No.6, Pak. Agric. Res. Council, Islamabad, Pakistan. Pg.1

Bailey, M. A., Mian, M. A. R., Carter Jr. T. E., Ashley D. A and Boerma H. R., (1997): Pod deshiscence of soybean: identification of quantitative trait loci. *J. Hered.* 88**,** 152-154.

Barr, R. A. (2009). Marker-assisted selection in theory and practice In: Plant Breeding and Farmer Participation. Edited by: Ceccarelli, S., Gumaraes, E. P., and Weltzien

Berner, D. K., Kling, J. G. and Singh B. B. (1994) *Striga* research and control: A perspective from Africa. *Plant Disease* 79: 652-60.

Bhartiya Anuradha, Aditya, J. P., Singh, G., Gupta, A., Agarwal, P. K. and Bhat, J. C. (2011). Assessment of genetic diversity in indigenous and exotic collection of black soybean(Glycine max (L.) Merrill) *SABRAO Journal of Breeding and Genetics* 43: 81-90.

Brown-Guedira, G. L., Thoson, J. A., R. L. Nelson, R.L., and M. L., Warburton. (2000). Evaluation of genetic diversity of soybean introductions and North American ancestors using RAPD and SSR markers. *Crop Science* 40**:**815-823.

Calson, J. B. (1973): Morphology. *In Soybeans*; Improvement, production and uses. Ed. Caldwell, B. E. Amer. Soc. Agron. Madison, Winsconsin, 19-95.

Carroll, K. K., Giovannetti, P. M., Huff, M. W., Moase, O., Roberts, D.C. and Wolf, B. M. (1978). Hypocholesterolemic effect of substituting soybean protein for animal protein in the diet of healthy young women. *The American Journal of Clinical Nutrition* Vol. 31:1312-1321.

Caviness, C. E., (1965). Effects of relative humidity on pod dehiscence in soybeans. *Crop Sci*. **5:** 511-513.

Chowdhury, A.K., Srinives, P., Tongpamnak, P. and Saksoong, P. (2001). Genetic Diversity Based on Morpholgy and RAPD Analysis in Vegetable Soybean. *Korean J. Crop Sci*. 46(2): 112-120.

Dayaman, V., Senthil, N., Raveendraan, M. Nadarajan, N., Shanmugasundaram, P., Balasubramanian, P. (2009). Diversity analysis in selected Indian soybean [*Glycine max* (L.) Merrill] using morphological and SSR markers. *International Journal of Integrative Biology* Vol. 5: 125-129.

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

Denwar, N. N. and Ofori, K. (2003). Variation of trap-crop capacity of soybean Genotypes for the control of *Striga hermonthica*. *International Journal of Agriculture and Biology*. Vol. 5: 504-506

Diwan, N. and Cregan, P. B. (1997). Automated sizing of fluorescent labeled simple sequence repeat (SSR) markers to assay genetic variation in soybean. *Theoretical and Applied Genetics* 95**:**723-733.

Doldi ML, Vollmann J and Lelley T. (1997). Genetic diversity in soybean as determined by RAPD and microsatellite analysis. *Plant Breeding* 116**:**331-335.

Embrapa (2003). *Technologias de produÇão de soja-região central do Brasil* 2004. Embrapa, Epamig e FundaÇao Triangulo, Londrina, 237p.

Feng, C., Hou, A., Chen, P., Cornelious, B., Shi, A., and Zhang, B. (2008). Genetic Diversity Among Popular Historical Southern U.S. soybean cultivars using AFLP markers. *Journal of Crop Improvement* Vol.22 (1) 31-46.

Food and Agriculture Organisation of the United Nations. FAO Statistics Division, FAOSTAT database (2012).

Funatsuki, H., Ishimoto, M.,Tsuji H., Kawaguchi, K., Hajika, M., and Fujino, N. (2006): Simple sequence repeat markers linked to a major QTL controlling pod shattering in soybean. *Plant Breeding* 125, 195-197.

Funatsuki, H., Hajika, M., Hagihara, S., Yamada, T., Tanaka, Y., Tsuji, H., Ishimoto, M. and Fujino, K. (2008). Confirmation of the location and the effects of a major QTL controlling pod dehiscence, *qPDH1*, in soybean. *Breeding Science* 58: 63-69.

Griffin, J. D., and Palmer, R. G. (1995). Variability of thirteen isozyme loci in the USDA soybean germplasm collections. *Crop Sci*. 35: 897-907.

Helms, T. C., (1994). Greenhouse and field evaluation of pod dehiscence in soybean. *Can. J. Plant Sci*. 74**:** 699-701

International Institute of Tropical Agriculture (IITA), (1986). A laboratory method of evaluating resistance to pod shattering on soybean. *Annual Report*.58-59. IITA, Ibadan-Nigeria.

Jarne, P. and Lagoda, P. J. L., (1996). Microsatellites, from molecules to populations and back. *Trends in Ecol. Evol* 11:424-429.

Jiang, J. L., Thseng, F.S., and Yeh M. S., (1991): Studies on the pod shattering in soybean (in Chinese with English summary). *J. Agric Assoc. China* 156**,** 15-23

Juvik, G. A., Bernard, R. L. and Kauffman, H.E. (1985). *Directory of Germplasm Collections*. II. Food Legumes (Soybean). International Board for Plant Genetic Resources. Rome, Italy.

Karp, A. Edward, K., Bruford, M. Vosman, B., Morgante, M., Serberg, O., Kremer, A. Boursot, P., Arctander, P., Tautz, D. and Hewitt, G., (1997). New molecular technologies for biodiversity evaluation: opportunites and challenges. *Nature Biotechnol*. 15: 625-628.

Kroschel, J. and Sauerborn, J., (1988). Training Manual for course on Biology and control of Parasitic Weeds. University of Hohenheim, Germany.

Lawson, I.Y.D., Mensah, E.A. and Yeboah, E. N. (2008). Improving the establishment and yield of soybean through planting depth and land Preparation Methods in northern Ghana. *West Africa Journal of Applied Ecolgy*. Vol. 14:1-8

Mardia, K. V., Kent, J. T. and Bibby, J. M. (1979). Multivariate analysis. Academic Press, London.

MAPA (2005). Ministério da Agricultura, Pecuária e Abastecimento. Available at http://www.agricultura.gov.br/pls/portal/docs/PAGE/MAPA/BALANCE-COMERCIAL/NOTA%20DEZEMBRO%202004. PDF. Assessed in 26 Agust 2005.

Maughan. P.J., A. A. Saghai Maroof, and G. LR. Buss, (1995). Microsatellite and amplified sequence length polymorphism in cultivated and wild soybean. *Genome* 38**:**715-723.

Mebatsion, H. K., Paliwal, J., Jayas, D. S. (2012). Evaluation of variation in the shape of grain types using principal component analysis of the elliptic Fourier descriptors. *Computers and Electronics in Agriculture* 80: 63-70.

Mebrahtu T. (2008). Analysis of Nutritional Contents in Vegetable Soybeans. *Journal of Crop Improvement*, Vol.21 (2): 157-170.

Messina J. M. (1999). Legumes and soybeans: Overview of their nutritional profiles and health effects. *The American Journal of Clinical Nutrition*. 70 (Suppl): 439S-50S.

Ministry of Food and Agriculture (MoFA), (2002). Food Crops Development Project: Survey Report on Soybean Processing Industries in Ghana.

Ministry of Food and Agriculture (MoFA), (2003a). Towards Integrated Soybean Industry in Ghana.

Ministry of Food and Agriculture (MoFA), (2003b) Food Crops Dev. Project: Report on National Soybean Workshop. CSIR-CRI, Fumesua-Kumasi.

Ministry of Food and Agriculture (MoFA), (2006). Food and Agriculture Sector Development Policy (FASDEP).

Ministry of Food and Agriculture (MoFA), (2011). Statistics Research and Information Directorate (SRID), May, 2011. *AGRICULTURE IN GHANA FACTS AND FIGURES (2010)*.

Mohammed H. (2010). Genetic Analysis of Resistance to pod Shattering in soybean (*Glycine max*. (L.) Merrill). MSc. Dissertation, KNUST, Kumasi.

Morgante M., Hanafey H, Powell W. (2002). Microsatellites are preferentially associated with nonrepetitive DNA in plant genome, *Nat. Genet*. 30**:**194-200.

Morgante, M., and Olivieri, A. M., (1994). Genetic Mapping and variability of seven soybean simple sequence repeat loci. *Genome* 37: 763-769.

Newell, C. A. and Hymowitz, T. (1980). A taxonomic revision of the genus *Glycine* subgenus *Glycine* (Leguminosae). *Brittoni*a 32**:** 63-69.

Nti, C. A., Plahar, W. A., Ellis, W. O., Wiad Ashitey, E. (2005). Soybean Recipes in Ghana. Food Crops Development Project, Ministry of Food and Agriculture.

Ojo, D. K., Ajayi, A. O. and Oduwaye, O. A. (2012). Genetic relationship among soybean accessions based on morphological and RAPDs Techniques. *Pertanika J. Tropical Agric Sc*. 35 (2) 237-248.

Palmer, R. G., Hymowitz and Nelson, R. L., (1996). Germplasm Diversity within Soybean. In: *Soybean, Genetics, Molecular Biology and Biotechnology*. Verma, D.P.S. and Shoemaker, R. C. (eds.)

Parkes, E. Y., Allotey, D. F.K., Lotsu, E. and Akuffo, E. A. (2012). Yield performance of five cassava genotypes under different fertilizer rates. International Journal of Agricultural Sciences Vol. 2 (5): 173-177.

Philbrook, B. and Oplinger, E. S. (1989). Soybean field losses as influenced by harvest delays. *Argon. J*. 81 : 251- 258.

Prioli, R. H. G., Mendes, Jr. C. T., Arantes, N. E. and Contel, E. P. B., (2002). Characterisation of Brazilian soybean cultivars using microsatellite markers. *Genetics Molecular Biology* 25: 185-193.

QIAGEN sciences, Maryland, 20874, U.S.A

Romkaew, J. and Umezaki, T., (2006). Pod dehiscence in soybean: assessing methods and varietal difference. *Plant Prod. Sci*. 9**:** 373-382.

Rongwen, J., M.S. Akkaya, A. A. Ghahwat, U. Lavi, and P. B. Cregan .(1995). The use of microsatellite DNA markers for soybean genotype identification. *Theor. Appl. Genet*. 90**:** 43-48.

SAS Institute (2008) SAS/STAT 9.2 user's guide. SAS Institute, Cary, NC.

SAS online support publication: http: // support .sas .com /publishing /pubcat /chaps/ 55129.pdf

Shadakshari, T. V., Kalaimagal, T., Senthil, N., Boranayaka, M. B., Kambe Gowda, R., and Rajesha, G. (2011). Genetic diversity studies in soybean [*Glycine max* (L.) Merrill] based on morphological characters. *Asian Journal of Bio Science* Vol. 6 Issue 1: 7-11.

Shirota, M., Hamada, Y., Nakajima, Y., Tani, T., Yoshida, T., Shaku, I., and Uchida, T., (2001). Study on actual condition of pre-harvest shatter loss on soybean cultivation in Aichi Prefecture. *Jpn. J. Crop Sci*. 70 (Suppl. 2): 23-24.

Singh K. Ram, Bhatia S. Virenda, Bhat K.V., Mohapatra Trilochan, Singh K. Nagendra, Bansal C. Kailash and Koundal K. R. (2010). SSR and AFLP based genetic diversity of soybean germplasm differing in photoperiod sensitivity. *GENETICS and MOLECULAR BIOLOGY* April-June 33 (2), 319-324.

Singh, R. J. and Hymowitz, T. (1985). The genomic relationships among six wild perennial species of the genus *Glycine* subgenus, *Glycine* Willd. *Theoretical and Applied Genetics* 71, 221-230.

Singh, R. J., Kollipara, K. P. and Hymowitz, T. (1988) Further data on the genomic relationships among wild perennial species (2n=40) of the genus Glycine Willd. *Genome* 30, 166-176.

Singh, R. J., Kollipara, K. P., Ahmad, F, and Hymowitz, T. (1992b). Putative diploid ancestors of 80-chromosome *Glycine tobacina* (Labill). Benth. *Genome* 35, 140-146.

Singh, R. J., Kollipara, K. P. and Hymowitz, T. (1992a). Genomic relationships among diploid wild perennial species of the genus Glycine Willd. Subgenus Glycine revealed by cytogenetics and seed protein electrophoresis. *Theoretical and Applied Genetics* 85, 276-282.

Singh, R. J., Kollipara, K.P. and Hymowitz, T. (1998). Ancestors of 80- and 78-chromosome *Glycine tomentella* Hayata intersubgeneric hybrids. *Genome* 32, 796-801.

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

Song, Q.J., Marek L. F., Shoemarker K. G., Lark, Concibido V. C., Delannay, X., Specht J. E., Cregan P. B., 2004: A new integrated genetic linkage map of the soybean. *Theor. Appl. Genet*. 109, 122-128.

Statistical Analysis System (SAS), 2007. The SAS system for Windows, version 9.2 SAS Institute Inc, Cary, NC, USA.

Stephenson, H. Frank, Calculations in Molecular Biology and Biotechnology*: A guide to Mathematics in the Laboratory*, 2003. Academic Press, London - UK.

Suzuki, M., Fujino, K., Nakamoto, Y., Ishimoto, M., Funatsuki, H. (2010). Fine Mapping and development of DNA markers for the *qPDH1* locus associated with pod dehiscence in soybean. *Mol. Breeding* 25: 407- 418.

Tantasawat, P., Trongchuen, J., Prajongjai, T., Jenweerawat, S., and Chaowiset, W., (2011). SSR analysis of soybean (*Glycine max* (L.) Merr.) genetic relationships and variety identification in Thailand. *Australian Journal of Crops Science* 5 (3) 283-290

Tatineni, V., Cantrell, R. G. and Davis, D. D., (1996). Genetic Diversity in elite cotton germplasm determined by morphological characteristics and RAPDs. *Crop Science*. 36: 186-192.

Tefera H., Kamara Y. A. Asafo-Adjei B., Dashiell E. K., (2010). Breeding progress for grain yield and associated traits in medium and late maturing promiscuous soybean in Nigeria. *Euphytica* 175: 251-260.

Tindale, M. D. (1984). Two new eastern Australian species of *Glycine* Willd. (*Fabaceae*). *Brunonia* 7: 207-213.

Tindale, M.D. (1986a). A new North Queensland species of *Glycine* Willd. (*Fabaceae*). *Brunonia* 9: 99-103.

Tindale, M.D. (1986b). Taxonomic notes on three Australian and Norfolk Island species of *Glycine* Willd. (Fabaceae, Phaseolae) including the choice of a neotype for *G. clandestine* Wendl. *Brunonia* 9: 179-191.

Tindale, M.D. and Craven. L.A. (1988). Three new species of *Glycine* (Fabaceae, Phaseolae) from North-Western Australia, with notes on the amphicarpy in the genus. *Australian Systematic Botany* 1:399-410.

Tindale, M.D. and Craven L.A. (1993) *Glycine pindanica* (Fabaceae, Phaseolae), a new species from West Kimberley, Western Australia. *Australian Systematic Botany* 6: 371-376.

Tsuchiya, T., (1986): Studies on shattering resistance in soybean breeding (in Japanese with English summary). *Rep. Hokkaido Prefect. Agric Exp. Stn*. 58: 1-53.

Tsuchiya, T., (1987): Physiological and genetic analysis of pod shattering in soybeans. *Jpn. Agrc. Res. Q*. 21**:** 166-175

Tukamuhabwa, P., Rubaihayo P., and Dashiell, K. E. (2002): Genetic components of pod shattering in soybean. *Euphtica* 125: 29-34.

USDA, ARS, National Genetic Resources Program (11 June 2010). *Germplasm Resources Information Network - (GRIN)*. [Online Database] National Germplasm Resources Laboratory, Beltsville, Maryland. Available:http://www.ars-grin.gov/cgi-bin/npgs/html/makeform.pl

Walsh, S. P., Fildes, J. N. and Reynolds, R. (1996). Sequence analysis and characterization of stutter products at the tetranucleotide repeats locus vWA. *Nucleic Acid Research* Vol. 24. 2807-2812.

Yamada, T., Funatsuki, H., Hagihara, S., Fujita, S., Tanaka, Y., Tsuji, H., Ishimoto, M., Fujino, K., and Hajika, M. (2009). A major QTL, *qPDH1*, is commonly involved in shattering resistance of soybean cultivars. *Breeding Science* 59: 435-440.

APPENDICES

Appendix 1: PCR Reagents (*Biolabs*) for Soybean SSR Amplification

PCR component	1X Reaction Vol.
Nuclease free water	6.15μl
10X buffer	1μl
25mM $MgCl_2$	0.9μl
10mM dNTPs	0.4μl
5 μM Primer (F/R)	0.5μl
5 U/μl *Taq* Polymerase	0.05μl
30ng/μl genomic DNA	1μl

Appendix 2: Thermocycling profile for amplification of soybean SSRs in diversity studies

95^{0}C	2 min.	Initial Denaturation	
92^{0}C	1min.	Denaturation	} 33 cycles
47^{0}C	1min.	Annealing	
72^{0}C	1min.	Extension	
72^{0}C	10min.	Final extension	

Appendix 3: Thermocycling profile for amplification of soybean SSRs in pod shattering studies

95^{0}C	2 min.	Initial Denaturation	
92^{0}C	1min.	Denaturation	} 33 cycles
55^{0}C	1min.	Annealing	
72^{0}C	1min.	Extension	
72^{0}C	10min.	Final extension	

Appendix 4: Composition of developer solution

1. Na_2CO_3 1.0g
2. 37% Formaldehyde 1.5ml

3. Stock solution (sodium thiosulphate 10mg/ml) 0.25ml
4. De-ionized autoclaved water 1000ml

Appendix 5: Silver Staining solution

1. Silver nitrate 1.0g
2. 37% formaldehyde 1.5g
3. De-ionized autoclaved water 1000ml

Appendix 6: Glacial Acetic Acid (1.5%) (Fixative and stopper solution)

Glacial acetic acid 15ml

De-ionized water 985ml

(These staining solutions were stored in refrigerator at 4^0C and used under chilled condition

Appendix 7: Composition of developer solution

1. Na_2CO_3 1.0g
2. 37% Formaldehyde 1.5ml
3. Stock solution (sodium thiosulphate 10mg/ml) 0.25ml
4. De-ionized autoclaved water 1000ml

Appendix 8: Preparation of 6% polyacrylamide gel (PAGE)

6% polyacrylamide non-denaturing gel. This was prepared using the following reagents in concentrations indicated below:

40% acrylamide Bis – 19:1	4.5ml
10x TBE	3.0ml
10% APS	300μl
TEMED	30μl
H_20 (sterilized)	22.5ml

Appendix 9: Qualitative traits

Seed shape			Cumulative	
Seed shape	Frequency	Percent	Frequency	Percent
Near Round	2	5.56	2	5.56
Oblong	21	58.33	23	63.89
Oval	13	36.11	36	100.00

Helium Colour			Cumulative	
Helium colour	Frequency	Percent	Frequency	Percent
Black	1	2.78	1	2.78
Brown	7	19.44	8	22.22
Brown with black outer ring	10	27.78	18	50.00
Dark brown	7	19.44	25	69.44
Light brown	11	30.56	36	100.00

Seed coat colour			Cumulative	
Seed coat colour	Frequency	Percent	Frequency	Percent
Brown	4	11.11	4	11.11
Greenish brown	8	22.22	12	33.33
Light brown	21	58.33	33	91.67
Light gray	3	8.33	36	100.00

Flower colour			Cumulative	
Flower colour	Frequency	Percent	Frequency	Percent
Purple	29	80.56	29	80.56
white	7	19.44	36	100.00

Pod colour			**Cumulative**	
Pod colour	Frequency	Percent	Frequency	Percent
Brown	5	13.89	5	13.89
Dark Brown	2	5.56	7	19.44
Light Brown	29	80.56	36	100.00

Maturity group			**Cumulative**	
Maturity group	Frequency	Percent	Frequency	Percent
Early	7	19.44	7	19.44
Late	1	2.78	8	22.22
Medium	28	77.78	36	100 .00

Pubescence colour			**Cumulative**	
Pubescence colour	Frequency	Percent	Frequency	Percent
Gray	7	19.44	7	19.44
Light tawny	5	13.89	12	3.33
Near Gray	19	52.78	31	86.11
Tawny	5	13.89	36	100.00

Pubescence density			**Cumulative**	
Pubescence density	Frequency	Percent	Frequency	Percent
Dense	1	2.78	1	2.78
Normal density	28	77.78	29	80.56
Semi dense	5	13.89	34	94.44
Sparsely dense	2	5.56	36	100.00

Pubescence form			**Cumulative**	
Pubescence density	Frequency	Percent	Frequency	Percent
Erect	1	2.78	1	2.78
Irregular	31	86.11	32	88.89
Semi appressed	4	11.11	36	100.00

Appendix 10: Key to Soybean accessions from the three countries used in PCA biplot

Table 4.4 Key to Soybean accessions from the three countries (Prefix: i = IITA-Nigeria; b = EMBRAPA-Brazil; g = Ghana) used in principal component analysis (biplot) in figure 4.3 below.

Cultivar / Line	Code
TGX1805_31F	i01
TGX1903_7F	i02
TGX1903_8F	i03
TGX1904_2F	i04
TGX1904_3F	i05
TGX1904_6F	i06
TGX1909_3F	i07
TGX1910_16F	i08
TGX1835_10E	i09
TGX1903_2F	i10
TGX1903_1F	i11
TGX1842_18E	i12
TGX1843_29E	i13
TGX1844_4E	i14
TGX1844_18E	i15
TGX1910_2F	i16
TGX1910_3F	i17
TGX1910_6F	i18
TGX1910_14F	i19
GMX92_6_10	i20
Boavista	b01
Celeste	b02
Tracaja	b03
Pirarara	b04
Flora	b05
Brazillia	b06
MG68	b07
Ramunda	b08
MG/BR46	b09
Sambaiba	b10
Jenguma	g01
Salintuya_1	g02
Salintuya_2	g03
Quarshie	g04
Nangbaar	g05
Anidaso	g06

Appendix 11: List of 20 soybean SSR primers with their respective sequences and core motifs

Primer	Sequence	5'——————————— 3'	Core motif
Sat_119	Forward primer	TAG GCT TTC AAT TTG CAG AAC T	$(AT)_{25}$
	Reverse primer	GTT AGG TGT CCC AAG CAA CTT A	
Sat_143	Forward primer	GAA GAT TGG GTA GAT ACT TCA ACA C	$(AT)_{13}$
	Reverse primer	GGA TGG ATG GTC CAT TGA TTC TTT	
Sat_147	Forward primer	GTG CGA CGT CAT GCC TTA CTC AAT	$(AT)_{12}$
	Reverse primer	GCG CTC CGT ACA CTT AAA AAA GAA	
Sat_150	Forward Primer	GCG CAC ATG CTC ACC AAG CAA AGT AT	$(AT)_{24}$
	Reverse Primer	GCG GTA GAG CGG ATT AAA CTT GTC	
Sat_151	Forward primer	GCT GCA TCA GAT CAC CCA TCC TTC	$(AT)_{13}$
	Reverse primer	CAT GCC ATG TTG TAT GTA TGT	
Sat_172	Forward primer	GCG TTC TAA TTT CCT GAC ACT GTT	$(AT)_{18}$
	Reverse primer	GCG GGA CGT AAA CGG ATA ATA AGG T	
Sat_192	Forward primer	GCG GAA TGG CAA TAG TTG ATG AGT A	$(AT)_{13}$
	Reverse primer	GCG GGA TGG GAT ATG AGA GTA AG	
Sat_219	Forward primer	GCG TCA TGC CAC GTG ATA TTT TAT	$(AT)_{26}$
	Reverse primer	GCG TGT GTC CCA AAT GTG ATT CA	
Sat_229	Forward primer	GCG TGT GCT ACT TCA CAT CTT GAG AGA AAG A	$(AT)_{21}$
	Reverse primer	GCG AGG GTT TAG AAA AAG ATT CAC CAA ATA T	
Sat_250	Forward primer	GCG GTT TTT GCT TTA GGA CAT TTT GAT A	$(AT)_{19}$
	Reverse primer	GCG TTG GGT ACA ACA TAT AAT ATT TTG GA	
Sat_260	Forward primer	GCG CCG TTA GTT GTC GAG GTG TCA ACC	$(AT)_{25}$
	Reverse primer	GCG TCG GTG ATT AAA AAT AAG TAT CAA AG	
Sat_299	Forward primer	GCG ACA AGG CAC TCA CAT CTC TTC TC	$(AT)_{23}$
	Reverse primer	GCG CTA CCC ATA ACA AAA AGT TCA AAT C	
Sat_237	Forward primer	GCG TTC CTG AAT TTT TCT TCT TTG TTG TA	$(AT)_{25}$
	Reverse primer	GCG TTT TGG TTT ACT TGC TAT TTA TCC T	
Satt522	Forward primer	GCG AAA CTG CCT AGG TTA AAA	$(ATT)_{16}$
	Reverse primer	TTA GGC GAA ATC AAC AAT	
Satt531	Forward primer	GCA TGC AAC TGA GGG AGC AGA T	$(ATT)_{14}$
	Reverse primer	GCC ACA AAT TAT GCA GAA TAT A	
Satt593	Forward primer	GCG GGG TTG TTG ATC TAT AAT GTA A	$(ATT)_{15}$
	Reverse primer	GCG GGT TTG GAT TTT ATA ATG TGA T	
Satt597	Forward primer	GCT GCA GCG TGT CTG TAG TAT	$(ATT)_{13}$
	Reverse primer	CGA GGC ACA ACC ATC ACC AC	
Satt631	Forward primer	GGT AGA TCC AGG AGC TTG AGT CAG	$(ATT)_{21}$
	Reverse primer	GCG CAT CTC ACT GCA CTT GAT TTT	
Sat_366	Forward primer	GCG GCA CAA GAA CAG AGG AAA CTA TT	$(AT)_{8}$
	Reverse primer	GCG GAC ATG GTA CAT CTA TAT ATT ACG AGT ATT	
BE475343	Forward primer	GCG TCT CCC TGT CTC TC	$(GT)_{11}$
	Reverse primer	GCG AGC TTA AAA CAA TCA TC	